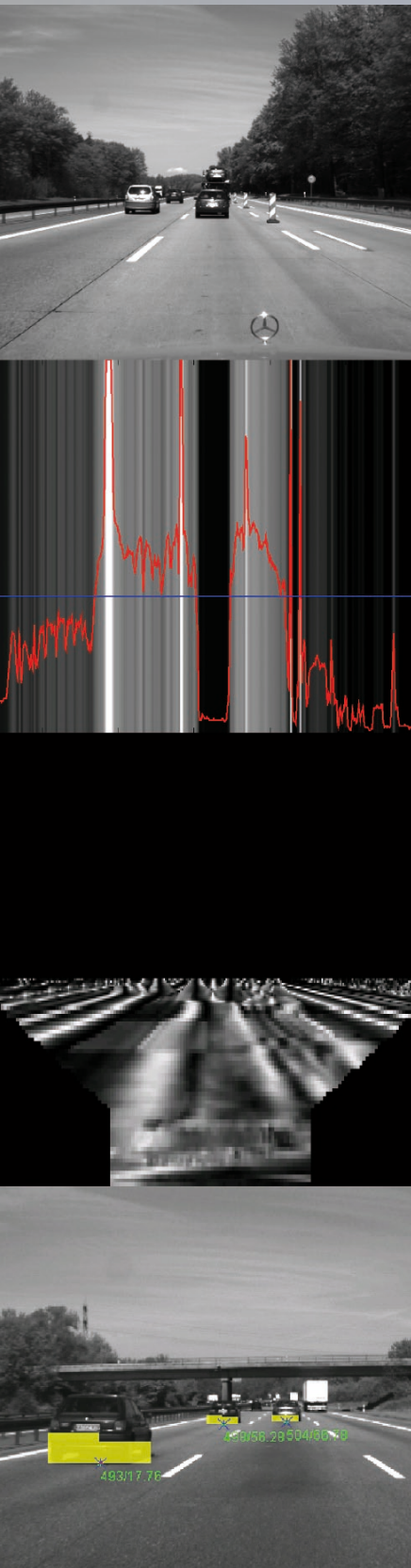


Institut für Mess- und Regelungstechnik
Universität Karlsruhe (TH)
Nr. 007

Dipl.-Ing. Christian Hoffmann

Fahrzeugdetektion durch Fusion monoskopischer Videomerkmale



Christian Hoffmann

Fahrzeugdetektion durch Fusion monoskopischer Videomerkmale

Schriftenreihe
Institut für Mess- und Regelungstechnik,
Universität Karlsruhe (TH)

Band 007

Fahrzeugdetektion durch Fusion monoskopischer Videomerkmale

von
Christian Hoffmann



universitätsverlag karlsruhe

Dissertation, Universität Karlsruhe (TH)
Fakultät für Maschinenbau, 2006

Impressum

Universitätsverlag Karlsruhe
c/o Universitätsbibliothek
Straße am Forum 2
D-76131 Karlsruhe
www.uvka.de



Dieses Werk ist unter folgender Creative Commons-Lizenz
lizenziert: <http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

Universitätsverlag Karlsruhe 2007
Print on Demand

ISSN: 1613-4214
ISBN: 978-3-86644-139-2

Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftlicher Angestellter am Institut für Mess- und Regelungstechnik der Universität Karlsruhe (TH). Herrn Prof. Dr.-Ing. Christoph Stiller danke ich für die Betreuung meiner Arbeit und den mir eingeräumten Freiraum, der für eine wissenschaftliche Arbeit wichtig ist. Unter seiner Leitung konnte ich viele Kenntnisse und Erfahrungen sammeln, die sich für meine weitere Laufbahn als wertvoll erweisen werden.

Herrn Prof. Dr.-Ing. Uwe D. Hanebeck danke ich für die Übernahme des Korreferats und das freundliche Interesse an meiner Arbeit.

Mein besonderer Dank gilt den Herren Dr.-Ing. Thao Dang und Dr.-Ing. Andreas Kapp für die Zusammenarbeit, für die vielen konstruktiven Anregungen in unzähligen interessanten Diskussionen und nicht zuletzt für die sorgfältige Durchsicht des Manuskripts.

Auch den Kollegen am Institut möchte ich für die angenehme Arbeitsatmosphäre danken, die eine wichtige Voraussetzung für die Durchführung einer solchen Arbeit ist. Hier möchte ich auch Herrn Prof. em. Dr.-Ing. Franz Mesch erwähnen, der mit vielen guten Ideen diese Atmosphäre maßgeblich gestaltet hat.

Für die Unterstützung in allen Belangen der Verwaltung und Organisation danke ich den Mitarbeitern des Sekretariats, insbesondere Frau Erna Nagler und Frau Sieglinde Klimesch.

Zu guter Letzt gilt mein besonderer Dank den Werkstätten unter der Leitung von Herrn Manfred Hauser und Herrn Marcus Hoffner. Erst ihre Kompetenz und Kreativität ermöglichten viele Experimente und Untersuchungen.

Karlsruhe, im Dezember 2006

Christian Hoffmann

Für Steffi

Kurzfassung

Moderne Fahrerassistenzsysteme haben das Ziel, den Fahrer zu entlasten und damit das Fahren angenehmer und sicherer zu machen. Dazu ist eine Erfassung des Fahrzeugumfeldes, insbesondere anderer Verkehrsteilnehmer, unerlässlich. Die in aktuellen Systemen eingesetzten Sensoren können alleine jedoch nicht die hohen Anforderungen an die Zuverlässigkeit und die Größe des Erfassungsbereiches erfüllen.

Die vorliegende Arbeit beschreibt ein Verfahren zur Fahrzeugdetektion anhand monoskopischer Videomerkmale. Aus den Bildern einer im Fahrzeug angebrachten Videokamera werden charakteristische Merkmale extrahiert, die einen Hinweis auf das Vorhandensein eines Fahrzeuges geben. Ein solches Merkmal ist der Schattenbereich, der unter Personen- und Lastkraftwagen auftritt. Die Auswertung dunkler Bildregionen liefert Informationen über mögliche Objektpositionen. Allerdings ist ein einzelnes Merkmal nicht ausreichend, um eine zuverlässige Detektion zu gewährleisten. Daher werden des Weiteren Bereiche analysiert, die sich durch eine Achsensymmetrie auszeichnen, welche sich gerade bei der Front- und Heckpartie von Fahrzeugen beobachten lässt.

Diese Beobachtungen werden fusioniert und zeitlich verfolgt. Da aufgrund der Natur der Merkmale ein hoher Anteil Fehlhypothesen enthalten ist, kommen *Joint Probabilistic Data Association*-Filter zum Einsatz. Allerdings lässt sich die Vielzahl der zu berücksichtigenden Verkehrssituationen nicht durch ein einziges Bewegungsmodell beschreiben. Daher werden mehrere Bewegungsmodelle in einem *Interacting Multiple Model*-Verfahren verknüpft.

Experimentelle Ergebnisse belegen, dass erst diese Kombination der Verfahren eine stabile Verfolgung von Fahrzeugen ermöglicht. Außerdem zeigt sich, dass selbst stark rauschbehaftete Sensoren die Zuverlässigkeit des Systems erhöhen können. Das beschriebene Verfahren wurde in einem Versuchsfahrzeug umgesetzt und ist in der Lage, Informationen über Position und Verhalten anderer Verkehrsteilnehmer in Videobildrate bereitzustellen.

Schlagnworte: Fahrerassistenzsysteme – Fahrzeugdetektion – monoskopische Videomerkmale – Schatten – Symmetrie – Fusion – Interacting Multiple Model Filter – Joint Probabilistic Data Association Filter

Abstract

Modern driver assistance systems have to ease the task of driving and thus increase traffic safety. To accomplish this goal, reliable perception of the surrounding environment, especially of other road users, is indispensable. However, sensors employed in current systems cannot satisfy the high demands on reliability and field of view.

This work presents a method for detecting vehicles using monoscopic video features. From the images of a video camera mounted in the car, characteristic cues are extracted that indicate the presence of a vehicle. The shadow that occurs below cars and trucks is such a cue. By analysing dark image regions, information about possible object locations is obtained. However, a single feature is not sufficient for reliable detection. Thus, by searching for regions with high symmetry about a vertical axis, additional observations are obtained as symmetry is a property of the front or rear of vehicles.

These observations are fused and tracked over time. As these features exhibit a high rate of false positives, *Joint Probabilistic Data Association* filters are employed. Due to the fact that the sensor has to cope with varying traffic situations, it is not possible to formulate a single motion model. Thus, several motion models are combined in an *Interacting Multiple Model* approach.

Experimental results prove that only the combination of these approaches yields stable tracking results. Additionally, it is shown that even noisy sensors can improve the reliability of the proposed system. The presented techniques have been implemented in an experimental vehicle and are capable of providing information about position and behaviour of other road users in video frame rate.

Keywords: Driver assistance systems – vehicle detection – monoscopic video features – shadow – symmetry – fusion – Interacting Multiple Model Filter – Joint Probabilistic Data Association Filter

Inhaltsverzeichnis

Symbolverzeichnis	X
1 Einleitung	1
1.1 Fahrerassistenzsysteme	2
1.2 Ziel der Arbeit	3
1.3 Konzeption des Sensors	4
1.3.1 Anforderungen an den Sensor	6
1.3.2 Merkmalsextraktion	7
1.3.3 Fusion	7
1.4 Struktur der Arbeit	8
2 Videobasierte Fahrzeugdetektion	10
2.1 Merkmale in Videosequenzen	11
2.1.1 Stereoskopie	11
2.1.2 Bewegungsbasierte Segmentierung	13
2.1.3 Monoskopische Einzelbildmerkmale	15
2.2 Entfernungsinformation	22
2.3 Schatten	25
2.3.1 Grauwert der Fahrbahn	26
2.3.2 Dunkle Zeilensegmente	27
2.3.3 Gruppieren der Zeilensegmente	29
2.3.4 Anpassen der Berandungen	30
2.3.5 Klassifikation	31
2.4 Symmetrie	37
2.4.1 Initialisieren der Symmetriehypothesen	39
2.4.2 Ausdehnen der Symmetrieachsen	41

2.4.3	Resultierende Symmetriehypothesen	42
3	Fusion	45
3.1	Grundlagen der Informationsfusion	45
3.1.1	Fusionsprinzipien	45
3.1.2	Verfahren zur zeitlichen Verfolgung	48
3.1.3	Assoziation	53
3.2	Detektion und zeitliche Verfolgung von Fahrzeugen	56
3.2.1	Modellierung des Objektverhaltens	56
3.2.2	Eigenschaften der Sensoren	58
3.2.3	Architektur der Sensordatenfusion	60
3.3	Probabilistische Datenassoziation	62
3.3.1	Validierung der Hypothesen	66
3.3.2	Joint Probabilistic Data Association-Filter	68
3.4	Interacting Multiple Model-Filter	74
3.5	Fusion der Daten mehrerer Sensoren	80
3.6	Trackkontrolle	81
3.6.1	Initialisieren von Tracks	81
3.6.2	Entfernen von Tracks	83
3.6.3	Zusammenführen von Tracks	84
4	Aufbau und Implementierung	86
4.1	Ausstattung und Aufbau des Versuchsträgers	86
4.1.1	Sensoren	86
4.1.2	Rechner	88
4.1.3	Kommunikation	88
4.1.4	Spannungsversorgung	88
4.2	Softwarearchitektur	89

5	Experimentelle Ergebnisse	92
5.1	Vergleich verschiedener Konfigurationen	93
5.1.1	Mehrere Sensoren, mehrere Modelle	93
5.1.2	Mehrere Sensoren, ein Modell	96
5.1.3	Ein Sensor, mehrere Modelle	98
5.2	Untersuchung der Detektionsgüte	98
6	Zusammenfassung	102
	Literaturverzeichnis	105

Symbolverzeichnis

Notationsvereinbarungen

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

Symbole

$\ \cdot\ $	euklidische Norm
\mathbf{a}	Normalenvektor der Hyperebene des Klassifikators
$\hat{\mathbf{a}}$	geschätzter Normalenvektor
\mathbf{A}	Transitionsmatrix
$\text{bel}(x)$	Belief-Funktion für x
B_F	Fahrzeugbreite
$\mathbf{c} = (c_u, c_v)^T$	Bildhauptpunkt
$c(\mathbf{y}_M)$	lineare Diskriminanzfunktion des Klassifikators
$\det(\mathbf{X})$	Determinante der Matrix \mathbf{X}
D_f	Anzahl der falschen Hypothesen
D_t	Anzahl der richtig erkannten Fahrzeuge
DR	Detektionsrate
$E(f)$	gerader Funktionsanteil von f
$E_n(f)$	mittelwertfreier gerader Funktionsanteil von f
$\mathbf{f} = (f_u, f_v)^T$	Brennweiten
FAR	Fehlalarmrate
$g(\mathbf{p})$	Grauwertbild
g_F	Grauwert der Fahrbahnoberfläche
h	Höhe der Kamera über der Fahrbahn
\mathbf{H}	Beobachtungsmatrix
\mathbf{I}	Einheitsmatrix

\mathbf{K}	Kalman-Gain
l	Grad der Überlappung zweier Zeilensegmente
L	Anzahl der Markierungen
$m(\mathbf{y}^k)$	Anzahl der falschen Klassifizierungen für \mathbf{y}^k
$M(k)$	Anzahl der validierten Beobachtungen zum Zeitschritt k
\mathcal{M}	Bildausschnitt
$O(f)$	ungerader Funktionsanteil von f
$p(x)$	Wahrscheinlichkeitsdichte von x
$\mathbf{p} = (u, v)^T$	Bildpunkt
P_D^t	Detektionswahrscheinlichkeit für Objekt t
$\mathcal{P}_{ik} = \{u_{l,i}, u_{r,k}\}$	Paar seitlicher Berandungspunkte
$P(X)$	Wahrscheinlichkeit für Ereignis X
\mathbf{P}	Kovarianzmatrix des Systemzustandes \mathbf{x}
\mathbf{Q}	Kovarianzmatrix des Prozessrauschens \mathbf{w}
$R_t = n$	Ereignis „Objekt k ist durch Modell n korrekt beschrieben“
\mathbf{R}	Kovarianzmatrix einer Beobachtung \mathbf{z}
S_A	Symmetriemaß
u_l	linke Berandung
u_r	rechte Berandung
u_S	u -Koordinate einer Symmetrieachse
v_u	untere Berandung
\mathbf{v}	Sensorrauschen
\mathbf{V}	Jacobi-Matrix
V	Messvolumen
$\text{Var}\{x\}$	Varianz von x
w_{ik}	Distanz der Berandungspunkte $u_{l,i}$ und $u_{r,k}$
w	Breite des untersuchten Bereiches
\mathbf{w}	Prozessrauschen
w_i	Gewichte des Klassifikators
\mathbf{x}^T	Transposition des Vektors \mathbf{x}
\bar{x}	empirischer Mittelwert von x
\mathbf{x}	Systemzustand
$\hat{\mathbf{x}}$	Schätzwert von \mathbf{x} , geschätzter Systemzustand

\mathbf{x}^-	Prädiktion von \mathbf{x} , prädizierter Systemzustand
\mathbf{x}^t	alle \mathbf{x} bis zum Zeitpunkt t
\mathbf{y}_M	Merkmalsvektor
\mathbf{y}	erweiterter Merkmalsvektor
\mathbf{y}^k	k -tes Element der Sequenz falsch klassifizierter Merkmale
$\mathbf{z} = (X, Y, Z)^T$	Punkt im Raum, Beobachtung
$\mathbf{Z}(k)$	Menge der Beobachtungen für Zeitschritt k
\mathbf{Z}^k	Menge der Beobachtungen bis einschließlich Zeitschritt k
α	positiver Skalierungsfaktor
β_m	Wahrscheinlichkeit, dass Beobachtung m wahr ist
$\delta_t(\Theta)$	Indikator: in Θ wurde eine Beobachtung mit Track t assoziiert
φ	Winkel der Fahrbahnoberfläche
Ω	Validierungsmatrix
$\hat{\Omega}$	Matrixdarstellung eines möglichen Verbundereignisses Θ
$\phi(\Theta)$	Anzahl der Beobachtungen, denen in Θ kein Track zugeordnet wurde
σ^2	Varianz
$\tau_m(\Theta)$	Indikator: Beobachtung m wurde in Θ mit einem Track assoziiert
Θ_m	Ereignis „Beobachtung m wahr“
Θ_{mt}	Ereignis „Beobachtung m von Track t ausgelöst“
Θ	Verbundereignis aus einer möglichen Kombination von Θ_{mt}

1 Einleitung

Allein im Gebiet der Europäischen Union kamen im Jahr 2005 bei Verkehrsunfällen 41 600 Menschen ums Leben. Darüber hinaus gab es 1,9 Mio. Verletzte. Zwar sinken diese Zahlen in den letzten Jahren kontinuierlich (2001 waren es noch 50 400 Verkehrstote), doch schreitet die Entwicklung in Anbetracht dieser erschreckend hohen Zahlen viel zu langsam voran¹.

Der allein im Jahr 2005 durch Straßenverkehrsunfälle entstandene volkswirtschaftliche Schaden wird auf 200 Mrd. € geschätzt. Dieser jährliche Schaden war 2001 ein weiterer Grund für die Europäische Kommission, das Ziel auszurufen, bis 2010 die Anzahl der Verkehrstoten um 25 000 zu senken, also quasi zu halbieren [Europäische Kommission 2001].

Während dieses ehrgeizige Ziel wohl nicht erreicht werden kann, so signalisiert es doch, wie hoch der Bedarf an Maßnahmen ist, die die Sicherheit aller Verkehrsteilnehmer verbessern. Gleichzeitig markiert der Aufruf der Europäischen Kommission den Start einer ganzen Reihe nationaler und internationaler Projekte, die sich mit dieser Thematik befassen.

Neben infrastrukturellen Maßnahmen gilt die Aufmerksamkeit dabei hauptsächlich den Fahrzeugen. Passive Sicherheitsmaßnahmen, wie z. B. Airbags, Seitenaufprallschutz, Design der Fahrgastzelle etc., wurden in den letzten Jahren konsequent weiterentwickelt und ausgebaut. Doch die Möglichkeiten sind hier begrenzt.

Eine weitere Steigerung der Sicherheit im Straßenverkehr versprechen aktive Eingriffe in die Fahrdynamik, die die Unfallschwere mindern oder Unfälle gänzlich vermeiden sollen. Erste Systeme sind bereits heute weit verbreitet. Neufahrzeuge sind inzwischen mit *ABS* (Anti-Blockier-System) ausgerüstet, das die Kontrollierbarkeit während eines Bremsmanövers verbessert. *ESP* (Elektronisches Stabilitäts-Programm) bremst gezielt einzelne Räder, um ein Schleudern zu verhindern.

Doch es ist abzusehen, dass eine signifikante Verbesserung erst durch Systeme erzielt wird, die nicht nur auf eine kritische Situation reagieren, sondern eingreifen, bevor es zu dieser kommt. Das bedeutet jedoch, dass sich die Wahrnehmungsfähigkeiten dieser Systeme nicht auf das eigene Fahrzeug beschränken können, sondern die Fahrzeugumgebung einbezogen werden muss. Damit wird die Erfassung von

¹Quelle: *European Road Safety Charter* (<http://ec.europa.eu/transport/roadsafety/charter.htm>)

Umgebungsinformation zu einem zentralen Bestandteil moderner Fahrerassistenzsysteme.

1.1 Fahrerassistenzsysteme

Die Bedeutung der Umfelderkennung spiegelt sich auch in Forschungsprojekten und Produkten wieder. Aus dem europäischen Forschungsprojekt PROMETHEUS (PROgram for European Traffic with Highest Efficiency and Unprecedented Safety) entstand das heute in Fahrzeugen der Oberklasse erhältliche ACC-System (Adaptive Cruise Control). Dabei handelt es sich um ein Komfortsystem zur Längsführung, das den Abstand zu einem vorausfahrenden Fahrzeug bestimmt und konstant hält. Die Regeleingriffe sind dabei moderat, sodass der Fahrer jederzeit durch Übersteuern die Kontrolle übernehmen kann.

Aufgrund der Eigenschaften des verwendeten Radarsensors ist es nur auf Autobahnen sinnvoll einsetzbar. Der eingeschränkte Sichtbereich des Sensors lässt eine Erfassung von Objekten bei stark gekrümmtem Straßenverlauf nicht zu. Zudem können stehende Objekte nicht als Hindernis erkannt werden.

Eine Weiterentwicklung dieses Systems, das ACC für den Ballungsraum, wurde im MoTiV-Projekt (MObilität und Transport im Intermodalen Verkehr) untersucht. Für diese Anwendung wird eine deutliche Erweiterung der Fähigkeiten zur Umfeldwahrnehmung notwendig, die auch benachbarte Fahrspuren mit einschließt. Ein weiteres Ziel des MoTiV-Projektes war die Umsetzung eines Spurwechselassistenten, der den Fahrer bei einem Spurwechsel warnt, falls sich andere Fahrzeuge auf der Zielspur befinden.

Bereits für Lastkraftwagen erhältlich ist ein Spurverlassenswarner. Über einen Videosensor mit angeschlossener Bildverarbeitung wird die Position des Fahrzeuges in der Fahrspur ermittelt. Kommt der Fahrer von dieser ab, so wird er akustisch oder haptisch gewarnt. Erweiterte Warnstrategien, die auch die Absicht des Fahrers berücksichtigen, sind Gegenstand aktueller Forschungen [Schmitz 2005].

Ein System, das einige Aspekte zukünftiger Fahrerassistenzsysteme schon heute demonstriert, ist das in Japan erhältliche CMS (Collision Mitigation Brake System). Es übernimmt nicht nur eine Komfort- und Warnfunktion, sondern ergreift bei einer drohenden Kollision Maßnahmen zur Minderung der Unfallschwere. So werden das Fahrzeug aktiv abgebremst und die Sicherheitsgurte gestrafft.

Während all diese Systeme auf bestimmte Szenarien spezialisiert sind, werden zukünftige Systeme immer komplexere Aufgaben übernehmen. Damit wachsen auch

die Anforderungen an die Umfeldwahrnehmung, sowohl bezüglich der Zuverlässigkeit und Genauigkeit als auch der Fülle der gelieferten Information.

1.2 Ziel der Arbeit

Um eine umfassende und zuverlässige Beschreibung des Fahrzeugumfeldes zu erhalten, ist ein einzelner Sensor nicht mehr ausreichend. Erst die Fusion der Daten verschiedener Sensoren kann hier zum Ziel führen.

Entsprechende Ansätze sind weit verbreitet. So werden häufig Radar- mit Videosensoren kombiniert. Im CHAUFFEUR-Projekt wurden ein Radarsensor zur Entfernungsmessung und eine videobasierte Fahrspurdetektion kombiniert, um eine verbesserte Abstandsregelung zu realisieren [Fritz u. a. 2004].

Einen ähnlichen Ansatz verwenden Langer und Jochem allgemein zur Erfassung von Hindernissen [Langer u. Jochem 1996]. Objekte werden auch hier mittels eines Radarsensors verfolgt. Ob es sich um ein relevantes Hindernis handelt, entscheidet eine Spurerkennung anhand von Bilddaten.

Während diese Verfahren eine strikte Aufgabentrennung der Sensoren vorsehen, existieren auch Ansätze zur kombinierten Detektion und Verfolgung. Kato et al. [Kato u. a. 2002] erfassen Objekte zunächst mit einem Radarsensor und bestimmen die entsprechenden Bildausschnitte. Die detektierten Objekte werden dann mit Methoden der Bildverarbeitung verfolgt.

Ein auf dem japanischen Markt erhältliches System wird in [Tokoro u. a. 2004] beschrieben. Ein Radarsensor wird hier durch einen Videosensor ergänzt. Anhand von Videoinformationen werden die Radardetektionen verifiziert. Zudem wird die Breite des erfassten Objektes ermittelt, um die Schwere einer Kollision abzuschätzen.

Der Vorteil der Berücksichtigung mehrerer Sensoren ergibt sich aus der Fusion diversitärer, sich ergänzender Messdaten. Aber auch eine einzelne Videokamera bietet ein breites Spektrum diversitärer Information. Diesen Ansatz verfolgen z. B. die Autoren in [ten Kate u. a. 2004], die die Merkmale Schatten, Symmetrie und Entropie analysieren, um Fahrzeuge zu detektieren.

Betrachtet man die vorgestellten Verfahren, so erkennt man eine gemeinsame Vorgehensweise. In einem ersten Schritt werden Hypothesen generiert, die im zweiten Schritt validiert und verfolgt werden. Die Berücksichtigung mehrerer Sensoren erfolgt fast ausschließlich erst im zweiten Schritt. Die Generierung der Hypothesen basiert jedoch auf den Messungen eines einzelnen Sensors bzw. eines einzelnen Videomerkmals.

Zweck dieser zweistufigen Vorgehensweise ist, die für die Bildverarbeitung benötigte Rechenleistung zu reduzieren, um die Echtzeitfähigkeit zu gewährleisten. Das führt jedoch dazu, dass die Detektionsgüte des Systems mit den Fähigkeiten eines einzelnen Sensors steht und fällt.

Dabei stellen heutige und zukünftige Plattformen deutlich mehr Ressourcen zur Verfügung. Prinzipiell ist es also möglich, diese Struktur aufzubrechen. Dies wird auch bereits von mehreren Gruppen demonstriert.

Ninomiya et al. [Ninomiya u. a. 1995] verwenden Kanteninformation, 2D-Verschiebungen und Stereoskopie zur Umfelderkennung. Dazu entwickelten sie spezielle Hardwarebausteine, die die Teilaufgaben übernehmen.

Auch Bensrhair et al. [Bensrhair u. a. 2001] nutzen eine Kombination aus monokularen Merkmalen und Stereoskopie zur Detektion von Fahrzeugen. Hier wurde der Stereoansatz zur Verfeinerung der Ergebnisse aus der monokularen Objekterkennung genutzt.

Diese Ansätze sind jedoch meist als geschlossene Einheiten für einen konkreten Anwendungsfall konzipiert. Sie kombinieren daher spezielle Eigenschaften der verwendeten Merkmale, um bei möglichst geringem Aufwand die Anforderungen dieser Anwendung zu erfüllen. Das macht sie allerdings ungeeignet für die Integration in einem übergeordneten Gesamtsystem zur Umfelderkennung.

Um diese Lücke zu füllen, wurde im Rahmen der vorliegenden Arbeit ein videobasierter Sensor zur Detektion vorausfahrender und entgegenkommender Fahrzeuge entwickelt, der die vorhandene diversitäre Information komplett nutzen kann. Hierbei war die Entwicklung eines erweiterbaren Konzeptes ein erklärtes Ziel, um einfach zusätzliche Merkmale und sogar andere Sensoren berücksichtigen zu können. Diese Anforderungen führten letztendlich auch zu einer Verallgemeinerung und Weiterentwicklung bestehender Verfahren zur Merkmalsfusion.

1.3 Konzeption des Sensors

Zu einer umfassenden Beschreibung der Fahrzeugumgebung gehören viele Teilaspekte. Hier sind zunächst statische Infrastrukturelemente wie Fahrbahn und Gebäude zu nennen. Hinzu kommen Straßenschilder, Ampeln und andere Objekte, die den Verkehr regeln.

Gerade für Systeme, die die Sicherheit der Fahrzeuginsassen und anderer Verkehrsteilnehmer erhöhen sollen, ist die Erfassung von Fahrzeugen, Fußgängern und Hindernissen von zentraler Bedeutung. Diese Arbeit betrachtet speziell vor-

ausfahrende und entgegenkommende Fahrzeuge, da diese Informationen eine Reihe wichtiger Fahrerassistenzfunktionen ermöglichen bzw. unterstützen.

Bisherige ACC-Systeme sind lediglich als Komfortsysteme ausgelegt. Die ihnen zur Verfügung stehende Information ist für eine sicherheitsrelevante Anwendung, wie z. B. eine automatische Notbremsung, nicht ausreichend. Zudem ist eine vorausschauende Planung von Ausweichmanövern nur möglich, wenn zusätzlich zu den Fahrzeugen auf der eigenen Spur auch noch diejenigen lokalisiert werden können, die sich auf benachbarten Spuren befinden.

Eine Videokamera, deren Sichtbereich sich mit dem des Fahrers deckt, kann hierbei wertvolle Informationen liefern. Darüber hinaus ist abzusehen, dass die nächste Generation von Fahrerassistenzsystemen mit eben solchen Kameras ausgestattet sein wird, um Funktionen wie eine unterstützende Querführung zu realisieren. Damit bietet sich dieser Sensor auch zur Detektion von entgegenkommenden und vorausfahrenden Fahrzeugen an.

Die Auswertung monokularer Einzelbildmerkmale erlaubt es, Fahrzeuge selbst in größerer Entfernung zu erfassen. Gleichzeitig sind solche Verfahren mit vertretbarem Rechenaufwand umsetzbar. Jedoch erst die Kombination mehrerer Merkmale kann die gewünschte Zuverlässigkeit gewährleisten. Daher ist die Integration in ein System erforderlich, das die Fusion der so erhaltenen Daten bewältigt.

Bild 1.1 veranschaulicht das zugrunde liegende Gesamtkonzept. Aus den Bildern der nach vorne gerichteten Kamera gewinnen die Merkmalsextraktoren Informationen über Fahrzeugpositionen und stellen diese als Beobachtungen zur Verfügung. Die untersuchten Merkmale sollen möglichst diversitär sein, d. h. verschiedene Eigenschaften von Fahrzeugen berücksichtigen.

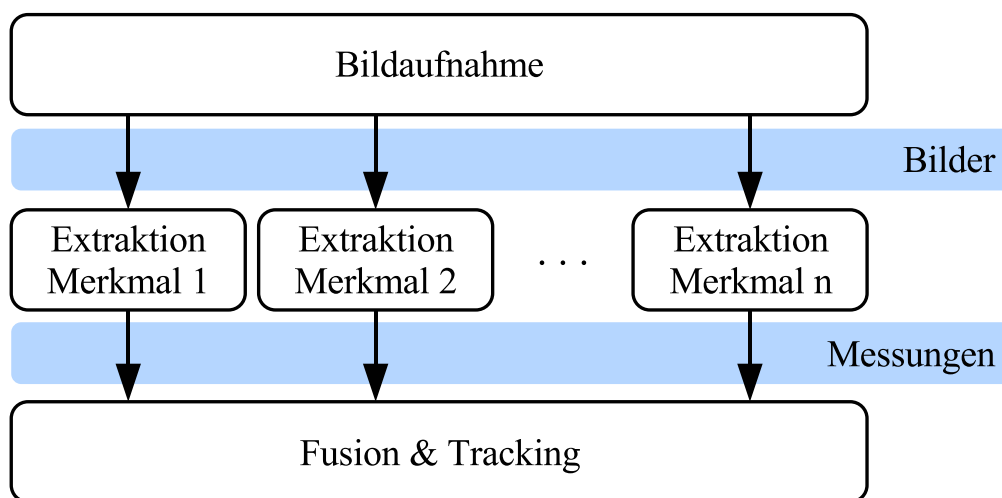


Bild 1.1: Elemente des Gesamtsystems.

Damit wird sichergestellt, dass eine zuverlässige Erkennung auch dann gewährleistet ist, wenn nicht alle Eigenschaften ausreichend ausgeprägt sind.

Aus Sicht der anschließenden Weiterverarbeitung handelt es sich bei den Merkmalsextraktoren um einzelne, unabhängige Sensoren. Aufgabe dieser Weiterverarbeitung ist es, aus den einzelnen Beobachtungen Objektbeschreibungen abzuleiten. Mit immer weiteren Messungen werden diese verfeinert. Damit ergibt sich eine zeitliche Verfolgung der Objekte, die weitere Aussagen möglich macht. So kann die Geschwindigkeit bestimmt und eine zuverlässigere Unterscheidung zwischen wirklichen Objekten und durch Rauschen verursachten Geisterobjekten vorgenommen werden. Indem die Beobachtungen aller Sensoren berücksichtigt werden, findet bei der zeitlichen Verfolgung auch die Fusion der einzelnen Merkmale statt.

1.3.1 Anforderungen an den Sensor

Eine mögliche Anwendung des vorgestellten Systems liegt in der Verbesserung bekannter ACC-Systeme. So kann es den heute verwendeten Radarsensor unterstützen oder sogar ersetzen.

Dafür ist eine zuverlässige Schätzung des Abstandes zum vorausfahrenden Fahrzeug notwendig. Zusätzlich muss beobachtet werden, ob sich dieses auf den Betrachter zu bewegt oder von diesem entfernt.

Auch zur Realisierung zukünftiger Sicherheitsfunktionen wie der automatischen Notbremsung oder autonomer Ausweichmanöver kann ein solches videobasiertes Sensorsystem einen Beitrag leisten. Für diese Anwendungen ist vor allem die sog. *Time To Collision* (TTC), die verbleibende Zeit bis zu einem Zusammenstoß, relevant. Sind der Abstand und die Relativgeschwindigkeit bekannt, so kann die TTC einfach bestimmt werden (vgl. Bild 1.2). Unter der Annahme, dass sich alle Beteiligten mit konstanter Geschwindigkeit bewegen, kann daraus der Zeitpunkt des Aufpralls ermittelt werden. Da dies im Allgemeinen nicht zutrifft, handelt es sich dabei um eine zeitlich veränderliche Schätzung.

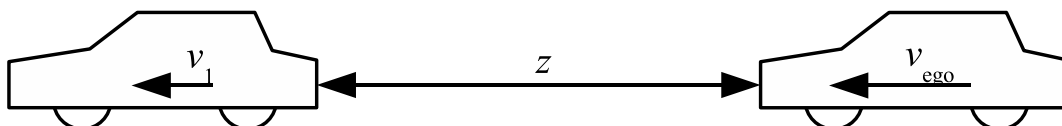


Bild 1.2: Abstand z und Relativgeschwindigkeit $v_1 - v_{ego}$ müssen zur Bestimmung der *Time To Collision* bekannt sein.

Andere Einsatzfelder, die z. B. eine Verbindung von Kartendaten mit bordeigener Sensorik oder die Kommunikation zwischen Fahrzeugen vorsehen, benötigen absolute Positions- und Geschwindigkeitsinformationen. Da in diesen Konfigurationen jedoch auch Sensoren zum Ermitteln der Eigenbewegung vorausgesetzt werden können, können diese einfach aus den relativen Größen errechnet werden.

Der in dieser Arbeit beschriebene Sensor wurde daher auf die Messung der Relativgeschwindigkeit und -position ausgelegt. Dabei wurde auf eine Erfassung von Fahrzeugen auch in großer Entfernung, d. h. bis etwa 100 m, Wert gelegt, um eine frühe Reaktion zu ermöglichen.

1.3.2 Merkmalsextraktion

Bei allen monokularen Einzelbildmerkmalen handelt es sich um schwache Merkmale, d. h. eine zuverlässige Erfassung von Fahrzeugen unter variierenden Bedingungen ist anhand eines einzelnen Merkmals nicht möglich. Erst die Kombination mehrerer Merkmale kann diese Anforderungen erfüllen. Die aus den Bilddaten zu extrahierenden Merkmale sollten möglichst diversitär sein, um einen gegen merkmalspezifische Störeinflüsse robusten Sensor zu gewährleisten und eine hohe Güte der Detektionen zu sichern.

Zwei Bildmerkmale wurden im Rahmen dieser Arbeit betrachtet: Schatten unter Fahrzeugen und Achsensymmetrie der Front- oder Heckansicht. Während ersteres Schwächen zeigt, sobald die Fahrbahn durch Regen dunkel erscheint und spiegelt, so ist die Symmetrieeigenschaft davon nicht betroffen. Diese wird jedoch durch starkes, seitlich einfallendes Licht gestört, was wiederum kaum Auswirkung auf den Fahrzeugschatten hat. Weitere Merkmale können ohne Eingriff in die vorhandene Struktur einfach hinzugefügt werden, um eine weitere Verbesserung der Fahrzeugdetektion zu erzielen.

1.3.3 Fusion

Aufgabe der Fusion ist es, die von den einzelnen Sensoren aufgenommenen Messungen zu vereinigen, um eine exaktere und zuverlässigere Schätzung der Objektparameter zu erhalten. Gleichzeitig werden diese zeitlich verfolgt, was zu einer weiteren Verbesserung führt.

Die Elemente der in dieser Arbeit entwickelten Vorgehensweise sind in Bild 1.3 skizziert. In einem ersten Schritt muss geprüft werden, ob eine neue Messung einem schon erfassten Objekt zuzuordnen ist. Diese Entscheidung lässt sich oft

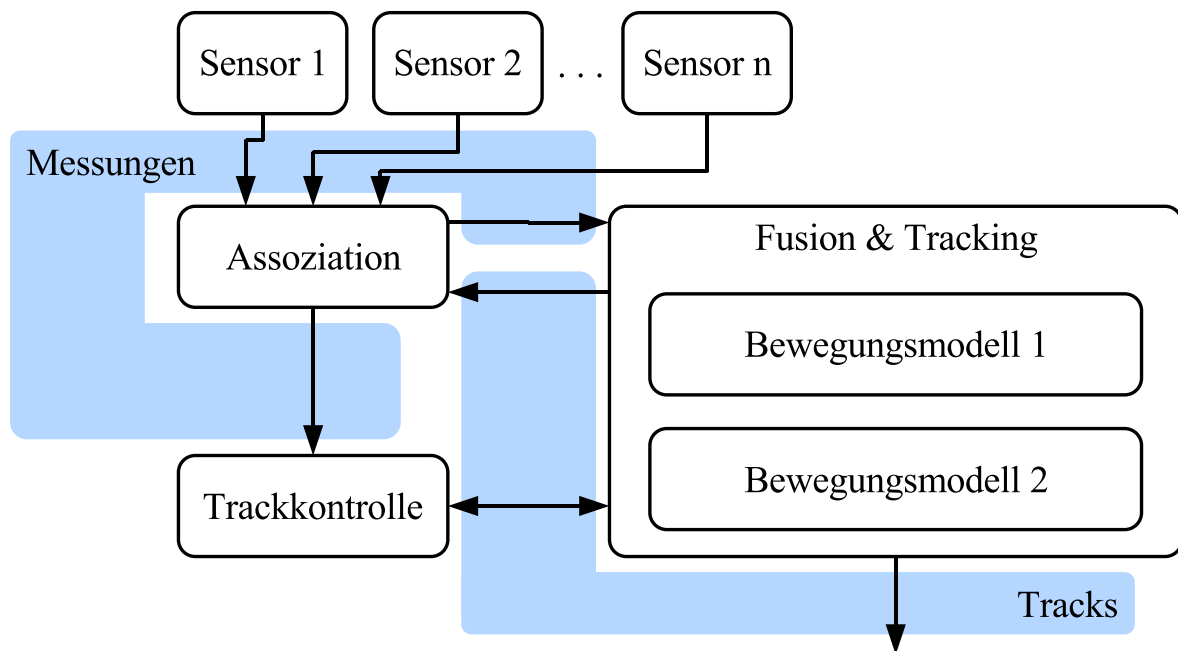


Bild 1.3: Überblick über die Fusion

jedoch nicht eindeutig treffen. Die Assoziationsstufe muss dies entsprechend berücksichtigen.

Das Verhalten erfasster Fahrzeuge wird durch dynamische Modelle beschrieben. Da nicht alle auftretenden Situationen mit einem umfassenden Modell abgedeckt werden können, werden mehrere parallel verfolgt. Dies führt entsprechend zu mehreren Schätzungen des Systemzustandes, die jeweils ein eigenes Systemmodell voraussetzen. Eine übergeordnete Instanz ermittelt daraus dann eine konsistente Schätzung der Objektparameter.

Die Trackkontrolle ist dafür zuständig, dass neue Objektbeschreibungen, sog. Tracks, initialisiert und andere, die z. B. den Erfassungsbereich des Sensors verlassen haben, entfernt werden. Sie stellt damit ein Kontrollorgan dar, das die Ergebnisse der Assoziation berücksichtigt und die Fusion steuert.

1.4 Struktur der Arbeit

Nach dieser Einleitung behandelt Kapitel 2 die Aspekte der videobasierten Merkmalsextraktion. Zunächst liefert Abschnitt 2.1 einen Überblick über Möglichkeiten, die benötigten Informationen aus Videosequenzen zu gewinnen. Abschnitt 2.2 befasst sich mit dem Problem der Entfernungsbestimmung aus zweidimensionalen

Kamerabildern. Diese ist eine Voraussetzung für die in den Abschnitten 2.3 und 2.4 beschriebenen Verfahren zur Extraktion der Schatten- bzw. Symmetriemerkmale.

Auf das zentrale Thema Fusion geht Kapitel 3 ein. Hier gibt Abschnitt 3.1 einen Überblick über wichtige Grundlagen, bevor in Abschnitt 3.2 die Forderungen an das zu entwickelnde Verfahren aufgestellt werden. Die Assoziation der Messdaten zu bereits erfassten Objekten und deren zeitliche Verfolgung beschreibt Abschnitt 3.3. Abschnitt 3.4 befasst sich mit dem *Interacting Multiple Model* Filter, das zur Kombination der Systemmodelle eingesetzt wird. Danach geht Abschnitt 3.5 auf Aspekte bei der Berücksichtigung mehrerer Sensoren ein. Das Kapitel wird von einer Beschreibung der Trackkontrolle in Abschnitt 3.6 abgeschlossen.

Die entwickelten Verfahren wurden im Versuchsfahrzeug des Instituts für Mess- und Regelungstechnik implementiert. Den Aufbau und die Umsetzung hierfür beschreibt Kapitel 4. Es gliedert sich in einen Abschnitt 4.1, der den Versuchsträger und seine Ausstattung an sich erklärt, und einen Abschnitt 4.2, der die Eigenschaften der verwendeten Software sowie die Module, die den Algorithmus bilden, umreißt.

Mit diesem Versuchsträger wurden zahlreiche Experimente durchgeführt, um sicherzustellen, dass die Algorithmen die gestellten Anforderungen erfüllen. Eine Gegenüberstellung der Ergebnisse bietet Kapitel 5, bevor Kapitel 6 die Arbeit mit einer Zusammenfassung und einem Ausblick abschließt.

2 Videobasierte Fahrzeugdetektion

Eine entscheidende Voraussetzung für eine exakte Beschreibung des Fahrzeugumfeldes ist eine zuverlässige Erfassung der zu detektierenden Objekte. Damit nimmt schon die Wahl des Sensors eine zentrale Rolle ein. Neben Ultraschallsensoren, die aufgrund ihrer geringen Reichweite nur für bestimmte Problemstellungen eingesetzt werden können, sind Radar- und Lidar-Sensoren (Light Detection And Ranging) für Fahrerassistentenanwendungen am weitesten verbreitet. Diese liefern bei geringem Rechenaufwand direkt die benötigten Messdaten. Sie senden in Form elektromagnetischer Wellen ein Signal aus und werten den vom zu beobachtenden Objekt reflektierten Anteil aus, um die gewünschte Information zu erhalten. Unproblematisch ist die Verwendung solcher Sensoren jedoch nicht. So treten bei Radarsensoren Schwierigkeiten auf, wenn sich mehrere Objekte mit geringem Abstand im Erfassungsbereich befinden. Die Funktion von Lidar-Sensoren wiederum wird durch Regen und Schnee stark gestört. Zudem erlauben diese Sensoren über die Bestimmung von Position und Geschwindigkeit hinaus nur eingeschränkt Aussagen über weitere Eigenschaften der beobachteten Objekte.

Hier bieten die in dieser Arbeit betrachteten Kameras eine sinnvolle Alternative. Videosensoren liefern eine Vielzahl verschiedener Informationen. Das spiegelt sich in dem Spektrum der Anwendungen wider, in denen sie zu finden sind. So werden sie außer zur Detektion von Fahrzeugen z. B. zur Erfassung der Fahrbahn [Goldbeck u. Huertgen 1999; Chapuis u. a. 2002; Rotaru u. a. 2004], zur Wahrnehmung von Fußgängern [Broggi u. a. 2000; Curio u. a. 2000] oder zur Erkennung von Verkehrsschildern [Fang u. a. 2003] eingesetzt. Darüber hinaus erlauben sie anhand der Bildinformation eine feine Klassifikation der detektierten Objekte. Allerdings ist hier ein deutlich höherer Aufwand bei der Signalauswertung notwendig, als dies bei Radar oder Lidar der Fall ist. Zudem haben bildgebende Sensoren gerade bei Nebel oder Nacht Nachteile gegenüber z. B. einem Radarsensor. Während dieser durch Aussenden des Signals weitgehend unabhängig von der Umgebungssituation ist, werden die Fähigkeiten der Kamera von der vorherrschenden Beleuchtung beschränkt. Zwar könnte auch hier in Form einer aktiven Beleuchtung der Szene diese Abhängigkeit reduziert werden, für die angestrebte Sichtweite des Sensors ist dies jedoch unpraktikabel.

2.1 Merkmale in Videosequenzen

Bei der Extraktion von Objektinformation aus den Bildern von Videokameras können vielfältige Verfahren zum Einsatz kommen. Einen Überblick über die wichtigsten Verfahren geben Sun et al. [Sun u. a. 2006]. Prinzipiell können dabei drei Klassen unterschieden werden: Stereoskopie, bewegungsbasierte Segmentierung und monoskopische Einzelbildmerkmale. Die folgenden Abschnitte gehen auf diese näher ein.

2.1.1 Stereoskopie

Stereoskopische Verfahren verwenden zwei Kameras, die synchron Bilder aufnehmen. Blicken die Kameras aus verschiedenen Positionen auf die selbe Szene, so kann aus ihren Bildern Information über die räumliche Struktur gewonnen werden [Faugeras 1996].

Dieser Sachverhalt kann an einem einfachen Beispiel verdeutlicht werden: Bild 2.1 zeigt zwei identische Lochkameras, deren optische Achsen parallel sind und den Abstand b haben. Die Bildsensoren liegen zudem in einer Ebene.

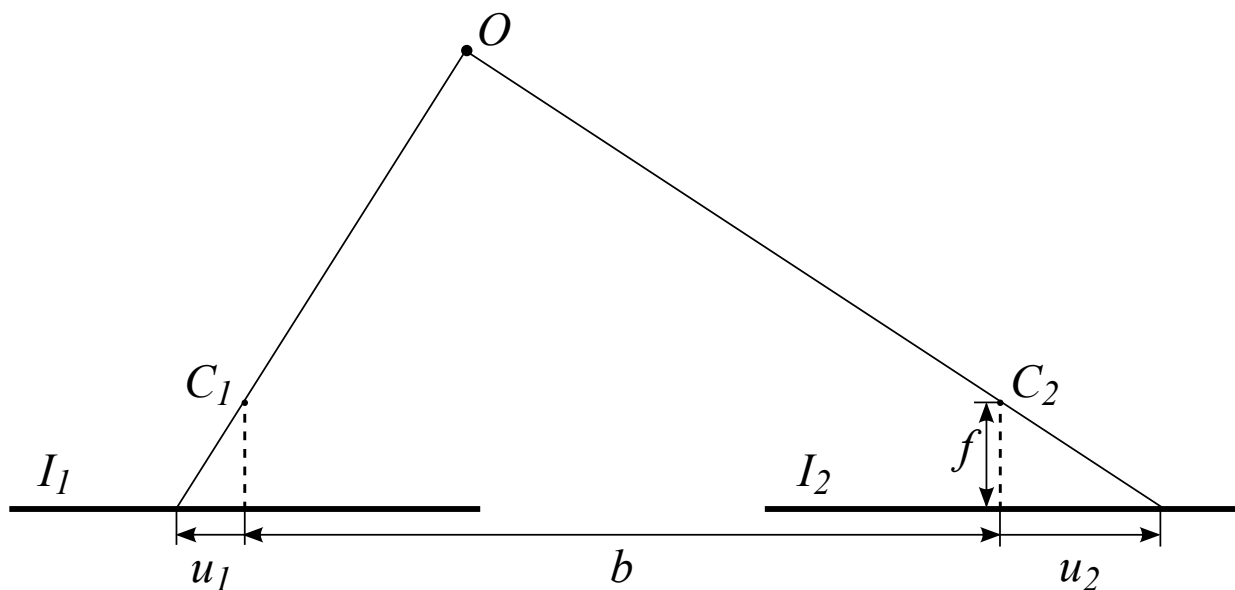


Bild 2.1: Zwei Kameras C_1 und C_2 in Stereoanordnung.

Der Körper O erscheint im Bild I_2 der Kamera C_2 verschoben gegenüber der Position im Bild I_1 der Kamera C_1 . Diese Verschiebung ist die sog. Disparität:

$$d = u_2 - u_1 . \quad (2.1)$$

Aus der Disparität kann bei Kenntnis der Brennweite f der Abstand z des Objektes ermittelt werden:

$$z = \frac{b}{d} \cdot f . \quad (2.2)$$

Im Allgemeinen entspricht die Anordnung der Kameras nicht diesem idealisierten Fall. Um dennoch eine 3D-Rekonstruktion durchführen zu können, müssen die extrinsischen Parameter, d. h. Rotation und Verschiebung der Kameras zueinander, bekannt sein. Auch die Annahme eines Lochkameramodells mit exakt bekannter Brennweite ist in der Realität verletzt, so dass noch zusätzliche intrinsische Parameter hinzukommen. Nur ein Teil dieser Unbekannten kann einmalig vorab bestimmt werden. Erschütterungen und starke Temperaturschwankungen im Fahrzeug machen eine nachführende Kalibrierung notwendig [Dang u. Hoffmann 2006].

Wurde mit einem Stereokamerasystem, für das alle relevanten Parameter bekannt sind, ein Bildpaar aufgenommen, so wird ein Disparitätenbild berechnet, wie es in Bild 2.2 beispielhaft gezeigt ist. Dazu müssen unter Verwendung von Korrelationsverfahren korrespondierende Punkte in den Bildern einander zugeordnet werden. In Bereichen mit geringer Textur ist dies nicht immer eindeutig möglich. Zudem führen Verdeckungen und periodische Strukturen dazu, dass nicht für jeden Bildpunkt eine Korrespondenz gefunden werden kann.

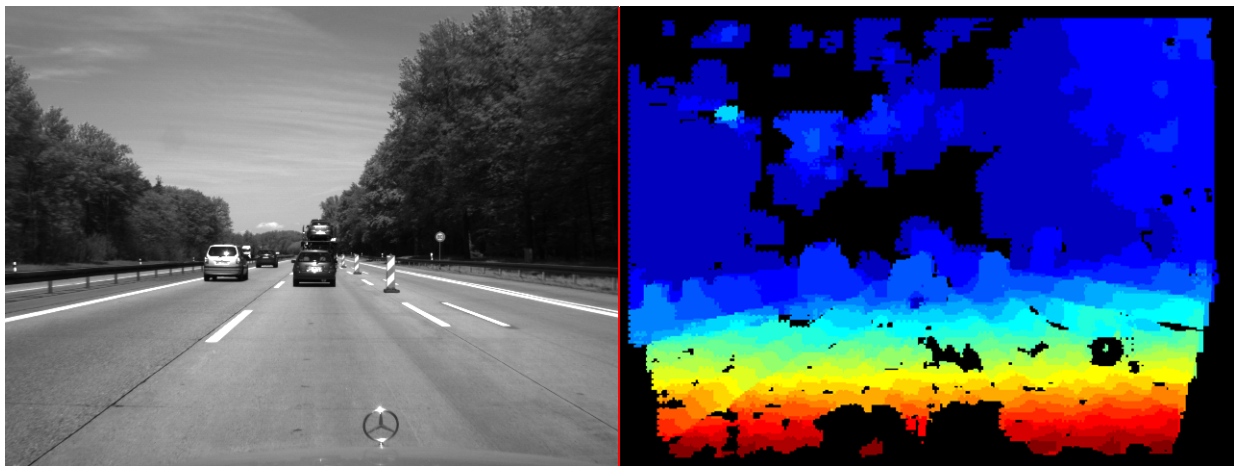


Bild 2.2: Links: Linkes Kamerabild. Rechts: Disparitätenbild in Falschfarbendarstellung (rot: nahe Bereiche, blau: entfernte Bereiche).

Dieser Prozess ist extrem aufwändig, wird jedoch handhabbar, wenn nicht eine vollständige Disparitätenkarte benötigt wird. Wurde ein Objekt schon anderweitig segmentiert, so ist es oft ausreichend, einige wenige Punkte zu untersuchen. Bertozzi et al. [Bertozzi u. a. 2000] verwenden daher Stereoinformation lediglich

zum Bestimmen der Entfernung von Objekten, die mittels monokularer Merkmale erfasst wurden.

Aber auch ohne die Berechnung der Disparitäten kann eine Stereoanordnung zur Objektdetektion verwendet werden. Dies ermöglicht die sog. *Inverse Perspective Mapping*-Technik. Hierbei wird angenommen, dass alle Bildpunkte Punkte in der Fahrbahnebene darstellen. Mit deren Kenntnis wird von beiden Stereobildern die Projektion in die Fahrbahnebene berechnet. Für Punkte, die tatsächlich diese Annahme erfüllen, ergibt sich für das rechte und linke Bild die identische Projektion. Objekte, die nicht in der Fahrbahnebene liegen, erscheinen jedoch in beiden Bildern unterschiedlich verzerrt. Durch Bilden der Differenz beider Projektionen erhält man damit hohe Werte dort, wo sich ein Objekt aus der Fahrbahnebene erhebt [Bertozzi u. Broggi 1998].

Gerade für Szenarien abseits befestigter Straßen bietet ein Stereoansatz die benötigten Informationen für eine Pfadplanung. So kann aus der Disparitäten- eine Höhenkarte errechnet werden, durch die ein optimaler Weg gesucht wird [van der Mark u. a. 2001]. Eine weitere Verbesserung kann durch Einbeziehen von monoskopischen Merkmalen erzielt werden. Einen entsprechenden Ansatz beschreiben Hummel et al. in [Hummel u. a. 2006].

Um aus einer Disparitätenkarte Rückschlüsse auf vorhandene Objekte ziehen zu können, ist eine Segmentierung und Klassifizierung notwendig. Eine umfassende Szenenbeschreibung ist hier nur mit großem Aufwand möglich. Für einzelne Anwendungen ist diese jedoch oft nicht erforderlich. Ein Beispiel hierfür ist der Ansatz von Franke und Kutzbach [Franke u. Kutzbach 1996]. Aus einem Disparitätenhistogramm wird der Abstand zum vorausfahrenden Fahrzeug direkt ermittelt, was eine Objektverfolgung ohne weiterführende Segmentierung erlaubt.

Die Stereorekonstruktion der beobachteten Szene bietet eine Fülle von Informationen, die weit über die reine Detektion von Fahrzeugen hinausgeht. Doch ist der Aufwand, und damit die benötigte Rechenleistung, zur Gewinnung dieser Information hoch. Neben dieser Beschränkung verhindert zur Zeit auch der Preis der benötigten Komponenten, dass in naher Zukunft solche Systeme Einzug in die Serienausstattung neuer Fahrzeuge halten.

2.1.2 Bewegungsbasierte Segmentierung

Die zweite Klasse von Verfahren zur videobasierten Objekterkennung stützt sich auf die durch Eigen- und Objektbewegung verursachten Verschiebungen in konsekutiven Bildern einer Kamera. Um diese Verschiebungsfelder zu bestimmen, wurde eine Vielzahl an Algorithmen entwickelt, die auf verschiedenen Ansätzen ba-

sieren. So kann z. B. durch Korrelieren von Bildausschnitten eine direkte Bestimmung der Verschiebungen vorgenommen werden oder ein Rückschluss von zeitlichen auf räumliche Grauwertgradienten erfolgen. Eine gute Übersicht und einen Vergleich von Methoden bieten Barron et al. [Barron u. a. 1994]. Ein Beispiel für ein solches 2D-Verschiebungsfeld ist in Bild 2.3 dargestellt.

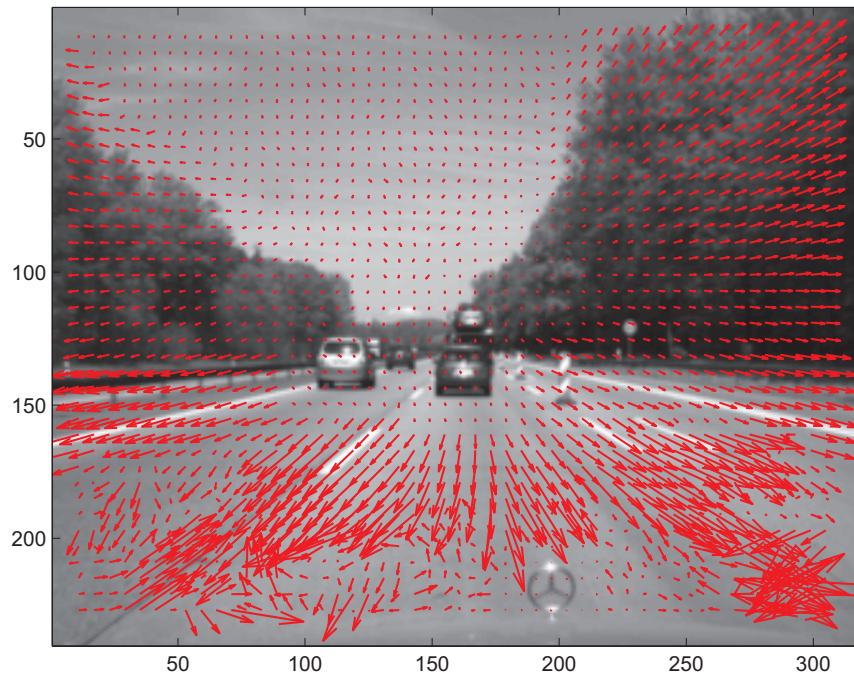


Bild 2.3: Beispiel für die 2D-Verschiebung in aufeinander folgenden Bildern.

Bewegt sich die beobachtende Kamera, so ist prinzipiell eine Rekonstruktion der räumlichen Szene analog zum Stereoansatz möglich. Der Basisabstand des Stereosystems ergibt sich nun aus der Eigenbewegung der Kamera. Eine Schwierigkeit ergibt sich aus dem Sachverhalt, dass Beobachtungs- und Bewegungsrichtung im Allgemeinen zusammenfallen. Damit ist gerade in dieser Richtung eine Vermessung nicht möglich, wodurch entfernte Objekte nur schwer erkannt werden können.

Dennoch zeigen verschiedene Umsetzungen, dass der Ansatz zur Objektdetektion eingesetzt werden kann. So ist in [Lorusso u. De Micheli 1996] ein Verfahren beschrieben, bei dem anhand des Verschiebungsfeldes Objekte segmentiert und verfolgt werden. Die Autoren in [Giachetti u. a. 1994] entwickelten einen Algorithmus, der robust gegen Vibrationen bei der Aufnahme der Bilder ist. Solche Vibrationen sind typisch für sich bewegende Fahrzeuge und erschweren die korrekte Bestimmung des Verschiebungsfeldes. Anhand der Verschiebungen werden dann die Eigenbewegung bestimmt und Objekte segmentiert.

Ist die Kamera nicht in Bewegungsrichtung ausgerichtet, so gilt diese Einschrän-

kung nicht. Daher eignet sich ein Ansatz zur Objektdetektion aus Verschiebungen gut zum Erkennen überholender Fahrzeuge mit einer Kamera, die auf die benachbarten Spuren blickt. Wang et al. entwickelten ein entsprechendes System, bei dem die Kamera schräg nach hinten gerichtet ist [Wang u. a. 2005].

Auch ohne eine komplette Bestimmung der Verschiebungen können Bewegungen zur Detektion von Objekten verwendet werden. Die Autoren in [Betke u. a. 1996] erfassen überholende Fahrzeuge, indem sie die Differenz zweier aufeinander folgender Bilder berechnen und Bereiche mit starken Änderungen segmentieren.

Die Verfahren dieser Klasse eignen sich besonders zur Erfassung von Objekten, die in der näheren Umgebung des Beobachters sind und sich nicht in Fahrtrichtung befinden. Zudem müssen sie eine deutliche Relativbewegung gegenüber der Kamera aufweisen. Zur Detektion vorausfahrender Fahrzeuge in größerer Entfernung sind diese Verfahren daher eher ungeeignet. Hier bieten sich monoskopische Einzelbildmerkmale an.

2.1.3 Monoskopische Einzelbildmerkmale

Monoskopische Einzelbildmerkmale setzen voraus, dass Vorwissen über das zu detektierende Objekt, also ein Objektmodell, vorliegt. Je detaillierter dieses Vorwissen ist, desto zuverlässiger kann das entsprechende Objekt identifiziert werden. Soll eine Menge von Objekten erkannt werden, so muss man die eingebrachte Vorabinformation auf die für diese Menge typischen Eigenschaften beschränken.

Bei der Detektion einer Objektklasse, beispielsweise Fahrzeugen, muss also für jede dieser Untermengen der Klasse ein Modell vorliegen. Je detaillierter diese Modelle sind, desto kleiner wird die durch sie beschriebene Untermenge, desto größer wird also die Zahl der benötigten Modelle. Abgesehen von Aspekten der benötigten Rechenleistung birgt dies die Gefahr, dass die vorhandenen Modelle die Klasse nicht komplett beschreiben und Objekte nicht erfasst werden.

Einfache Modelle beschreiben das Auftreten grundlegender Merkmale. Sie haben den Vorteil, dass sie mit geringem Aufwand schnell zu bestimmen sind. Um die Detektionsgüte zu erhöhen, werden oft Kombinationen einfacher Merkmale gewählt, um komplexere Modelle zu erstellen.

2.1.3.1 Kanten

Ein einfaches Merkmal sind Grauwertgradienten im Bild. Diese werden von Strukturen, die Helligkeitsunterschiede verursachen, hervorgerufen. Kantendetektoren werten die Gradienten aus, um ein Kantenbild zu erstellen.

Die im Kantenbild enthaltene Information ist jedoch lediglich ein Hinweis auf ein Objekt: es handelt sich hierbei um Grauwertkanten, die nicht zwangsläufig mit Objektkanten übereinstimmen müssen. Dennoch kann dieses Merkmal einen Beitrag zur Objektdetektion leisten.

In Abbildung 2.4 ist eine Verkehrsszene und das Ergebnis einer Kantendetektion abgebildet. Schatten und Fahrbahnmarkierungen zeichnen sich deutlich im Kantenbild ab, da sich hier starke Grauwertgradienten finden. Es sind jedoch auch Objekte, wie z.B. die Fahrzeuge, erkennbar.

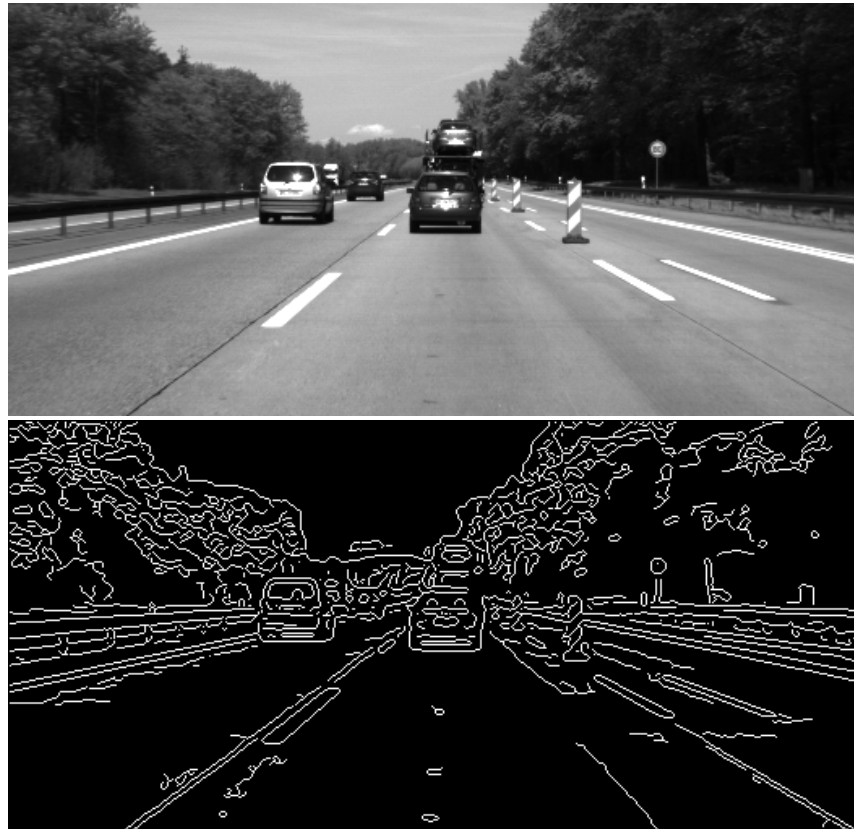


Bild 2.4: Grauwert- und Kantenbild einer Verkehrsszene.

Letztere zeichnen sich aufgrund ihrer Stoßfänger und der Windschutz- bzw. Heckscheibe durch ausgeprägte vertikal und horizontal verlaufende Kanten aus. Diese Eigenschaft nutzen Betke et al. in [Betke u. a. 1996], um Fahrzeughypothesen zu generieren.

Verbreitete Kantendetektoren sind der Sobel-Operator [Jähne 1989] und der Canny-Detektor [Canny 1986]. Doch auch speziell entwickelte Operatoren wie der LOC-Algorithmus (*Local Orientation Coding*) finden Anwendung [Goerick 1994]. Diesen wenden Handmann et al. in [Handmann u. a. 1998] in Kombination mit weiteren Merkmalen wie Entropie und Linien zur Detektion von Fahrzeugen

an.

Ein weiterer Ansatz, der Kanteninformation zur Generierung von Fahrzeughypothesen nutzt, ist in [Srinivasa 2002] beschrieben. Hier werden Radarinformationen mit diesen Hypothesen fusioniert, um eine Steigerung der Detektionsgüte zu erreichen.

2.1.3.2 Form

Die Form ist eine nahezu eindeutige Objekteigenschaft. Ihre Projektion in das zweidimensionale Kamerabild erzeugt dort den Umriss, der für einen Vergleich herangezogen werden kann. Auch anhand der Form kann ein Objekt daher weitgehend klassifiziert werden. Liegt eine geschlossene, segmentierte Kontur vor, so können rotations- und skalierungsinvariante Ähnlichkeitsmaße zwischen einer Vorlage und dem betrachteten Objekt bestimmt werden [Liu u. Srinath 1990].

Da im Fahrzeugumfeld stetig wechselnde Verkehrsszenen betrachtet werden, kann eine vollständige Segmentierung jedoch nicht vorausgesetzt werden. Andererseits ist die Ausrichtung der zu erfassenden Objekte im Regelfall bekannt. Dies ermöglicht es, im Ursprungs- oder einem vorverarbeiteten Bild nach Übereinstimmungen mit einem Satz von Vorlagen zu suchen (vgl. auch Bild 2.5).



Bild 2.5: Form in einem Kantenbild.

Die Distanztransformation ist ein Hilfsmittel zur Bewertung der Ähnlichkeit zweier nicht-deckungsgleicher Kantenbilder [Borgefors 1988]. Da die Korrelation eines Kantenbildes mit einer Vorlage nur bei einer exakten Übereinstimmung zum gewünschten Ziel führt, bestimmt die Distanztransformation zu jedem Bildpunkt den nächsten Kantenpunkt. Die entsprechende Entfernung in Bildkoordinaten wird in der Distanztransformierten an der Position des Bildpunktes eingetragen. Die Korrelation zwischen der Transformierten und einer Vorlage hat dann an der Stelle

der besten Übereinstimmung ein Minimum. Diese Vorgehensweise findet auch im Bereich der Umfelderkennung für Fahrerassistenzsysteme Anwendung. So setzen Gavrilu und Philomin sie zur Erkennung von Fußgängern und Verkehrszeichen ein [Gavrilu u. Philomin 1999].

Auch zur Lokalisierung von Fahrzeugen gibt es Form-basierte Ansätze. Hilario et al. entwickeln in [Hilario u. a. 2005] einen Energieterm, der unter anderem den Abstand von Kantenpunkten zu den entsprechenden Punkten einer Vorlage sowie Gradienteninformation berücksichtigt. Minimieren des Energieterms liefert die gesuchten Fahrzeugpositionen.

Eine Schwierigkeit dieser Verfahren liegt darin, dass der Umriss von der Orientierung des zu detektierenden Objekts abhängt. Daher verwenden Fleischer et al. dreidimensionale Modelle, deren Projektion in das Kamerabild mit Kanteninformation verglichen wird [Fleischer u. a. 2002]. Dies ermöglicht es, zusätzlich zur Position die Ausrichtung des Objekts zu bestimmen. Allerdings ist diese Vorgehensweise wiederum mit einer erhöhten Rechenkomplexität verbunden.

Ansätze, die das Merkmal Form auswerten, sind im Allgemeinen auf eine umfassende Datenbasis angewiesen, die für alle zu erkennenden Objekte eine entsprechende Vorlage enthält. Dazu kommen, je nach Algorithmus, für jede Vorlage mehrere Skalierungsstufen, da die Entfernung des Objektes, und damit seine Größe im Bild, nicht vorab bekannt sind. Dies führt zu einem hohen Rechenaufwand. Ein Vorteil dieser Verfahren ist aber, dass sie, abhängig von den zur Verfügung stehenden Vorlagen, verschiedene Objekttypen detektieren und klassifizieren können.

2.1.3.3 Farbe

Zwar ist Farbe kein Merkmal, das die Detektion von Fahrzeugen direkt unterstützt, da es nicht charakteristisch für diese Objekte ist, dennoch erweist sich Farbinformation als wertvoll für Fahrerassistenzanwendungen. So kann sie z. B. bei der Erkennung von Rücklichtern und Scheinwerfern bei Dunkelheit einen Beitrag leisten [Cabani u. a. 2005].

Gerade für die Erfassung von Infrastrukturelementen, wie z.B. Schildern, Ampeln, der Fahrbahn und Fahrspuren, kann sie ein wichtiges Merkmal sein [He u. a. 2004]. Damit dient Farbinformation wiederum indirekt der Segmentierung relevanter Bereiche, was eine starke Hilfe bei der Detektion von Fahrzeugen mit weiteren Merkmalen ist.

2.1.3.4 Textur

Obgleich auf den ersten Blick kaum eine allgemeine Aussage über die Textur von Fahrzeugen getroffen werden kann, die eine Detektion erlaubt, so verursacht der typische Fahrzeugaufbau doch charakteristische Merkmale. Zwar erlauben diese keine exakte Lokalisation, werden jedoch zur Segmentierung relevanter Bildbereiche angewandt.

Ein Texturmerkmal ist die lokale Entropie, ein Informationsmaß, das auf der ursprünglich von Shannon entwickelten Informationstheorie basiert [Shannon 1948]. In der Bildverarbeitung beschreibt die lokale Entropie $H(u, v)$ den Informationsgehalt eines Bereiches [Kalinke u. von Seelen 1996]. Sie bestimmt sich zu

$$H(u, v) = - \sum_{k=1}^K p(g(u, v) = k) \log(k). \quad (2.3)$$

Hierbei ist g das Grauwertsignal, das die Zustände $k = 1, \dots, K$ annehmen kann. Um die lokale Entropie für einen Bildpunkt (u, v) zu bestimmen, wird die Verteilung der Grauwerte $p(g(u, v))$ durch das Histogramm einer kleinen Umgebung dieses Punktes angenähert. Bild 2.6 stellt die lokale Entropie der Szene aus Bild 2.4 dar.

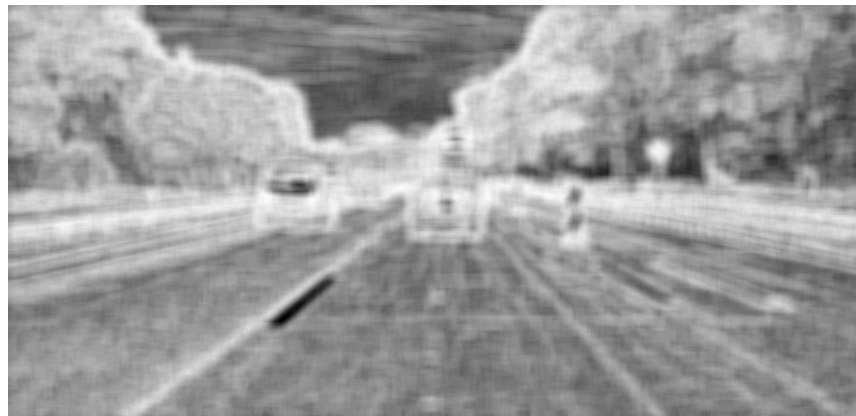


Bild 2.6: Lokale Entropie, berechnet in einer 9×9 -Umgebung. Helle Bereiche weisen eine hohe, dunkle Bereiche eine niedrige Entropie auf.

Eine mögliche Anwendung zeigen die Autoren in [ten Kate u. a. 2004]. Sie verwenden die Merkmale Schatten, Symmetrie und lokale Entropie zur Fahrzeugdetektion.

2.1.3.5 Schatten

Ein mögliches Modell zur Detektion von Fahrzeugen anhand monoskopischer Einzelbildmerkmale ergibt sich aus der Beobachtung, dass die zu detektierenden Fahrzeuge einen Schatten auf die Fahrbahnoberfläche werfen. Dabei wird lediglich der dunkle Bereich zwischen den Rädern berücksichtigt, welcher durch den diffusen Beleuchtungsanteil entsteht (siehe Abbildung 2.7). Der Schlagschatten, der bei schräg einfallendem Sonnenlicht seitlich des Objektes auftritt, wird aufgrund der Abhängigkeit von der schwer zu erfassenden Beleuchtungskonstellation nicht als Merkmal betrachtet. Er stellt jedoch eine Störquelle dar, die ein robustes Detektionsverfahren nötig macht.

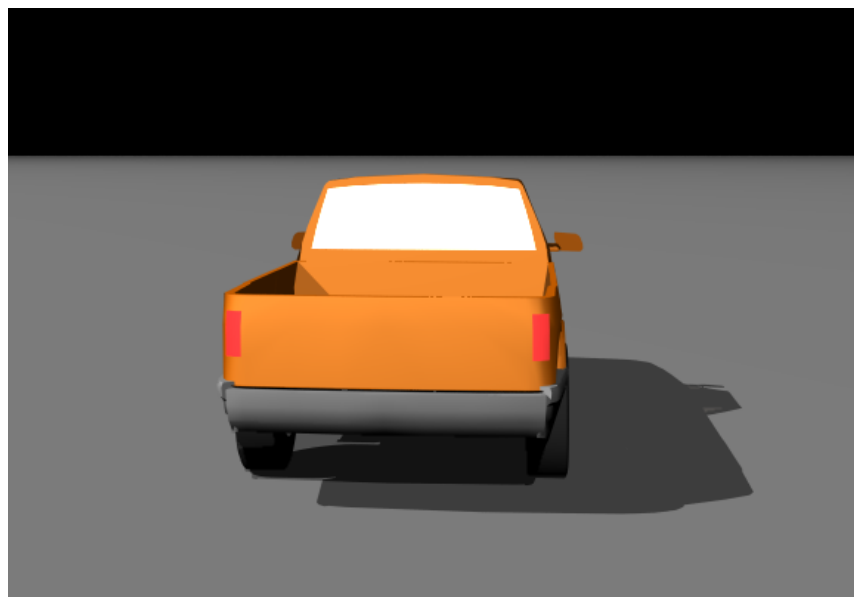


Bild 2.7: Dunkler Bereich unter dem Fahrzeug und Schlagschatten.

Gerade weil dies ein sehr allgemeines Modell ist, wurde es schon von verschiedenen Gruppen verwendet. Denn aufgrund dieser Allgemeinheit ermöglicht es die Detektion vieler verschiedener Fahrzeugtypen. Zudem ist es mit vertretbarem Rechenaufwand zu implementieren, was für Echtzeitanwendungen entscheidend ist.

So finden sich schon früh entsprechende Untersuchungen in der Literatur. Mori und Charkari untersuchten in [Mori u. Charkari 1993] die Grauwertverteilung unter Fahrzeugen und der Straßenoberfläche für verschiedene Beleuchtungsbedingungen und stellten fest, dass die Schattenregion stets deutlich dunkler als die Fahrbahn ist.

Erste Anwendungen prüfen anhand des Schattenmerkmals, ob in einer bestimmten Bildregion ein Fahrzeug vorhanden ist [Charkari u. Mori 1993]. Eine weitere

Anwendung findet man im Versuchsfahrzeug *VaMoRs-P* [Thomanek u. a. 1994]. Es zeigt eindrucksvoll, dass eine zuverlässige Detektion alleine schon mit diesem Merkmal möglich ist. Andere Autoren verwenden Schatten als Merkmal zur Verifikation der Messungen eines weiteren Sensors oder zur Bestimmung von Parametern wie z. B. der Fahrzeugbreite bei bekannter Position des Fahrzeuges. So wird z. B. in [Kim u. a. 2005] die Kombination eines Sonar-Sensors mit videobasierter Objektdetektion anhand des Schattens und anderer Merkmale untersucht.

In [Tzomakas u. von Seelen 1998] beschreiben die Autoren ein Verfahren zur Detektion von Fahrzeugen, das auf der Extraktion von Schatten beruht. Dabei wird eine Detektion der befahrbaren Fläche durchgeführt, um dann den mittleren Grauwert und die Varianz der Fahrbahnoberfläche bestimmen zu können. Die gefundenen Schattenregionen dienen allerdings nur als Hypothesen für weitere Verarbeitungsschritte.

Auch in dieser Arbeit wird das Merkmal Schatten zur Detektion von Fahrzeugen eingesetzt. Der entwickelte Algorithmus ist in Kapitel 2.3 eingehend beschrieben, weshalb hier nicht weiter darauf eingegangen wird.

2.1.3.6 Symmetrie

Betrachtet man das Heck eines vorausfahrenden Fahrzeugs, so kann man eine ausgeprägte Symmetrie bezüglich einer senkrechten Achse feststellen (vgl. Bild 2.8). Gleiches gilt für ein Automobil, das sich auf den Betrachter zu bewegt. Diese Eigenschaft kann zur Detektion von Fahrzeugen herangezogen werden [Kuehnle 1991; Zielke u. a. 1993].

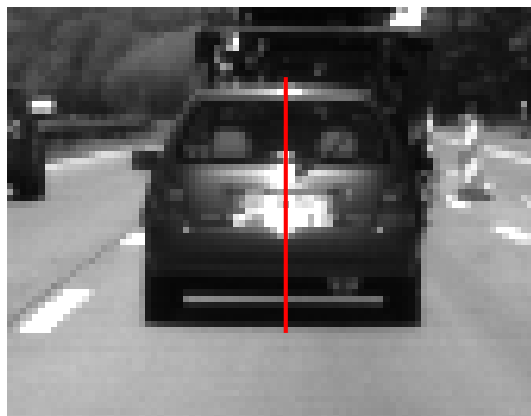


Bild 2.8: Vertikale Symmetrieachse der Heckansicht eines Fahrzeugs.

Beleuchtungsinhomogenitäten können jedoch die Symmetrie im Grauwertbild stören. So führt z. B. seitlicher Lichteinfall dazu, dass eine Seite des beobachteten

Fahrzeuges heller erscheint. Daher bietet sich eine Analyse im Kantenbild an [Du u. Papanikolopoulos 1997]. Dieses ist aufgrund des Hochpasscharakters des Kantendetektors weniger empfindlich gegenüber vielen solchen niederfrequenten Störungen. Dennoch führen etwa Spiegelungen zu asymmetrischen Strukturen im Bild, die auch auf diese Weise nicht eliminiert werden können.

Zur Bewertung der Symmetrie in einem Bildausschnitt können auch direkt die Grauwerte in einer Region um eine hypothetische Symmetrieachse betrachtet werden. Gegenüber kantenbasierten Verfahren steht damit mehr Information zur Verfügung, was wiederum zu einer erhöhten Detektionsgüte führen kann. Dieser Ansatz wird in der vorliegenden Arbeit verfolgt. Ein entsprechender Algorithmus wird in Kapitel 2.4 beschrieben.

2.2 Entfernungsinformation

In naher Zukunft halten Assistenzfunktionen, die als Sensor eine Videokamera einsetzen, Einzug in Fahrzeuge der Oberklasse. Beispiel hierfür sind Spurverlassenswarner, wie sie schon jetzt in Lastkraftwagen eingesetzt werden. Aus Kostengründen werden jedoch noch keine Stereosysteme verfügbar sein.

In dieser Arbeit wurde gezielt ein Sensor entwickelt, der auf der Basis der vorhandenen Komponenten die Realisierung weiterer Assistenzfunktionen ermöglicht. Zur Gewinnung von Stereoinformation aus Bewegung ist die Auswertung mehrerer aufeinander folgender Bilder notwendig, zur sicheren Erfassung von entfernten Objekten ist sogar eine Berücksichtigung mehrerer Bilder erforderlich. Objektdetektion mittels monoskopischer Einzelbildmerkmale hat demgegenüber den Vorteil, in den Sichtbereich des Sensors eintretende Objekte schnell erfassen zu können. Daher werden hier Verfahren dieser Kategorie genutzt. Von den im vorangegangenen Abschnitt beschriebenen Möglichkeiten zur videobasierten Objekterkennung wurden die monoskopischen Einzelbildmerkmale Schatten und Symmetrie gewählt.

Wie alle monoskopischen Einzelbildmerkmale erlauben diese eine Aussage über die Position von Objekten im Bild der Kamera. Für Fahrerassistenzanwendungen ist es jedoch entscheidend, die relative Position dieser Objekte in Weltkoordinaten zu kennen.

Die Projektion eines Punktes $\mathbf{z} = (X, Y, Z)^T$ in das Bild der Kamera liefert die Bildposition $\mathbf{p} = (u, v)^T$. Dabei liegt der Ursprung des Koordinatensystems, in dem der Punkt \mathbf{z} beschrieben ist, im optischen Zentrum, die Koordinatenrichtungen von X und Y sind parallel zu denen von u und v angeordnet. Unter Annahme eines

Lochkameramodells gilt dann der Zusammenhang

$$u = \frac{f_u}{Z} \cdot X + c_u \quad \text{bzw.} \quad v = \frac{f_v}{Z} \cdot Y + c_v. \quad (2.4)$$

Hierbei sind $\mathbf{c} = (c_u, c_v)^T$ die Koordinaten des Bildhauptpunktes. Da die Bildpunkte des Sensors nicht zwingend quadratisch sind, ergeben sich auch unterschiedliche Brennweiten $\mathbf{f} = (f_u, f_v)^T$ für die beiden Koordinatenrichtungen.

An Gleichung (2.4) wird deutlich, dass die Projektion nicht ein-eindeutig ist. Der Übergang von einem dreidimensionalen in ein zweidimensionales Koordinatensystem ist eine Projektion in einen niederdimensionalen Raum: um die X - und Y -Koordinaten für einen gegebenen Bildpunkt bestimmen zu können, muss zusätzlich der Abstand Z bekannt sein.

Um also zu einer detektierten Bildposition die 3D-Position bestimmen zu können, muss die Entfernung des entsprechenden Objektes gegeben sein. Diese Anforderung kann z. B. durch die Annahme einer Fahrbahnebene mit bekannten Parametern erfüllt werden (vgl. Bild 2.9). Genauere Ergebnisse liefert eine Fahrbahnerkennung, die zur Bestimmung der Parameter herangezogen wird [Behringer u. a. 1992; Kim 2006; Aufrère u. a. 2000].

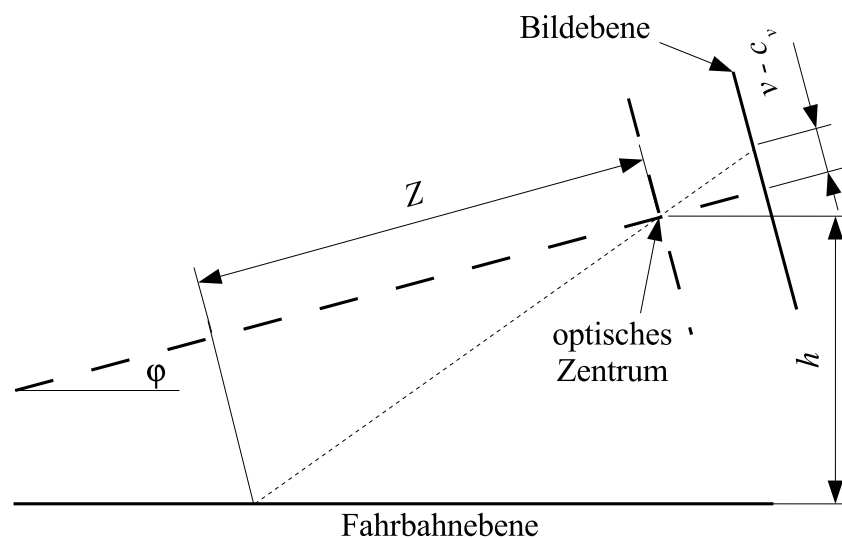


Bild 2.9: Bestimmung der Distanz Z zu einer Bildposition v bei gegebener Fahrbahnorientierung φ und Höhe der Kamera h .

Wird ein Punkt auf der Fahrbahnoberfläche betrachtet und sind die Höhe der Kamera h und die Orientierung der Fahrbahn φ bekannt, so ergibt sich für die 3D-

Position $\mathbf{z} = \mathbf{l}(u, v, h, \varphi)$:

$$X = \frac{h \cdot \frac{u-c_u}{f_u}}{\sin(\varphi) + \frac{v-c_v}{f_v} \cdot \cos(\varphi)} \quad (2.5)$$

$$Y = \frac{h}{\frac{f_v}{v-c_v} \cdot \sin(\varphi) + \cos(\varphi)} \quad (2.6)$$

$$Z = \frac{h}{\sin(\varphi) + \frac{v-c_v}{f_v} \cdot \cos(\varphi)} \quad (2.7)$$

Eine Aussage über die Unsicherheit dieser 3D-Position liefert die Kovarianzmatrix \mathbf{R} . Da die Gleichungen (2.5) bis (2.7) nichtlinear sind, folgt nach einer Linearisierung um den Bildpunkt für kleine Abweichungen:

$$\mathbf{R} = \mathbf{V} \mathbf{R}_{2D} \mathbf{V}^T \quad (2.8)$$

In Gleichung (2.8) beschreibt \mathbf{R}_{2D} die Unsicherheiten der Bildpositionen sowie der Ergebnisse der Fahrbahndetektion. Sind diese Parameter unkorreliert, so ist \mathbf{R}_{2D} eine Diagonalmatrix:

$$\mathbf{R}_{2D} = \begin{bmatrix} \sigma_u^2 & 0 & 0 & 0 \\ 0 & \sigma_v^2 & 0 & 0 \\ 0 & 0 & \sigma_h^2 & 0 \\ 0 & 0 & 0 & \sigma_\varphi^2 \end{bmatrix} \quad (2.9)$$

\mathbf{V} bezeichnet die Jacobi-Matrix von $\mathbf{z} = \mathbf{l}(u, v, h, \varphi)$:

$$\mathbf{V} = \frac{\partial \mathbf{l}(u, v, h, \varphi)}{\partial ((u, v, h, \varphi)^T)} \quad (2.10)$$

Damit ist sichergestellt, dass nicht nur die Unsicherheiten in der Merkmalsdetektion, sondern auch in der vorangehenden Fahrbahnebenenschätzung berücksichtigt werden.

Ist die 3D-Position eines Punktes auf der Fahrbahnoberfläche, der durch den Bildpunkt \mathbf{p} beschrieben wird, von Interesse, so kann diese aus den Gleichungen (2.5) bis (2.7) bestimmt werden. Dies ist für die Untersuchung von Fahrzeugschatten der Fall. Liegt der Punkt nicht auf der Fahrbahnoberfläche, sondern im Abstand h_o darüber, so ergibt sich der entsprechende Zusammenhang durch die Substitution $h \rightarrow h - h_o$. Dies entspricht einer Parallelverschiebung der Fahrbahnebene.

2.3 Schatten

Für den im Rahmen dieser Arbeit entwickelten Sensor wurde neben der Symmetrie das Merkmal Schatten herangezogen. Im Gegensatz zu den bisher bekannten Verfahren sollte der Algorithmus jedoch in der Lage sein, komplett eigenständig zu arbeiten, d. h. es stehen weder Hypothesen eines vorgeschalteten Sensors zur Verfügung noch ist eine anschließende Verfeinerung notwendig. Dennoch wird der Vorteil der Kombination mit anderen Sensoren bzw. weiteren Merkmalen genutzt, sofern diese vorhanden sind. Dies ist Aufgabe der Fusion, die in Kapitel 3 beschrieben ist.

Vorausgesetzt werden kann jedoch, dass bei der Merkmalsextraktion die Orientierung der Fahrbahn bekannt ist, da diese auch zur Gewinnung der 3D-Information aus 2D-Bildpunkten benötigt wird (vgl. Abschnitt 2.2). Diese liegt in Form einer Zuordnung von Entfernungswerten zu Bildpunkten vor unter der Annahme, dass der betrachtete Punkt auf der Fahrbahnoberfläche liegt. Hierdurch wird zwar der Suchbereich auf den Bereich unter dem Horizont eingeschränkt, jedoch bedeutet dies keine laterale Abgrenzung. Damit ist es möglich, auch Fahrzeuge zu erfassen, die sich auf benachbarten Spuren befinden, entgegenkommen oder einen Spurwechsel durchführen.

Der entwickelte Algorithmus liefert als eigenständiger Sensor Informationen über die Position von Fahrzeugen anhand ihrer Schatten und besteht aus folgenden Einzelschritten (siehe dazu auch [Hoffmann u. a. 2004; Hoffmann 2006]):

1. Bestimmen des Grauwertes der Fahrbahnoberfläche,
2. Extrahieren dunkler Zeilensegmente,
3. Gruppieren der Zeilensegmente zu Rechteckhypothesen,
4. Anpassen der Berandungen,
5. Bestimmen der Ausprägung der Submerkmale und
6. Klassifikation der Hypothesen.

Die Schritte 1 bis 3 dienen dazu, mittels einfacher Operationen Hypothesen zu gewinnen, die mögliche Schattenregionen beschreiben. Diese Hypothesen werden im vierten Schritt verfeinert, indem eine Analyse der seitlichen und des unteren Gradienten durchgeführt wird. Schließlich werden Merkmale bestimmt, die bei der Klassifikation dazu verwendet werden, unwahrscheinliche Hypothesen wieder zu verwerfen.

Diese Vorgehensweise vermeidet es, aufwändige Rechenschritte für jeden Bildpunkt durchzuführen. Stattdessen werden sie nur für die zuvor gewonnenen Hypothesen notwendig, was eine effiziente Implementierung begünstigt. Die einzelnen Schritte werden im Folgenden eingehend erläutert.

2.3.1 Grauwert der Fahrbahn

Der Schatten unter einem Fahrzeug ist dadurch gekennzeichnet, dass er dunkler als die umgebende Fahrbahn ist. Diese Feststellung ist für einen Menschen trivial; zur automatisierten Auswertung muss jedoch eine geeignete Beschreibung für „dunkler“ gefunden werden. Nimmt man den Grauwert der freien Fahrbahnoberfläche $g(\mathbf{p}_F)$ als konstant an, also

$$g(\mathbf{p}_F) \approx g_F, \quad (2.11)$$

so erhält man die einfache Beziehung

$$g(\mathbf{p}_S) < g_F, \quad (2.12)$$

wobei $g(\mathbf{p}_S)$ der Grauwert eines Schattenpunktes ist. Die Annahme 2.11 ist unter realen Bedingungen nur näherungsweise erfüllt, da der Grauwert der Fahrbahnoberfläche von Lichteinfallswinkel und damit von Sonnenstand und Beobachterposition abhängt. Zudem ist auch der Fahrbahnbelag selbst nicht homogen, sodass kein einfaches Modell zur genaueren Beschreibung gefunden werden kann. Dennoch ist die Bedingung hinreichend genau erfüllt, um erste Schattenhypothesen zu extrahieren. Diese sind noch unsicherheitsbehaftet und müssen in weiteren Schritten verfeinert werden. Zudem müssen Regionen, die fälschlicherweise als Schatten klassifiziert wurden, verworfen werden.

Um g_F zu bestimmen, wird ein Bildausschnitt \mathcal{M} untersucht, der einen Teil der Fahrbahnoberfläche vor dem Fahrzeug (vgl. Bild 2.10) wiedergibt. Mit dem Mittelwert

$$g_F = \overline{g(\mathbf{p})}, \quad \mathbf{p} \in \mathcal{M} \quad (2.13)$$

erhält man einen Schätzwert für den Grauwert der Fahrbahn. Dabei muss jedoch sichergestellt werden, dass \mathcal{M} keine Störungen wie z. B. Fahrspurmarkierungen oder Fahrzeuge enthält. Die Varianz der Grauwerte der Fahrbahnoberfläche ist gewöhnlich sehr gering und kann empirisch ermittelt werden. Ein Anstieg von $\text{Var}\{g(\mathbf{p})\}, \mathbf{p} \in \mathcal{M}$ liefert also einen Hinweis auf Störobjekte.

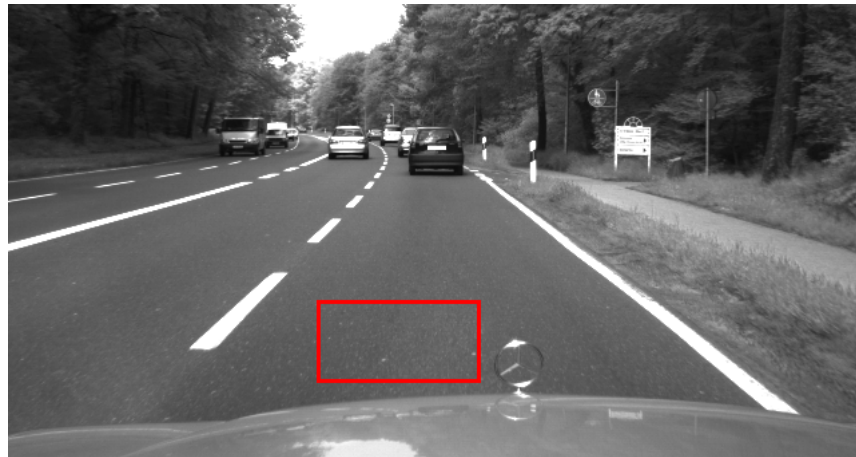


Bild 2.10: Bereich \mathcal{M} zur Bestimmung des Fahrbahngrauwertes.

2.3.2 Dunkle Zeilensegmente

Ist der Grauwert der Fahrbahn bekannt, so wird nach Zeilensegmenten gesucht, die zu einem Schatten gehören können. Dabei wird vorausgesetzt, dass die Zeilen des Kamerabildes näherungsweise parallel zur Fahrbahnebene verlaufen. Für die relevanten Fahrsituationen ist dies leicht zu gewährleisten, da lediglich vernachlässigbar kleine Wankbewegungen auftreten.

Um die gesuchten Zeilensegmente zu bestimmen, werden zunächst seitliche Berandungspunkte extrahiert. Dies sind Bildpunkte, an denen in horizontaler Richtung ein Übergang von hell nach dunkel bzw. dunkel nach hell stattfindet. Hierzu wird der in Abschnitt 2.3.1 bestimmte Grauwert der Fahrbahnoberfläche herangezogen.

Betrachtet man eine einzelne Bildzeile v , so erhält man einen Grauwertverlauf, wie er in Bild 2.11 dargestellt ist. Die Schnittpunkte des Grauwertverlaufes mit dem Straßengrauwert repräsentieren nun eine Vielzahl von Hypothesen für linke bzw. rechte Berandungspunkte $u_{l,i}$, $i = 1, \dots, N_l$ und $u_{r,k}$, $k = 1, \dots, N_r$ möglicher Schatten. Indem man diesem Schritt eine Tiefpassfilterung, beispielsweise mit einem 3×3 Gauß-Filter, voranstellt, kann man unerwünschte Störungen, die durch sehr feine Strukturen im Bild entstehen, unterdrücken. Bild 2.12 zeigt die so erhaltenen Hypothesen für linke und rechte Randpunkte eines kompletten Bildes.

Im nächsten Schritt werden Punktpaare $\mathcal{P}_{ik} = \{u_{l,i}, u_{r,k}\}$ gewählt, deren Abstand $w_{ik} = u_{r,k} - u_{l,i}$ einer für Fahrzeuge typischen Breite entspricht. Dabei wird vorausgesetzt, dass die Breite B_F der betrachteten Fahrzeuge durch eine unimodale Verteilung $p(B_F)$ beschrieben ist. Zwar ist diese Annahme aufgrund typischer Abmaße verschiedener Fahrzeugtypen nicht streng erfüllt, jedoch kann die wahre

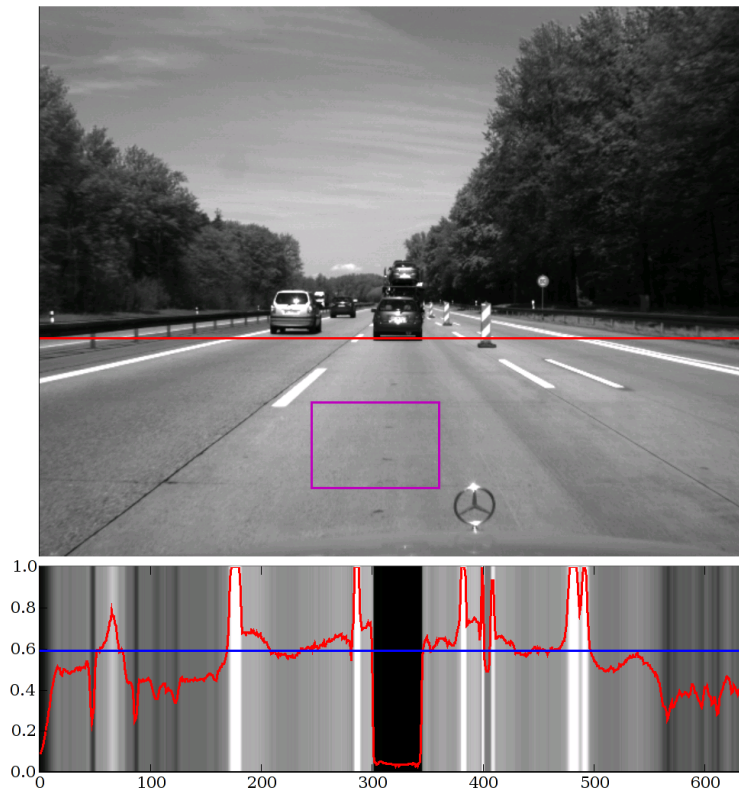


Bild 2.11: Oben: Eine zu untersuchende Bildzeile in einem Bild einer Autobahnsequenz. Unten: Grauwertverlauf (rot) in dieser Zeile und Grauwert g_F der Fahrbahnoberfläche (blau).

Verteilung gut durch eine unimodale angenähert werden. Setzt man des Weiteren voraus, dass die Messung w der Breite erwartungstreu ist, vorausgesetzt an der betrachteten Stelle befindet sich ein Fahrzeug, so gilt

$$E \{p(w|F(u, v))\} = B_F . \quad (2.14)$$

Dabei repräsentiert $F(u, v)$ das Ereignis, dass sich an der durch den Bildpunkt (u, v) beschriebenen Position ein Fahrzeug befindet, d. h. dass der Punkt auf der senkrechten Mittellinie eines Fahrzeugschattens liegt. Schließlich wird angenommen, dass in dem Fall, dass kein Fahrzeug vorhanden ist, die gemessene Breite gleichverteilt ist. Dies bedeutet, dass die durch Fehlereinflüsse hervorgerufenen Beobachtungen keine Vorzugsbreite aufweisen. Unter diesen Bedingungen folgt anschaulich, dass die Wahrscheinlichkeit dafür, dass es sich bei der Beobachtung um ein Fahrzeug handelt größer ist, je kleiner die Abweichung von der erwarteten Fahrzeugbreite $E \{B_F\}$ ist. Damit lassen sich die Hypothesen \mathcal{P}_{ik} bewerten: je kleiner die Distanz $|w_{ik} - B_F|$ ist, desto wahrscheinlicher beschreibt die Hypothese ein Fahrzeug.

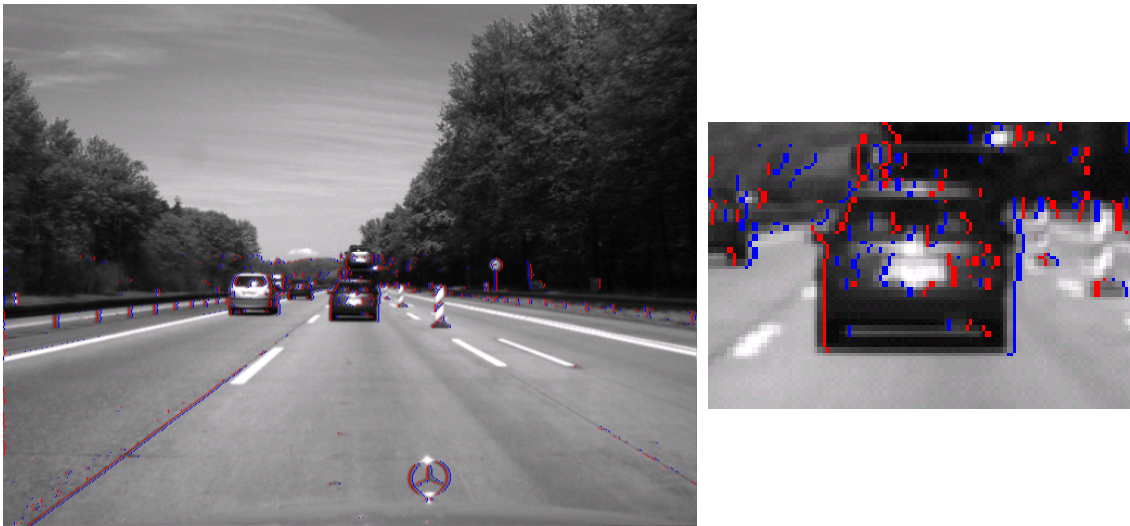


Bild 2.12: Mögliche linke (rot) und rechte (blau) Berandungspunkte von Schatten. Links: ganzes Bild, rechts: vergrößerter Ausschnitt.

Daher darf der Abstand der Berandungspunkte $w_{ik} = u_{r,k} - u_{l,i}$ maximal um einen Anteil a von der vorausgesetzten Fahrzeugbreite B_F abweichen, andernfalls wird das Punktpaar verworfen. Um bei der weiteren Betrachtung berücksichtigt zu werden, muss es also die Bedingung

$$a \geq \frac{|w_{ik} - B_F|}{B_F} \quad (2.15)$$

erfüllen.

2.3.3 Gruppieren der Zeilensegmente

Bisher wurde nur untersucht, wie sich ein Fahrzeugschatten auf den Grauwertverlauf einer einzelnen Bildzeile auswirkt. Ein Schattenbereich erstreckt sich im Allgemeinen jedoch über mehrere Bildzeilen. Daher werden nun die in Abschnitt 2.3.2 extrahierten Zeilensegmente, die einen zusammenhängenden Bereich bilden, zu Gruppen verknüpft.

Die erwartete Form kann grob durch ein Rechteck beschrieben werden. Daher werden Zeilensegmente, die sich stark überlappen, einander zugeordnet. Der Grad der Überlappung zweier Zeilensegmente \mathcal{P}_{ik} und \mathcal{P}_{lm} wird hierfür definiert zu:

$$l = \frac{\min\{u_{r,k}, u_{r,m}\} - \max\{u_{l,i}, u_{l,l}\}}{\min\{(u_{r,k} - u_{l,i}), (u_{r,m} - u_{l,l})\}} \quad (2.16)$$

Die Bedeutung der Überlappung wird in Bild 2.13 deutlich. Der Grad der Überlap-

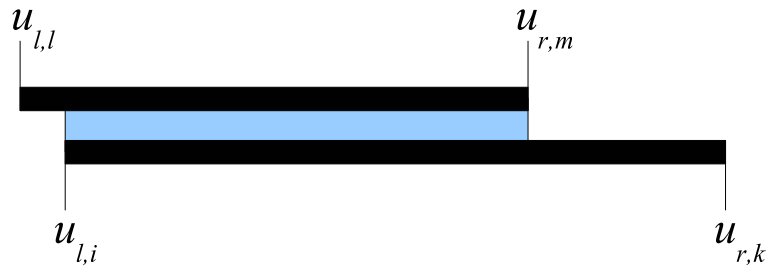


Bild 2.13: Zwei sich überlappende Zeilensegmente. Der Überlappungsbereich ist blau markiert.

Die Überlappung l ist also die Länge des Überlappungsbereiches bezogen auf die Länge des kürzeren Zeilensegments.

Weisen zwei Zeilensegmente einen Überlappungsgrad auf, der einen bestimmten Schwellwert überschreitet, so werden beide einer Gruppe zugeordnet. Dabei müssen folgende Fälle beachtet werden:

- War bisher keines der beiden Zeilensegmente einer Gruppe zugeordnet, so wird eine neue Gruppe angelegt.
- War schon eines der beiden Zeilensegmente einer Gruppe zugeordnet, so wird auch das andere dieser Gruppe zugeordnet.
- Waren schon beide Zeilensegmente Gruppen zugeordnet, so werden diese Gruppen vereinigt.

Die umschreibenden Rechtecke der so erhaltenen Gruppen dienen dann als erste Schattenhypothesen, die weiter analysiert werden. Bild 2.14 zeigt die so erhaltenen Hypothesen für das oben angeführte Beispiel.

2.3.4 Anpassen der Berandungen

Da diese Hypothesen aus den umschreibenden Rechtecken der Liniensegmente bestimmt wurden, sind sie noch recht ungenau. In den weiteren Schritten werden sie verfeinert.

Zunächst wird die genaue Position der Unterkante bestimmt. Dazu wird ein kleiner Bildbereich \mathcal{M} der Breite w um die Unterkante der Hypothese betrachtet (siehe Bild 2.15). Für die Grauwerte in diesem Bereich wird zeilenweise der Mittelwert

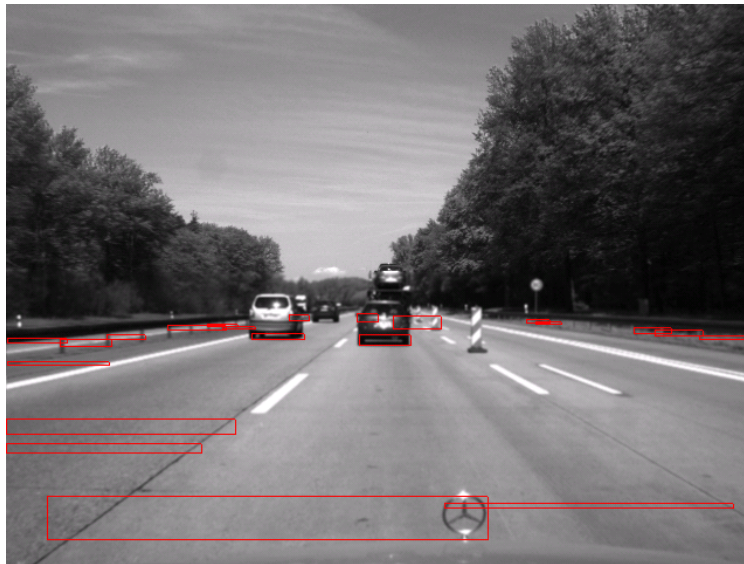


Bild 2.14: Ergebnis der Gruppierung: dargestellt sind die umschreibenden Rechtecke.

bestimmt:

$$\bar{g}_u(v) = \frac{1}{w} \sum_{u \in \mathcal{M}} g(u, v) \quad (2.17)$$

Für das Beispiel ist dies in Bild 2.16 gezeigt. Die verfeinerte Position v_u der Unterkante wird dann durch eine Betrachtung des Gradienten bestimmt. Da mehrere Gradienten auftreten können, die auf einen entsprechenden dunkel/hell-Übergang hinweisen, wird derjenige gewählt, der der Unterkante des betrachteten Bereiches am nächsten liegt. Eine analoge Vorgehensweise führt danach zu verfeinerten linken und rechten Berandungen u_l bzw. u_r .

Die Ausprägung der Gradienten gibt auch einen Hinweis darauf, ob es sich wirklich um einen Schattenbereich handelt. Daher werden die Gradientenwerte, zusammen mit anderen Merkmalen, nochmals im Klassifikationsschritt zur Bewertung der Hypothesen herangezogen.

2.3.5 Klassifikation

Durch das Anpassen der Berandungen wird sichergestellt, dass Schattenbereiche von den Hypothesen exakt umschrieben werden. Allerdings werden auch viele Strukturen erfasst, bei denen es sich nicht um Fahrzeugschatten handelt. So werden gerade im Leitplankenbereich und durch Büsche und Bäume viele falsche Hypo-



Bild 2.15: Anfängliche Hypothese (rot) und zur Bestimmung der Unterkante betrachteter Bereich (gelb).

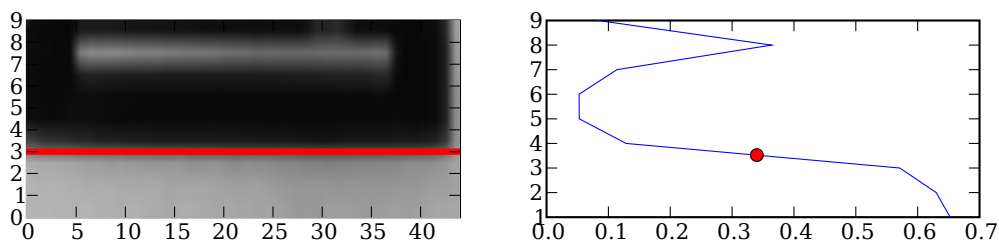


Bild 2.16: Links: Vergrößerter Bereich zur Bestimmung der Unterkante. Rechts: der Mittelwerte der Grauwerte wird zeilenweise bestimmt.

thesen hervorgerufen. Ein abschließender Klassifikationsschritt dient dazu, solche Fehldetektionen zurückzuweisen.

Basis der Klassifikation ist der siebendimensionale Merkmalsvektor y_M . Er setzt sich zusammen aus:

- dem Grauwertgradienten in vertikaler Richtung an der Unterkante,
- dem Mittelwert der beiden seitlichen Grauwertgradienten in horizontaler Richtung,

- dem Verhältnis von Höhe zu Breite der Hypothese,
- dem mittleren Grauwert und der Varianz der Grauwerte im Bereich der Hypothese und
- dem Grauwert und der Varianz der Grauwerte im Bereich unter der Hypothese.

Zur Klassifikation wird eine lineare Diskriminanzfunktion $c(\mathbf{y}_M)$ herangezogen:

$$c(\mathbf{y}_M) = w_0 + \mathbf{w}^T \mathbf{y}_M. \quad (2.18)$$

Die Zuordnung einer Hypothese i zu einer Klasse geschieht dann aufgrund des Vorzeichens der Diskriminanzfunktion. Ist der Ausdruck $c(\mathbf{y}_{Mi}) > 0$, so wird sie als Schattenmessung weiterverarbeitet. Ist er hingegen negativ, so wird diese Hypothese verworfen. Die Parameter \mathbf{w} und w_0 beschreiben somit den Klassifikator.

Für die Diskriminanzfunktion (2.18) wird eine vereinfachte Schreibweise eingeführt. Erweitert man den Merkmalsvektor um einen Eintrag mit dem Wert 1 und fasst die Parameter w_0 und \mathbf{w} in einem Gewichtsvektor \mathbf{a} zusammen, so ergibt sich

$$c(\mathbf{y}) = \mathbf{a}^T \mathbf{y}, \quad (2.19)$$

mit

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{pmatrix} = \begin{pmatrix} 1 \\ y_{M1} \\ y_{M2} \\ y_{M3} \\ y_{M4} \\ y_{M5} \\ y_{M6} \\ y_{M7} \end{pmatrix} \quad \text{und} \quad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \end{pmatrix} = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \end{pmatrix}. \quad (2.20)$$

Anschaulich ist \mathbf{a} der Normalenvektor einer Hyperebene durch den Ursprung des Merkmalsraums, die die Klassen trennt. Bild 2.17 veranschaulicht diesen Sachverhalt anhand eines Beispiels mit eindimensionalen Merkmalen. Durch die vereinfachte Schreibweise wird der Merkmalsvektor zweidimensional, die Trennebene geht durch den Ursprung des Koordinatensystems.

Die Wahl des Gewichtsvektors \mathbf{a} spielt eine entscheidende Rolle. Daher wird er aus einem Trainingsdatensatz $\mathcal{Y} = \{\mathbf{y}_i | i = 1, \dots, I\}$ ermittelt. Dieser Trainingsdatensatz wurde gewonnen, indem speziell für diesen Zweck aufgenommene Sequenzen ausgewertet und die Hypothesen manuell als wahr oder falsch annotiert wurden.

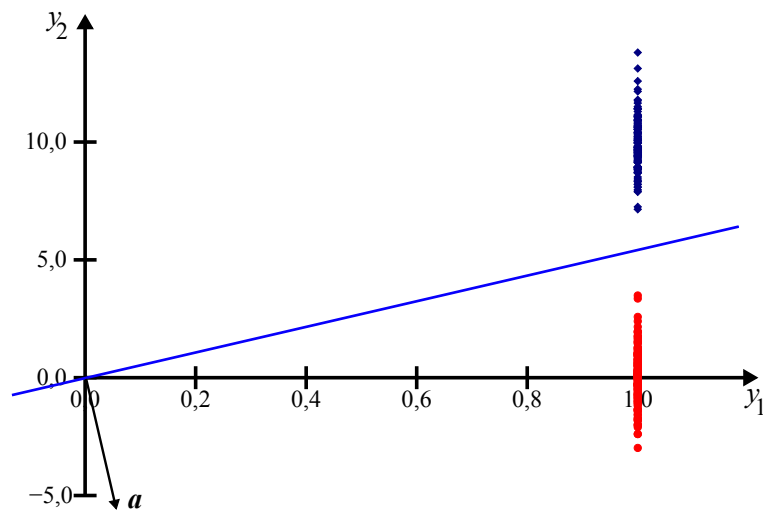


Bild 2.17: Merkmale, die zwei Klassen zuzuordnen sind, und die zugehörige trennende Hyperebene mit Gewichtsvektor a .

Ist die Menge der Merkmalsvektoren y im Trainingsdatensatz linear separierbar, so kann ein einfacher Perzeptron-Algorithmus ([Duda u. a. 2000]) angewandt werden.

Zur Vereinfachung wurden alle Merkmalsvektoren y des Datensatzes, die als falsch annotiert waren, negiert. Dadurch kann einfach entschieden werden, ob eine Hypothese korrekt klassifiziert wurde: für alle Merkmalsvektoren, die korrekt einer Klasse zugeordnet wurden, ergibt die Diskriminanzfunktion einen positiven Wert. Ein negativer Wert bedeutet, dass eine Fehlzuordnung stattgefunden hat.

Das iterative Verfahren korrigiert dann den geschätzten Gewichtsvektor \hat{a} schrittweise anhand der falsch klassifizierten Merkmalsvektoren. Es werden dabei alle Merkmalsvektoren des Trainingsdatensatzes einzeln betrachtet. Dieser Datensatz wird so lange wiederholt durchlaufen, bis keine Fehlklassifikationen mehr auftreten. Wird eine betrachtete Hypothese falsch klassifiziert, so wird der Korrekturschritt

$$\hat{a}(k+1) = \hat{a}(k) + y^k \quad (2.21)$$

ausgeführt. Bei y^k handelt es sich um das k -te Element der Sequenz falsch klassifizierter Merkmalsvektoren. Nach jedem Korrekturschritt wird $\hat{a}(k+1)$ normiert, um ein stetiges Anwachsen zu vermeiden.

Leider handelt es sich jedoch nicht um ein vollständig linear separierbares Problem. Schon der Trainingsdatensatz enthält Hypothesen, die als falsch annotiert sind, jedoch tatsächlich anhand der bestimmten Merkmale nicht von einer wahren Hypothese unterschieden werden können. Auch der umgekehrte Fall, also als

wahr annotierte Hypothesen, deren Merkmalsvektoren jedoch auch bei optimaler Wahl des Gewichtsvektors zur Ablehnung führen, kann nicht ausgeschlossen werden, da der gewählte Merkmalsvektor nur eine Untermenge der zur vollständigen Beschreibung der Fahrzeugschatten benötigten Information darstellt.

Dieser Umstand führt dazu, dass der einfache Perzeptron-Algorithmus nicht mehr konvergiert. Daher wurde eine adaptive Korrektur gewählt:

$$\hat{\mathbf{a}}(k+1) = \hat{\mathbf{a}}(k) + \frac{\mathbf{y}^k}{m(\mathbf{y}^k)}. \quad (2.22)$$

Hier gibt $m(\mathbf{y})$ an, wie oft der Vektor \mathbf{y} bisher falsch klassifiziert wurde. Somit tragen Hypothesen, die häufig falsch klassifiziert wurden, weniger zur Korrektur bei. Zwar kann auch hierfür die Konvergenz nicht vollständig analytisch nachgewiesen werden, für den linear separierbaren Fall ist dies jedoch möglich. Der nicht linear separierbare Fall kann dann plausibilisiert und mittels mehrfacher Durchläufe mit unterschiedlicher Initialisierung verifiziert werden.

Zunächst wird die Konvergenz der Modifikation aus Gleichung (2.22) für den linear separierbaren Fall betrachtet. Wie bei der Variante aus Gleichung (2.21) lässt sich diese für ausreichend lange Lösungsvektoren zeigen. Dies ist keine Einschränkung, da sich aus der vereinfachten Schreibweise eine besondere Eigenschaft des Gewichtsvektors ergibt. Der Betrag, d. h. die Länge, des Gewichtsvektors hat keinen Einfluss auf die Klassifikation. Ist also \mathbf{a} eine Lösung, so trennt auch $\alpha\mathbf{a}$ die Klassen korrekt, wobei α ein positiver Skalierungsfaktor ist. Der Nachweis der Konvergenz für ausreichend lange Gewichtsvektoren ist also hinreichend.

Nähert sich der geschätzte Gewichtsvektor $\hat{\mathbf{a}}$ einer Lösung $\alpha\mathbf{a}$ mit jedem Schritt weiter an, so konvergiert das Verfahren:

$$\|\hat{\mathbf{a}}(k+1) - \alpha\mathbf{a}\|^2 \leq \|\hat{\mathbf{a}}(k) - \alpha\mathbf{a}\|^2 \quad (2.23)$$

Um zu zeigen, dass diese Bedingung erfüllt ist, wird der Ausdruck

$$\hat{\mathbf{a}}(k+1) - \alpha\mathbf{a} \quad (2.24)$$

betrachtet. Mit der Vorschrift aus Gleichung 2.22 folgt dafür:

$$\hat{\mathbf{a}}(k+1) - \alpha\mathbf{a} = \left(\hat{\mathbf{a}}(k) + \frac{\mathbf{y}^k}{m(\mathbf{y}^k)} \right) - \alpha\mathbf{a} \quad (2.25)$$

und damit

$$\|\hat{\mathbf{a}}(k+1) - \alpha\mathbf{a}\|^2 = \left\| \left(\hat{\mathbf{a}}(k) + \frac{\mathbf{y}^k}{m(\mathbf{y}^k)} \right) - \alpha\mathbf{a} \right\|^2 \quad (2.26)$$

$$= \|\hat{\mathbf{a}}(k) - \alpha\mathbf{a}\|^2 + 2(\hat{\mathbf{a}}(k) - \alpha\mathbf{a})^T \frac{\mathbf{y}^k}{m(\mathbf{y}^k)} + \left\| \frac{\mathbf{y}^k}{m(\mathbf{y}^k)} \right\|^2. \quad (2.27)$$

Da es sich bei \mathbf{y}^k um eine falsch klassifizierte Hypothese handelt, ist der Ausdruck $\hat{\mathbf{a}}(k)^\top \frac{\mathbf{y}^k}{m(\mathbf{y}^k)}$ immer kleiner oder gleich Null. Damit folgt

$$\|\hat{\mathbf{a}}(k+1) - \alpha \mathbf{a}\|^2 \leq \|\hat{\mathbf{a}}(k) - \alpha \mathbf{a}\|^2 - 2\alpha \mathbf{a}^\top \frac{\mathbf{y}^k}{m(\mathbf{y}^k)} + \frac{\|\mathbf{y}^k\|^2}{m(\mathbf{y}^k)^2}. \quad (2.28)$$

Das Verfahren konvergiert also, wenn

$$2\alpha \mathbf{a}^\top \frac{\mathbf{y}^k}{m(\mathbf{y}^k)} \geq \frac{\|\mathbf{y}^k\|^2}{m(\mathbf{y}^k)^2} \quad (2.29)$$

ist. Da es sich bei \mathbf{a} um einen Lösungsvektor handelt und zuvor alle als falsch annotierten Merkmalsvektoren negiert wurden, ist $\mathbf{a}^\top \mathbf{y}^k \geq 0$. Da zudem immer $m(\mathbf{y}^k) \geq 1$ ist, gilt:

$$2\alpha \mathbf{a}^\top \frac{\mathbf{y}^k}{m(\mathbf{y}^k)} \geq 2\alpha \mathbf{a}^\top \frac{\mathbf{y}^k}{m(\mathbf{y}^k)^2}. \quad (2.30)$$

Daher ist die Bedingung (2.29) auch erfüllt, wenn

$$2\alpha \mathbf{a}^\top \frac{\mathbf{y}^k}{m(\mathbf{y}^k)^2} \geq \frac{\|\mathbf{y}^k\|^2}{m(\mathbf{y}^k)^2} \quad (2.31)$$

gilt. Damit ergibt sich die gleiche Konvergenzbedingung wie für den Perzeptron-Algorithmus ohne Modifikation [Duda u. a. 2000]:

$$2\alpha \mathbf{a}^\top \mathbf{y}^k \geq \|\mathbf{y}^k\|^2. \quad (2.32)$$

Ist β die maximale Länge, die ein Merkmalsvektor erreichen kann, also

$$\beta^2 = \max_{\mathbf{y}} \|\mathbf{y}_i\|^2, \quad (2.33)$$

und γ der kleinste Wert, den das Skalarprodukt $\mathbf{a}^\top \mathbf{y}_i$ erreicht,

$$\gamma = \min_{\mathbf{y}} [\mathbf{a}^\top \mathbf{y}_i] > 0, \quad (2.34)$$

so ergibt sich, dass das Verfahren konvergiert für

$$2\alpha \gamma \geq \beta^2, \quad (2.35)$$

d. h.

$$\alpha \geq \frac{\beta^2}{2\gamma}. \quad (2.36)$$

Damit ist gezeigt, dass das Verfahren für den linear separierbaren Fall zu einer Lösung führt.

Enthält der Trainingsdatensatz wenige Merkmalsvektoren, die nicht korrekt mit einer Hyperebene separierbar sind, so führt die Gewichtung der Korrektur mit $\frac{1}{m(y^k)}$ dazu, dass diese Ausreißer mit steigender Anzahl der Iterationsschritte ausgeblendet werden. Dass dies wirklich der Fall ist, wurde experimentell untersucht. Dazu wurden mehrere Durchläufe gestartet, wobei der Gewichtsvektor $\mathbf{a}(0)$ zufällig initialisiert wurde. In allen Fällen wurden Lösungen erzielt, die die gleiche korrekte Klassifikation bewirkten.



Bild 2.18: Schattenmessungen nach der Klassifikation (vgl. auch Bild 2.14).

Der so trainierte Klassifikator wird auf alle Hypothesen angewendet, um Fehlhypothesen weitgehend zurückzuweisen. Bild 2.18 zeigt die Hypothesen, die beim oben aufgeführten Beispiel als Schattenmessungen ausgegeben werden. Obwohl noch immer Fehlhypothesen vorhanden sind, hat sich ihre Zahl durch den Klassifikationsschritt stark verringert. Somit können Ausreißer in der Schattendetektion zwar nicht vollständig unterdrückt, aber stark reduziert werden.

2.4 Symmetrie

Auch die Symmetrieeigenschaft der Front- bzw. Heckansicht ist ein generisches Merkmal für Fahrzeuge. Zwar weisen auch andere Objekte, wie z. B. Gebäude und Schilder, eine starke Achsensymmetrie auf, jedoch ist sie für Fahrzeuge aller Klassen ein typisches Merkmal. Damit bietet sie sich, obwohl sie ungeeignet für

eine eigenständige Detektion und Klassifikation ist, als Merkmal in einem Fusionsansatz an.

Der entwickelte Algorithmus nutzt die Grauwertsymmetrie bezüglich einer als bekannt vorausgesetzten Symmetrieachse zur Detektion von Fahrzeugen. Die Orientierung der Symmetrieachse ergibt sich aus der Annahme, dass sich das zu erfassende Fahrzeug und der Beobachter auf der selben Ebene bewegen. Durch die starre Verbindung von Kamera und Fahrzeug kann also einfach gewährleistet werden, dass die Symmetrieachsen im Bild in Spaltenrichtung verlaufen.

Diese Annahme wird zum einen durch Wankbewegungen der Fahrzeuge verletzt. Des Weiteren entspricht auch die reale Fahrbahn nicht einer idealen Ebene. Um die Auswirkungen dieser Störeinflüsse zu verringern, wird auf das zu untersuchende Kamerabild ein Tiefpassfilter in Spaltenrichtung angewandt. In der praktischen Anwendung hat sich hierfür ein 5×1 Gauß-Kern bewährt.

Verläuft eine Symmetrieachse in Spaltenrichtung, so sind die Grauwerte $g(u, v)$ in jeder Zeile v symmetrisch zu u_S . Dies bedeutet, dass $g(u_S + u^*, v)$, mit u^* als Laufvariable, eine gerade Funktion ist. Natürlich liegt in der Praxis keine ideale Symmetrie vor. Ein symmetrischer Bereich zeichnet sich jedoch durch einen hohen geraden Funktionsanteil aus.

Der gerade und ungerade Anteil einer Funktion lässt sich einfach mit

$$E(u^*, u_S, v) = \frac{1}{2} [g(u_S + u^*, v) + g(u_S - u^*, v)] \quad (2.37)$$

und

$$O(u^*, u_S, v) = \frac{1}{2} [g(u_S + u^*, v) - g(u_S - u^*, v)] \quad (2.38)$$

bestimmen. Naturgemäß handelt es sich beim ungeraden Anteil $O(u^*, u_S, v)$ um eine mittelwertfreie Funktion, für $E(u^*, u_S, v)$ gilt dies nicht. Betrachtet man einen Bereich der Breite $2w$, so kann für diesen der mittelwertbereinigte gerade Anteil bestimmt werden:

$$E_n(u^*, u_S, v, w) = E(u^*, u_S, v) - \frac{1}{2w} \int_{-w}^w E(u^*, u_S, v) du^* \quad (2.39)$$

Zur Beurteilung der Signifikanz des geraden und ungeraden Anteils in diesem Bereich schlagen die Autoren in [Zielke u. a. 1992] vor, die entsprechende Signalenergie heranzuziehen. Daraus wird analog einer Kontrastfunktion ein Symmetriemaß bestimmt:

$$S(u_S, w, v) = \frac{\int_{-w}^w E_n(u^*, u_S, v, w)^2 du^* - \int_{-w}^w O(u^*, u_S, v)^2 du^*}{\int_{-w}^w E_n(u^*, u_S, v, w)^2 du^* + \int_{-w}^w O(u^*, u_S, v)^2 du^*} \quad (2.40)$$

Dieses Symmetriemaß nimmt einen Wert zwischen -1 und 1 an.

Um eine Aussage über die Signifikanz einer Symmetrieachse treffen zu können, wird das Maß S_A verwendet:

$$S_A(u_S, w, w_{\max}, v) = \frac{w}{2w_{\max}} (S(u_S, w, v) + 1) . \quad (2.41)$$

Dieses kann Werte zwischen 0 und 1 annehmen. Um große Bereiche stärker zu bewerten, wird hier der Quotient $\frac{w}{w_{\max}}$ berücksichtigt, wobei w_{\max} die maximale Breite eines symmetrischen Bereiches angibt.

2.4.1 Initialisieren der Symmetriehypothesen

Das Signifikanzmaß $S_A(u_S, w, w_{\max}, v)$ wird in der vorliegenden Arbeit zur Bestimmung einer Symmetriekarte eingesetzt. Dazu wird S_A für jeden Bildpunkt $\mathbf{p} = (u_S, v)$ berechnet. Es werden jedoch nur Bereiche untersucht, deren Breite der erwarteten Breite eines Fahrzeuges entsprechen. Aus dieser Bedingung ergibt sich, dass nur noch Werte für $w = w_{\max}$ untersucht werden müssen. Gleichung (2.41) vereinfacht sich damit zu

$$S_A(u_S, w, v) = \frac{1}{2} (S(u_S, w, v) + 1) . \quad (2.42)$$

Die Breite w ergibt sich aus der angenommenen Breite eines Fahrzeuges. Setzt man, wie in Kapitel 2.3, voraus, dass der Bildpunkt einen Punkt auf der Fahrbahnoberfläche wiedergibt, so kann die Breite eines Fahrzeuges an dieser Position in Pixeln errechnet werden. Allerdings zeigt sich, dass dieser Punkt zur Initialisierung einer Symmetriesuche ungeeignet ist, da er im Fahrzeugschatten liegt. Gerade hier führt schräg einfallendes Licht häufig zu niedrigen Symmetriewerten.

Zudem sind dies oft Bereiche mit homogenen Grauwerten. Auch die Fahrbahnoberfläche weist homogene Grauwerte auf. Da dies einen hohen Gleichanteil impliziert und sich damit auch für diese Bereiche hohe Symmetriewerte ergeben, muss ein weiteres Kriterium zur Initialisierung einer Symmetriehypothese gefunden werden.

Da aufgrund der komplexen Form die Grauwerte auf einem Fahrzeug stark variieren, bietet sich die Varianz als weitere Bedingung an. Eine neue Symmetriehypothese wird daher nur erzeugt, wenn der betrachtete Bildbereich neben einem hohen Symmetriemaß auch eine hohe Grauwertvarianz aufweist.

Folglich dürfen keine Punkte auf der Fahrbahnoberfläche betrachtet werden. Oberhalb des Schattens, im Bereich des Stoßfängers oder des Nummernschildes, sind

beide Bedingungen erfüllt. Daher wird bei der Ermittlung der Breite w angenommen, dass sich der betrachtete Punkt im Abstand h_o über der Fahrbahn befindet (vgl. Abschnitt 2.2).

Bild (2.19) zeigt ein Beispiel einer so bestimmten Symmetriekarte. Helle Punkte stehen für hohe, dunkle Punkte für niedrige Symmetriewerte. Da nur Punkte betrachtet werden können, für die die gesamte Umgebung der Breite $2w$ sichtbar ist, entstehen im unteren Bildbereich dreieckförmige Aussparungen. Auch der Bereich über dem Horizont wird nicht berücksichtigt, indem keine Punkte untersucht werden, die ein Fahrzeug in extrem weiter Entfernung repräsentieren würden.

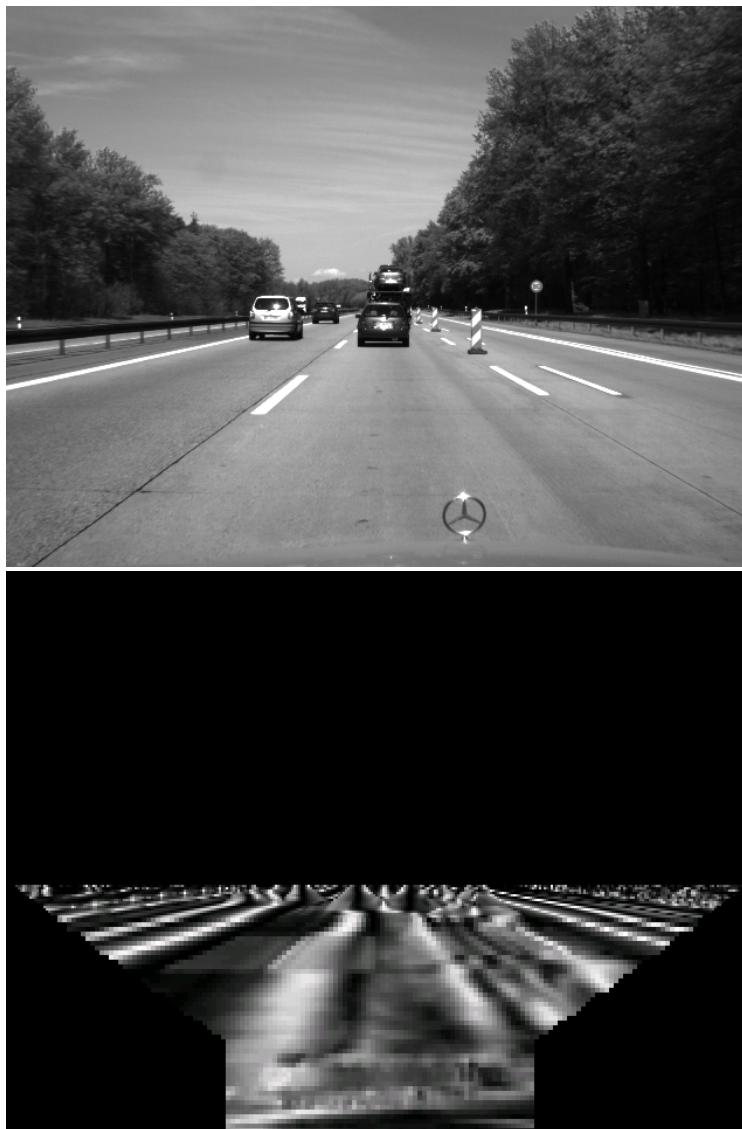


Bild 2.19: Oben: Originalbild aus einer Autobahnsequenz. Unten: Symmetriekarte zur Initialisierung der Suche nach Symmetrieachsen.

Da die Berechnung des Symmetriemaßes S_A eine große Umgebung um den be-

trachteten Bildpunkt berücksichtigt, ist der Rechenaufwand entsprechend hoch. Hier bietet sich die Bearbeitung auf reduzierten Auflösungsstufen an. Dazu wird eine Gauß-Pyramide des Ursprungsbildes aufgebaut [Jähne 1989]. Die Auflösungsstufe wird dann entfernungsabhängig gewählt. Zur Darstellung wurden die Ebenen wieder auf den Ursprungsmaßstab skaliert.

2.4.2 Ausdehnen der Symmetrieachsen

Da die Werte in der Symmetriekarte unter der Annahme eines festen Abstandes h_0 über der Fahrbahn bestimmt wurden, sind sie zur Detektion von Symmetrieachsen nicht ausreichend. Stattdessen dienen sie zur Initialisierung einer weitergehenden Suche. Als Startpunkte werden die zeilenweise gebildeten lokalen Maxima der Symmetriekarte herangezogen, wobei nur Punkte einbezogen werden, deren Umgebung das Varianzkriterium erfüllt. Von diesen ausgehend werden, unter Beibehaltung der Breite $2w$, erneut die Symmetriewerte dieser Spalte über und unter dem Startpunkt berechnet. Auf diese Weise können die Startpunkte zu Symmetrieachsen ausgedehnt werden.

Während nach oben ein einfacher Schwellwert für das Symmetriemaß als Abbruchkriterium der Suche ausreicht, ist das Eingrenzen der unteren Berandung deutlich schwieriger. Da sich hier üblicherweise Fahrbahn anschließt, die auch hohe Symmetriewerte aufweist, führt dieser Ansatz nicht zu einem zuverlässigen Ergebnis. Auch die Forderung einer minimalen Grauwertvarianz führt nicht zum Ziel, da auch damit nicht der Bereich zwischen den Rädern von der Fahrbahnoberfläche unterschieden werden kann.

Die Unterkante des Fahrzeugs, d. h. die Berührungslinie zwischen den Rädern und der Fahrbahn, stellt jedoch die einzige Möglichkeit dar, die Entfernung des Fahrzeuges zu bestimmen. Daher ist gerade eine exakte Bestimmung der Unterkante essentiell für die Genauigkeit der Messungen.

Bild (2.20) zeigt einen vergrößerten Ausschnitt der Symmetriekarte für ein Fahrzeug. Man erkennt, dass die Symmetrieachse des Fahrzeugs von Bereichen mit niedrigen Werten eingegrenzt wird, die vom Fahrzeugumriss herrühren. Während S_A auf der Symmetrieachse beim Übergang vom Fahrzeug auf die Fahrbahnoberfläche leicht absinkt, steigen die entsprechenden Werte im Abstand w von der Symmetrieachse deutlich an, da im Bereich des Fahrzeugumrisses keine Symmetrie vorliegt, auf der homogenen Fahrbahnoberfläche jedoch hohe Werte auftreten. Daher wird das Verhältnis der Symmetriewerte auf der Symmetrieachse und den Fahrzeugberandungen zur Bestimmung der Unterkante herangezogen:

$$\Delta S(u_S, w, v) = \frac{2S_A(u_S, w, v)}{S_A(u_S - w, w, v) + S_A(u_S + w, w, v)} . \quad (2.43)$$

Betrachtet man eine Hypothese an der Stelle u_S mit der Breite $2w$, so fällt $\Delta S(u_S, w, v)$ an der Unterkante des Fahrzeuges deutlich ab. Diese Eigenschaft wird daher zur Eingrenzung der Unterkante der Hypothese verwendet.



Bild 2.20: Oben: Vergrößerter Ausschnitt aus Bild (2.19), der ein zu detektierendes Fahrzeug zeigt. Unten: Entsprechender Ausschnitt der Symmetriekarte.

2.4.3 Resultierende Symmetriehypothesen

Auf diese Weise erhält man Symmetriehypothesen, wie sie in Bild (2.21) dargestellt sind. Da jedes lokale Maximum der Symmetriekarte zur Initialisierung einer Symmetriehypothese führen kann, ergeben sich üblicherweise mehrere Hypothesen für eine Symmetrieachse. Dies erkennt man im Bild an den verschiedenen Starthöhen, die für die Symmetrieachse des Fahrzeuges eingezeichnet sind. Dennoch führen alle zu den selben oberen und unteren Begrenzungen. Daher sind diese

Messungen als identisch zu betrachten und werden zu einer Symmetriehypothese zusammengefasst.

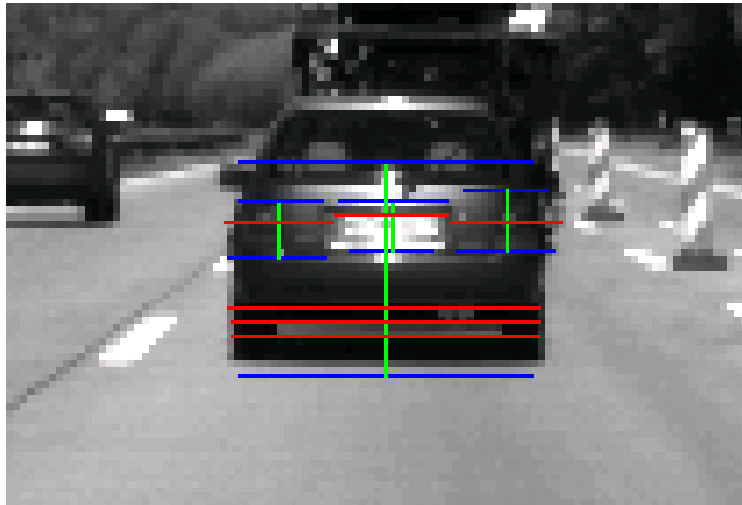


Bild 2.21: Resultierende Symmetriehypothesen. Zu den Symmetrieachsen (grün) sind jeweils Oberkante, Unterkante (beide blau) und Starthöhe der Symmetriesuche (rot) eingezeichnet.

Eine weitere Eigenschaft des Verfahrens wird am Beispiel deutlich: Die Hypothese beschreibt gut die Symmetrieachse und die untere Berandung des Fahrzeuges, die obere Begrenzung liegt aber nicht auf dem Dach des Fahrzeuges. Letzteres ist jedoch nicht relevant, da diese Information nicht weiterverwendet wird. Zur Positionsbestimmung sind die beiden erstgenannten Parameter entscheidend.

Neben den korrekten Fahrzeugpositionen liefert der Algorithmus allerdings auch einige Fehldetektionen. Diese lassen sich hauptsächlich auf die folgenden Phänomene zurückführen:

- *Substrukturen:* Im Bereich des Nummernschildes und der Rücklichter befinden sich weitere symmetrische Bereiche, die die Anforderungen an eine Hypothese erfüllen. Diese lassen sich jedoch durch eine einfache Betrachtung der Verdeckungen eliminieren: Da diese Hypothesen deutlich oberhalb der Fahrbahnoberfläche liegen, erscheinen sie als weiter entfernte Objekte. Damit wären sie jedoch nicht sichtbar, da sie vom Fahrzeug, das diese Hypothesen ausgelöst hat, verdeckt würden.
- *Gassen:* Auch die Gasse zwischen zwei nebeneinander fahrenden Fahrzeugen kann zu Fehldetektionen führen. Diese auszuschließen ist allein auf Basis einer Auswertung der Symmetrie nicht möglich.

- *Fahrbahnmarkierungen*: Schließlich können Fahrbahnmarkierungen Fehldetektionen verursachen. Da die Markierungen im Allgemeinen jedoch nicht senkrecht verlaufen, ist die Länge der daraus resultierenden Hypothesen stark begrenzt. Die Forderung eines minimalen Höhen- zu Breiten-Verhältnisses schließt diese Fehlerart weitgehend aus. Lediglich beim Überfahren einer Markierung, z. B. beim Spurwechsel oder bei Richtungsmarkierungen in der Mitte der Fahrspur, verbleiben falsche Detektionen.

Wie bei der Schattensuche verbleiben auch bei der Symmetriesuche Fehldetektionen, die nicht anhand dieses einen Merkmals auszuschließen sind. Erst über eine Fusion mehrerer Merkmale und eine zeitliche Verfolgung der Objekte kann eine weitere Verbesserung erzielt werden.

3 Fusion

Für Fahrerassistentenanwendungen, auf die der vorgestellte Sensoraufbau ausgerichtet ist, werden neben der Position detektierter Objekte weitere Informationen, wie deren Bewegungsrichtung und Geschwindigkeit, benötigt. Zudem sind hohe Anforderungen an die Genauigkeit und Zuverlässigkeit des Sensors zu erfüllen. Die reinen Detektionsschritte, wie sie in Kapitel 2 beschrieben wurden, können diese Kriterien nicht erfüllen. Erst eine Fusion und zeitliche Verfolgung der Merkmale führt hier zum Ziel.

Zentraler Bestandteil des Gesamtsystems ist somit die Fusion und zeitliche Verfolgung der Merkmale. Dieses Kapitel erarbeitet daher zunächst einige Grundlagen zur Informationsfusion, bevor es auf die Elemente des entwickelten Sensors eingeht.

3.1 Grundlagen der Informationsfusion

Die stetige Fortentwicklung sensorischer Fähigkeiten führt auch zu einem wachsenden Bedarf an Strategien, die erfassten Informationen zu kombinieren. Abhängig von der Anwendung, den Anforderungen und den zur Verfügung stehenden Sensoren wurden in der Literatur zahlreiche Ansätze vorgestellt.

3.1.1 Fusionsprinzipien

Aufgabe der Fusion ist, die von mehreren Sensoren zur Verfügung gestellte Information zu vereinigen, um eine genauere Schätzung des Zustandes beobachteter Objekte zu ermöglichen [Hall u. Llinas 2001]. Fusionsansätze können u. a. anhand der schon von den Sensoren geleisteten Vorverarbeitung klassifiziert werden [Klein 2004]. Weit verbreitet ist eine Einteilung in

- Fusion auf Signalebene,
- Fusion auf Merkmalebene und
- Fusion auf Entscheidungsebene.

Bei der *Fusion auf Signalebene* werden die nur minimal vorverarbeiteten Sensordaten direkt verknüpft. Werden Videosensoren eingesetzt, so spricht man auch häufig von einer Fusion auf Pixelebene. Dabei kann z. B. eine Verbesserung der Bildqualität das Ziel sein [Punkte León 1999]. Lassen sich die Messdaten schon auf diesem Niveau kombinieren, so kann man spezielle Eigenschaften der Sensoren nutzbar machen. Allerdings bedeutet dies auch, dass bei der Fusion Wissen über die Sensoren berücksichtigt wird. Dies erschwert den Aufbau einer flexiblen und erweiterbaren Architektur.

Extrahiert der Sensor bereits Informationen anhand der Eigenschaften z. B. der zu detektierenden Objekte, so spricht man von einer *Fusion auf Merkmalsebene*. Hier hat schon eine deutliche Verdichtung der vorhandenen Information stattgefunden.

Schließlich ist auch eine *Fusion auf Entscheidungsebene* möglich. Der Sensor führt bereits eine Klassifikation, oft verbunden mit einer zeitlichen Verfolgung, durch. Hier sind die Eingangsdaten und die Ergebnisse der Fusion von gleichem Abstraktionsgrad, z.B. kann es sich um Objekthypothesen handeln.

Dies ist nur eine Möglichkeit einer Einordnung verschiedener Ansätze. Oft lässt sich eine eindeutige Zuordnung zu einer dieser Klassen nicht vornehmen.

Der in dieser Arbeit beschriebene Aufbau ist zwischen der Fusion auf Merkmals- und der auf Entscheidungsebene einzuordnen. Die Sensoren, in diesem Fall also die Extraktion der Schatten- und Symmetriehypothesen, führen bereits eine Klassifikation durch, indem sie mögliche Objektpositionen bestimmen. Jedoch geht die Fusionsstufe noch einen Schritt weiter: hier werden diese Messungen validiert, Objekte zeitlich verfolgt und dabei weitere Parameter wie die Objektgeschwindigkeit bestimmt.

Die Schritte zur Merkmalsextraktion, im Weiteren werden sie als Sensoren bezeichnet, liefern also abstrahierte Merkmale. Dies hat den Vorteil, dass die Fusion keine Information über Sensoreigenschaften haben muss. Fällt ein Sensor aus, sinkt evtl. die Genauigkeit der Schätzung, doch bleibt die Funktion des Gesamtsystems erhalten. Umgekehrt können damit weitere Sensoren, sofern sie die gleiche Abstraktion verwenden, hinzugefügt werden.

Eine weitere Möglichkeit zur Kategorisierung von Fusionsansätzen ist die Reihenfolge, in der die Messdaten berücksichtigt werden. Um diese Differenzierung zu erläutern, soll zunächst der Ablauf einer rekursiven Zustandsschätzung, wie sie hier zum Einsatz kommt, erläutert werden (vgl. Bild 3.1).

Bevor die Schätzung des Systemzustandes eines verfolgten Objektes mit den Messwerten eines Sensors aktualisiert werden kann, muss ein zeitlicher Abgleich von Objektbeschreibung und Messungen erfolgen. Dazu wird der Systemzustand auf den Zeitpunkt prädiert, zu dem die Messwerte vorliegen.

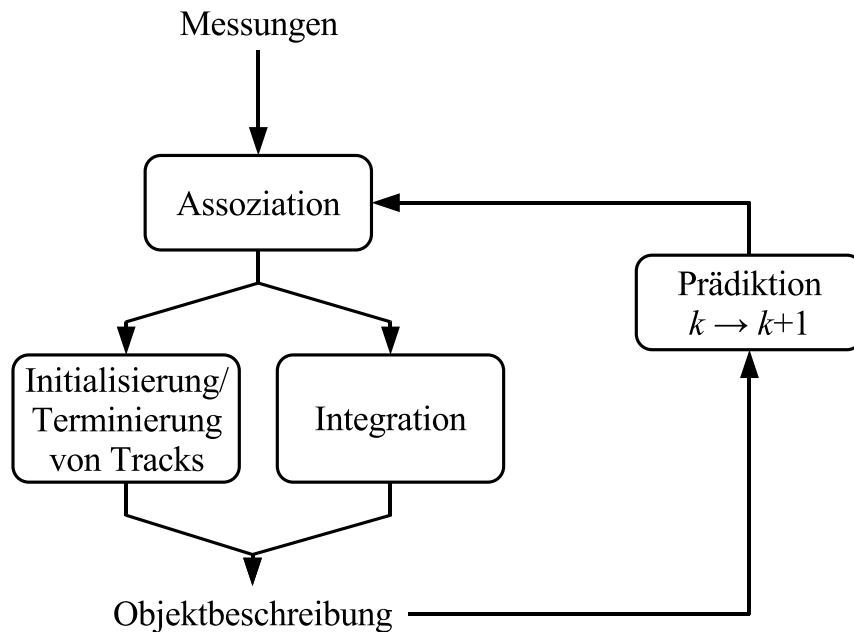


Bild 3.1: Ablauf der rekursiven Zustandsschätzung.

Dann werden bei der Assoziation den Objektbeschreibungen Messwerte zugeordnet. Mit diesen erfolgt die Integration, d. h. die Aktualisierung der Zustandsschätzung anhand der Messwerte. Konnten Messungen nicht assoziiert werden, können sie zur Initialisierung neuer Objektbeschreibungen, sogenannter Tracks, herangezogen werden. Liegen mehrere Objektbeschreibungen für ein Objekt vor oder kann ein Objekt nicht länger verfolgt werden, so muss eine Zusammenführung bzw. Terminierung erfolgen.

Werden die Messdaten mehrerer Sensoren in einem einzigen Iterationsschritt gleichzeitig berücksichtigt, so spricht man von *paralleler Fusion*. Dies bietet sich gerade bei der Verwendung synchroner Sensoren, also Sensoren, die zu gemeinsamen Zeitpunkten Daten liefern, an.

Liegen asynchrone Sensoren vor, so bietet sich eine *sequentielle Fusion* an. Die Messdaten mehrerer Sensoren werden dabei der Reihe nach verarbeitet.

Ogleich die in dieser Arbeit verwendeten Sensoren aufgrund der gemeinsamen Bildrohdaten als synchron anzusehen sind, liegen die Messdaten nicht zur gleichen Zeit vor. Da die Schritte zur Merkmalsextraktion unterschiedlich aufwändig sind, stehen die Messungen mit unterschiedlich großer Verzögerung zur Verfügung. Um dennoch schon möglichst früh eine korrigierte Schätzung zu erhalten, wurde eine sequentielle Verarbeitungsweise gewählt. Zudem erleichtert dies die Integration weiterer Sensoren.

Bei dieser Beschreibung wurde schon vorweggenommen, dass die Fusion eine zeitliche Verfolgung in Form einer rekursiven Zustandsschätzung vornimmt. Einen Überblick über verschiedene Ansätze dazu gibt der folgende Abschnitt.

3.1.2 Verfahren zur zeitlichen Verfolgung

Grundlage vieler gebräuchlichen Techniken zur rekursiven Zustandsschätzung ist das Bayes-Filter [Fox u. a. 2003]. Es ist ein Werkzeug zur Schätzung des Zustandes eines dynamischen Systems aus verrauschten Messwerten. Die Unsicherheit in der Schätzung des Zustandes $\mathbf{x}(k)$ des zu beobachtenden Systems zum Zeitpunkt k wird dabei mit einer in der Literatur oft Belief-Funktion genannten Wahrscheinlichkeitsdichtefunktion beschrieben:

$$\text{bel}(\mathbf{x}(k)) = p(\mathbf{x}(k)|\mathbf{z}^k) . \quad (3.1)$$

Dies ist die Wahrscheinlichkeitsdichte des Zustandes \mathbf{x} zum Zeitpunkt k bei gegebenen Beobachtungen \mathbf{z}^k . Mit \mathbf{z}^k sind alle Messungen bis zum Zeitpunkt k , einschließlich vorangegangener Messungen, bezeichnet.

Die Notation des Zeitpunktes k symbolisiert, dass eine zeitdiskrete Betrachtung durchgeführt wird. So bezeichnet Zeitschritt k im Folgenden den Zeitpunkt t_k .

Die Belief-Funktion wird in zwei Schritten aktualisiert [Thrun u. a. 2005]:

Prädiktion: In diesem Schritt wird die Belief-Funktion auf den nächsten Zeitschritt k prädiziert. Der vorangegangene Zeitschritt ist hier mit $k - 1$ gekennzeichnet:

$$\text{bel}^-(\mathbf{x}(k)) = \int p(\mathbf{x}(k)|\mathbf{x}(k-1)) \text{bel}(\mathbf{x}(k-1)) d\mathbf{x}(k-1) . \quad (3.2)$$

Die Wahrscheinlichkeitsdichte $p(\mathbf{x}(k)|\mathbf{x}(k-1))$ beschreibt die Dynamik des Systems, d. h. sie modelliert, welche Werte der Zustandsvektor zum Zeitpunkt k annimmt bei gegebenem Zustandsvektor des vorangegangenen Zeitpunktes. Dabei wird vorausgesetzt, dass das System die Markov-Bedingung erfüllt. Dies bedeutet, dass ein Zustand $\mathbf{x}(k)$ alle relevante Information enthält, d. h. die Kenntnis vorheriger Beobachtungen und Systemzustände liefert keine zusätzliche Information:

$$p(\mathbf{x}(k)|\mathbf{x}^{k-1}, \mathbf{z}^{k-1}) = p(\mathbf{x}(k)|\mathbf{x}(k-1)) . \quad (3.3)$$

Damit ist zur Prädiktion lediglich der vorangegangene Zustand $\mathbf{x}(k-1)$ erforderlich.

Korrektur: Nun wird die Schätzung mit den Messwerten korrigiert:

$$\text{bel}(\mathbf{x}(k)) = \alpha(k) p(z(k)|\mathbf{x}(k)) \text{bel}^-(\mathbf{x}(k)) . \quad (3.4)$$

Der Term $p(z(k)|\mathbf{x}(k))$ beschreibt die Eigenschaften des Sensors. Er gibt die Wahrscheinlichkeitsdichte für die Messung $z(k)$ an, vorausgesetzt, der Systemzustand ist $\mathbf{x}(k)$. Bei $\alpha(k)$ handelt es sich um einen Normierungsfaktor, der sich aus der Regel von Bayes ergibt. Er dient lediglich dazu, dass das Integral über die korrigierte Belief-Funktion wieder zu Eins wird.

Verschiedene bekannte Verfahren, wie z. B. das Partikel-Filter oder das Kalman-Filter, können als Varianten des Bayes-Filters aufgefasst werden. Sie unterscheiden sich vor allem durch die unterschiedlichen Modellierungen und Approximationen der Belief-Funktion und der Transitionsfunktion $p(\mathbf{x}(k)|\mathbf{x}(k-1))$. Hier kann man grundsätzlich nicht-parametrische und parametrische Ansätze unterscheiden [Fox u. a. 2003].

3.1.2.1 Nicht-parametrische Ansätze

Bei den nicht-parametrischen Ansätzen wird die Belief-Funktion diskret approximiert. Sie implizieren keine bekannte analytische Beschreibung der Verteilung des Systemzustandes und sind damit flexibel einsetzbar. Der Nachteil dieser Verfahren gerade bei mehrdimensionalen Zustandsvektoren liegt gewöhnlich im erhöhten Rechenaufwand. Im Folgenden werden zwei Vertreter dieser Gattung kurz vorgestellt.

Gitterbasierte Techniken

Eine Möglichkeit ist die Diskretisierung des kontinuierlichen Zustandsraumes in eine Gitterstruktur. Jedes Feld des Gitters enthält einen Belief-Wert. Damit lassen sich beliebige Verteilungen darstellen, wodurch dieses Filter für ein breites Spektrum an Anwendungen interessant wird. Nachteilig sind allerdings der hohe Rechenaufwand und Speicherplatzbedarf, die mit wachsender Anzahl der Dimensionen des Zustandsvektors schnell steigen.

Partikel-Filter

Eine deutliche Verbesserung diesbezüglich ist das Partikel-Filter. Anstatt mittels eines festen Gitters wird die Belief-Funktion hier durch n so genannte Partikel repräsentiert:

$$\text{bel}(\mathbf{x}(k)) \approx \left\{ \left\langle \mathbf{x}^{[i]}(k), w^{[i]}(k) \right\rangle \middle| i = 1, \dots, n \right\} . \quad (3.5)$$

Jedem Partikel $\mathbf{x}^{[i]}(k)$ ist ein Gewichtungsfaktor $w^{[i]}(k)$ zugeordnet, der die Wahrscheinlichkeitsdichte dieses Partikels beschreibt.

Gegenüber einer gitterbasierten Technik benötigt das Partikel-Filter deutlich weniger Ressourcen. Dennoch besteht auch hier das Problem, dass gerade bei hochdimensionalen Zustandsräumen ein immenser Rechenaufwand entsteht.

3.1.2.2 Parametrische Ansätze

Weit verbreitet sind Techniken mit parametrischer Darstellung. Dabei wird eine bekannte Verteilungsfunktion, häufig eine Gauß-Verteilung, angenommen. Lediglich deren Parameter stehen damit zur Modellierung der Belief-Funktion zur Verfügung. Damit benötigen parametrische Ansätze bedeutend weniger Ressourcen und werden daher gerade in Anwendungen, denen wenig Rechenleistung zur Verfügung steht, häufig eingesetzt. Allerdings stellen diese Verfahren sehr strikte Anforderungen an die Wahrscheinlichkeitsverteilung des beobachteten Zustandsvektors, weshalb ihre Anwendbarkeit im Einzelfall geprüft werden muss.

Kalman-Filter

Der wohl bekannteste Vertreter der parametrischen Ansätze ist das Kalman-Filter [Brown 1983]. Da seine Konzepte für Teile der in dieser Arbeit entwickelten Algorithmen relevant sind, soll es hier näher erläutert werden. Dabei wird im Gegensatz zur originären, verteilungsannahmefreien Herleitung, das Filter für stochastische Zustandsvariablen als Sonderfall eines Bayes-Filters betrachtet.

Die Belief-Funktion wird beim Kalman-Filter als mehrdimensionale Normalverteilung mit dem Erwartungswert $\hat{\mathbf{x}}(k)$ und der Kovarianz $\mathbf{P}(k)$ modelliert:

$$\text{bel}(\mathbf{x}(k)) \approx \frac{1}{\sqrt{\det(2\pi\mathbf{P}(k))}} \exp \left\{ -\frac{1}{2}(\mathbf{x}(k) - \hat{\mathbf{x}}(k))^T \mathbf{P}^{-1}(k)(\mathbf{x}(k) - \hat{\mathbf{x}}(k)) \right\} . \quad (3.6)$$

Diese Darstellung ergibt sich aus den System- und Sensoreigenschaften, die vorausgesetzt werden:

Systemmodell: Das beobachtete System ist durch eine lineare stochastische Differenzgleichung beschreibbar:

$$\mathbf{x}(k) = \mathbf{A}\mathbf{x}(k-1) + \mathbf{B}\mathbf{u}(k) + \mathbf{w}(k) . \quad (3.7)$$

Das zeitliche Verhalten des Systems, also der Übergang von $k - 1$ nach k , wird durch die Transitionsmatrix A beschrieben. Das Prozessrauschen w ist normalverteilt, mittelwertfrei und weiß mit der Kovarianz Q . Die Matrix B verbindet den Steuereingang u mit dem System.

Messmodell: Der Sensor bildet den Systemzustand x linear auf die Messungen z ab, wobei auch hier eine stochastische Störung berücksichtigt wird:

$$z(k) = Hx(k) + v(k) . \quad (3.8)$$

Die Abbildung wird von der Beobachtungsmatrix H beschrieben. Auch das Sensorrauschen v ist normalverteilt, mittelwertfrei und weiß mit der Kovarianz R . Zudem wird vorausgesetzt, dass es unabhängig vom Prozessrauschen w ist.

Natürlich bedeutet diese Modellierung eine signifikante Einschränkung des Einsatzbereiches, führt aber dazu, dass sowohl der Prädiktions- als auch der Korrekturschritt wieder normalverteilte Systemzustände ergeben.

Die Prädiktions- und Korrekturschritte des Kalman-Filters lassen sich sehr effizient berechnen [Welch u. Bishop 2001]:

Prädiktion:

$$\hat{x}^-(k) = A\hat{x}(k-1) + Bu(k-1) , \quad (3.9)$$

$$P^-(k) = AP(k-1)A^T + Q . \quad (3.10)$$

Hier werden zunächst der geschätzte Systemzustand und die zugehörige Kovarianz prädiziert. Diese beiden Parameter beschreiben die zugrunde gelegte Verteilung, der Schritt entspricht also der Bestimmung von $\text{bel}^-(x(k))$ in Gleichung (3.2).

Korrektur:

$$K(k) = P^-(k)H^T (HP^-(k)H^T + R)^{-1} , \quad (3.11)$$

$$\hat{x}(k) = \hat{x}^-(k) + K(k)(z(k) - H\hat{x}^-(k)) , \quad (3.12)$$

$$P(k) = (I - K(k)H)P^-(k) . \quad (3.13)$$

In Gleichung (3.11) wird zunächst die so genannte *Kalman-Verstärkung* K berechnet. Diese wird in den folgenden beiden Gleichungen benötigt, um die Prädiktionen zu korrigieren.

Auch für nichtlineare System- oder Messmodelle kann das Kalman-Filter erfolgreich angewandt werden. Hier kommen dann erweiterte Varianten zum Einsatz. So linearisiert das *Extended Kalman-Filter* zur Bestimmung der Kovarianzmatrix und der Kalman-Verstärkung die System- und Messgleichungen um die aktuelle Schätzung, wodurch die oben beschriebene Vorgehensweise erhalten bleibt. Eine andere Möglichkeit stellt das *Unscented Kalman-Filter* [Julier u. Uhlmann 1997] dar, das den Linearisierungsfehler noch weiter verkleinert. Dazu bedient es sich der *Unscented-Transformation*: Anstatt eine Gleichung direkt zu linearisieren und daraus die statistischen Parameter zu bestimmen, werden sog. Sigma-Punkte der Verteilung gewählt und mit dem nichtlinearen Zusammenhang transformiert. Die statistischen Parameter der gesuchten Verteilung werden dann so bestimmt, dass sie die transformierten Punkte beschreiben.

Obwohl die diesen Verfahren zugrunde liegende Modellierung der Wahrscheinlichkeitsverteilungen nur in Ausnahmefällen wirklich erfüllt ist, sind sie in vielen praktischen Anwendungen einsetzbar. Liegen System- und Messgleichungen vor, die linear oder gut linearisierbar sind, und sind das Prozess- und Messrauschen unimodal verteilt, so erweisen sich die Verfahren als unempfindlich gegenüber Modellverletzungen. Diese Voraussetzungen sind auch in der vorliegenden Arbeit erfüllt.

Gaußsche Mischverteilung

Mit der Modellierung als Gauß-Verteilung sind das Kalman-Filter und die erwähnten Varianten auf unimodale Belief-Funktionen beschränkt. Dies ist jedoch oft nicht ausreichend. Sollen z. B. die Positionen mehrerer Objekte im Raum anhand einer einzigen Verteilung beschrieben werden, so muss eine andere Darstellung gewählt werden. Auch wenn eine Beobachtung mehrere Interpretationsmöglichkeiten bietet, also den Systemzustand nicht eindeutig beschreibt, so ergeben sich multimodale Verteilungen. Eine Erweiterung für solche multimodale Verteilungen verwendet Mischungen von Gauß-Verteilungen:

$$\text{bel}(\mathbf{x}(k)) = \frac{1}{\sum_l \Theta_l(k)} \sum_l \frac{\Theta_l(k)}{\sqrt{|2\pi\mathbf{P}_l(k)|}} e^{-\frac{1}{2}(\mathbf{x}(k) - \hat{\mathbf{x}}_l(k))^T \mathbf{P}_l^{-1}(k) (\mathbf{x}(k) - \hat{\mathbf{x}}_l(k))} . \quad (3.14)$$

Bei den $\Theta_l(k)$ handelt es sich um Mischungsparameter, die anhand der Likelihoods, mit der die Messungen bei gegebener Verteilung l auftreten, ermittelt werden.

Neben dem höheren Rechenaufwand besteht die Schwierigkeit dieser Verfahren darin zu entscheiden, wann ein Mischkomponente hinzugefügt bzw. entfernt werden muss.

Probabilistic Data Association Filter

Während die Multi-Hypothesen-Verfolgung die Position mehrerer Objekte in Form einer einzigen, gemischten Verteilung beschreiben kann, geht das *Probabilistic Data Association*-Filter (PDAF) einen anderen Weg. Hier wird für jedes Objekt ein eigenes Filter benötigt. Für jedes Filter werden die Messungen, die innerhalb einer Validierungsregion liegen, gewichtet addiert. Die so erhaltene kombinierte Messung wird zum Aktualisieren des geschätzten Systemzustandes herangezogen [Bar-Shalom u. Fortmann 1988].

Das *Joint Probabilistic Data Association*-Filter (JPDAF) stellt eine Erweiterung des PDAF dar. Anstatt die Messungen anhand von Validierungsregionen auszuwählen, werden alle Messungen verwendet. Der Algorithmus bestimmt dabei die Relevanz jeder Messung auch unter Berücksichtigung der Assoziationen zu Filtern, die andere Objekte beschreiben.

Interacting Multiple Model-Filter

Bei der Beobachtung von Systemen mit wechselnden Systemmodellen werden häufig Instanzen eingesetzt, die Modellwechsel erkennen und eine Modellschaltung vornehmen. Allerdings ist das Erkennen eines Modellwechsels oft nicht trivial und wird von verrauschten Messungen gestört. Zudem werden dazu im Allgemeinen mehrere Messzyklen benötigt, was zu Verzögerungen führt, die negativen Einfluss auf die Schätzung haben.

Das *Interacting Multiple Model*-Filter (IMM) kennt diese Nachteile nicht. Es verfolgt alle möglichen Modelle in Form von einzelnen Filtern parallel und bestimmt die aktualisierte Schätzung aus den gewichteten modellspezifischen Schätzungen.

Nachteil des Verfahrens ist allerdings der erhöhte Rechenaufwand, da für jedes Modell ein eigenes Filter initialisiert werden muss.

3.1.3 Assoziation

Werden mehrere Objekte verfolgt, indem für jedes Objekt ein eigenes Filter initialisiert wird, so müssen die eintreffenden Messungen den entsprechenden Filtern zugeordnet werden. Dieser Schritt wird Assoziation genannt. Da hier die Entscheidung fällt, mit welchen Messungen die Schätzung des Systemzustandes aktualisiert wird, ist eine korrekte Assoziation maßgeblich für die Qualität der Ergebnisse.

3.1.3.1 Nächster Nachbar-Zuordnung

Eine eindeutige Zuordnung ist oft nicht möglich, da die verwendeten Sensoren fehlerbehaftet sind. Sie liefern einen nicht zu vernachlässigenden Anteil an falschen Hypothesen, der im Fusionsprozess wiederum zu Tracks führt, die kein reales Objekt beschreiben.

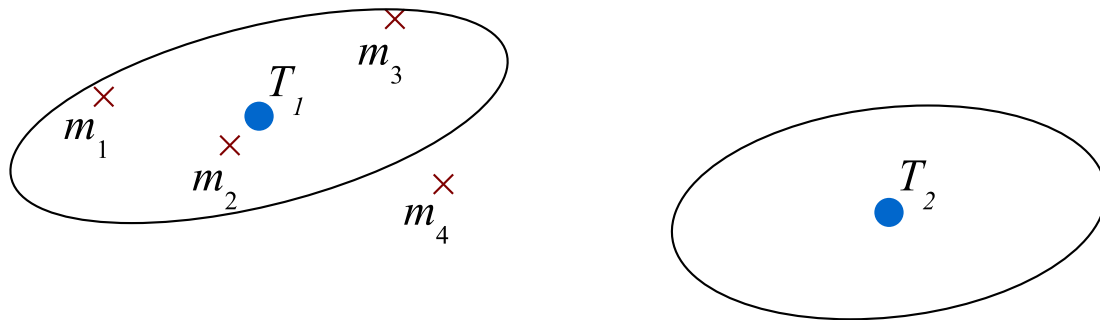


Bild 3.2: Tracks T_t , Messungen m_i und Validierungsregionen.

Bild 3.2 zeigt für den Fall eines zweidimensionalen Messvektors zwei Tracks T_1 und T_2 , denen neue Messungen $m_i, i \in 1, \dots, 4$ zugeordnet werden sollen. Eine einfache Möglichkeit ist, jedem Track die Messung zuzuordnen, die ihm am nächsten liegt [Dasarathy 1991]. Dabei sind jedoch die in den folgenden Abschnitten dargelegten Probleme zu berücksichtigen.

Objekt ohne Messung

Tritt für ein Objekt keine Messung auf (z. B. sensorbedingt oder weil es sich um einen Track handelt, der kein reales Objekt beschreibt), so würde eine unplausible Zuordnung erfolgen. Im vorliegenden Beispiel ergäben sich die Zuordnungen T_1/m_2 und T_2/m_4 , wobei letztere unerwünscht ist. Eine Lösung ist es, nur validierte Messungen zu berücksichtigen, d. h. Messungen, die in der Validierungsregion des entsprechenden Tracks liegen. In Bild 3.2 sind beispielhaft elliptische Validierungsregionen skizziert.

Fehlzuordnung aufgrund von Rauschen

Handelt es sich bei Messung m_2 um eine Fehlhypothese, so führt die binäre Entscheidung für den nächsten Nachbarn dazu, dass die möglicherweise validierte wahre Messung verworfen wird. Daher ist ein Assoziationsalgorithmus zu bevorzugen, der alle validierten Messungen einbezieht. Dies verringert den Einfluss von Störungen auf den Trackverlauf.

Unsicherheit bei der Zuordnung zu benachbarten Tracks

Befinden sich mehrere Tracks in unmittelbarer Nachbarschaft zueinander, so ergibt sich das in Bild 3.3 skizzierte Problem.

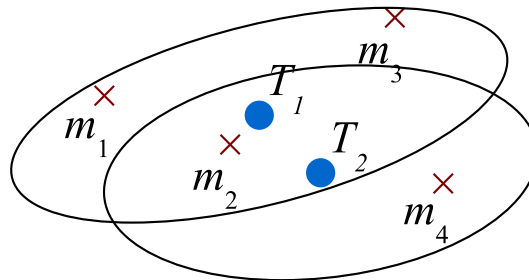


Bild 3.3: Tracks mit überlappenden Validierungsregionen.

Hier kann die Messung m_2 nicht sicher dem einen oder anderen Track zugeordnet werden. Auch dieser Umstand sollte vom Assoziationsalgorithmus berücksichtigt werden.

3.1.3.2 Multi-Hypothesen-Assoziation

Diesen Problemen begegnet die Multi-Hypothesen-Assoziation [Blake u. Yuille 1992]. Bei den Verfahren dieser Klasse wird nicht sofort eine harte Entscheidung für oder gegen eine Assoziation getroffen, sondern es werden mehrere Hypothesen parallel verfolgt.

Der ursprüngliche Algorithmus von Reid [Reid 1997] versucht nicht, die Beobachtungen mit vorhandenen Tracks zu assoziieren, sondern ordnet den Messungen Tracks zu. Jede mögliche Zuordnung wird dabei in einem Baum abgelegt und weiterverfolgt. Da dabei auch der Fall berücksichtigt wird, dass eine Messung von einem zuvor nicht erfassten Objekt ausgelöst wurde, ist eine separate Initialisierung von Tracks nicht notwendig.

Eine andere Variante arbeitet Track-orientiert: bestehen Mehrdeutigkeiten bei der Zuordnung von Messungen zu Tracks, so werden mehrere Hypothesen erzeugt. Die Entscheidung darüber, welche Hypothese weiterverfolgt wird, geschieht zu einem späteren Zeitschritt [Blackman u. Popoli 1999].

Allerdings sind diese Verfahren von sehr hoher Komplexität. Intelligente Strategien, um unwahrscheinliche Hypothesen früh aus dem Baum zu entfernen, sind entscheidend für eine sinnvolle Anwendung dieser Techniken.

3.1.3.3 Alle Nachbarn-Assoziation

Für viele Anwendungen stellen Techniken, die lediglich alle möglichen Zuordnungen des aktuellen Zeitschrittes berücksichtigen, eine gute Alternative dar. Sie sind deutlich robuster als Assoziationsverfahren, die nur einen Nachbarn berücksichtigen.

Bekannte Vertreter dieser Gattung sind das *Probabilistic Data Association*-Filter und Variationen. Diese Techniken finden auch in dieser Arbeit Anwendung und sind daher in Abschnitt 3.3 ausführlich erläutert.

3.2 Detektion und zeitliche Verfolgung von Fahrzeugen

Zur Detektion und Verfolgung von Fahrzeugen mittels der in den Kapiteln 2.3 und 2.4 beschriebenen monokularen Videomerkmale wurde eine auf die speziellen Anforderungen angepasste Fusionsarchitektur entwickelt.

Die folgenden beiden Abschnitte beschreiben zunächst diese Anforderungen, bevor auf die Architektur eingegangen wird. Eine detaillierte Beschreibung ist auch in [Hoffmann u. Dang 2006] zu finden.

3.2.1 Modellierung des Objektverhaltens

Wie bereits in Abschnitt 1.3.1 ausgeführt, ist für die Anwendungsbereiche, für die der vorgestellte Sensoraufbau konzipiert wurde, eine Beschreibung der Relativbewegung der beobachteten Fahrzeuge gegenüber dem Betrachter notwendig. Die Bewegung der beobachteten Fahrzeuge im Raum lässt sich aufgrund ihrer physikalischen Eigenschaften durch lineare stochastische Differentialgleichungen modellieren. Da sich auch der Beobachter in einem entsprechenden Fahrzeug befindet, lässt sich wiederum die Relativbewegung durch lineare stochastische Differentialgleichungen beschreiben.

Da die Verarbeitung der Sensordaten zeitdiskret geschieht, ergeben sich lineare stochastische Differenzgleichungen, wie sie auch schon in Abschnitt 3.1.2.2 beim Kalman-Filter eingeführt wurden:

$$\mathbf{x}(k) = \mathbf{A}(k)\mathbf{x}(k-1) + \mathbf{w}(k-1). \quad (3.15)$$

Hierbei bezeichnet \mathbf{x} den Zustandsvektor, \mathbf{A} ist die entsprechende Transitionsmatrix, bei \mathbf{w} handelt es sich um den stochastischen Anteil.

Bei einer üblichen Folgefahrt auf der Autobahn kann die Relativbewegung hinreichend genau als Bewegung mit konstanter Geschwindigkeit beschrieben werden. Zwar erscheint diese Annahme bei Betrachtung eines längeren Zeitraumes grob verletzt, doch ist für die Modellierung anhand der Differenzgleichung (3.15) entscheidend, ob sich die Geschwindigkeit in einem Zeitraum der Größenordnung eines Iterationsschrittes $\Delta t = t(k) - t(k-1)$ als quasi konstant betrachten lässt.

Aufgrund der Bildrate von 25 Hz, mit der die verwendete Kamera aufzeichnet, stehen auch die Messdaten mit $\Delta t = 40\text{ms}$ zur Verfügung. Damit lässt sich obige Annahme an einem Beispiel plausibilisieren: Der Beobachter bewege sich mit konstanter Geschwindigkeit, das beobachtete Fahrzeug beschleunigt konstant mit $1,5\text{m/s}^2$. Die Geschwindigkeitsänderung während eines Iterationsschrittes beträgt also

$$\Delta \dot{X} = 1,5\text{m/s}^2 \cdot 0,04\text{s} = 0,06\text{m/s} . \quad (3.16)$$

Diese Änderung kann noch durch den Rauschterm aufgefangen werden. Bei betragsmäßig größeren Beschleunigungen, wie sie z.B. beim Bremsen auftreten, liegt jedoch eine zu große Modellverletzung vor.

Experimentelle Ergebnisse zeigen, dass eine Kombination der folgenden beiden Modelle alle relevanten Situationen abdecken können:

Konstante Relativgeschwindigkeit: Mit diesem Modell wird ein Großteil aller auftretenden Situationen abgedeckt. Neben Folgefahrten werden z. B. auch entgegenkommende Fahrzeuge, die mit konstanter Geschwindigkeit fahren, erfasst.

Konstante Relativbeschleunigung: Manöver, die durch das erste Modell nicht ausreichend genau beschrieben sind, lassen sich hinreichend als Bewegungen mit konstanter Beschleunigung annähern. Dazu zählen vor allem Bremsmanöver und abrupte Spurwechsel.

Beide Modelle werden durch Gleichungen der Form von (3.15) beschrieben:

$$\mathbf{x}^n(k) = \mathbf{A}^n(k)\mathbf{x}^n(k-1) + \mathbf{w}^n(k-1) \quad (3.17)$$

Mit n wird nun das verwendete Modell bezeichnet, \mathbf{x}^n ist damit der Zustandsvektor des Modells n , \mathbf{A}^n ist die entsprechende Transitionsmatrix. Das Prozessrauschen \mathbf{w}^n wird als normalverteiltes, mittelwertfreies, weißes Rauschen mit der Kovarianz \mathbf{Q}^n angenommen.

Für das Modell der konstanten Relativgeschwindigkeit lautet der zugehörige Zustandsvektor \mathbf{x}^1

$$\mathbf{x}^1(k) = (X, Y, Z, \dot{X}, \dot{Y}, \dot{Z})^T . \quad (3.18)$$

Die zugehörige Transitionsmatrix A^1 , die den deterministischen Teil der Zustandsänderung während eines Zeitschrittes Δt beschreibt, ist

$$A^1(k) = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3.19)$$

Der Zeitschritt Δt muss nicht konstant sein, auf eine Notation als $\Delta t(k)$ wird hier jedoch verzichtet. Auch werden im Folgenden die Transitionsmatrix und Kovarianzmatrizen ohne das Argument k notiert.

Für das zweite Modell, also die Annahme konstanter Relativbeschleunigung, lauten Zustandsvektor und Transitionsmatrix entsprechend:

$$\mathbf{x}^2(k) = (X, Y, Z, \dot{X}, \dot{Y}, \dot{Z}, \ddot{X}, \ddot{Y}, \ddot{Z})^T, \quad (3.20)$$

$$A^2 = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3.21)$$

Obwohl in dieser Arbeit exemplarisch diese beiden Modelle zum Einsatz kommen, kann der Ansatz um beliebige Modelle erweitert werden. Daher wird im Weiteren stets auf n Modelle, $n \in N$, Bezug genommen.

3.2.2 Eigenschaften der Sensoren

Um aus den zweidimensionalen Videomeerkmalen eine dreidimensionale Positionsschätzung zu erhalten, wird die Entfernung eines Objektes mit Hilfe der Fahrbahnebene bestimmt. Diese Vorgehensweise wurde in Abschnitt 2.2 beschrieben.

Als Messwert erhält die Fusions- und Verfolgungsstufe daher einen Messvektor $\mathbf{z}(k)$ der Form

$$\mathbf{z}(k) = (X, Y, Z)^T, \quad (3.22)$$

der die Position des erfassten Objektes in 3D-Koordinaten bezüglich eines einheitlichen, beobachterfesten Koordinatensystems beschreibt.

Da auch die Zustandsvektoren der Einzelfilter die Objektposition in Weltkoordinaten zuzüglicher weiterer Ableitungen beschreiben, sind die Messgleichungen linear und haben die Form

$$\mathbf{z}(k) = \mathbf{H}^n \mathbf{x}^n(k) + \mathbf{v}(k) . \quad (3.23)$$

Auch hier wird für das Messrauschen $\mathbf{v}(k)$ vorausgesetzt, dass es normalverteilt, mittelwertfrei und weiß ist. Die zugehörige Kovarianzmatrix $\mathbf{R}(k)$ wird durch Gleichung (2.8) bereitgestellt. Die Beobachtungsmatrix \mathbf{H}^n beschreibt den Zusammenhang zwischen Zustandsvektor und Messung.

Die Ähnlichkeit der Mess- und Zustandsvektoren spiegelt sich auch in den Beobachtungsmatrizen wider. Die Beobachtungsmatrix für das Modell der konstanten Relativgeschwindigkeit lautet:

$$\mathbf{H}^1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} . \quad (3.24)$$

Da das Modell der konstanten Relativbeschleunigung das erste Modell um eine Ordnung erweitert, gleicht die Beobachtungsmatrix der aus (3.24), wird jedoch in den drei hinzugekommenen Spalten mit 0-Einträgen aufgefüllt:

$$\mathbf{H}^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} . \quad (3.25)$$

Erwähnt werden müssen noch einige Eigenschaften, die die verwendeten Video-merkmale aufweisen:

- Für ein Objekt wird maximal eine wahre Beobachtung generiert. Es kann jedoch nicht davon ausgegangen werden, dass ein vorhandenes Objekt immer eine Beobachtung verursacht.
- Aufgrund der Struktur der zu detektierenden Objekte treten gerade in der Umgebung der Fahrzeuge weitere Beobachtungen auf, die jedoch als falsch anzusehen sind. Ein Beispiel hierfür sind die Nummernschilder, die ähnliche geometrische Eigenschaften aufweisen wie ein weiter entferntes Fahrzeug.
- Eine Beobachtung ist entweder falsch oder wurde von genau einem Objekt ausgelöst.

Bei der zeitlichen Verfolgung muss diesen Besonderheiten Rechnung getragen werden.

3.2.3 Architektur der Sensordatenfusion

An und für sich schließt die Modellierung der Relativbewegung als Bewegung mit konstanter Beschleunigung den Fall der konstanten Relativgeschwindigkeit ein. Dabei handelt es sich lediglich um den Spezialfall, dass die Relativbeschleunigung Null ist:

$$\begin{pmatrix} \ddot{X} \\ \ddot{Y} \\ \ddot{Z} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (3.26)$$

Daher könnten alle beschriebenen Situationen durch ein einziges Modell abgedeckt werden. Modelle höherer Ordnung sind jedoch anfällig bei Fehldetektionen. Sie folgen diesen schnell, was letztlich zum Verlust des Tracks führt. Eine Kombination des robusten Modells niedriger Ordnung mit einem flexiblen Modell höherer Ordnung ist daher wünschenswert.

Das *Interacting Multiple Model-Filter (IMM)* [Bar-Shalom u. Blair 2000] macht dies möglich. Mehrere Modelle werden parallel verfolgt und danach beurteilt, wie gut sie die Messungen beschreiben. Dann wird unter Berücksichtigung dieser Güte ein gemeinsamer Zustandsvektor bestimmt. Verschiedenen Manövern kann also durch Bereitstellen entsprechender Modelle Rechnung getragen werden.

Das klassische *IMM-Filter* verwendet Kalman-Filter zur Verfolgung der einzelnen Modelle. Aufgrund der im letzten Abschnitt beschriebenen Eigenschaften der Sensoren müssen jedoch spezielle Anforderungen an die Assoziation der Beobachtungen zu den Objektrepräsentationen gestellt werden. Da für ein Objekt mehrere Beobachtungen auftreten können, von denen jedoch maximal eine wahr ist, führt eine harte Zuordnung einer einzigen Messung, wie sie die Nächste Nachbar-Assoziation vornimmt, nicht zum Ziel. Eine Multi-Hypothesen-Assoziation ist aufgrund der Vielzahl der Messungen pro Zeitschritt und der kurzen Zeitschritte extrem ressourcenintensiv. Die genannten Eigenschaften legen daher den Einsatz einer probabilistischen Assoziationstechnik nahe.

Das *Joint Probabilistic Data Association-Filter* berücksichtigt, dass im Allgemeinen gleichzeitig mehrere Fahrzeuge in einer Szene verfolgt werden. Daher wurde in dieser Arbeit ein Ansatz gewählt, der die Techniken des *Interacting Multiple Model-Filters* und des *Joint Probabilistic Data Association-Filters* vereint.

Bild 3.4 veranschaulicht die Zusammenhänge. Zunächst werden die Beobachtungen der angeschlossenen Sensoren, in diesem Fall der Schritte zur Merkmalsextraktion, anhand der Objektbeschreibungen validiert. Damit wird eine grobe Vorauswahl der zu berücksichtigenden Messungen getroffen.

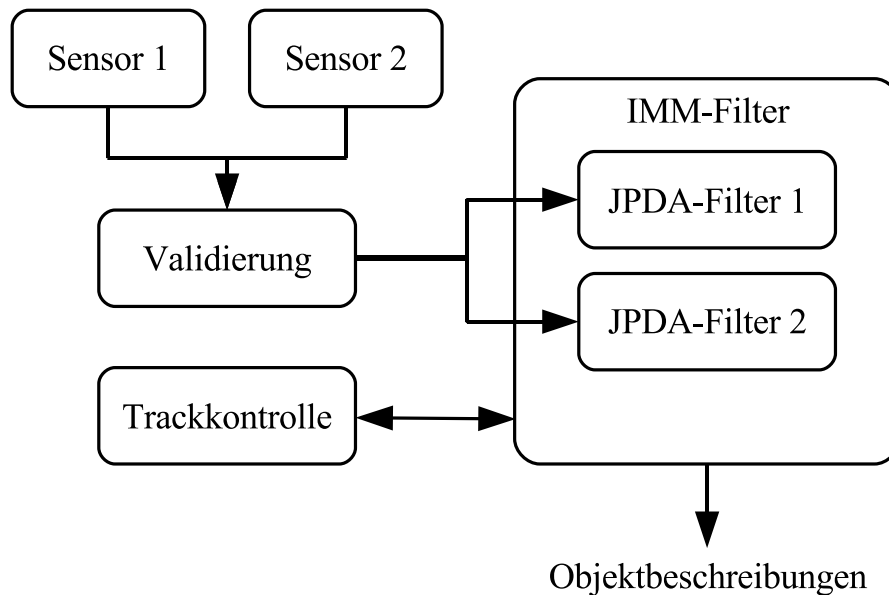


Bild 3.4: Überblick über die Fusion.

Zwei JPDA-Filter repräsentieren die beiden Systemmodelle. Die validierten Beobachtungen werden beiden Filtern übergeben, welche unabhängig voneinander eine vollständige Iteration durchführen.

Mittels der *Interacting Multiple Model*-Technik werden die modellbezogenen Filter gekoppelt. Daraus resultiert eine kombinierte Zustandsbeschreibung des Objektes.

Kann für eine Beobachtung keine geeignete Assoziation gefunden werden, so legt die Trackkontrolle eine neue Instanz eines IMM-Filters an. Zudem ist sie dafür verantwortlich, ungültige Tracks zu verwerfen. Dies wird z. B. notwendig, wenn ein beobachtetes Objekt den Erfassungsbereich des Sensors verlässt.

Abschnitt 3.3 geht nun zunächst auf die probabilistische Datenassoziation ein, bevor in Abschnitt 3.4 ihre Integration in die *Interacting Multiple Model*-Technik ausgeführt wird.

3.3 Probabilistische Datenassoziation

Bevor auf die Besonderheiten der gleichzeitigen Verfolgung mehrerer Objekte eingegangen wird, soll die grundlegende Vorgehensweise bei der probabilistischen Datenassoziation erläutert werden. Zunächst wird daher davon ausgegangen, dass lediglich ein einziges Objekt vorliegt, für das ein Filter initialisiert wurde [Bar-Shalom u. Fortmann 1988].

Ein Sensor liefert mehrere Beobachtungen, von denen maximal eine wahr ist. Eine vorgeschaltete Validierung wählt $M(k)$ davon aus, die für eine weitere Betrachtung in Frage kommen. Die Menge der validierten Beobachtungen sei

$$\mathbf{Z}(k) = \{z_m(k)\}, m \in M(k). \quad (3.27)$$

Die Menge aller Beobachtungen, die bis zum Zeitpunkt k einschließlich $\mathbf{Z}(k)$ aufgetreten ist, soll im Folgenden mit \mathbf{Z}^k bezeichnet werden.

Für die Analyse der möglichen Assoziationen werden die folgenden beiden Ereignisse betrachtet:

$\Theta_m(k)$: für $m = 1, \dots, M(k)$ das Ereignis, dass Beobachtung $z_m(k)$ wahr ist, also vom verfolgten Objekt ausgelöst wurde.

$\Theta_0(k)$: das Ereignis, dass keine Beobachtung wahr ist, es sich also bei allen um Fehldetektionen handelt.

Diese Ereignisse treten mit den Wahrscheinlichkeiten $\beta_m(k)$ auf:

$$\beta_m(k) = P(\Theta_m(k) | \mathbf{Z}^k). \quad (3.28)$$

Hier kann m auch den Wert 0 annehmen, womit die Wahrscheinlichkeit dafür gekennzeichnet wird, dass keine vom Objekt stammende Beobachtung aufgetreten ist. In der Ermittlung der Wahrscheinlichkeiten $\beta_m(k)$ liegt die eigentliche Assoziationsaufgabe, auf die in Abschnitt 3.3.2 eingegangen wird. Da eines der genannten Ereignisse eintreten muss und sie sich gegenseitig ausschließen, ist die Summe aller Wahrscheinlichkeiten Eins.

Der gemeinsame geschätzte Systemzustand kann damit nach dem Theorem der totalen Wahrscheinlichkeit bestimmt werden:

$$\hat{\mathbf{x}}(k) = E\{\mathbf{x}(k) | \mathbf{Z}^k\} = \sum_{m=0}^{M(k)} E\{\mathbf{x}(k) | \Theta_m(k), \mathbf{Z}^k\} P(\Theta_m(k) | \mathbf{Z}^k). \quad (3.29)$$

Beim ersten Term des Produkts handelt es sich um den geschätzten Systemzustand unter der Bedingung, dass Ereignis $\Theta_m(k)$ vorliegt. Damit erhält man die gemeinsame Schätzung aus der gewichteten Summe der Einzelschätzungen:

$$\hat{\mathbf{x}}(k) = \sum_{m=0}^{M(k)} \hat{\mathbf{x}}_m(k) \beta_m(k). \quad (3.30)$$

Der in [Bar-Shalom u. Fortmann 1988] beschriebene Weg über ein kombiniertes Residuum ist hier nicht zulässig, da nicht die Annahme getroffen werden kann, dass alle Beobachtungen mit der gleichen Kovarianz behaftet sind. Stattdessen wird die gemeinsame Schätzung aus den korrigierten Einzelschätzungen bestimmt. Da für jede Einzelschätzung vorausgesetzt wird, dass die entsprechende Beobachtung m wahr ist, kann sie analog zum Kalman-Filter (siehe Gleichungen (3.11) und (3.12)) bestimmt werden:

$$\mathbf{K}_m(k) = \mathbf{P}^-(k) \mathbf{H}^T (\mathbf{H} \mathbf{P}^-(k) \mathbf{H}^T + \mathbf{R}_m)^{-1}, \quad (3.31)$$

$$\hat{\mathbf{x}}_m(k) = \hat{\mathbf{x}}^-(k) + \mathbf{K}_m(k) (z_m(k) - \mathbf{H} \hat{\mathbf{x}}^-(k)). \quad (3.32)$$

Die hier auftretende prädizierte Schätzung des Systemzustandes und die zugehörige prädizierte Kovarianz ergeben sich analog zum Kalman-Filter wie in (3.9) und (3.10):

$$\hat{\mathbf{x}}^-(k) = \mathbf{A} \hat{\mathbf{x}}(k-1), \quad (3.33)$$

$$\mathbf{P}^-(k) = \mathbf{A} \mathbf{P}(k-1) \mathbf{A}^T + \mathbf{Q}. \quad (3.34)$$

Für den Fall Θ_0 , also den Fall, dass keine der vorhandenen Messungen wahr ist, ist der neue geschätzte Systemzustand gleich dem prädizierten:

$$\hat{\mathbf{x}}_0(k) = \hat{\mathbf{x}}^-(k) \quad (3.35)$$

Für die korrigierte Kovarianz der Schätzung ist eine gesonderte Betrachtung notwendig. Auch hier muss wieder berücksichtigt werden, dass sich die Kovarianzen der Beobachtungen unterscheiden, wodurch die Vorgehensweise des ursprünglichen PDA-Filters nicht übernommen werden kann.

Bei der Kovarianz handelt es sich um das zweite zentrale Moment der Verteilung von $\mathbf{x}(k)$:

$$\mathbf{P}(k) = \mathbb{E} \{ [\mathbf{x}(k) - \mathbb{E} \{ \mathbf{x}(k) \}] [\mathbf{x}(k) - \mathbb{E} \{ \mathbf{x}(k) \}]^T \} \quad (3.36)$$

$$= \mathbb{E} \{ [\mathbf{x}(k) - \hat{\mathbf{x}}(k)] [\mathbf{x}(k) - \hat{\mathbf{x}}(k)]^T \} \quad (3.37)$$

$$= \int [\mathbf{x}(k) - \hat{\mathbf{x}}(k)] [\mathbf{x}(k) - \hat{\mathbf{x}}(k)]^T p(\mathbf{x}(k)) d\mathbf{x}(k). \quad (3.38)$$

Unter Verwendung des Theorems der totalen Wahrscheinlichkeit ist die Wahrscheinlichkeitsdichte

$$p(\mathbf{x}) = \sum_{m=0}^M p(\mathbf{x}|\Theta_m) \cdot P(\Theta_m) = \sum_{m=0}^M p(\mathbf{x}_m) \cdot \beta_m, \quad (3.39)$$

wobei hier wie im Folgenden zur besseren Lesbarkeit die Angabe des Zeitschrittes k weggelassen wird. In (3.38) ergibt dies

$$\mathbf{P} = \int \sum_{m=0}^M [\mathbf{x} - \hat{\mathbf{x}}] [\mathbf{x} - \hat{\mathbf{x}}]^T p(\mathbf{x}_m) \cdot \beta_m d\mathbf{x} \quad (3.40)$$

$$= \sum_{m=0}^M \int [\mathbf{x} - \hat{\mathbf{x}}] [\mathbf{x} - \hat{\mathbf{x}}]^T p(\mathbf{x}_m) d\mathbf{x} \cdot \beta_m \quad (3.41)$$

$$= \sum_{m=0}^M E\{[\mathbf{x} - \hat{\mathbf{x}}] [\mathbf{x} - \hat{\mathbf{x}}]^T | \Theta_m\} \cdot \beta_m \quad (3.42)$$

$$= \sum_{m=0}^M E\{\mathbf{x}\mathbf{x}^T - \mathbf{x}\hat{\mathbf{x}}^T - \hat{\mathbf{x}}\mathbf{x}^T + \hat{\mathbf{x}}\hat{\mathbf{x}}^T | \Theta_m\} \cdot \beta_m \quad (3.43)$$

$$= \sum_{m=0}^M \{E\{\mathbf{x}\mathbf{x}^T | \Theta_m\} - E\{\mathbf{x}\hat{\mathbf{x}}^T | \Theta_m\} - E\{\hat{\mathbf{x}}\mathbf{x}^T | \Theta_m\} + E\{\hat{\mathbf{x}}\hat{\mathbf{x}}^T | \Theta_m\}\} \cdot \beta_m \quad (3.44)$$

Im Folgenden werden die Erwartungswerte aus Gleichung (3.44) einzelnen betrachtet.

$$E\{\mathbf{x}\mathbf{x}^T | \Theta_m\} = E\{[\hat{\mathbf{x}}_m + (\mathbf{x} - \hat{\mathbf{x}}_m)][\hat{\mathbf{x}}_m + (\mathbf{x} - \hat{\mathbf{x}}_m)]^T | \Theta_m\} \quad (3.45)$$

$$= E\{\hat{\mathbf{x}}_m \hat{\mathbf{x}}_m^T + \hat{\mathbf{x}}_m (\mathbf{x} - \hat{\mathbf{x}}_m)^T + (\mathbf{x} - \hat{\mathbf{x}}_m) \hat{\mathbf{x}}_m^T + (\mathbf{x} - \hat{\mathbf{x}}_m)(\mathbf{x} - \hat{\mathbf{x}}_m)^T | \Theta_m\} \quad (3.46)$$

Da es sich bei $\hat{\mathbf{x}}_m$ um die Schätzung des Systemzustandes \mathbf{x} handelt, gilt $E\{\hat{\mathbf{x}}_m\} = \hat{\mathbf{x}}_m$. Damit ergibt sich

$$E\{\mathbf{x}\mathbf{x}^T | \Theta_m\} = \hat{\mathbf{x}}_m \hat{\mathbf{x}}_m^T + \hat{\mathbf{x}}_m E\{\mathbf{x} - \hat{\mathbf{x}}_m | \Theta_m\}^T + E\{\mathbf{x} - \hat{\mathbf{x}}_m | \Theta_m\} \hat{\mathbf{x}}_m^T + E\{(\mathbf{x} - \hat{\mathbf{x}}_m)(\mathbf{x} - \hat{\mathbf{x}}_m)^T | \Theta_m\} \quad (3.47)$$

Ist das zugrunde liegende Kalman-Filter erwartungstreu, d. h. $\hat{\mathbf{x}}_m = E\{\mathbf{x} | \Theta_m\}$, so gilt

$$E\{\mathbf{x} - \hat{\mathbf{x}}_m | \Theta_m\} = E\{\mathbf{x} | \Theta_m\} - E\{\hat{\mathbf{x}}_m | \Theta_m\} = \hat{\mathbf{x}}_m - \hat{\mathbf{x}}_m = 0. \quad (3.48)$$

Voraussetzung für die Erwartungstreue ist zum einen, dass Messung m wahr ist. Da der Erwartungswert unter der Bedingung Θ_m betrachtet wird, ist dies gegeben.

Zudem muss das Systemmodell das Objektverhalten korrekt beschreiben. Für die Anwendung in einem *Interacting Multiple Model*-Ansatz trifft dies jedoch nicht zu, da hier mehrere Modelle parallel zum Einsatz kommen, die unterschiedliches Objektverhalten modellieren. Hier nimmt man, im Gegensatz zum einfachen Kalman-Filter, explizit in Kauf, dass das Objektverhalten nicht durch alle Systemmodelle korrekt beschrieben wird. Aufgabe des *IMM* ist es dann, das beste Modell durch entsprechende Gewichtung zu bevorzugen. Dabei behalten jedoch alle anderen Modelle einen Einfluss, auch wenn dieser geringer ausfällt.

Dennoch wird hier für alle Filter Erwartungstreue vorausgesetzt. Dies deckt sich mit den Annahmen, die auch das *Interacting Multiple Model*-Filter trifft. Da es eine Kombination der Systemzustände der Einzelfilter bestimmt, ist auch für das *IMM*-Filter streng genommen Erwartungstreue nicht gegeben.

Mit Gleichung (3.48) kann Gleichung (3.47) weiter vereinfacht werden:

$$\mathbb{E}\{\mathbf{x}\mathbf{x}^T|\Theta_m\} = \hat{\mathbf{x}}_m\hat{\mathbf{x}}_m^T + \mathbb{E}\{(\mathbf{x} - \hat{\mathbf{x}}_m)(\mathbf{x} - \hat{\mathbf{x}}_m)^T|\Theta_m\} \quad (3.49)$$

$$= \hat{\mathbf{x}}_m\hat{\mathbf{x}}_m^T + \mathbf{P}_m \quad (3.50)$$

Für den Fall Θ_0 , d. h. keine der Messungen ist wahr, ergibt sich entsprechend zu Gleichung (3.35)

$$\mathbf{P}_0(k) = \mathbf{P}^-(k), \quad (3.51)$$

für Θ_m mit $m \neq 0$ wird eine Kalman-Aktualisierung der Kovarianz durchgeführt:

$$\mathbf{P}_m(k) = [\mathbf{I} - \mathbf{K}_m\mathbf{H}] \mathbf{P}^-(k), \quad m = 1, \dots, M. \quad (3.52)$$

Nun soll der zweite Term in (3.44) untersucht werden.

$$\mathbb{E}\{\mathbf{x}\hat{\mathbf{x}}^T|\Theta_m\} = \mathbb{E}\{\mathbf{x}|\Theta_m\}\hat{\mathbf{x}}^T \quad (3.53)$$

$$= \hat{\mathbf{x}}_m\hat{\mathbf{x}}_m^T \quad (3.54)$$

Der dritte Term wird analog zu (3.54) umgeformt:

$$\mathbb{E}\{\hat{\mathbf{x}}\mathbf{x}^T|\Theta_m\} = \hat{\mathbf{x}}\mathbb{E}\{\mathbf{x}^T|\Theta_m\} \quad (3.55)$$

$$= \hat{\mathbf{x}}\hat{\mathbf{x}}_m^T \quad (3.56)$$

Die Erwartungswertbildung des vierten Terms reduziert sich einfach zu

$$E\{\hat{\mathbf{x}}\hat{\mathbf{x}}^T|\Theta_m\} = \hat{\mathbf{x}}\hat{\mathbf{x}}^T. \quad (3.57)$$

Setzt man diese Teilergebnisse nun wieder in (3.44) ein, so erhält man

$$\mathbf{P} = \sum_{m=0}^M \left\{ \hat{\mathbf{x}}_m \hat{\mathbf{x}}_m^T + \mathbf{P}_m - \hat{\mathbf{x}}_m \hat{\mathbf{x}}^T - \hat{\mathbf{x}} \hat{\mathbf{x}}_m^T + \hat{\mathbf{x}} \hat{\mathbf{x}}^T \right\} \cdot \beta_m \quad (3.58)$$

$$= \sum_{m=0}^M \left\{ \hat{\mathbf{x}}_m \hat{\mathbf{x}}_m^T + \mathbf{P}_m \right\} \cdot \beta_m - \left[\sum_{m=0}^M \hat{\mathbf{x}}_m \cdot \beta_m \right] \hat{\mathbf{x}}^T - \hat{\mathbf{x}} \sum_{m=0}^M \hat{\mathbf{x}}_m^T \cdot \beta_m + \hat{\mathbf{x}} \hat{\mathbf{x}}^T \quad (3.59)$$

$$= \sum_{m=0}^M \left\{ \hat{\mathbf{x}}_m \hat{\mathbf{x}}_m^T + \mathbf{P}_m \right\} \cdot \beta_m - \hat{\mathbf{x}} \hat{\mathbf{x}}^T - \hat{\mathbf{x}} \hat{\mathbf{x}}^T + \hat{\mathbf{x}} \hat{\mathbf{x}}^T \quad (3.60)$$

Da sich die beiden letzten Terme aufheben, ergibt sich für die korrigierte Kovarianz der Schätzung

$$\mathbf{P} = \sum_{m=0}^M \left\{ \hat{\mathbf{x}}_m \hat{\mathbf{x}}_m^T + \mathbf{P}_m \right\} \cdot \beta_m - \hat{\mathbf{x}} \hat{\mathbf{x}}^T. \quad (3.61)$$

Auf die Bestimmung der Assoziationswahrscheinlichkeiten $\beta_m(k)$ für den Fall, dass nur ein Objekt vorliegt, wird hier nicht näher eingegangen. Stattdessen wird der hier relevante Fall mehrerer verfolgter Objekte untersucht. Zunächst wird jedoch noch die oben schon erwähnte Validierung der Beobachtungen beschrieben.

3.3.1 Validierung der Hypothesen

Das Beispiel aus Abschnitt 3.1.3.1 veranschaulicht die Bedeutung der Hypothesenvalidierung für eine sinnvolle Assoziation. Entscheidend ist dabei die Wahl der Validierungsregion. Zu berücksichtigen sind hierbei:

- **Messunsicherheit:** Da Messungen, die mit einer hohen Unsicherheit behaftet sind, eine größere Distanz zum erfassten Objekt aufweisen können, sollte die Ausdehnung der Validierungsregion von der Messunsicherheit abhängen.
- **Kovarianz der Schätzung:** Da auch die geschätzte von der tatsächlichen Objektposition abweicht, ist die Kovarianz der Schätzung mit einzubeziehen.

Eine verbreitete Vorgehensweise, die die Wahrscheinlichkeit der Zuordnung einer Messung zu einem Track berücksichtigt, wird in [Bar-Shalom u. Blair 2000] beschrieben. Um einen Track mit einer Messung vergleichen zu können, muss zunächst ein zeitlicher Abgleich erfolgen. Dies geschieht durch Prädiktion des geschätzten Systemzustandes und seiner Kovarianz:

$$\hat{\mathbf{x}}^-(k) = \mathbf{A} \hat{\mathbf{x}}(k-1) \quad (3.62)$$

$$\mathbf{P}^-(k) = \mathbf{A} \mathbf{P}(k-1) \mathbf{A}^T + \mathbf{Q} \quad (3.63)$$

Daraus ergibt sich die prädizierte Messung

$$\hat{\mathbf{z}}^-(k) = \mathbf{H} \hat{\mathbf{x}}^-(k), \quad (3.64)$$

um die die Validierungsregion gelegt werden soll.

Der Abstand zwischen prädizierter und tatsächlicher Messung wird durch das Residuum beschrieben:

$$\mathbf{v}(k) = \mathbf{z}(k) - \mathbf{H} \hat{\mathbf{x}}^-(k). \quad (3.65)$$

Die Kovarianz des Residuums ergibt sich zu

$$\mathbf{S}(k) = \mathbf{H} \mathbf{P}^-(k) \mathbf{H}^T + \mathbf{R}. \quad (3.66)$$

Es wird angenommen, dass das Residuum normalverteilt ist mit

$$p(\mathbf{v}(k)) = \frac{1}{\sqrt{\det(2\pi\mathbf{S}(k))}} \exp\left\{-\frac{1}{2}\mathbf{v}^T(k)\mathbf{S}^{-1}(k)\mathbf{v}(k)\right\}. \quad (3.67)$$

Die Mahalanobis-Distanz

$$d^2 = \mathbf{v}^T(k)\mathbf{S}^{-1}(k)\mathbf{v}(k) \quad (3.68)$$

unterliegt also einer χ^2 -Verteilung mit drei Freiheitsgraden, was der Dimension des Messvektors entspricht. Unter Vorgabe einer geforderten minimalen Zuordnungswahrscheinlichkeit P_Z wird nun eine maximale Distanz d_{\max} so bestimmt, dass gilt:

$$P(d \leq d_{\max}) = P_Z. \quad (3.69)$$

Zur Validierung einer Hypothese muss dann

$$d \leq d_{\max} \quad (3.70)$$

erfüllt sein.

Diese Bedingung beschreibt ein Ellipsoid um die prädizierte Messung. Liegt eine Beobachtung innerhalb dieses Ellipsoids, so gilt sie als validiert.

3.3.2 Joint Probabilistic Data Association-Filter

Im Gegensatz zum bisher vorgestellten *Probabilistic Data Association Filter* berücksichtigt das *Joint Probabilistic Data Association-Filter* [Bar-Shalom u. Fortmann 1988] bei der Bestimmung der Assoziationswahrscheinlichkeiten $\beta_m(k)$, dass T Objekte gleichzeitig vorhanden sind. Bei der Untersuchung der Ereignisse $\Theta_m(k)$ muss daher auch das betrachtete Objekt $t = 1, \dots, T$ mit einbezogen werden. Im Folgenden werden sie daher definiert als

$\Theta_{mt}(k)$: Das Ereignis, dass Beobachtung m von Objekt t ausgelöst wurde. Hierbei ist mit $m = 1, \dots, M(k)$ eine der $M(k)$ aufgetretenen Beobachtungen bezeichnet. Das Objekt wird mit $t = 0, \dots, T$ indiziert, wobei $t = 0$ bedeutet, dass die Beobachtung von keinem Objekt verursacht, d. h. durch Fehldetektionen hervorgerufen wurde.

Damit können die Verbundereignisse $\Theta(k)$ definiert werden:

$$\Theta(k) = \bigcap_{m=0}^{M(k)} \Theta_{mt_m}(k) \quad (3.71)$$

Das Verbundereignis $\Theta(k)$ repräsentiert also eine mögliche Kombination von Assoziationen aller Messungen zu jeweils einem zugeordneten Objekt t_m . Ist ein $t_m = 0$, so bedeutet dies, dass es sich bei der Beobachtung m um eine Fehldetektion handelt.

Um die Auswertung der Zuordnungen handhabbar zu machen, wird bei der Berechnung der Wahrscheinlichkeiten der Einfluss der Validierung vernachlässigt. Dies bedeutet, dass keine beschränkte Umgebung betrachtet wird, wodurch keine Beschneidung der Gauß-Verteilungen der Residuen beachtet werden muss. Da eine Berücksichtigung aller Beobachtungen für einen Track jedoch zu einem erhöhten Rechenaufwand führen würde, wird eine Validierung bei der Auswahl der untersuchten Ereignisse berücksichtigt. Sind die Validierungsregionen so gewählt, dass lediglich Beobachtungen, die einen vernachlässigbar kleinen Einfluss auf die Schätzung haben, ausgeschlossen werden, so ist auch der dadurch entstehende Fehler vernachlässigbar.

Die Validierung der Beobachtungen dient also dazu, mögliche Assoziationen zu bestimmen. Die zu untersuchenden Assoziationen werden dann aus der Menge der möglichen gewählt. Dazu wird zunächst die Validierungsmatrix Ω bestimmt. Deren Zeilennummer gibt die Beobachtung an, die Spaltennummer steht für ein Objekt. Enthält das entsprechende Element eine Eins, so handelt es sich um eine für dieses Objekt validierte Beobachtung. Zu beachten ist, dass die erste Spalte

das Objekt $t = 0$ repräsentiert, also den Fall, dass die Messung von keinem Objekt stammt. Damit enthält die erste Spalte der Validierungsmatrix in jedem Element eine Eins. Dies wird an dem Beispiel mit zwei Objekten und vier Beobachtungen veranschaulicht, das schon in Bild 3.3 dargestellt wurde (im Folgenden wird k weggelassen):

$$\mathbf{\Omega} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}. \quad (3.72)$$

Beobachtung $m = 1$ ist also z. B. für Objekt $t = 1$ validiert, während Beobachtung $m = 2$ für beide Objekte validiert ist. Beobachtung $m = 3$ ist für Objekte $t = 1$ und Beobachtung $m = 4$ schließlich für Objekt $t = 2$ validiert. Da es sich bei allen Beobachtungen auch um Fehlhypothesen handeln kann, sind alle für das spezielle Objekt $t = 0$ validiert.

Ein Verbundereignis Θ kann analog zu dieser Vorgehensweise auch als Matrix $\hat{\mathbf{\Omega}}(\Theta)$ mit den Elementen $\hat{\omega}_{mt}$ dargestellt werden:

$$\hat{\mathbf{\Omega}}(\Theta) = [\hat{\omega}_{mt}]. \quad (3.73)$$

Ist eine mögliche Assoziation Θ_{mt} im Verbundereignis Θ enthalten, so ist das entsprechende Matrixelement $\hat{\omega}_{mt} = 1$, andernfalls ist es 0. So beschreibt

$$\hat{\mathbf{\Omega}}(\Theta_l) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.74)$$

folgendes Verbundereignis Θ_l :

- Beobachtung 1 ist eine Fehldetektion, d. h. wurde von Objekt 0 ausgelöst,
- Beobachtung 2 wurde von Objekt 1 ausgelöst,
- Beobachtung 3 ist eine Fehldetektion, und
- Beobachtung 4 stammt von Objekt 2.

Die Validierungsmatrix $\mathbf{\Omega}$ dient als Hilfsmittel zur Auswahl der $\hat{\mathbf{\Omega}}(\Theta)$. Dazu wird ein $\hat{\mathbf{\Omega}}(\Theta)$ so gewählt, dass in $\mathbf{\Omega}$ in jeder Zeile alle Elemente bis auf eines 0 gesetzt sind. Zudem darf in jeder Spalte höchstens ein Element 1 sein. Lediglich in der

ersten Spalte, die Objekt 0 beschreibt, können mehrere Elemente gesetzt sein, da mehrere Fehldetektionen gleichzeitig auftreten können.

Für die Ableitung der Assoziationswahrscheinlichkeiten werden noch einige Abkürzungen definiert.

Wurde ein Objekt t beobachtet, d. h. wurde eine Beobachtung mit ihm assoziiert, so ist $\delta_t(\Theta) = 1$, andernfalls 0:

$$\delta_t(\Theta) = \sum_{m=1}^M \hat{\omega}_{mt}(\Theta), \quad t = 1, \dots, T \quad (3.75)$$

Entsprechend gibt $\tau_m(\Theta)$ an, ob der Beobachtung ein Objekt zugeordnet wurde:

$$\tau_m(\Theta) = \sum_{t=1}^T \hat{\omega}_{mt}(\Theta), \quad m = 1, \dots, M \quad (3.76)$$

Die Anzahl der Beobachtungen, denen kein Objekt zugeordnet wird, bestimmt sich damit zu

$$\phi(\Theta) = \sum_{m=1}^M (1 - \tau_m(\Theta)) . \quad (3.77)$$

Um die Assoziationswahrscheinlichkeiten β_{mt} zu erhalten, müssen zunächst die Wahrscheinlichkeit für die Verbundereignisse bestimmt werden:

$$P(\Theta(k)|Z^k) = P(\Theta(k)|Z(k), Z^{k-1}) . \quad (3.78)$$

Mit der Regel von Bayes lässt sich Gleichung (3.78) umformulieren:

$$P(\Theta(k)|Z(k), Z^{k-1}) = \frac{p(Z(k)|\Theta(k), Z^{k-1}) P(\Theta(k)|Z^{k-1})}{p(Z(k)|Z^{k-1})} . \quad (3.79)$$

Da der Nenner von $\Theta(k)$ unabhängig ist und das Verbundereignis nicht von zuvor aufgetretenen Beobachtungen abhängt, ergibt sich

$$P(\Theta(k)|Z^k) = \frac{1}{c} p(Z(k)|\Theta(k), Z^{k-1}) P(\Theta(k)) . \quad (3.80)$$

Zunächst wird der Term für die Wahrscheinlichkeitsdichte der Beobachtungen $Z(k)$ betrachtet. Diese ist das Produkt der Wahrscheinlichkeitsdichten der einzelnen Beobachtungen, die als unabhängig voneinander vorausgesetzt werden:

$$p(Z(k)|\Theta(k), Z^{k-1}) = \prod_{m=1}^{M(k)} p(z_m(k)|\Theta_{mt_m}(k), Z^{k-1}) . \quad (3.81)$$

Tatsächlich können bei den eingesetzten Verfahren zur Merkmalsextraktion Fehldetektionen als voneinander unabhängig angesehen werden. Zwar lösen tatsächlich vorhandene Fahrzeuge aufgrund von Substrukturen, wie z. B. Rücklichtern und Nummernschildern, gelegentlich Fehlhypothesen aus. Hier besteht demnach eine Abhängigkeit. Allerdings erscheinen diese Fehlhypothesen als weiter entfernte Objekte (vgl. Abschnitt 2.4.3) und werden daher nicht für den wahren Track validiert. Die Annahme der Unabhängigkeit der Beobachtungen ist daher hinreichend gut erfüllt.

Wurde Beobachtung m von einem Objekt ausgelöst, d. h. $t_m \neq 0$, so ist der Messwert entsprechend der Voraussetzungen normalverteilt um die prädierte Messung \hat{z}^{-t_m} . Beobachtungen, die durch Rauschen verursacht wurden, werden als gleichverteilt über das Messvolumen V angenommen:

$$p\left(\mathbf{z}_m(k)|\Theta_{m t_m}(k), \mathbf{Z}^{k-1}\right) = \begin{cases} \mathcal{N}\left\{\mathbf{z}_m(k); \hat{\mathbf{z}}^{-t_m}, \mathbf{S}^{t_m}(k)\right\} & \text{falls } \tau_m(\Theta(k)) = 1 \\ \frac{1}{V} & \text{falls } \tau_m(\Theta(k)) = 0 \end{cases} \quad (3.82)$$

Das Messvolumen V umschließt den Erfassungsbereich des Sensors. Die Annahme gleichverteilter Fehldetektionen bringt zum Ausdruck, dass keine Information darüber vorhanden ist, wo diese bevorzugt auftreten.

Mit (3.82) in (3.81) ergibt sich

$$p\left(\mathbf{Z}(k)|\Theta(k), \mathbf{Z}^{k-1}\right) = \frac{1}{V^{\phi(\Theta)}} \prod_{m=1}^{M(k)} \left[\mathcal{N}\left\{\mathbf{z}_m(k); \hat{\mathbf{z}}^{-t_m}, \mathbf{S}^{t_m}(k)\right\} \right]^{\tau_m(\Theta)}. \quad (3.83)$$

Nun muss noch der zweite Term aus (3.80), $P(\Theta(k))$, bestimmt werden. Dazu wird der Vektor $\delta(\Theta)$ eingeführt, der aus den T Elementen $\delta_t(\Theta)$ besteht. Er gibt also an, welchen Objekten eine Beobachtung zugeordnet wurde. Für das oben in (3.74) aufgeführte Beispiel lautet er

$$\delta(\Theta_I) = (1, 1)^T, \quad (3.84)$$

da beide Objekte eine Beobachtung auslösten.

Da das Verbundereignis $\Theta(k)$ die Zuordnungen festlegt, sind damit sowohl $\delta(\Theta)$ als auch die Anzahl der falschen Beobachtungen $\phi(\Theta)$ beschrieben. Damit gilt

$$P(\Theta(k)) = P(\Theta(k), \delta(\Theta), \phi(\Theta)). \quad (3.85)$$

Dann lässt sich damit formulieren:

$$P(\Theta(k)) = P(\Theta(k)|\delta(\Theta), \phi(\Theta)) P(\delta(\Theta), \phi(\Theta)). \quad (3.86)$$

Der erste Term, $P(\Theta(k)|\delta(\Theta), \phi(\Theta))$, gibt die Wahrscheinlichkeit für eine bestimmte Kombination von Assoziationen an unter der Voraussetzung, dass bekannt ist,

- welche Objekte detektiert wurden und
- wie viele Messungen durch Rauschen verursacht wurden.

Diese Wahrscheinlichkeit ist, da alle Zuordnungen gleich wahrscheinlich sind, der Kehrwert der Anzahl möglicher Kombinationen der Zuordnung der Beobachtungen zu diesen detektierten Objekten. Da von M Beobachtungen $M - \phi(\Theta)$ einem Objekt zuzuordnen sind, ergeben sich $\frac{M!}{\phi(\Theta)!}$ Kombinationen. Damit wird der erste Term

$$P(\Theta(k)|\delta(\Theta), \phi(\Theta)) = \left(\frac{M!}{\phi(\Theta)!} \right)^{-1} = \frac{\phi(\Theta)!}{M!}. \quad (3.87)$$

Der zweite Term aus Gleichung (3.86) setzt sich zusammen aus der Wahrscheinlichkeit für das Auftreten von $\phi(\Theta(k))$ falschen Beobachtungen und der Wahrscheinlichkeit für die Detektion bzw. Nicht-Detektion der durch $\delta(\Theta(k))$ beschriebenen Objekte:

$$P(\delta(\Theta), \phi(\Theta)) = P(\phi(\Theta(k))) \prod_{t=1}^T (P_D^t)^{\delta_t(\Theta)} (1 - P_D^t)^{1-\delta_t(\Theta)} \quad (3.88)$$

Bei P_D^t handelt es sich um die Wahrscheinlichkeit für die Detektion von Objekt t , $1 - P_D^t$ ist damit die Wahrscheinlichkeit, keine Beobachtung von diesem Objekt zu erhalten.

Setzt man nun (3.88) und (3.87) in (3.86) ein, so erhält man

$$P(\Theta(k)) = \frac{\phi(\Theta)!}{M!} P(\phi(\Theta(k))) \prod_{t=1}^T (P_D^t)^{\delta_t(\Theta)} (1 - P_D^t)^{1-\delta_t(\Theta)}. \quad (3.89)$$

Damit lässt sich auch die Wahrscheinlichkeit der Verbundereignisse angeben, indem man (3.89) und (3.83) in (3.80) einsetzt, wobei auf der rechten Seite zur besseren Lesbarkeit das Argument Θ weggelassen wird:

$$P(\Theta(k)|Z^k) = \frac{1}{c} \frac{1}{V_\phi} \prod_{m=1}^{M(k)} [\mathbb{N}\{z_m(k)\}]^{\tau_m} \cdot \frac{\phi!}{M!} P(\phi) \prod_{t=1}^T (P_D^t)^{\delta_t} (1 - P_D^t)^{1-\delta_t} \quad (3.90)$$

Geht man von einem nicht-parametrischen Rauschmodell aus, so ist $P(\phi)$ konstant. Auch die Anzahl der Beobachtungen ist, im Rahmen eines Iterationsschrittes k betrachtet, konstant. Damit können sie mit der Konstante c kombiniert werden, d. h. $c_2 = \frac{P(\phi)}{cM!}$:

$$P(\Theta(k)|Z^k) = \frac{c_2 \phi!}{V_\phi} \prod_{m=1}^{M(k)} [\mathbb{N}\{z_m(k)\}]^{\tau_m} \prod_{t=1}^T (P_D^t)^{\delta_t} (1 - P_D^t)^{1-\delta_t} \quad (3.91)$$

Nun können daraus die Assoziationswahrscheinlichkeiten β_{mt} bestimmt werden:

$$\beta_{mt}(k) = P(\Theta_{mt}(k)|Z^k) = \sum_{\Theta(k)} P(\Theta(k)|Z^k) \hat{\omega}_{mt}(\Theta(k)), \quad \begin{array}{l} m = 1, \dots, M(k) \\ t = 0, \dots, T \end{array} \quad (3.92)$$

Zur Berechnung der Assoziationswahrscheinlichkeiten β_{mt} muss die Konstante c_2 nicht vorgegeben werden. Da die Summe der Wahrscheinlichkeiten zu Eins werden muss, kann sie durch Normieren ermittelt werden.

3.3.2.1 Cheap Joint Probabilistic Data Association

Der oben beschriebene JPDA-Algorithmus zur Bestimmung der Assoziationswahrscheinlichkeiten ist aufgrund der Vielzahl der zu untersuchenden Kombinationen sehr rechenaufwändig. Daher wird in der im Versuchsträger eingesetzten Implementierung die Näherung nach Fitzgerald eingesetzt [Bar-Shalom 1990]. Dennoch werden, gerade bei der Integration in das *Interacting Multiple Model*-Filter, die Eigenschaften des vollständigen Algorithmus vorausgesetzt. Dies ist notwendig, um eine konsistente Integration der modellspezifischen Filter in das *Interacting Multiple Model*-Verfahren zu gewährleisten.

Der *Cheap Joint Probabilistic Data Association*-Algorithmus bestimmt zunächst für jede mögliche Zuordnung einer validierten Messung $m \in \{1, \dots, M\}$ zu einem Track $t \in \{1, \dots, T\}$ die Likelihood Λ_{mt} entsprechend Gleichung (3.82):

$$\Lambda_{mt} = p(z_m(k)|\Theta_{mt}, Z^{k-1}) = \frac{1}{\sqrt{|2\pi S_{mt}(k)|}} \exp \left\{ -\frac{1}{2} \mathbf{v}_{mt}^T(k) S_{mt}^{-1}(k) \mathbf{v}_{mt}(k) \right\} \quad (3.93)$$

Diese ist ein Maß dafür, wie gut Track und Messung zusammenpassen, und wurde daher auch schon bei der Validierung verwendet (vgl. 3.67).

Nun folgt die näherungsweise Bestimmung der Wahrscheinlichkeiten für die Ereignisse Θ_{mt} :

$$\beta_{mt} = P(\Theta_{mt} | \mathbf{Z}^k) = \frac{\Lambda_{mt}}{T_t + M_m - \Lambda_{mt} + B}. \quad (3.94)$$

Mit T_t ist hier die Summe aller Likelihoods für Track t bezeichnet, M_m ist die Summe aller Likelihoods für Messung m :

$$T_t = \sum_{m=1}^{M_t} \Lambda_{mt} \quad (3.95)$$

$$M_m = \sum_{t=1}^T \Lambda_{mt}. \quad (3.96)$$

Mit der Konstanten B wird dem Umstand Rechnung getragen, dass möglicherweise keine der assoziierten Beobachtungen wahr ist. Die entsprechende Wahrscheinlichkeit wird durch Gleichung (3.97) angenähert:

$$\beta_{t0} = \frac{B}{T_t + B}. \quad (3.97)$$

Diese von Fitzgerald entwickelte Näherung führt zu einer deutlichen Reduzierung des Rechenaufwandes gegenüber der vollständigen Bestimmung im *Joint Probabilistic Data Association*-Filter, was für eine Echtzeitanwendung entscheidend ist. Die weitere Vorgehensweise stimmt jedoch mit diesem überein.

3.4 Interacting Multiple Model-Filter

Das *Interacting Multiple Model*-Filter [Bar-Shalom u. Blair 2000] basiert auf dem Gedanken, dass der Zustand eines beobachteten Systems zwischen den Iterationsschritten durch eines von mehreren Modellen korrekt beschrieben wird. Zu jedem ganzen Zeitschritt kann das beschreibende Modell wechseln, was durch Markov-Übergangswahrscheinlichkeiten π_{ij} ausgedrückt wird. Dabei ist π_{ij} die Wahrscheinlichkeit für einen Wechsel des Systems von Modell i im letzten Zeitschritt zu Modell j . Die Wahrscheinlichkeit, dass das System weiterhin durch Modell i beschrieben wird, ist π_{ii} . Blom stellte den *IMM*-Algorithmus 1984 vor [Blom 1984]. Dabei kamen *Kalman*-Filter zur Beschreibung der Modelle zum Einsatz.

Das im vorangegangenen Abschnitt beschriebene *JPDA*-Filter ist dem *Kalman*-Filter unter Berücksichtigung der besonderen Eigenschaften der hier gewählten

Sensoren jedoch überlegen. Daher liegt es nahe, das *JPDA*-Filter für die Beschreibung der Modelle in einem *Interacting Multiple Model*-Ansatz zu verwenden. Dazu werden nun mehrere dieser Filter nach Abschnitt 3.3.2 parallel eingesetzt. Sie arbeiten weitgehend unabhängig voneinander, lediglich die Validierung der Messungen erfolgt gemeinsam. Dies gewährleistet, dass in allen Filtern die selben Messdaten verwendet und filterspezifische Assoziationen vermieden werden.

Dazu wird der Validierungsschritt zunächst für jedes Filter $n \in N$ durchgeführt, was zu N Assoziationsmengen D_n führt. Die Vereinigungsmenge

$$D = D_1 \cup D_2 \cup \dots \cup D_N \quad (3.98)$$

dient dann für alle Filter gleichermaßen zur Validierung der Hypothesen.

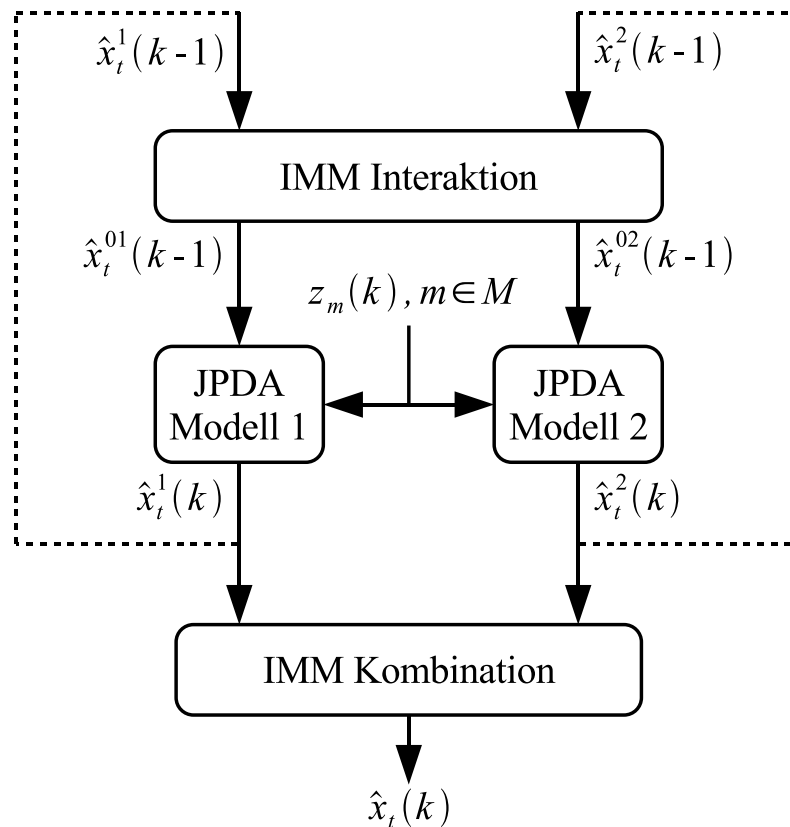


Bild 3.5: Das *IMM*-Filter greift durch Modifikation der Systemzustände und Kovarianzen in den Ablauf der *JPDA*-Filter ein

Wie in Bild 3.5 veranschaulicht, durchlaufen die *JPDA*-Filter jeweils einen kompletten Prädiktions- und Korrekturschritt. Der zu prädizierende Systemzustand ist dabei jedoch nicht gleich dem geschätzten Systemzustand aus dem letzten Zeitschritt. Hier greift die Interaktionsstufe des *IMM*-Algorithmus ein, die aus den

Systemzuständen $\hat{\mathbf{x}}_t^n(k-1)$ der N JPDA-Filter jeweils einen neuen Systemzustand $\hat{\mathbf{x}}_t^{0n}(k-1)$ bildet, der als Ausgangspunkt für den nächsten Filterschritt dient.

Bei $\hat{\mathbf{x}}_t^n(k-1)$ handelt es sich um den Erwartungswert der Verteilung

$$p(\mathbf{x}_t(k-1)|R_t(k-1) = n, \mathbf{Z}_t^{k-1}), \quad (3.99)$$

d. h. der Verteilung des Systemzustandes \mathbf{x}_t zum Zeitpunkt $k-1$ unter der Voraussetzung, dass zum Zeitpunkt $k-1$ das Modell n zutrifft und die Beobachtungen \mathbf{Z}_t^{k-1} vorliegen. $R_t(k-1)$ bezeichnet also das zu diesem Zeitpunkt gültige Modell, \mathbf{Z}_t^{k-1} die akkumulierten Beobachtungen bis einschließlich Zeitschritt $k-1$.

Als Basis für den nächsten Filterschritt ist jedoch die Verteilung unter der Voraussetzung, dass zum *aktuellen* Zeitpunkt k Modell n zutrifft, entscheidend:

$$p(\mathbf{x}_t(k-1)|R_t(k) = n, \mathbf{Z}_t^{k-1}) = \sum_{i=1}^N p(\mathbf{x}_t(k-1)|R_t(k-1) = i, \mathbf{Z}_t^{k-1}) \mu_t^{in}(k-1). \quad (3.100)$$

Die bedingte Modellwahrscheinlichkeit μ_t^{in} gibt die Wahrscheinlichkeit an, mit der Track t im vorangegangenen Zeitschritt durch Modell i beschrieben wurde, unter der Voraussetzung, dass jetzt Modell n gültig ist.

Sind die Wahrscheinlichkeitsdichten $p(\mathbf{x}_t(k-1)|R_t(k-1) = n, \mathbf{Z}_t^{k-1})$ wie vorausgesetzt normalverteilt, so ist die Verteilung aus Gleichung (3.100) eine Gaußsche Mischverteilung. Diese Gaußsche Mischverteilung wird wiederum durch eine Normalverteilung mit dem Erwartungswert $\hat{\mathbf{x}}_t^{0n}(k-1)$ und der Kovarianz $\mathbf{P}_t^{0n}(k-1)$ angenähert. Auf diese Weise berücksichtigt die Interaktionsstufe die Möglichkeit eines Modellwechsels:

$$\hat{\mathbf{x}}_t^{0n}(k-1) = \sum_{i=1}^N \mu_t^{in}(k-1) \hat{\mathbf{x}}_t^i(k-1) \quad (3.101)$$

$$\mathbf{P}_t^{0n}(k-1) = \sum_{i=1}^N \mu_t^{in}(k-1) \left\{ \mathbf{P}_t^i(k-1) + [\hat{\mathbf{x}}_t^i(k-1) - \hat{\mathbf{x}}_t^{0n}(k-1)] [\hat{\mathbf{x}}_t^i(k-1) - \hat{\mathbf{x}}_t^{0n}(k-1)]^T \right\}. \quad (3.102)$$

Die bedingten Modellwahrscheinlichkeiten $\mu_t^{in}(k-1)$ erhält man nach dem Theorem der totalen Wahrscheinlichkeit aus

$$\mu_t^{in}(k-1) = \frac{1}{\mu_t^{n-}(k)} \pi_{in} \mu_t^i(k-1). \quad (3.103)$$

Die prädizierten Modellwahrscheinlichkeiten $\mu_t^{n-}(k)$ ergeben sich aus

$$\mu_t^{n-}(k) = P\left(R_t(k) = n | \mathbf{Z}_t^{k-1}\right) = \sum_{i=1}^N \pi_{in} \mu_t^i(k-1). \quad (3.104)$$

Mit den Ergebnissen der *IMM*-Interaktion, $\hat{\mathbf{x}}_t^{0n}(k-1)$ und $\mathbf{P}_t^{0n}(k-1)$, durchlaufen die *JPDA*-Filter nun einen Filterschritt, was zu den geschätzten Systemzuständen bzw. Kovarianzen $\hat{\mathbf{x}}_t^n(k)$ und $\mathbf{P}_t^n(k)$ führt. Um die nächste Interaktion der Filter zu ermöglichen, werden noch die zugehörigen Modellwahrscheinlichkeiten $\mu_t^n(k)$ benötigt. Beim einfachen *IMM*-Filter werden dazu, unter der Annahme normalverteilter Filterresiduen, Likelihoods für das Auftreten der Messungen bei gegebenem Modell und gegebenen Messungen bis zum Zeitpunkt $k-1$ bestimmt:

$$\Lambda_t^n(k) = p\left(\mathbf{z}_t(k) | R_t(k) = n, \mathbf{Z}_t^{k-1}\right) \quad (3.105)$$

$$= \frac{1}{\sqrt{|2\pi\mathbf{S}_t^n(k)|}} \exp\left\{-\frac{1}{2}\mathbf{v}_t^n(k)^T \mathbf{S}_t^n(k)^{-1} \mathbf{v}_t^n(k)\right\}. \quad (3.106)$$

Aufgrund der Verwendung von *Joint Probabilistic Data Association*-Filtern ist es jedoch, wie in Abschnitt 3.3.2 ausgeführt, nicht möglich, eine einzelne Messung $\mathbf{z}_t(k)$ anzugeben. Hier müssen alle Messungen, die einen Beitrag zu den modell-spezifischen Filtern leisten, berücksichtigt werden.

Die Modellwahrscheinlichkeit ergibt sich so zu

$$\mu_t^n(k) = P\left(R_t(k) = n | \mathbf{Z}_t^k\right) = P\left(R_t(k) = n | \mathbf{Z}_t(k), \mathbf{Z}_t^{k-1}\right) \quad (3.107)$$

$$= \frac{p\left(\mathbf{Z}_t(k) | R_t(k) = n, \mathbf{Z}_t^{k-1}\right) P\left(R_t(k) = n | \mathbf{Z}_t^{k-1}\right)}{\sum_{i=1}^N \left\{p\left(\mathbf{Z}_t(k) | R_t(k) = i, \mathbf{Z}_t^{k-1}\right) P\left(R_t(k) = i | \mathbf{Z}_t^{k-1}\right)\right\}}. \quad (3.108)$$

Aus Gleichung (3.104) ist $P\left(R_t(k) = n | \mathbf{Z}_t^{k-1}\right)$ bekannt. Es wird also noch die Wahrscheinlichkeitsdichte $p\left(\mathbf{Z}_t(k) | R_t(k) = n, \mathbf{Z}_t^{k-1}\right)$ benötigt, für die sich mit dem Theorem der totalen Wahrscheinlichkeit ergibt:

$$p\left(\mathbf{Z}_t(k) | R_t(k) = n, \mathbf{Z}_t^{k-1}\right) = \sum_{m=0}^{M_t} \left\{p\left(\mathbf{Z}_t(k) | \Theta_{mt}(k), R_t(k) = n, \mathbf{Z}_t^{k-1}\right) \cdot P\left(\Theta_{mt}(k) | R_t(k) = n, \mathbf{Z}_t^{k-1}\right)\right\}. \quad (3.109)$$

Der Ausdruck $p\left(\mathbf{Z}_t(k) | \Theta_{mt}(k), R_t(k) = n, \mathbf{Z}_t^{k-1}\right)$ beschreibt die Wahrscheinlichkeitsdichte für das Auftreten der Messungen $\mathbf{Z}_t(k) = \{z_{1t}(k), \dots, z_{M_t t}(k)\}$ für Objekt

t unter der Voraussetzung, dass Messung m wahr und das System durch Modell n beschrieben ist. Zudem werden die Messungen bis zum Zeitpunkt $k - 1$ berücksichtigt, d. h. nicht die aktuellen Messungen $\mathbf{Z}(k)$.

Diese Wahrscheinlichkeitsdichte ist das Produkt der Wahrscheinlichkeitsdichten der einzelnen Messungen unter den genannten Voraussetzungen:

$$p\left(\mathbf{Z}_t(k)|\Theta_{mt}(k), R_t(k) = n, \mathbf{Z}_t^{k-1}\right) = \prod_{i=1}^{M_t} p\left(z_{ti}(k)|\Theta_{mt}(k), R_t(k) = n, \mathbf{Z}_t^{k-1}\right). \quad (3.110)$$

Für den Fall $i = m$, also für eine wahre Messung, ergibt sich aufgrund der vorausgesetzten Normalverteilung der Residuen

$$p\left(z_{mt}(k)|\Theta_{mt}(k), R_t(k) = n, \mathbf{Z}_t^{k-1}\right) = \mathbf{N}\{z_t^-(k), \mathbf{S}_{mt}(k)\}, \quad (3.111)$$

wobei $z_t^-(k)$ die prädizierte Messung für Objekt t und $\mathbf{S}_{mt}(k)$ die Kovarianz des Residuums kennzeichnet.

Das Auftreten von Fehldetektionen wird als gleichverteilt über das gesamte Messvolumen angesehen. Damit ergibt sich für $i \neq m$

$$p\left(z_{it}(k)|\Theta_{mt}(k), R_t(k) = n, \mathbf{Z}_t^{k-1}\right) = \frac{1}{V}, \quad i \neq m. \quad (3.112)$$

Bei diesen Betrachtungen wurde der Einfluss der Validierung nicht berücksichtigt. Diese beschneidet zum einen die Normalverteilung (3.111), zum anderen beschränkt sie das relevante Messvolumen. Beide Effekte ließen sich kompensieren, so wie es auch beim *PDA* der Fall ist. Jedoch wäre die Bestimmung des Volumens der Validierungsregionen sehr aufwändig, da es sich hierbei um Vereinigungen mehrerer Ellipsoide handelt. Daher wird auf diesen Schritt verzichtet. Dies deckt sich auch mit der Vorgehensweise des *JPDA* [Bar-Shalom u. Fortmann 1988]. Auch hier wird der Algorithmus ohne Validierung hergeleitet. Um die benötigte Rechenleistung zu beschränken, wird in der praktischen Anwendung aber dennoch eine Validierung durchgeführt. Schließt diese lediglich Hypothesen aus, die keinen signifikanten Einfluss auf die Schätzung haben, so kann man den dadurch entstehenden Fehler vernachlässigen.

Mit (3.111) und (3.112) in (3.110) ergibt sich dann

$$p\left(\mathbf{Z}_t(k)|\Theta_{mt}(k), R_t(k) = n, \mathbf{Z}_t^{k-1}\right) = \begin{cases} \frac{1}{V^{M-1}} \mathbf{N}\{z_t^-(k), \mathbf{S}_{mt}(k)\} & m = 1, \dots, M \\ \frac{1}{V^M} & m = 0 \end{cases}. \quad (3.113)$$

Um (3.109) zu komplettieren, wird noch der zweite Term, die Wahrscheinlichkeit $P(\Theta_{mt}(k)|R_t(k) = n, \mathbf{Z}_t^{k-1})$, benötigt. Dies ist die Wahrscheinlichkeit, dass Messung m wahr ist, wobei nur die Messungen bis zum Zeitpunkt $k - 1$ berücksichtigt werden. Diese a-priori-Wahrscheinlichkeit setzt sich aus der Detektionswahrscheinlichkeit P_D und der Anzahl der Messungen zusammen:

$$P(\Theta_{mt}(k)|R_t(k) = n, \mathbf{Z}_t^{k-1}) = \frac{P_D}{M} \quad m = 1, \dots, M. \quad (3.114)$$

Die Wahrscheinlichkeit, dass keine der erhaltenen Messungen wahr ist, ist

$$P(\Theta_{0t}(k)|R_t(k) = n, \mathbf{Z}_t^{k-1}) = 1 - P_D. \quad (3.115)$$

Damit können aus (3.108) die Modellwahrscheinlichkeiten bestimmt werden. Mit diesen liefert ein Kombinationsschritt schließlich den geschätzten Systemzustand des IMM-Filters:

$$\hat{\mathbf{x}}_t(k) = \sum_{n=1}^N \mu_t^n(k) \hat{\mathbf{x}}_t^n(k) \quad (3.116)$$

$$\mathbf{P}_t(k) = \sum_{n=1}^N \mu_t^n(k) \left\{ \mathbf{P}_t^n(k) + [\hat{\mathbf{x}}_t^n(k) - \hat{\mathbf{x}}_t(k)] [\hat{\mathbf{x}}_t^n(k) - \hat{\mathbf{x}}_t(k)]^T \right\}. \quad (3.117)$$

Hier stellen $\hat{\mathbf{x}}_t(k)$ und $\mathbf{P}_t(k)$ die Ausgabe des Filters dar, intern finden sie allerdings keine Verwendung.

Im Kombinationsschritt ist zu beachten, dass die Schätzungen $\hat{\mathbf{x}}_t^n(k)$ und Kovarianzmatrizen $\mathbf{P}_t^n(k)$ für die unterschiedlichen Bewegungsmodelle unterschiedliche Dimensionen aufweisen. Daher ist es notwendig, sie vorher in ein gemeinsames System zu überführen [Blackman u. Popoli 1999].

Im vorliegenden Fall beschreibt $\hat{\mathbf{x}}_t^2(k)$ eine höherdimensionale Variante von $\hat{\mathbf{x}}_t^1(k)$. Da die Anforderungen an den Sensor Informationen über Position und Geschwindigkeit vorsehen, ist die Beschreibung durch einen Zustandsvektor im System von $\hat{\mathbf{x}}_t^1$ ausreichend. Damit gestaltet sich auch die Transformation einfach. Der entsprechende Zusammenhang lässt sich formulieren als:

$$\hat{\mathbf{x}}_t^{21}(k) = \mathbf{A}_{21} \hat{\mathbf{x}}_t^2(k) \quad (3.118)$$

mit

$$\mathbf{A}_{21} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \quad (3.119)$$

Für den Kombinationsschritt wird also die Ordnung von $\hat{\mathbf{x}}_t^2(k)$ auf die Ordnung von $\hat{\mathbf{x}}_t^1(k)$ reduziert, indem die Beschleunigungskomponenten entfernt werden.

Genauso einfach gestaltet sich die Transformation der Kovarianzmatrix:

$$\mathbf{P}_t^{21}(k) = \mathbf{A}_{21} \mathbf{P}_t^2(k) \mathbf{A}_{21}^T. \quad (3.120)$$

Auch dies bedeutet im vorliegenden Fall, dass die Komponenten, die den Beschleunigungen entsprechen, weggestrichen werden.

3.5 Fusion der Daten mehrerer Sensoren

In den vorangegangenen Abschnitten wurde eine Vorgehensweise zur Verfolgung mehrerer Objekte unter Verwendung mehrerer Systemmodelle vorgestellt. Nicht beachtet wurde jedoch, dass die Messungen von mehreren Sensoren geliefert werden. Obwohl die unterschiedliche Kovarianz der Messungen berücksichtigt wurde, ist es nicht zulässig, die Messungen mehrerer Sensoren in einem Schritt zusammenzuführen. Vereinigt man die Messungen mehrerer Sensoren zu einer Gesamtmenge, so verletzt man die in Abschnitt 3.3.2 getroffene Annahme, dass höchstens eine der Messungen wahr ist.

Daher werden die Sensoren sequentiell abgearbeitet. Zunächst wird mit den Messungen des ersten Sensors ein vollständiger *IMM*-Iterationsschritt durchgeführt. Dabei findet eine Prädiktion der Systemzustände vom Zeitpunkt $t - \Delta t$ auf den Zeitpunkt t statt. Dann werden die Messungen des nächsten Sensors berücksichtigt. Wieder wird ein *IMM*-Iterationsschritt vollzogen, wobei nun eine Prädiktion der Systemzustände nicht notwendig ist. Hier beschreiben die geschätzten Systemzustände schon die Systemzustände zum Zeitpunkt t , sodass statt (3.62) und (3.63) nun einfach gilt:

$$\hat{\mathbf{x}}^-(k) = \hat{\mathbf{x}}(k-1) \quad (3.121)$$

$$\mathbf{P}^-(k) = \mathbf{P}(k-1) \quad (3.122)$$

Auf diese Weise können beliebig viele Sensoren berücksichtigt werden. Dabei ist jedoch zu beachten, dass die Reihenfolge der Sensoren Einfluss auf die Schätzung hat: Die Korrektur eines geschätzten Systemzustandes mit den Messungen eines Sensors ändert auch die Kovarianz des Systemzustandes. Dies bedeutet gleichzeitig einen Einfluss auf

- die Validierungsregion für die Messungen des nächsten Sensors aufgrund von Gleichung (3.67) und

- die Modellwahrscheinlichkeiten aus Gleichung (3.108) aufgrund von Gleichung (3.111).

Dies lässt sich zwar umgehen, indem man zunächst für die Messungen aller Sensoren die Validierung und die Bestimmung der Modellwahrscheinlichkeiten durchführt und dann mit diesen die einzelnen Filterschritte durchläuft. Allerdings führt dies nicht zu einer konsistenten Schätzung. Zudem kann man argumentieren, dass durch die Verfeinerung der Schätzung aufgrund der Messungen des ersten Sensors die Validierung für die Messungen des zweiten Sensors genauer durchgeführt werden kann. Daher wurde darauf verzichtet, diesen Effekt zu berücksichtigen.

3.6 Trackkontrolle

Der zur zeitlichen Verfolgung der Objekte eingesetzte *Joint Probabilistic Data Association*-Ansatz geht davon aus, dass für alle vorhandenen Objekte eine Filterinstanz erzeugt wurde und dass alle Filterinstanzen wirklich vorhandene Objekte beschreiben (vgl. 3.3.2). Nur unter dieser Voraussetzung erfolgt die Bestimmung der Assoziationswahrscheinlichkeiten β_{mt} und damit letztendlich die Korrektur des geschätzten Systemzustandes korrekt.

Daher nimmt die Trackkontrolle eine zentrale Rolle im Gesamtprozess ein. Sie gewährleistet, dass diese Voraussetzungen erfüllt bleiben. Dazu muss sie

- neue Filterinstanzen, so genannte Tracks, erzeugen, sobald ein neues Objekt erfasst wird,
- mehrere Tracks zusammenführen, falls sie das selbe Objekt beschreiben und
- Tracks entfernen, die kein real existierendes Objekt beschreiben.

In der Praxis zeigt sich, dass der bei den verwendeten Merkmalen auftretende große Anteil falscher Beobachtungen hohe Anforderungen an die Trackkontrolle stellt.

Die einzelnen Aufgabenbereiche werden daher im Folgenden näher betrachtet.

3.6.1 Initialisieren von Tracks

Erfassen die Sensoren ein Objekt, für das bisher noch keine Filterinstanz vorliegt, so treten Beobachtungen auf, die nicht zugeordnet werden können. Dies bedeutet, dass die entsprechenden Beobachtungen für kein Objekt validiert wurden.

Die Trackkontrolle muss in diesem Fall eine neue Filterinstanz erzeugen. Dies geschieht einfach, indem für jede nicht-validierte Beobachtung ein Track gestartet wird. Da vorausgesetzt wird, dass jeder Sensor lediglich eine wahre Beobachtung pro Objekt liefert, wird dadurch auch nur eine Filterinstanz pro Objekt erzeugt.

Allerdings wird auch für jede nicht-validierte Beobachtung, bei der es sich um eine Fehldetektion handelt, eine Instanz erzeugt. Da Beobachtungen im Umfeld von nicht erfassten Objekten in aller Regel auch nicht für ein anderes Objekt validiert wurden, kann dadurch eine Vielzahl von Filterinstanzen auftreten, die kein real existierendes Objekt beschreiben. Diese müssen, da sie einen negativen Einfluss auf andere Tracks haben, schnellstmöglich wieder verworfen werden. Damit beschäftigt sich Abschnitt 3.6.2.

Da die verwendeten Sensoren Positionen, jedoch keine Geschwindigkeits- oder Beschleunigungsinformationen liefern, lässt sich aus den Beobachtungen eines Zeitschrittes noch keine vollständige Filterinstanz initialisieren. Da sowohl vorausfahrende als auch entgegenkommende Fahrzeuge erfasst werden sollen, kann auch keine sinnvolle Annahme über die fehlenden Elemente der Zustandsvektoren getroffen werden.

Um dieses Problem zu lösen, wird eine zweischrittige Initialisierung durchgeführt. Mit den Beobachtungen aus einem Zeitschritt t_i werden zunächst vorläufige Filterinstanzen erzeugt. Trifft zum nächsten Zeitschritt t_{i+1} eine Beobachtung ein, die der ersten zugeordnet werden kann, so können die Geschwindigkeitskomponenten der Zustandsvektoren näherungsweise nach

$$\dot{z} \approx \frac{z(t_{i+1}) - z(t_i)}{t_{i+1} - t_i} \quad (3.123)$$

bestimmt werden.

Die Beschleunigungskomponenten des Systemmodells, das eine Bewegung mit konstanter Relativbeschleunigung beschreibt, werden zu Null gesetzt. Zwar ließen sich diese mit einem weiteren Schritt annähern, jedoch sind die dabei entstehenden Fehler so groß, dass die Annahme, die Bewegung verlaufe unbeschleunigt, bei der Initialisierung keine Verschlechterung bedeutet.

Eine Schwierigkeit besteht noch darin, den vorläufigen Filterinstanzen im darauf folgenden Zeitschritt Beobachtungen zuzuordnen. Eine vollständige Validierung und Assoziation ist nur für vollständig initialisierte Filterinstanzen möglich. Daher wird folgendes einfache Schema angewandt:

- Alle Beobachtungen, die einen vorgegebenen Maximalabstand zum vorläufigen Track unterschreiten, werden für diesen validiert.

- Von den validierten Beobachtungen wird lediglich der nächste Nachbar betrachtet.

Werden im zweiten Zeitschritt keine Beobachtungen für die vorläufige Filterinstanz validiert, so wird sie verworfen. Dies ist notwendig, um die Anzahl der neu entstehenden Tracks zu minimieren.

3.6.2 Entfernen von Tracks

Doch nicht nur durch Rauschen oder fehlerhafte Initialisierung können Filterinstanzen auftreten, die kein Objekt beschreiben. Eine weitere Möglichkeit ist, dass ein verfolgtes Objekt den Erfassungsbereich der Sensoren verlässt. Dies muss erkannt und der entsprechende Track gelöscht werden.

Zudem können auch die beiden eingeführten Systemmodelle nicht alle Situationen korrekt beschreiben, gerade wenn über mehrere Iterationsschritte das Objekt von keinem Sensor erfasst wird. Dann können die Systemzustände nicht entsprechend korrigiert werden. So kommt es durchaus vor, dass sich ein Track von seinem Objekt „ablöst“. Zwar wird dann, sobald die Sensoren das Objekt wieder erfassen, eine neue Filterinstanz erzeugt, doch liegt weiterhin eine Instanz vor, die kein Objekt mehr beschreibt.

Liegen viele Tracks vor, die kein real existierendes Objekt beschreiben, so hat das wiederum einen negativen Einfluss auf Tracks, die wirklich ein Objekt beschreiben. Dies verdeutlicht Bild 3.6.

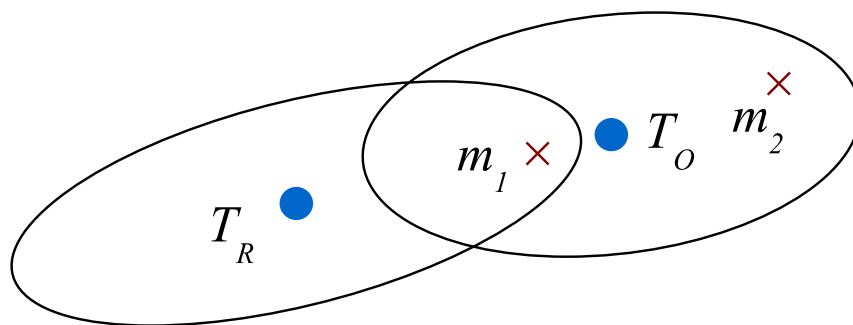


Bild 3.6: Ein Track T_R , der durch Fehldetektionen ausgelöst wurde, beeinflusst Track T_O , der ein real existierendes Objekt beschreibt: Die wahre Beobachtung m_1 wird auch mit T_R assoziiert. Damit verliert sie Gewicht in der Assoziation zu T_O .

Im Beispiel liegen zwei Tracks vor, der durch Fehldetektionen ausgelöste Track T_R und der ein existierendes Objekt beschreibende Track T_O . Ohne T_R wäre die Assoziationswahrscheinlichkeit von m_1 zu T_O höher als wenn m_1 auch zu T_R assoziiert

wird. Auf diese Weise berücksichtigt das *Joint Probabilistic Data Association*-Filter die Unsicherheit bei der Zuordnung.

Daher ist es entscheidend, solche falschen Filterinstanzen möglichst schnell zu entfernen. Eine Konsistenzprüfung aller Tracks zu jedem Iterationsschritt stellt dies sicher.

Zunächst wird die Kovarianz betrachtet. Das *Interacting Multiple Model*-Filter bestimmt im Kombinationsschritt sowohl eine kombinierte Schätzung $\hat{\mathbf{x}}_t(k)$ als auch eine kombinierte Kovarianz $\mathbf{P}_t(k)$ (vgl. auch Gleichungen (3.116) und (3.117)). Zur Beurteilung sollten jedoch nur die Unsicherheiten der Positionsschätzungen berücksichtigt werden. Gerade wenn ein betrachtetes Objekt einem Modellwechsel unterliegt, wachsen die Unsicherheiten der Geschwindigkeitskomponenten kurzzeitig an. Jedoch soll das nicht dazu führen, dass der Track gelöscht wird.

Daher wird die Kovarianz in den Raum der Messvektoren transformiert. Da sie vorher entsprechend dem Systemmodell für konstante Relativgeschwindigkeit vorlag, wird hierzu die schon aus Gleichung (3.24) bekannte Beobachtungsmatrix \mathbf{H}^1 herangezogen:

$$\mathbf{P}_{t,z}(k) = \mathbf{H}^1 \mathbf{P}_t(k) (\mathbf{H}^1)^T \quad (3.124)$$

Übersteigt der größte Eigenwert von $\mathbf{P}_{t,z}(k)$ einen Maximalwert, so wird die Schätzung als zu unsicher angesehen und verworfen.

Zudem wird ein weiteres Kriterium zur Überprüfung angewendet. Wurden dem Track über einen Zeitraum Δt_g keine Beobachtungen zugeordnet, so wird er gelöscht. Diese Schranke greift bei Tracks, die zuvor eine geringe Kovarianz aufwiesen, oft vor dem letztgenannten Kriterium. Gerade durch periodische Strukturen im Leitplankenbereich treten immer wieder Tracks auf, die kein real existierendes Objekt beschreiben, aber geringe Kovarianz aufweisen. Sie können mit diesem Kriterium schneller entfernt werden.

3.6.3 Zusammenführen von Tracks

Häufig werden auch Objekte mehrfach initialisiert, sodass mehrere Instanzen für ein einzelnes Objekt vorliegen. Dieser Effekt kann auch auftreten, wenn ein durch Rauschen ausgelöster Track von einem Objekt „eingefangen“ wird. Da sich durch die gegenseitige Beeinflussung der Filterinstanzen dadurch für beide eine Verschlechterung ergibt, müssen überzählige Repräsentationen entfernt werden.

Dazu wird zunächst die räumliche Nähe aller Tracks untersucht, um solche herauszufinden, die möglicherweise das gleiche Objekt beschreiben. Ist so ein Fall entdeckt, so werden die entsprechenden Objektpositionen für den nächsten Zeitschritt

prädiziert. Liegen auch die Prädiktionen nahe beieinander, so wird die jüngere der beiden Objektrepräsentationen entfernt.

Damit sind die notwendigen Kontrollmechanismen vorhanden, die eine stabile Funktion der zeitlichen Verfolgung gewährleisten.

4 Aufbau und Implementierung

Die in dieser Arbeit entwickelten Verfahren wurden für das Versuchsfahrzeug des Instituts für Mess- und Regelungstechnik der Universität Karlsruhe (TH) implementiert. Zwar ist die Erfassung von Verkehrsteilnehmern für Fahrerassistenzsysteme, wie z. B. ein erweitertes ACC, der Grundgedanke, ein Eingriff in die Fahrdynamik erfolgt jedoch nicht. Ziel war hier eine Beschreibung der umgebenden Fahrzeuge und eine Darstellung der Messergebnisse in Echtzeit.

4.1 Ausstattung und Aufbau des Versuchsträgers

Das verwendete Versuchsfahrzeug verfügt, neben der für diese Arbeit benötigten Kamera, über weitere Sensorik, die in anderen Projekten Anwendung findet. Daher ist der Gesamtaufbau von verschiedenen Quellen beeinflusst.

Um den verschiedenen Anforderungen gerecht zu werden, wurde ein modularer Aufbau mit definierten Schnittstellen gewählt. Daraus ergeben sich vier grundlegende Funktionsgruppen:

- Sensoren,
- Rechner,
- Kommunikation und
- Spannungsversorgung.

Einen Überblick über die Komponenten gibt Bild 4.1.

4.1.1 Sensoren

Hinter der Windschutzscheibe ist in der Nähe des Innenspiegels das Kamerasystem angebracht. Obwohl in dieser Arbeit lediglich monoskopische Merkmale ausgewertet werden, handelt es sich um ein Stereosystem aus zwei Kameras, die in Fahrtrichtung blicken. Beide Kameras sind in CCD-Technik ausgeführt, verfügen

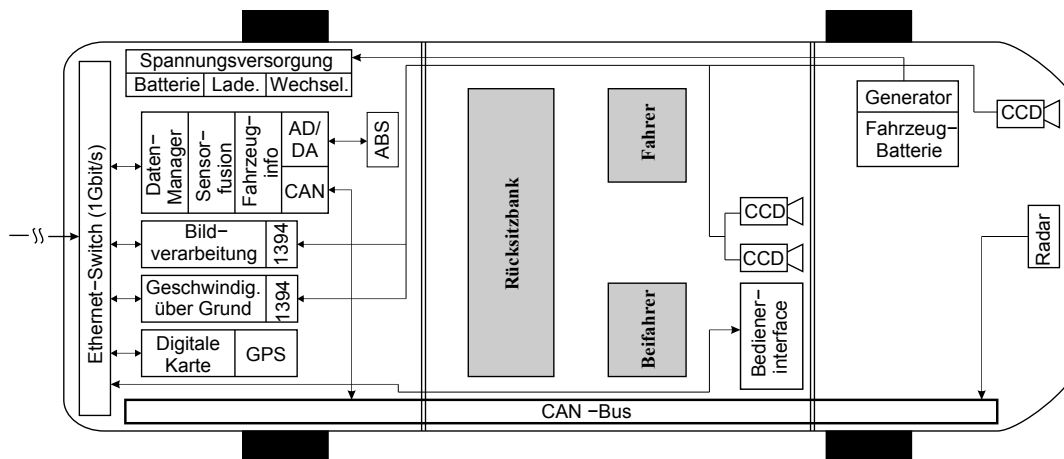


Bild 4.1: Aufbau des Versuchsträgers.

über eine IEEE1394-Schnittstelle und liefern 25 Vollbilder pro Sekunde. Bei einer Brennweite von 8mm verfügen sie über einen Öffnungswinkel von ca. 45°. Zur Detektion der Schatten- und Symmetriemerkmale wurde die Kamera auf der Fahrerseite verwendet, die auch den einzigen hier verwendeten Sensor darstellt.

Zur Realisierung anderer Projekte sind weitere Sensoren montiert, die hier noch kurz erwähnt werden. So verfügt das Testfahrzeug über einen GPS-Sensor (Global Positioning System), der in Verbindung mit einer digitalen Karte für Ortungsaufgaben eingesetzt wird.

Ein Radarsensor, wie er auch in aktuellen ACC-Systemen (Adaptive Cruise Control) Verwendung findet, dient zur Erfassung vorausfahrender Fahrzeuge. Er stellt einen weiteren Baustein zu einer umfassenden Umgebungsbeschreibung dar.

Zur Messung der Geschwindigkeit über Grund ist im Bereich des vorderen Stoßfängers eine weitere Kamera angebracht, die nach unten gerichtet ist. Eine detaillierte Beschreibung hierzu ist bei Horn [Horn 2005] zu finden.

Des Weiteren steht mit den Raddrehzahlmessern des ABS-Systems eine grobe Geschwindigkeitsinformation zur Verfügung. Obwohl die im Fahrzeug montierte Sensorik auch die merkmalsbasierte Objektdetektion unterstützen könnte, z. B. durch Kenntnis der Eigenbewegung, wird diese Information hier nicht genutzt. Dies stellt sicher, dass das System unabhängig von weiteren Komponenten und damit leichter auf andere Fahrzeuge übertragbar ist.

4.1.2 Rechner

Für die Verarbeitung der Sensorsignale stehen zwei im Kofferraum untergebrachte, aus handelsüblichen Komponenten aufgebaute Personal Computer zur Verfügung. Für die Extraktion der Merkmale, die Fusion und die zeitliche Verfolgung wird lediglich einer dieser Rechner eingesetzt. Dieser verfügt über einen 3GHz Pentium-Prozessor und ist mit dem Betriebssystem Linux ausgestattet.

Vor dem Beifahrersitz ist ein Laptop-Computer montiert. Dieser fungiert als Bediener-Schnittstelle und zur Darstellung der Ergebnisse. Über ihn können Parameter gewählt und die Ausführung der Algorithmen gestartet und angehalten werden.

4.1.3 Kommunikation

Verschiedene Bus-Systeme verbinden die Komponenten im Fahrzeug. Die beschriebenen Rechner sind über eine 1Gbit-Ethernet-Schnittstelle verbunden. Damit ist eine flexible Architektur gewährleistet, da Geräte einfach hinzugefügt oder ersetzt werden können. Zudem besteht ein Anschluss, der das stehende Fahrzeug für Administrationszwecke mit einem Netz verbindet. Daneben ist noch eine Wireless-LAN-Verbindung vorhanden, die für Projekte zur Fahrzeug-Fahrzeug-Kommunikation eingesetzt wird.

Die Kameras liefern die Bilddaten über eine IEEE1394-Verbindung an die Rechner. Da hier mit IIDC (Instrumentation & Industrial Digital Camera) eine anerkannte Spezifikation für die Kommunikation vorliegt, lassen sich die Kameras leicht austauschen.

Schließlich befindet sich noch ein CAN-Bus (Controller Area Network) im Fahrzeug, über den die Messdaten des Radarsensors übertragen werden. In modernen Fahrzeugen dient ein CAN-Bussystem zur Kommunikation vieler Komponenten, die zur Steuerung und Datenerfassung dienen.

4.1.4 Spannungsversorgung

Da im mobilen Einsatz keine Netzspannung zur Verfügung steht, kommt der Spannungsversorgung des Testaufbaus eine besondere Bedeutung zu. So wurden alle Komponenten mit speziellen Netzteilen ausgestattet, die einen Betrieb an 12V Gleichspannung erlauben.

Diese wird von zwei Gel-Batterien zur Verfügung gestellt, die während der Fahrt über den Generator gespeist werden. Steht das Fahrzeug über einen längeren Zeitraum, so ersetzt ein fest integriertes Ladegerät den Generator.

Der Anschluss weiterer Komponenten, die nicht umgerüstet werden können, kann über einen 230V-Wechselrichter erfolgen, der in die Spannungsversorgung integriert ist.

4.2 Softwarearchitektur

Der komplette Algorithmus wurde für das im Fahrzeug eingesetzte MRT-Vision-System implementiert. Dabei handelt es sich um eine modulare Software, die den einzelnen funktionalen Komponenten eine gemeinsame Infrastruktur zur Kommunikation und Datensynchronisation bietet.

Diese Komponenten wurden in der Sprache C++ implementiert und können vom Benutzer interaktiv verknüpft werden. Dies vereinfacht den Austausch einzelner Elemente. Für den in dieser Arbeit entwickelten Algorithmus werden die in Bild 4.2 dargestellten Module eingesetzt, die im Folgenden kurz erläutert werden sollen.

Camera: Dieses Modul empfängt die Bilder vom Stereo-Kamerasystem und stellt sie zur weiteren Verarbeitung zur Verfügung.

Shadow: Hier findet die Extraktion der Schattenhypothesen statt. Das Modul besitzt je einen Eingang für die Bilddaten und Fahrbahnebeneninformation. Ein weiterer, optionaler Eingang wird im Zusammenhang mit der Klassifikation der Hypothesen (siehe Abschnitt 2.3.5) genutzt. Während ein Ausgang zur Darstellung dient und normalerweise ungenutzt bleibt, liefert der zweite Beobachtungen in 3D-Koordinaten.

Symmetry: Analog zum Shadow-Modul werden hier die vorhandenen Bilder auf Symmetrieregionen untersucht. Bis auf den Eingang für den Klassifikationsschritt entsprechen sich die Schnittstellen.

Lane: Das Lane-Modul liefert Informationen über die Lage und Orientierung der Fahrbahnebene. Auf die dazu verwendeten Verfahren wird hier nicht näher eingegangen, da sie nicht Gegenstand dieser Arbeit sind. Sie wurden von Herrn Duchow zur Verfügung gestellt und basieren auf der Verfolgung der Fahrbahnmarkierungen. Hier kann auch ein Modul von Herrn Dang zum Einsatz kommen, das die benötigten Informationen aus einer Auswertung

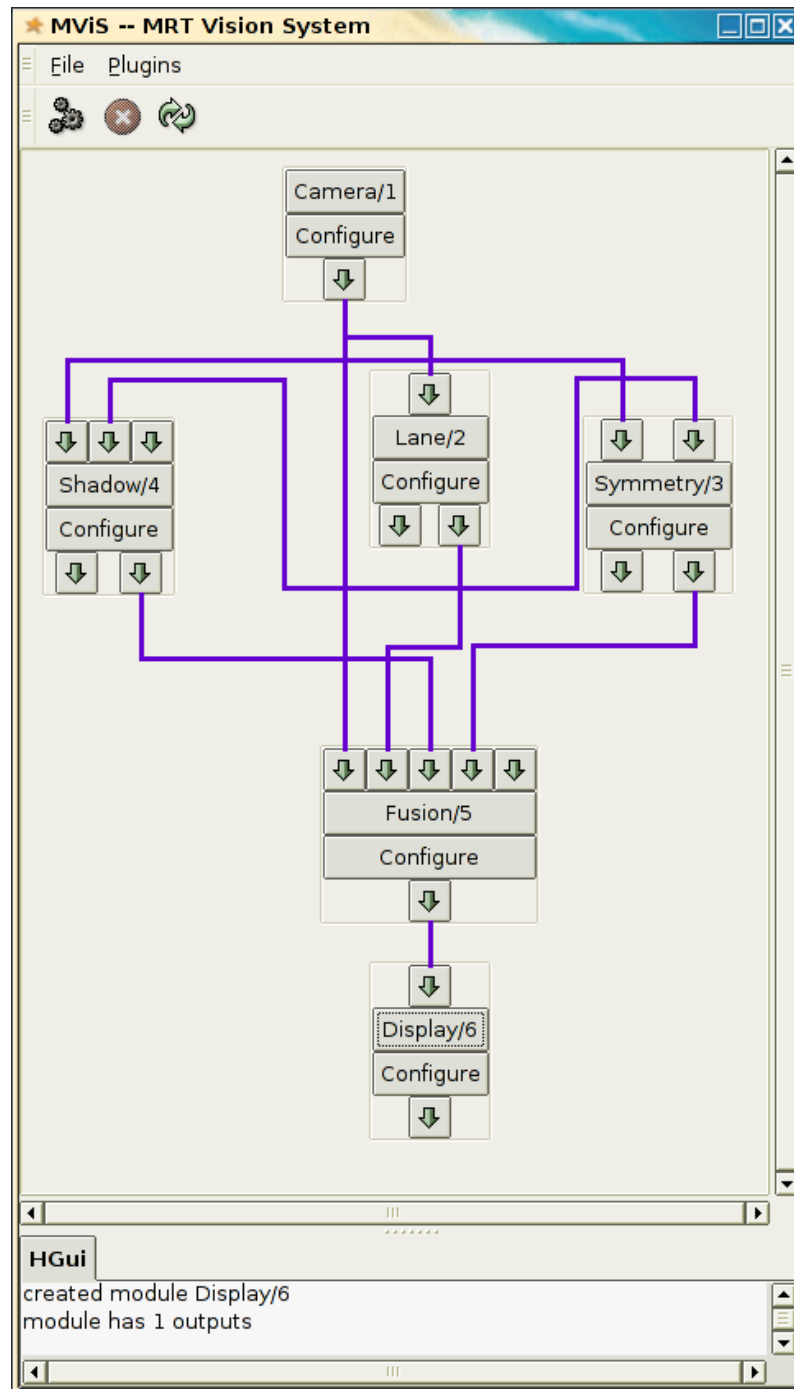


Bild 4.2: Der Algorithmus im MRT Vision System.

der Stereodaten gewinnt. Zwar ist dieses Verfahren genauer, benötigt jedoch die Bilder beider Kameras, was den Zielen der Arbeit widerspricht.

Fusion: Die Algorithmen für das *Interacting Multiple Model*-Filter und die *Cheap Joint Probabilistic Data Association*-Filter sind hier zusammengefasst. Neben einem Eingang für die Bilddaten, die lediglich zur Darstellung der Er-

gebnisse verwendet werden, verfügt das Modul über Eingänge für Fahrbahneninformation und die Beobachtungen der Schatten- und Symmetriedetektoren. Über einen weiteren Eingang können Beobachtungen weiterer Sensoren hinzugefügt werden. Da keine Aktorik angesteuert wird, werden die Ergebnisse in die Bilddaten eingezeichnet und ausgegeben.

Display: Das Display-Modul dient schließlich zur Anzeige von Bildern und Messdaten.

Alle Module verfügen über Dialogfelder, die eine gezielte Parametrierung erlauben. Diese kann bequem vom in der Fahrgastzelle untergebrachten Laptop-Rechner durchgeführt werden, da sowohl die in Bild 4.2 abgebildete Benutzerschnittstelle mit den Dialogen als auch die Ergebnisse auf diesem dargestellt werden.

5 Experimentelle Ergebnisse

Das in dieser Arbeit entwickelte Verfahren wurde auf einem der Rechner des Versuchsfahrzeuges implementiert. Die Umsetzung setzt keine festen Zeitschritte voraus und ist damit nicht von einer festen Bildrate abhängig bzw. davon, dass alle gelieferten Bilder bearbeitet werden. Dennoch ist die Implementierung selbst in komplexen Verkehrsszenarien in der Lage, alle 25 Bilder pro Sekunde auszuwerten.

Im Folgenden wird beispielhaft eine etwa dreieinhalb Minuten langen Sequenz analysiert, die eine typische Fahrt auf einer Autobahn wiedergibt. Während diese aufgenommen wurde, wurde das Versuchsfahrzeug von einem anderen Fahrzeug überholt, das einen Spurwechsel von der linken auf die mittlere Spur durchführte. Bild 5.1 zeigt dieses Manöver. Nach weiteren etwa 25 Sekunden wechselte das Fahrzeug in größerer Entfernung zurück auf die linke Spur, um weitere Fahrzeuge zu überholen.



Bild 5.1: Das überholende Fahrzeug wechselt von der linken auf die mittlere Spur.

5.1 Vergleich verschiedener Konfigurationen

Die Ergebnisse für die Erfassung und Verfolgung dieses Fahrzeuges sollen nun analysiert werden, um die Eigenschaften des entwickelten Verfahrens zu demonstrieren. Um die verschiedenen Ansätze vergleichen zu können, wurden die Berechnungen auf den Daten dieser zuvor aufgenommenen Sequenz, und damit stets unter gleichen Bedingungen, ausgeführt.

Für den Vergleich sollen drei verschiedene Konfigurationen betrachtet werden:

1. Der vollständige Ansatz, also Fusion der Daten zweier Sensoren (Schatten und Symmetrie) in einem *Interacting Multiple Model*-Filter mit zwei Systemmodellen, die durch *Cheap Joint Probabilistic Data Association*-Filter repräsentiert werden. Das erste Modell beschreibt eine Bewegung mit konstanter Relativgeschwindigkeit, das zweite Modell nimmt eine konstante Relativbeschleunigung an. Dies ist das in dieser Arbeit entwickelte Verfahren.
2. Die Daten der beiden genannten Sensoren werden zur Detektion und Verfolgung mittels eines einzelnen *Cheap Joint Probabilistic Data Association*-Filters fusioniert.
3. Das Fahrzeug wird mit dem *Interacting Multiple Model*-Filter aus der erstgenannten Konfiguration, jedoch nur mit den Daten eines Sensors, verfolgt.

5.1.1 Mehrere Sensoren, mehrere Modelle

Zunächst wurde die Sequenz mit dem in dieser Arbeit entwickelten Verfahren analysiert. In Bild 5.2 ist die Entfernung des beobachteten Fahrzeuges zum Testfahrzeug über der Zeit aufgetragen. Es wird früh erfasst und stabil über 25 Sekunden verfolgt. Als es eine Entfernung von etwa 65 Metern erreicht, scheitert schließlich die weitere Detektion.

Bild 5.3 zeigt die seitliche relative Abweichung des verfolgten Fahrzeuges. Während der ersten fünf Sekunden nach seiner Erfassung wechselt das Auto von der linken auf die mittlere Fahrspur, auf der sich auch das Testfahrzeug befindet. Daher ändert sich die seitliche Abweichung von etwa -3 m auf etwa Null. Hier verweilt es für ungefähr 15 Sekunden, während es sich weiter entfernt. Dann führt es wieder einen Spurwechsel nach links aus, um weitere Fahrzeuge zu überholen.

Da Spurwechselmanöver bei Annahme konstanter Relativgeschwindigkeit eine Modellverletzung darstellen, sollten sie sich im Verlauf der Modellwahrschein-

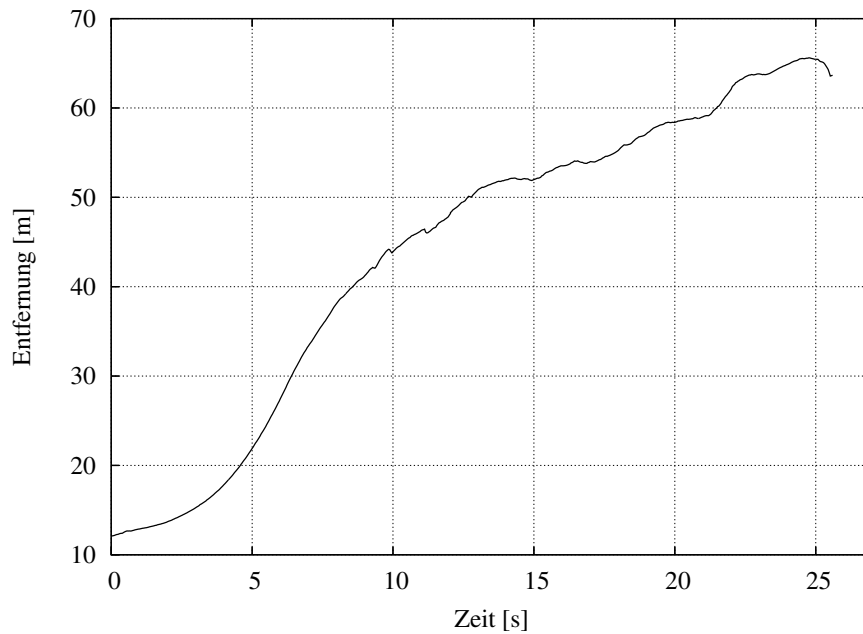


Bild 5.2: Entfernung des verfolgten Fahrzeuges über der Zeit bei Verwendung mehrerer Sensoren und mehrerer Modelle.

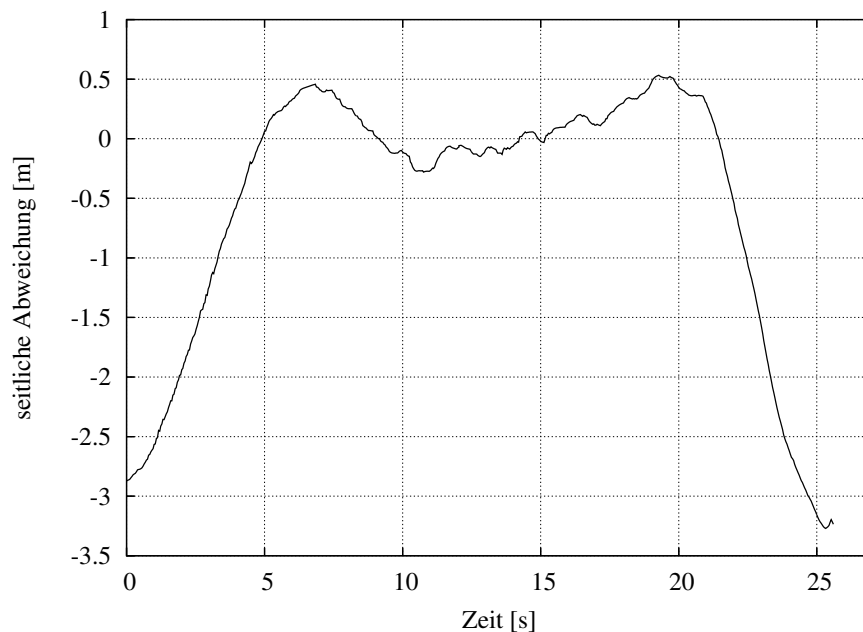


Bild 5.3: Laterale Abweichung des beobachteten Fahrzeuges relativ zum Testfahrzeug. Negative Werte beschreiben Positionen links vom Betrachter, positive Werte sind Positionen auf der rechten Seite.

lichkeiten μ_t^n des *Interacting Multiple Model*-Filters widerspiegeln. Tatsächlich erkennt man dieses Verhalten in Bild 5.4. Hier ist die Modellwahrscheinlichkeit μ_1

für das Modell der konstanten Relativgeschwindigkeit über der Zeit aufgetragen. Vergleicht man diesen Verlauf nun mit Bild 5.3, so erkennt man, dass während der beiden Spurwechselmanöver μ_1 absinkt, womit gleichzeitig das Modell der konstanten Relativbeschleunigung an Bedeutung gewinnt. Damit stellt dieses Modell sicher, dass auch in diesen Situationen eine robuste Verfolgung möglich ist.

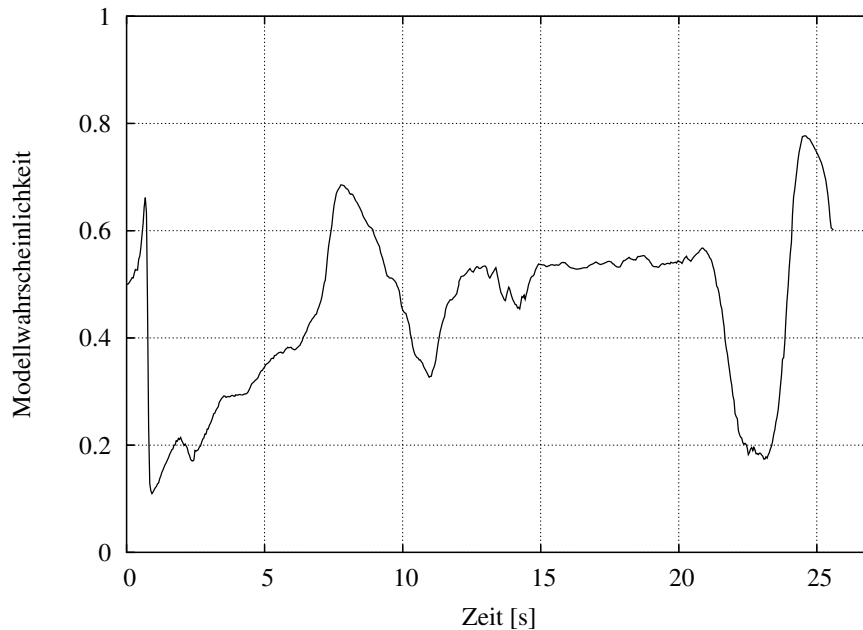


Bild 5.4: Modellwahrscheinlichkeit μ_1 für das Modell der konstanten Relativgeschwindigkeit. Hohe Werte bedeuten, dass die Bewegung des verfolgten Objektes gut durch dieses Modell beschrieben wird, niedrige Werte sind gleichbedeutend mit hohen Wahrscheinlichkeiten für das Modell der konstanten Relativbeschleunigung.

Obwohl sich das verfolgte Fahrzeug mit nahezu konstanter Geschwindigkeit bewegt, nähert sich die Modellwahrscheinlichkeit für das Modell der konstanten Relativgeschwindigkeit nie dem Wert 1,0, was bedeuten würde, dass das Objektverhalten perfekt durch dieses Modell beschrieben ist. Dieser Effekt kommt dadurch zustande, dass Messfehler, vor allem den Nickwinkel betreffend, dazu führen, dass die Tracks ständig leichten Positionsschwankungen unterliegen. Das Modell der konstanten Relativbeschleunigung kann diese Störungen besser kompensieren, wodurch es durchgehend einen relevanten Anteil an der Objektverfolgung behält.

5.1.2 Mehrere Sensoren, ein Modell

Um die Vorteile eines Multi-Modell-Ansatzes gegenüber einem einzelnen Systemmodell zu verdeutlichen, wurde die gleiche Auswertung mit nur einem Systemmodell, das eine Bewegung mit konstanter Relativbewegung beschreibt, durchgeführt. Da sich das *Interacting Multiple Model*-Verfahren mit nur einem Filter zur Beschreibung des Systemverhaltens auf eben dieses Filter reduziert, ist dies äquivalent zu einem *Cheap Joint Probabilistic Data Association*-Filter, das konstante Relativgeschwindigkeit des verfolgten Objektes voraussetzt.

Da keine Konfigurationsparameter verändert wurden, wurden dem Filter die gleichen Messungen wie in dem in Abschnitt 5.1.1 beschriebenen Fall zugeführt. In Bild 5.5 ist wieder die Entfernung des verfolgten Fahrzeuges über der Zeit aufgetragen. Mit nur diesem einen Systemmodell kann das beobachtete Objekt nicht zuverlässig verfolgt werden. Insgesamt werden vier Tracks initialisiert, jedoch kann keiner davon die wirkliche Objektbewegung beschreiben.

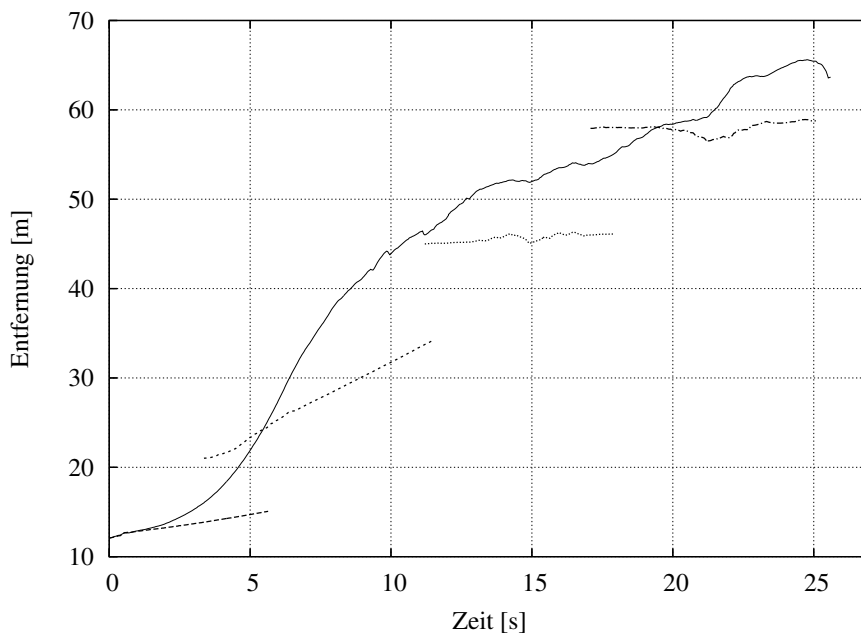


Bild 5.5: Entfernung des verfolgten Fahrzeuges über der Zeit. Die vier nicht zusammenhängenden, gestrichelten Linien erhält man bei Verwendung nur eines Systemmodells, das konstante Relativgeschwindigkeit beschreibt. Zum Vergleich ist noch das Ergebnis des vollständigen Verfahrens aus Bild 5.2 wiedergegeben (durchgezogene Linie).

Offensichtlich ist das Filter mit der Modellannahme der konstanten Relativgeschwindigkeit zu träge. Dies ließ auch schon der letzte Versuch unter Verwendung

beider Systemmodelle vermuten. Hier zeigte sich, dass die Modellwahrscheinlichkeit für das Modell der konstanten Relativbeschleunigung mehrfach ansteigt. Zwar kann der Effekt durch eine Anpassung des Systemrauschens ausgeglichen werden, doch führt dies wiederum zu Trackverlusten.

Eine Bewegung mit konstanter Relativgeschwindigkeit lässt sich jedoch auch als Sonderfall der Bewegung mit konstanter Relativbeschleunigung auffassen, wobei die Beschleunigung Null ist. Aus dieser Überlegung heraus stellt sich die Frage, ob nicht ein einzelnes Filter mit dem Modell der konstanten Relativbeschleunigung hinreichend ist. Daher wurde eine weitere Auswertung mit einem einzelnen Modell durchgeführt, diesmal dem der konstanten Relativbeschleunigung. Bild 5.6 gibt das Ergebnis wieder. Zwar sind die Tracks nicht mehr zu träge, jedoch reagieren sie zu stark auf Störungen durch falsche Beobachtungen. Somit entfernt sich die Schätzung vom wahren Objekt, wodurch keine korrekten Beobachtungen mehr assoziiert werden können. Dies führt schließlich wieder zum Trackverlust.

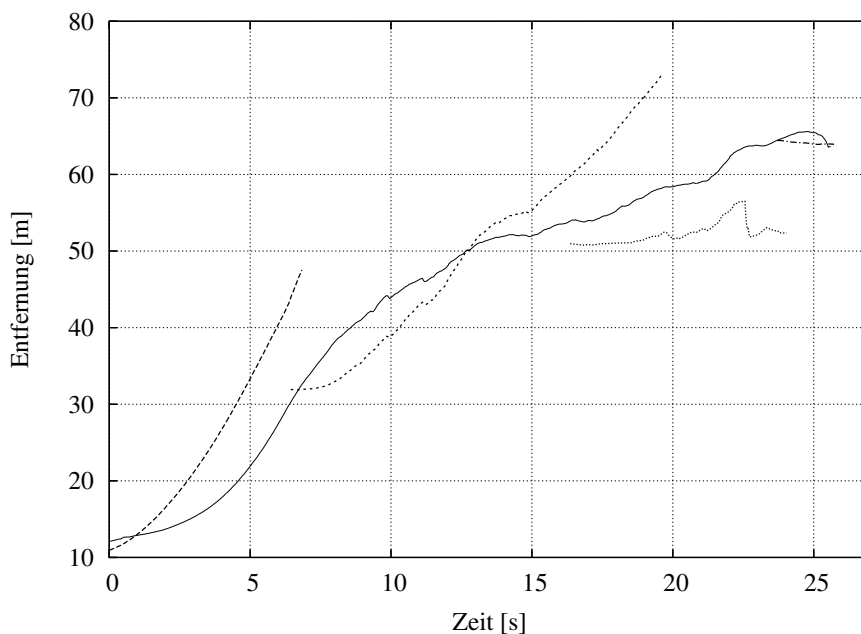


Bild 5.6: Entfernung des verfolgten Fahrzeuges über der Zeit. Hier wurde ein einzelnes Systemmodell, das eine konstante Relativbeschleunigung beschreibt, eingesetzt. Wieder sind zum Vergleich die Ergebnisse des vollständigen Verfahrens aus Bild 5.2 aufgetragen.

Diese Beispiele veranschaulichen die Vorteile des *Interacting Multiple Model*-Verfahrens bei der Verfolgung von Fahrzeugen in wechselnden Situationen.

5.1.3 Ein Sensor, mehrere Modelle

Bisher wurde veranschaulicht, dass das entwickelte Multi-Modell-Verfahren Vorteile gegenüber einem Verfahren mit nur einem Systemmodell hat. Der Gewinn durch die Fusion der Informationen mehrerer Sensoren wurde dabei jedoch nicht betrachtet. Um diesen Punkt zu beleuchten, wurde ein Sensor von der in Abschnitt 5.1.1 beschriebenen Konfiguration entfernt.

Eine Auswertung der von den Sensoren aufgezeichneten Beobachtungen zeigt schnell, dass das Schattenmerkmal dem Symmetriemerkmal deutlich überlegen ist. Dies kommt daher, dass die Bestimmung der Unterkante aufgrund von Symmetrie schwieriger ist als beim Schattenmerkmal, was zu falschen Entfernungswerten führt. Zudem ist der Rauschanteil in den Beobachtungen des Symmetriemerkmals höher als in denen des Schattenmerkmals. Um zu zeigen, dass das Symmetriemerkmal dennoch einen positiven Beitrag leistet, wurde es hier entfernt.

Bild 5.7 zeigt, dass die Objektverfolgung ohne das Symmetriemerkmal weniger robust ist, obwohl dieses Merkmal weniger zuverlässig ist. Insgesamt ergaben sich sieben verschiedene Tracks, da das beobachtete Objekt mehrfach verloren wurde.

Dazu kommt es vor allem, wenn das Schattenmerkmal keine lückenlosen Beobachtungen liefern kann. Tracks werden dann durch Rauschen so stark verfälscht, dass sie die verfolgten Objekte nicht mehr korrekt beschreiben. In solchen Fällen kann selbst ein weniger zuverlässiger Sensor, wie der Symmetriedetektor, die Lücken schließen und den Trackverlust verhindern.

5.2 Untersuchung der Detektionsgüte

Um ein quantitatives Maß zur Beurteilung des Verfahrens zu erhalten, wurden in jedem zehnten Bild der 220 Sekunden langen Sequenz alle Fahrzeuge von Hand markiert. Dies entspricht einem Vergleichsdatensatz von 550 Bildern. Die Sequenz wurde während einer Autobahnfahrt bei starker Sonneneinstrahlung und mittlerer Verkehrsdichte aufgezeichnet. Anschließend wurden diese Markierungen mit den Ergebnissen des beschriebenen Verfahrens verglichen. Auf diese Weise ließen sich Detektions- und Fehlalarmrate bestimmen.

Die Detektionsrate DR ist definiert als das Verhältnis der Anzahl der wahren Detektionen D_t und der Anzahl der Markierungen L, d. h. der vorhandenen Objekte:

$$DR = \frac{D_t}{L}. \quad (5.1)$$

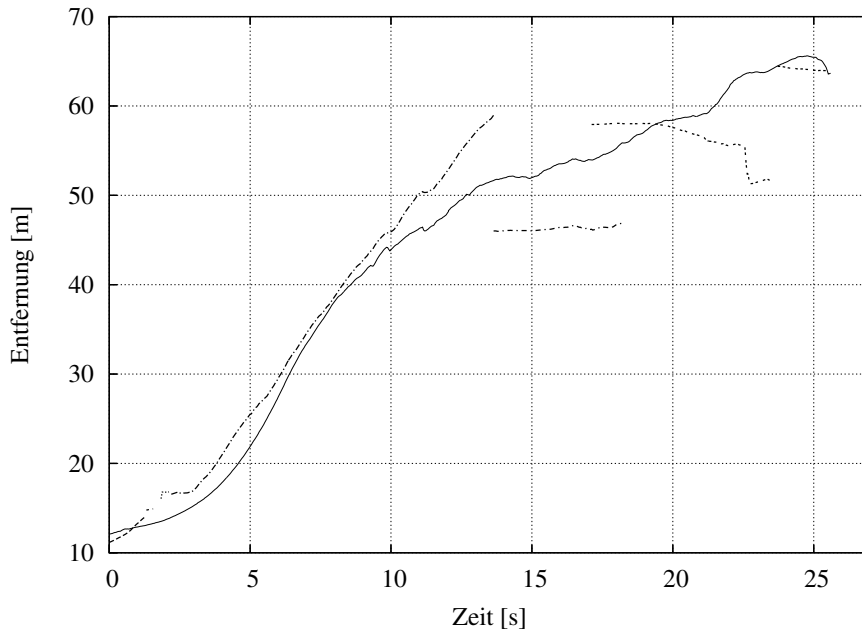


Bild 5.7: Entfernung des beobachteten Fahrzeuges über der Zeit. Ohne Symmetriemerkmale kann das Objekt nicht durchgehend erfasst werden. Wieder sind die Ergebnisse der vollständigen Verfahrens zum Vergleich abgebildet.

Würden alle Fahrzeuge korrekt erkannt, so würde dies eine Detektionsrate von $DR = 1,0$ zur Folge haben. Dies allein ist jedoch noch nicht ausreichend, um eine Aussage über die Güte des Verfahrens zu treffen. An ein Verfahren zur Detektion von Fahrzeugen für Fahrerassistenzsysteme wird nicht nur die Anforderung gestellt, möglichst alle Fahrzeuge zu erfassen. Es soll auch möglichst wenige falsche Hypothesen liefern, d. h. Objektbeschreibungen, die keinem real existierenden Objekt entsprechen.

In der Praxis lassen sich solche Fehlalarme nicht völlig ausschließen. Daher wird eine möglichst niedrige Fehlalarmrate FAR angestrebt:

$$FAR = \frac{D_f}{D_t + D_f} . \quad (5.2)$$

D_f ist hier die Anzahl der falschen Hypothesen. Wäre diese genauso groß wie die Anzahl der wahren Detektionen, so ergäbe sich eine Fehlalarmrate $FAR = 1,0$.

Dem implementierten Verfahren stehen keine Informationen über die Begrenzungen der Fahrbahn zur Verfügung. Daher wird das komplette Bild bis zur Horizontlinie untersucht. Dadurch wird auch eine bedeutende Anzahl falscher Beobachtungen ausgelöst, die durch Infrastrukturelemente, wie z. B. Leitplanken, oder Büsche und Bäume am Fahrbahnrand hervorgerufen werden. Bild 5.8 zeigt beispielhaft einige wahre und falsche Detektionen.



Bild 5.8: Detektionsbeispiele: Tracks sind durch zwei gelbe Balken gekennzeichnet, die die Modellwahrscheinlichkeiten, einem Balkendiagramm gleich, wiedergeben. Die Höhe des linken Balkens repräsentiert das Modell der konstanten Relativgeschwindigkeit, die des rechten Balkens das der konstanten Relativbeschleunigung. Im mittleren Bild sind an der Leitplanke zwei Fehldetektionen zu erkennen.

Andererseits sind, abhängig von der konkreten Anwendung, nur bestimmte Regionen von Interesse. So ist z. B. für Systeme, die den Abstand zum vorausfahrenden Fahrzeug konstant halten sollen, vor allem die eigene Fahrspur von Interesse.

Ausschnitt	Detektionsrate	Fehlalarmrate
eigene Fahrspur	95,6%	3,4%
befahrbare Fläche	86,2%	8,0%
alles	71,5%	47,4%

Tabelle 5.1: Detektions- und Fehlalarmraten

Daher wurden Detektions- und Fehlalarmrate für drei verschiedene Bereiche bestimmt:

- lediglich die Fahrspur, auf der sich das beobachtende Fahrzeug befindet,
- die komplette befahrbare Fläche und
- der ganze Sichtbereich bis zum Horizont.

Da die Informationen über die eigene Fahrspur und die über die befahrbare Fläche nicht vorliegen, wurden auch sie von Hand eingetragen.

Tabelle 5.1 gibt die Ergebnisse für diese Gebiete wieder. 95,6% aller Fahrzeuge auf der eigenen Fahrspur wurden von dem System erkannt, wobei 3,4% aller Detektionen Fehlalarme waren. Mit wachsender Suchregion fällt die Detektionsrate, während die Fehlalarmrate steigt.

Diese Ergebnisse zeigen, dass einfache Fahrerassistenzsysteme schon alleine auf Basis des entwickelten Verfahrens zur Fusion monoskopischer Videomerkmale realisiert werden können. Allerdings bereiten ungünstige Witterungsbedingungen wie z. B. Regen, Schnee oder Dunkelheit Probleme. Daher ist für eine fortgeschrittene Anwendung eine Kombination mit anderen Sensoren notwendig. In einem solchen Gesamtsystem kann das Verfahren einen hilfreichen Beitrag leisten.

6 Zusammenfassung

Im Rahmen dieser Arbeit wurde ein Verfahren zur videobasierten Detektion und Verfolgung von Fahrzeugen für Fahrerassistenzsysteme entwickelt. Es bietet eine Alternative zu den in aktuellen Systemen eingesetzten Radarsensoren, kann aber auch zu deren Ergänzung eingesetzt werden.

Fahrerassistenzsysteme unterstützen den Fahrer bei seinen Aufgaben, um durch diese Entlastung seine Konzentrations- und Reaktionsfähigkeit zu steigern. Eine weitere Aufgabe besteht darin, den Fahrer in kritischen Situationen frühzeitig zu warnen, um eine rechtzeitige Reaktion zu ermöglichen. Sie ermöglichen z. B. eine automatisierte Längsführung oder signalisieren ein unbeabsichtigtes Verlassen der Fahrspur.

All diese Systeme benötigen dazu eine exakte und umfassende Beschreibung der Fahrzeugumgebung. Hier bietet sich die Verwendung eines Videosensors an. Das in der vorliegenden Arbeit beschriebene Verfahren wertet monoskopische Einzelbildmerkmale aus, um vorausfahrenden und entgegenkommenden Verkehr zu erfassen. Dazu untersucht es charakteristische Merkmale in den Bildern einer Videokamera, die ein Indikator für die Präsenz eines Fahrzeuges sind. Entsprechende Extraktionsalgorithmen wurden für die Merkmale Schatten und Symmetrie entwickelt. Eine zeitliche Verfolgung dieser Merkmale erlaubt es, die Zuverlässigkeit des Verfahrens zu erhöhen. Zudem werden so weitere Objektparameter wie Geschwindigkeit und Bewegungsrichtung bestimmt.

Einen Überblick über bestehende Systeme gibt Kapitel 1, bevor es auf die Anforderungen an das entwickelte Verfahren eingeht. Daraus wird das Sensorconcept und der schematische Aufbau abgeleitet. Dieser umfasst, neben zusätzlichen Elementen wie z. B. der Bildaufnahme und der Trackkontrolle, die Hauptkomponenten Merkmalsextraktion, Fusion und zeitliche Verfolgung.

So stellt Kapitel 2 zunächst Möglichkeiten zur Detektion von Fahrzeugen in Videosequenzen vor und beschreibt dann detailliert die hier entwickelten Verfahren zur Merkmalsextraktion. Die Merkmale Schatten und Symmetrie bieten sich deshalb an, weil sie trotz des einfachen zugrunde liegenden Modells gute Erkennungsraten ermöglichen. Gerade die einfache Modellierung bringt den Vorteil mit sich, dass von ihr nicht nur bestimmte Fahrzeugtypen, sondern generisch nahezu alle relevanten Fahrzeuge repräsentiert werden. Allerdings erwächst daraus auch das

Problem, dass die so erhaltenen Beobachtungen eine Vielzahl falscher Hypothesen beinhalten.

Die Lokalisierung von monoskopischen Merkmalen liefert jedoch lediglich zweidimensionale Positionsinformationen in Bildkoordinaten. Die für Fahrerassistenzsysteme essentielle Entfernungsinformation muss somit auf andere Weise gewonnen werden. Daher wurde die Information aus einem externen Modul zur Erfassung der Fahrbahnorientierung mit den Schritten zur Merkmalsextraktion kombiniert, um dreidimensionale Objektkoordinaten bestimmen zu können.

Die Fusion und zeitliche Verfolgung werden in Kapitel 3 behandelt. Da die gewonnenen Beobachtungen einen hohen Rauschanteil enthalten, kommt ihrer Assoziation zu den vorhandenen Objektbeschreibungen eine besondere Bedeutung zu. Die eingesetzten *Joint Probabilistic Data Association*-Filter nehmen eine probabilistische Zuordnung der Beobachtungen zu den vorhandenen Objektbeschreibungen vor. Mit der hier entwickelten Erweiterung können damit auch Beobachtungen mit uneinheitlicher Fehlerkovarianz verarbeitet werden. Damit werden alle Unsicherheiten, die bei der Erfassung der Fahrzeuge auftreten, durchgängig berücksichtigt.

Allerdings stellt sich eine weitere Herausforderung: Da das System in einem weiten Spektrum an Verkehrssituationen und Manövern korrekte Ergebnisse liefern muss, ist die Formulierung eines einzelnen, umfassenden dynamischen Modells für die Objektbewegung nicht möglich. Ein solches Modell ist jedoch Kernstück des *Joint Probabilistic Data Association*-Filters. Daher werden mehrere dynamische Modelle formuliert, die jeweils durch ein eigenes Filter repräsentiert sind, und in einem Multi-Modell-Verfahren verknüpft.

Hier findet ein *Interacting Multiple Model*-Ansatz Verwendung. Die Modellwahrscheinlichkeiten geben an, wie gut die Schätzung eines Filters eines Bewegungsmodells das Objektverhalten beschreibt. Sie werden dann bei der Berechnung einer kombinierten Schätzung des Systemzustandes berücksichtigt. Eine Besonderheit des Verfahrens ist, dass es eine Klassifikation der beobachteten Situation erlaubt: Aus den Modellwahrscheinlichkeiten lassen sich Rückschlüsse auf das vorliegende Bewegungsmodell bzw. Fahrmanöver ziehen.

Allerdings müssen bei der Integration in einen *Interacting Multiple Model*-Ansatz die Eigenschaften der modellspezifischen Filter berücksichtigt werden. Die vorliegende Arbeit führt diese Integration für das *Joint Probabilistic Data Association*-Filter konsistent durch.

Um alle relevanten Verkehrssituationen erfassen zu können, werden zwei Systemmodelle eingesetzt. Übliche Folgefahrten auf Autobahnen oder Landstraßen werden durch eine Bewegung mit konstanter Relativgeschwindigkeit beschrieben. Für Situationen, die nach einer höheren Dynamik verlangen, wird das Objektverhalten

als Bewegung mit konstanter Relativbeschleunigung modelliert. Dank der leichten Integrierbarkeit weiterer Modelle kann das Verfahren auch für bisher unberücksichtigte Situationen angepasst werden.

Die Fusion der Merkmale Schatten und Symmetrie geschieht durch sequentielles Einbeziehen der entsprechenden Beobachtungen. Dadurch kann das Verfahren leicht um andere Merkmale erweitert werden. Selbst andere Sensoren, die die vorgegebenen Anforderungen erfüllen, können integriert werden. Damit bietet es eine flexibel anpassbare Plattform zur Detektion und Verfolgung von Fahrzeugen.

Das entwickelte Verfahren wurde für den Versuchsträger des Instituts für Mess- und Regelungstechnik der Universität Karlsruhe implementiert. Sowohl der Aufbau des Versuchsträgers als auch die Softwarearchitektur werden in Kapitel 4 umrissen.

Um die Stärken des Verfahrens in der praktischen Anwendung zu verdeutlichen, vergleicht Kapitel 5 anhand von aufgezeichneten Testsequenzen dessen Ergebnisse mit denen eines Ein-Modell-Ansatzes und denen eines Ein-Sensor-Ansatzes. Dabei wird deutlich, dass sowohl die Berücksichtigung mehrerer Modelle als auch mehrerer Sensoren für eine zuverlässige Erkennung und Verfolgung unerlässlich ist.

Das vorliegende Gesamtsystem zeigt, wie zuverlässig eine Detektion und Verfolgung von Fahrzeugen selbst mit schwachen Merkmalen sein kann. So ließe sich schon mit ihm alleine eine Komfortfunktion wie z. B. eine automatische Längsführung realisieren. Dies demonstriert eindrucksvoll die Vorteile der Kombination mehrerer Merkmale. Doch sind der videobasierten Fahrzeugdetektion Grenzen gesetzt. Gerade bei Regen, Schnee oder Dunkelheit ist sie nur noch eingeschränkt einsetzbar. Erst die Kombination mit anderen Sensoren kann hier Abhilfe schaffen. Aufgrund seines großen Erfassungsbereiches ist der entwickelte Sensor eine wertvolle Ergänzung für andere Sensoren in einer solchen erweiterten Anwendung.

Literaturverzeichnis

Aufrère u. a. 2000

Aufre`re, R. ; Ch`errier, R. ; Ch`errier, F.: A Fast and Robust Vision Based Road Following Algorithm. In: *Proc. IEEE Intelligent Vehicles Symposium* (2000), S. 192–197

Bar-Shalom 1990

Bar-Shalom, Yaakov (Hrsg.): *Multitarget-Multisensor Tracking: Advanced Applications*. Artech House, 1990

Bar-Shalom u. Blair 2000

Bar-Shalom, Yaakov (Hrsg.) ; Blair, William D. (Hrsg.): *Multitarget-Multisensor Tracking: Applications and Advances*. Bd. 3. Artech House, 2000

Bar-Shalom u. Fortmann 1988

Bar-Shalom, Yaakov ; Fortmann, Thomas E.: *Tracking and Data Association*. Academic Press, 1988

Barron u. a. 1994

Barron, J. L. ; Fleet, D. J. ; Beaumont, S. S.: Performance of Optical Flow Techniques. In: *International Journal of Computer Vision* 12 (1994), Nr. 1, S. 43–77

Behringer u. a. 1992

Behringer, R. ; V. H., V. ; D., D.: Road and Relative Ego-state Recognition. In: *Proc. IEEE Intelligent Vehicles Symposium* (1992), S. 385–390

Bensrhair u. a. 2001

Bensrhair, A. ; Bensrhair, M. ; Bensrhair, A. ; M., P. ; M., S. ; T., G.: A Cooperative Approach to Vision-based Vehicle Detection. In: *Proc. IEEE Intelligent Transportation Systems* (2001), August, S. 207–212

Bertozzi u. Broggi 1998

Bertozzi, M. ; Broggi, A.: GOLD: A Parallel Real-time Stereo Vision System for Generic Obstacle and Lane Detection. In: *IEEE Transactions on Image Processing* 7 (1998), Januar, Nr. 1, S. 62–81

Bertozzi u. a. 2000

Bertozzi, M.; Broggi, A.; Frazzetta, A.; Nanni, S.: Stereo Vision-based Vehicle Detection. In: *Proc. of the IEEE Intelligent Vehicles Symposium 2000* (2000), S. 39–44

Betke u. a. 1996

Betke, Margit; Hough, Esin; Durrant-Webb, Larry S.: Multiple Vehicle Detection and Tracking in Hard Real-Time. In: *Proc. IEEE Intelligent Vehicles Symposium* (1996), September, S. 351–356

Blackman u. Popoli 1999

Blackman, Samuel; Popoli, Robert: *Design and Analysis of Modern Tracking Systems*. Artech House, 1999

Blake u. Yuille 1992

Blake, Andrew (Hrsg.); Yuille, Alan (Hrsg.): *Active Vision*. The MIT Press, 1992

Blom 1984

Blom, H. A. P.: An Efficient Filter for Abruptly Changing Systems. In: *Proceedings of 23rd Conference on Decision and Control* (1984), Dezember, S. 656–658

Borgefors 1988

Borgefors, Gunilla: Hierarchical Chamfer Matching: A Parametric Edge Matching Algorithm. In: *Transactions on Pattern Analysis and Machine Intelligence* 10 (1988), November, Nr. 6

Broggi u. a. 2000

Broggi, A.; Bertozzi, M.; Frazzetta, A.; Sisti, M.: Shape-based Pedestrian Detection. In: *IEEE Intelligent Vehicles Symposium* (2000), S. 215–220

Brown 1983

Brown, Robert G.: *Introduction to Random Signal Analysis and Kalman Filtering*. John Wiley & Sons, 1983

Cabani u. a. 2005

Cabani, I.; Tognoni, G.; Bertozzi, A.: Color-based Detection of Vehicle Lights. In: *IEEE Proc. Intelligent Vehicles Symposium 2005* (2005), Juni, S. 278–283

Canny 1986

Canny, John: A Computational Approach to Edge Detection. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-8* (1986), November, Nr. 6, S. 679–698

Chapuis u. a. 2002

C , Roland ; A , Romuald ; C , Frédéric: Accurate Road Following and Reconstruction by Computer Vision. In: *IEEE Transactions on Intelligent Transportation Systems* 3 (2002), 4 Dezember, Nr. 4, S. 261–270

Charkari u. Mori 1993

C , N. M. ; M , H.: A New Approach for Real Time Moving Vehicle Detection. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems* (1993), S. 273–278

Curio u. a. 2000

C , C. ; E , J. ; K , T. ; T , C. ; S , W. von: Walking Pedestrian Recognition. In: *IEEE Transactions on Intelligent Transportation Systems* 1 (2000), September, Nr. 3, S. 155–163

Dang u. Hoffmann 2006

D , T. ; H , C.: Tracking Camera Parameters of an Active Stereo Rig. In: *28th Annual Symposium of the German Association for Pattern Recognition (DAGM 2006)* (2006), September

Dasarathy 1991

D , Belur V. (Hrsg.): *Nearest Neighbor (NN) Norms : NN Pattern Classification Techniques*. IEEE Computer Society Press, 1991

Du u. Papanikolopoulos 1997

D , Y. ; P , N. P.: Real-Time Vehicle Following through a Novel Symmetry-Based Approach. In: *Proc. of the IEEE 1997 International Conference on Robotics and Automation* (1997), April, S. 3160–3165

Duda u. a. 2000

D , R. O. ; H , P. E. ; S , D. G.: *Pattern Classification*. zweite Auflage. Wiley, 2000

Europäische Kommission 2001

E " K : *Weissbuch – Die europäische Verkehrspolitik bis 2010: Weichenstellungen für die Zukunft*. ISBN: 92-894-0339-X : Luxemburg: Amt für Amtliche Veröffentlichungen der Europäischen Gemeinschaften, 2001

Fang u. a. 2003

F , Chiung-Yao ; C , Sei-Wang ; F , Chiou-Shann: Road-sign Detection and Tracking. In: *IEEE Transactions on Vehicular Technology* 52 (2003), September, Nr. 5, S. 1329–1341

Faugeras 1996

Faugeras, Olivier: *Three-Dimensional Computer Vision: A Geometric Viewpoint*. zweite Auflage. The MIT Press, 1996

Fleischer u. a. 2002

Fleischer, K.; Nishikawa, H.-H.; Rasmussen, T.M.: 3D-Model-Based-Vision for In-nercity Driving Scenes. In: *Proc. of the IEEE Intelligent Vehicles Symposium 2002 2* (2002), S. 477–482

Fox u. a. 2003

Fox, D.; Hightower, J.; Lathrop, L.; Sridharan, D.; Borriello, G.: Bayesian Filtering for Location Estimation. In: *IEEE Pervasive Computing 2* (2003), Nr. 3, S. 24–33

Franke u. Kutzbach 1996

Franke, U.; Kutzbach, I.: Fast Stereo Based Object Detection for Stop&Go Traffic. In: *Proc. IEEE Intelligent Vehicles Symposium* (1996), September, S. 339–344

Fritz u. a. 2004

Fritz, H.; Giese, A.; Sauer, H.; Baur, C.: CHAUFFEUR Assistant: A Driver Assistance System for Commercial Vehicles Based on Fusion of Advanced ACC and Lane Keeping. In: *Proc. IEEE Intelligent Vehicles Symposium* (2004), Juni, S. 495–500

Gavrila u. Philomin 1999

Gavrila, D. M.; Philomin, V.: Real-time Object Detection for Smart Vehicles. In: *Proc. IEEE International Conference on Computer Vision* (1999)

Giachetti u. a. 1994

Giachetti, A.; Cacciatore, M.; Scalet, R.; Scalet, A.: The Recovery of Optical Flow for Intelligent Cruise Control. In: *Proc. IEEE Intelligent Vehicles Symposium* (1994), S. 91–96

Goerick 1994

Goerick, C.: Local Orientation Coding and Adaptive Thresholding for Real Time Early Vision / Institut für Neuroinformatik, Ruhr-Universität Bochum. 1994 (IRINI 94-05). – Internal Report

Goldbeck u. Huertgen 1999

Goldbeck, J.; Huertgen, B.: Lane Detection and Tracking by Video Sensors. In: *IEEE Intelligent Transportation Systems Conference* (1999), S. 74–79

Hall u. Llinas 2001

H , David L. (Hrsg.) ; L , James (Hrsg.): *Handbook of Multisensor Data Fusion*. CRC Press, 2001

Handmann u. a. 1998

H , Uwe ; L , Gesa ; S , Werner von: Fusion von Basisalgorithmen zur Segmentierung von Straßenverkehrsszenen. In: *20. DAGM-Symposium* (1998), S. 101–108

He u. a. 2004

H , Yinghua ; W , Hong ; Z , Bo: Color-based Road Detection in Urban Traffic Scenes. In: *IEEE Transactions on Intelligent Transportation Systems* 5 (2004), Dezember, Nr. 4, S. 309–318

Hilario u. a. 2005

H , C. ; C , J. M. ; A , J. M. ; E , A. de l.: Pyramidal Image Analysis for Vehicle Detection. In: *Proc. IEEE Intelligent Vehicles Symposium* (2005), S. 88–93

Hoffmann 2006

H , C.: Fusing Multiple 2D Visual Features for Vehicle Detection. In: *Proceedings of the IEEE Intelligent Vehicles Symposium*, 2006, S. 406–411

Hoffmann u. Dang 2006

H , C. ; D , T.: Cheap Joint Probabilistic Data Association Filters in an Interacting Multiple Model Design. In: *Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, 2006, S. 197–202

Hoffmann u. a. 2004

H , C. ; D , T. ; S , C.: Vehicle Detection Fusing 2D Visual Features. In: *Proceedings of the IEEE Intelligent Vehicles Symposium*, 2004, S. 280–285

Horn 2005

H , J.: Bildbasierte vektorielle Geschwindigkeitsmessung texturierter Oberflächen. In: *tm – Technisches Messen* Vol. 72 (2005), Nr. 10, S. 556–565

Hummel u. a. 2006

H , B. ; K , S. ; D , T. ; D , C. ; S , C.: Vision-based Path Planning in Unstructured Environments. In: *Proc. IEEE Intelligent Vehicles Symposium* (2006), Juni, S. 176–181

Jähne 1989

Jähne, Bernd: *Digitale Bildverarbeitung*. Springer-Verlag, 1989

Julier u. Uhlmann 1997

Julier, S.; Uhlmann, J.: A New Extension of the Kalman Filter to Nonlinear Systems. In: *Int. Symposium Aerospace/Defense Sensing, Simulation and Controls* (1997)

Kalinke u. von Seelen 1996

Kalinke, T.; von Seelen, W.: Entropie als Maß des lokalen Informationsgehalts in Bildern zur Realisierung einer Aufmerksamkeitssteuerung. In: *Mustererkennung 1996* (1996), S. 627–634

Kato u. a. 2002

Kato, T.; Nishimura, Y.; Mizuno, I.: An Obstacle Detection Method by Fusion of Radar and Motion Stereo. In: *IEEE Transactions on Intelligent Transportation Systems* 3 (2002), September, Nr. 3

Kim u. a. 2005

Kim, SamYong; Oh, Se-Young; Kim, JeongKwan; Rhee, YoungWoo; Kim, Kwangsoo; Park, Sang-Cheol; Park, KyongHa: Front and Rear Vehicle Detection and Tracking in the Day and Night Times Using Vision and Sonar Sensor Fusion. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems* (2005), S. 2173 – 2178

Kim 2006

Kim, Z.: Realtime Lane Tracking of Curved Local Roads. In: *Proc. IEEE Intelligent Transportation Systems Conference* (2006), S. 1149–1155

Klein 2004

Klein, Lawrence A.: *Sensor and Data Fusion – A Tool for Information Assessment and Decision Making*. SPIE Press, 2004

Kuehnle 1991

Kuehnle, A.: Symmetry-based Recognition of Vehicle Rears. In: *Pattern Recognition Letters* 12 (1991), Nr. 4, S. 249–258

Langer u. Jochem 1996

Langer, D.; Jochem, T.: Fusing Radar and Vision for Detecting, Classifying and Avoiding Roadway Obstacles. In: *Proc. IEEE Intelligent Vehicles Symposium* (1996), September, S. 333–338

Liu u. Srinath 1990

Liu, Hong-Chih; Srinath, Mandyam D.: Partial Shape Classification Using

Contour Matching in Distance Transformation. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12 (1990), November, Nr. 11, S. 1072–179

Lorusso u. De Micheli 1996

Lorusso, A.; De Micheli, E.: An Approach to Obstacle Detection and Steering Control from Optical Flow. In: *Proc. IEEE Intelligent Vehicles Symposium* (1996), September, S. 357–362

Mori u. Charkari 1993

Mori, H.; Charkari, N. M.: Shadow and Rhythm as Sign Patterns of Obstacle Detection. In: *IEEE International Symposium On Industrial Electronics* (1993), Juni, S. 271–277

Ninomiya u. a. 1995

Ninomiya, Y.; Mori, S.; Ohtsuka, M.; Hara, Y.; Saito, T.: A Real-time Vision for Intelligent Vehicles. In: *Proc. IEEE Intelligent Vehicles Symposium* (1995), September, S. 315–320

Puente León 1999

Puente León, Fernando: *Automatische Identifikation von Schußwaffen*. ISBN: 3-18-378708-3 : VDI-Verlag, 1999 (787)

Reid 1997

Reid, D.: An Algorithm for Tracking Multiple Targets. In: *IEEE Transactions on Automatic Control* 24 (1997), Dezember, Nr. 6, S. 843–854

Rotaru u. a. 2004

Rotaru, C.; Gavruta, T.; Zlatos, J.: Extracting Road Features from Color Images Using a Cognitive Approach. In: *IEEE Intelligent Vehicles Symposium* (2004), S. 298–303

Schmitz 2005

Schmitz, Carsten: *Adaptiver Spurverlassenswarner mit fahrerabsichts- und fahrerzustandsabhängiger Warnstrategie*. Shaker Verlag, 2005

Shannon 1948

Shannon, C. E.: A Mathematical Theory of Communication. In: *The Bell System Technical Journal* 27 (1948), Juli, S. 379–423, 623–656

Srinivasa 2002

Srinivasa, N.: Vision-based Vehicle Detection and Tracking Method for Forward Collision Warning in Automobiles. In: *IEEE Intelligent Vehicle Symposium* Vol. 2 (2002), S. 626–631

Sun u. a. 2006

Sun, Zehang ; Bhanu, George ; Mueggler, Ronald: On-Road Vehicle Detection: A Review. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (2006), Mai, Nr. 5, S. 694–711

ten Kate u. a. 2004

ten Kate, T. K. ; Lichtenhan, M. B. ; Mueggler, S. E. ; Durrant-Webster, B. J. F. ; Vaidyanathan, A. H. G. ; Gage, F. C. A.: Mid-range and Distant Vehicle Detection With a Mobile Camera. In: *IEEE Intelligent Vehicles Symposium* (2004), Juni, S. 72–77

Thomanek u. a. 1994

Thomanek, F. ; Durrant-Webster, E.D. ; Durrant-Webster, D.: Multiple Object Recognition and Scene Interpretation for Autonomous Road Vehicle Guidance. In: *Proceedings of the Intelligent Vehicles Symposium* (1994), Oktober, S. 231–236

Thrun u. a. 2005

Thrun, S. ; Burgard, W. ; Fox, D.: *Probabilistic Robotics*. MIT Press, 2005

Tokoro u. a. 2004

Tokoro, S. ; Mueggler, K. ; Kamekura, T. ; Nishida, T. ; Arai, K. ; Fujita, K.: Sensor Fusion System for Pre-crash Safety System. In: *Proc. IEEE Intelligent Vehicles Symposium* (2004), Juni, S. 945–950

Tzomakas u. von Seelen 1998

Tzomakas, C. ; Seelen, W. von: *Vehicle Detection in Traffic Scenes Using Shadows*. Internal Report IRINI 98-06, Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany, 1998

van der Mark u. a. 2001

van der Mark, W. ; Gage, F. C. A. ; Heisterkamp, J. C.: Stereo Based Navigation in Unstructured Environments. In: *Proc. of the IEEE Instrumentation and Measurement Technology Conference 2001* (2001), S. 2038–2043

Wang u. a. 2005

Wang, J. ; Bhanu, G. ; Mueggler, R.: Overtaking Vehicle Detection Using Dynamic and Quasi-Static Background Modeling. In: *IEEE Computer Vision and Pattern Recognition* 3 (2005)

Welch u. Bishop 2001

Welch, Greg ; Bishop, Gary: An Introduction to the Kalman Filter. In: *ACM SIGGRAPH Course 8* (2001)

Zielke u. a. 1992

Zielke, Thomas ; Bredemeyer, Michael ; Sauer, Werner von: CARTRACK: Computer Vision-Based Car-Following. In: *IEEE Workshop on Applications of Computer Vision* (1992), S. 156–163

Zielke u. a. 1993

Zielke, Thomas ; Bredemeyer, Michael ; Sauer, Werner von: Intensity and Edge-based Symmetry Detection with an Application to Car-Following. In: *CVGIP: Image Understanding* 58 (1993), September, Nr. 2, S. 177–190

Institut für Mess- und Regelungstechnik Universität Karlsruhe (TH)

Moderne Fahrerassistenzsysteme haben das Ziel, den Fahrer zu entlasten und damit das Fahren angenehmer und sicherer zu machen. Dazu ist eine Erfassung des Fahrzeugumfeldes, insbesondere anderer Verkehrsteilnehmer, unerlässlich. Die in aktuellen Systemen eingesetzten Sensoren können alleine jedoch nicht die hohen Anforderungen an die Zuverlässigkeit und die Größe des Erfassungsbereiches erfüllen.

Die vorliegende Arbeit beschreibt ein Verfahren zur Fahrzeugdetektion anhand monoskopischer Videomerkmale. Aus den Bildern einer im Fahrzeug angebrachten Videokamera werden charakteristische Merkmale extrahiert, die einen Hinweis auf das Vorhandensein eines Fahrzeuges geben. Ein solches Merkmal ist der Schattenbereich, der unter Personen- und Lastkraftwagen auftritt. Die Auswertung dunkler Bildregionen liefert Informationen über mögliche Objektpositionen. Allerdings ist ein einzelnes Merkmal nicht ausreichend, um eine zuverlässige Detektion zu gewährleisten. Daher werden des Weiteren Bereiche analysiert, die sich durch eine Achsensymmetrie auszeichnen, welche sich gerade bei der Front- und Heckpartie von Fahrzeugen beobachten lässt.

Diese Beobachtungen werden fusioniert und zeitlich verfolgt. Da aufgrund der Natur der Merkmale ein hoher Anteil Fehlhypothesen enthalten ist, kommen Joint Probabilistic Data Association-Filter zum Einsatz. Allerdings lässt sich die Vielzahl der zu berücksichtigenden Verkehrssituationen nicht durch ein einziges Bewegungsmodell beschreiben. Daher werden mehrere Bewegungsmodelle in einem Interacting Multiple Model-Verfahren verknüpft.

Experimentelle Ergebnisse belegen, dass erst diese Kombination der Verfahren eine stabile Verfolgung von Fahrzeugen ermöglicht. Außerdem zeigt sich, dass selbst stark rauschbehaftete Sensoren die Zuverlässigkeit des Systems erhöhen können. Das beschriebene Verfahren wurde in einem Versuchsfahrzeug umgesetzt und ist in der Lage, Informationen über Position und Verhalten anderer Verkehrsteilnehmer in Videobildrate bereitzustellen.

ISSN: 1613-4214

ISBN: 978-3-86644-139-2