

Modellbasierter Ansatz zur probabilistischen Interpretation von Fahrsituationen

zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften

von der Fakultät für Informatik
des Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Thomas Schamm

aus Friedrichshafen

Tag der mündlichen Prüfung: 17. Januar 2014

Erster Gutachter: Prof. Dr.-Ing. J. Marius Zöllner

Zweiter Gutachter: Prof. Dr.-Ing. Christoph Stiller

Danksagung

Diese Arbeit entstand während meiner Zeit als wissenschaftlicher Mitarbeiter in der Abteilung Technisch Kognitive Assistenzsysteme (TKS) am FZI Forschungszentrum Informatik. Während dieser Zeit habe ich von vielen Personen Unterstützung bekommen, für die ich mich an dieser Stelle besonders bedanken möchte.

Mein besonderer Dank gilt meinem Doktorvater Prof. J. Marius Zöllner, der es mir ermöglicht hat, in einer unvergleichbaren Umgebung zu promovieren und mich dabei stets voll unterstützt hat. Weiterhin möchte ich Prof. Christoph Stiller für die Übernahme des Korreferats, sowie Prof. Rüdiger Dillmann und Prof. Jürgen Beyerer für die Mitwirkung als Prüfer herzlich danken.

Die Abteilung TKS wurde zu Beginn meiner Promotion am FZI etabliert. Insbesondere den Abteilungsleitern Marco, Thilo und Marcus gilt mein Dank für den Aufbau und die Leitung der Gruppe sowie dem organisatorischen Geschick, in kürzester Zeit eine einmalige Arbeitsumgebung aufzubauen.

Besonders herzlich danken möchte ich den beiden TKS-Urgesteinen Dennis und Thomas: für stets fröhliche und interessante Gespräche, viele fruchtbare Diskussionen bei Eis oder Kaffee über die Konzepte dieser Arbeit sowie für den gemeinsamen Aufbau von grundlegenden Software-Bibliotheken; Dennis für eine tolle Zeit im gemeinsamen Büro, zwei schöne Reisen in die USA und für seine Fähigkeit, Probleme schneller als sein Schatten zu lösen; und Thomas für die Geduld, die ich ihm in manchen Gesprächen über Gott und die Welt ganz sicherlich abverlangt habe.

Weiterer Dank gilt all meinen jetzigen und ehemaligen Kollegen der Abteilung TKS und der Nachbarabteilung IDS. Tobias für die Fortführung der Arbeiten zur Fahrerbeobachtung; Ralf für sein Durchhaltevermögen bei der Kalibrierung von Kameras, sein Wissen über Koordinatentransformationen und für die Zeit im gemeinsamen Büro; Florian für seine wertvollen Beiträge zu dieser Arbeit als Hiwi, während seiner Diplomarbeit und schlussendlich als Kollege. Steffen danke ich für seine gute Freundschaft, die gemeinsam betriebenen Hobbys seit Beginn unserer Studienzeit sowie viele kreative Gespräche; Arne für die konstruktive Zusammenarbeit im letzten Jahr und die Diskussionen über Sinn und Zweck von Bürokratie. Den weiteren Kollegen danke ich für interessante Diskussionen bei unzähligen Kaffees und Kuchen, für Montagsrunden und für ihre spannenden Arbeiten mit Lauf-, Schlangen-, Flug- und mobilen Robotersystemen. Besonders danken möchte ich auch den Damen des Sekretariats für die unermüdliche Unterstützung bei organisatorischen und bürokratischen Dingen.

Mein herzlichster Dank gilt ebenfalls all meinen Studenten, die mich im Rahmen von Studien-, Diplom-, Bachelor- und Masterarbeiten sowie als Hilfwissenschaftler unterstützt haben. Ich danke insbesondere Jürgen, Ruoqing, Christoph, Michael, Ning, Dominik und Hugo. Ihr unermüdliches Engagement erleichterte die Durchführung dieser Arbeit ungemein.

Großen Dank schulde ich auch meinen Korrekturlesern Eike, Steffen, Ralf, Ruth, Reinhold und Kirstin, die trotz teilweiser Fachkenntnis durchhielten (Zitat Eike: „Ich weiß, dass ich gerade etwas gelesen habe ... nur was, das kann ich nicht sagen“) und somit zur Gesamtqualität beitrugen.

Meinen Eltern Ruth und Reinhold gilt mein besonderer Dank. Sie haben mich während der gesamten Ausbildungszeit unterstützt, mir Rückhalt gegeben und mich stets ermuntert, entschlossen und zuversichtlich diese Arbeit durchzuführen. Das Gleiche gilt für meine Schwester Tina, die stets ein offenes Ohr für meine Anliegen hat.

Der größte Dank geht an meine Frau Kirstin. Nur durch ihre bedingungslose Unterstützung konnte diese Arbeit entstehen.

Karlsruhe, im Januar 2014

Thomas Schamm

Kurzfassung

Intelligente Fahrerassistenzsysteme unterstützen den Benutzer bei der Vermeidung von Gefahren, empfehlen effiziente Fahraktionen oder stellen Hinweise über den zukünftigen Fahrtverlauf bereit. Die Erkennung der aktuell vorherrschenden Fahrsituation sowie die Prädiktion der sich daraus entwickelnden Situationen stellt für die Zuverlässigkeit und Akzeptanz dieser Systeme eine wichtige Voraussetzung dar. Fahrsituationen ergeben sich aus dem Zustand des Eigenfahrzeugs, dem Zustand relevanter Objekte im Fahrzeugumfeld, sowie der Veränderung dieser Zustände. In der Interpretation der Fahrsituation liegt eine große Herausforderung, insbesondere in der Analyse der Objektzustände und -relationen. Viele dieser Zustände können nicht exakt bestimmt werden, sondern sind aufgrund von Limitierungen der Sensorik oder der verarbeitenden Algorithmen mit Unsicherheiten behaftet. Eine robuste Situationsanalyse muss den dynamischen Veränderungen des Fahrzeugumfelds flexibel begegnen. Ebenso muss diese mit Fehlern und Unsicherheiten der Wahrnehmung sowie der Prädiktion umgehen können und somit Aussagen über die Fahrsituation probabilistisch bewerten.

Die Entwicklung eines Systems zur Interpretation der vorherrschenden Fahrsituation erfordert es, Methoden zur Wahrnehmung, Klassifikation und Verfolgung von Objekten mit Methoden zur probabilistischen Analyse von Informationen zu verknüpfen. Im Rahmen dieser Arbeit wurde eine Methodik entwickelt, welche die Bestimmung der Fahrsituation auf Basis der Objekte im Fahrzeugumfeld ermöglicht. Hierzu wird eine formale Repräsentation der Umgebung und des Hintergrundwissens aufgebaut. Durch den entwickelten Ansatz wird aus gemessenen Objektinformationen der Umfeldsensorik sowie der formalen Beschreibung ein graphisches Modell erstellt, aus welchem Informationen über die Fahrsituation abgeleitet werden. Der vorgestellte Ansatz berücksichtigt die dynamischen Veränderungen des Fahrzeugumfelds bei der Modellerstellung. Die erzeugten graphischen Modelle liefern eine Bestimmung der Situationsaussage.

Abstract

Intelligent driver assistance systems support the driver by avoiding dangerous situations, by recommending efficient driving actions or by providing hints about the future route. The assessment of the current driving situation and the prediction of evolving situations influence the reliability and the user acceptance of such a system in a great manner. A driving situation is specified by the state of the ego vehicle, the state of relevant objects in the surrounding and the variation of those states. The interpretation of driving situations is still a major challenge, in particular the analysis of the object states and relations. Many of these states can not exactly be determined, and due to limitations of sensor capabilities or because of errors in the processing algorithms the information is affected by uncertainty. A robust approach for situation analysis must behave flexible, to cope with the dynamic manner of the driving environment. Further, it must handle errors and uncertainties of the perception and prediction steps to reason about the driving situation in a probabilistic manner.

The development of an approach for driving situation interpretation requires a combination of methods for perception, classification and tracking of objects with methods for the probabilistic analysis of information. This work presents a methodical approach, which allows to estimate the current driving situation based on the objects in the surrounding. Therefore, knowledge about the environment and background information is stored in a formal representation. By combining perceived object information and formal knowledge, a graphical model is constructed, on which reasoning about the driving situation is performed. Thus, during model construction, the presented approach allows to consider for dynamic changes in the surrounding. The resulting graphical models then provide a definition of the situation statement.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	2
1.2	Problemstellung und Zielsetzung	3
1.3	Begriffsbildung	5
1.3.1	Wissensmodellierung	5
1.3.2	Analyse unter Berücksichtigung von Unsicherheiten	5
1.3.3	Interpretation der Fahrsituation	6
1.4	Einordnung und wissenschaftlicher Beitrag	6
1.5	Aufbau der Arbeit	7
2	Grundlagen	9
2.1	Sensorsysteme zur Umfelderkennung	9
2.1.1	Bildgebende Systeme im Automobil	10
2.1.2	Radar- und Lidar-Systeme	14
2.1.3	Eignung zur Umfeldwahrnehmung	15
2.2	Stochastik, Logik und Inferenz	16
2.2.1	Bedingte Wahrscheinlichkeit	17
2.2.2	Aussagenlogik und Logik erster Ordnung	19
2.2.3	Wahrscheinlichkeitsbehaftete Logik	19
2.2.4	Inferenz	20
2.3	Graphische Modelle	20
2.3.1	Bayes'sche Netze	21
2.3.2	Lernen in graphischen Modellen	22
2.3.3	Erweiterungen von graphischen Modellen	22
2.4	Statistisches Relationales Lernen	24
2.4.1	Taxonomie der Verfahren	24
2.4.2	Probabilistische Relationale Modelle	25
2.4.3	Objektorientierte Probabilistische Relationale Modelle	27
3	Stand der Forschung	29
3.1	Sensordatenverarbeitung	29
3.1.1	Koordinatensystem-Konvention	30
3.1.2	Vorverarbeitung	32
3.1.3	Bildung von Objekthypothesen	32
3.1.4	Verifikation von Objekthypothesen	33
3.1.5	Weiterführende Datenverarbeitung	34
3.2	Umfeldwahrnehmung	34

3.2.1	Erkennung dynamischer Objekte	34
3.2.2	Erkennung von Infrastruktur-Objekten	38
3.3	Temporale Informationsfusion	38
3.3.1	Datenassoziation	39
3.3.2	Filterung des Systemzustandes	40
3.3.3	Kalman Filter	41
3.3.4	Interacting Multiple Model Filter	42
3.4	Situationsanalyse	45
3.4.1	Situationsanalyse ohne logische Ausdrücke	46
3.4.2	Ontologiebasierte Ansätze	48
3.4.3	Wissensbasierte Ansätze	51
3.4.4	Diskussion	53
3.5	Forschungsbedarf	54
4	Wahrnehmung des Fahrzeugumfelds	59
4.1	Konzept der Umfelderkennung	59
4.2	Objekterkennung aus Tiefendaten	61
4.2.1	Objektsegmentierung durch Regionenwachstum	62
4.2.2	Histogrammbasierte Objektsegmentierung	64
4.3	Videobasierte Objekterkennung	72
4.3.1	Fahrzeugeterkennung bei Tag	72
4.3.2	Fahrzeugeterkennung bei Nacht	76
4.3.3	Fahrzeugeterkennung auf Nebenspuren	80
4.4	Klassifikation	84
4.4.1	Entzerrung flächiger Objekte	85
4.4.2	Klassifikation von Fahrzeugen	86
4.4.3	Klassifikation von Fußgängern	88
4.5	Temporale Zustandsschätzung	88
4.6	Diskussion	91
5	Modellbasierte Datenfusion zur Situationsinterpretation	93
5.1	Methodisches Konzept der wissensbasierten Interpretation	93
5.1.1	Formale Beschreibung	93
5.1.2	Probabilistisches Modell	96
5.1.3	Modellerstellung mittels Frame-Slot Konzepten	97
5.1.4	Logischer Ablauf und Anbindung der Umfeldwahrnehmung	99
5.2	Ablauf der Modellierung von Fahrsituationen	100
5.3	Anwendungsbeispiel der Methodik	104
6	Evaluation und Ergebnisse	111
6.1	Grundlagen der Evaluation	111
6.2	Evaluierungsplattformen	112
6.3	Evaluation der Umfelderkennung	115
6.3.1	Erkennungsleistung von ToF-Kameras	115

6.3.2	Bewertung der Objektsegmentierung durch Regionenwachstum	125
6.3.3	Bewertung der Histogrammbasierten Objektsegmentierung	130
6.3.4	Evaluation der videobasierten Fahrzeugerkennung	132
6.4	Evaluation der wissensbasierten Interpretation	138
6.4.1	Verfahren zur Risikoschätzung	138
6.4.2	Bestimmung einer Blockade-Situation durch Vorausschau-Information	147
6.4.3	Einsatz zur automatisierten Längsregelung	148
6.4.4	Bewertung der wissensbasierten Interpretation	153
7	Zusammenfassung und abschließende Betrachtung	155
7.1	Zusammenfassung	155
7.2	Bewertung der Prozesskette	156
7.3	Ausblick	158
	Abbildungsverzeichnis	159
	Glossar	167
	Abkürzungsverzeichnis	169
	Quellennachweise	173

1 Einleitung

„Der Wagen denkt“ – so titelt ZEIT ONLINE einen Artikel über zukünftige Fahrerassistenzsysteme (FAS) (engl. auch *Advanced Driver Assistance Systems (ADAS)*) [ZEI11]. Diese Aussage stellt dar, welche Ansprüche an zukünftige FAS gestellt werden. Fahrzeuge sollen mit Hilfe von elektronischen Zusatzgeräten das eigene Fahrverhalten mit der Fahrintention des Fahrers, dem Fahrzeugumfeld und dem Verhalten des Fahrers abgleichen. Bereits heute leisten FAS einen wichtigen Beitrag zur Minderung von Verkehrsunfällen und erhöhen zusätzlich den Komfort während der Fahrt. Dafür greifen FAS gezielt in die Steuerung des Fahrzeugs ein und erfüllen so verschiedenste Funktionen. Die ersten FAS wurden vor 40 Jahren entwickelt. Seither werden immer komplexere FAS auf den Markt gebracht.

Die ersten FAS bestanden im Wesentlichen aus einfachen, abgeschlossenen Systemen mit einem Steuergerät (ECU) und Sensoren zur Durchführung von Regelungsaufgaben. Das Anti-Blockier-System (ABS) beispielsweise besteht aus Drehzahlsensoren an den Rädern, einer ECU zur Verarbeitung der Daten sowie Magnetventilen zur Regelung des Bremsdrucks, um das Blockieren einzelner Räder zu verhindern. Moderne FAS verarbeiten eine größere Menge an Informationen. Das Elektronische Stabilitätsprogramm (ESP) baut beispielsweise auf ABS auf und verwendet zusätzlich Informationen über den Lenkwinkel, die Querschleunigung sowie die Drehgeschwindigkeit des Fahrzeugs. Unter Berücksichtigung eines mathematischen Modells werden die Informationen in einem leistungsfähigen Mikrocomputer verarbeitet, um die Abweichung zwischen gewünschtem und tatsächlichem Fahrverhalten festzustellen und somit Schleuderverhalten entgegenzuwirken. Noch einen Schritt weiter geht das autonome Notbremsystem (AEB), welches zusätzlich zur Auswertung der Fahrzeugsensoren Informationen über das Fahrzeugumfeld nutzt. Es bestimmt zunächst Abstand und Differenzgeschwindigkeit zu vorausfahrenden Fahrzeugen und führt im Notfall ein vollautomatisches Abbremsen des Fahrzeugs durch, um einen Aufprall zu verhindern oder die Aufprallschwere zu reduzieren. Moderne FAS reflektieren somit nicht nur den internen Fahrzeugzustand, sondern ebenfalls das Fahrzeugumfeld sowie den Bediener des Fahrzeugs. Dadurch können sie teilautomatisiert oder autonom eingreifen oder dem Fahrer wichtige Informationen signalisieren.

Zur Steigerung der Sicherheit und des Komforts greifen die Systeme zunehmend in die Längs- beziehungsweise Querregelung des Fahrzeugs ein. Hierbei wird zwischen drei Automatisierungsgraden unterschieden: teil-, hoch- und vollautomatisiert [Bun12]. Im Fokus steht dabei, wie stark die Aufmerksamkeit des Fahrers auf das Verkehrsgeschehen gerichtet ist. Die Automatisierungsgrade hoch- und vollautomatisiert sind bislang verkehrsrechtlich noch nicht legitimiert.

Der abstandsregelnde Tempomat (ACC) gilt als das erste FAS, welches aktiv in die Längsführung eingreift. Kombiniert mit dem Spurhalteassistenten (LKS) werden damit erste teilautomatisierte Fahrzeugfunktionen in den kommenden Jahren in Serie produziert. Dies soll auch zu einer signifikanten Reduktion der Unfallzahlen führen (bis hin zur sogenannten „Vision

Zero“). Beispielsweise werden etwa ein Fünftel der Lastwagen-Unfälle dem unbeabsichtigten Verlassen der Fahrspur zugerechnet [Con13]. Inzwischen widmen sich viele Forschungsgruppen und Unternehmen dem teil- beziehungsweise vollautomatisierten Fahren in städtischer Umgebung oder auf Autobahnen [DAR07; Mar10; HAV11; Deb13].

Die Kernprobleme zukünftiger FAS liegen in der Zuverlässigkeit und der Nutzerakzeptanz. Während des Betriebs dieser Systeme dürfen keine oder nur sehr wenige fehlerhafte Auslösungen oder fehlerhaft ausbleibende Eingriffe auftreten. Anforderungen an die Systemgenauigkeit und an die Nutzerakzeptanz führen dazu, dass die eingesetzten Verfahren sehr genau aufeinander abgestimmt und durch langwierige Testzyklen verifiziert und optimiert werden müssen. Zur Gewährleistung der Zuverlässigkeit muss das System auf Grundlage der wahrgenommenen Informationen nachvollziehbare Entscheidungen treffen können. Bis heute gibt es nur sehr wenige Assistenzsysteme, die auf die Sensordaten anderer Assistenzsysteme zurückgreifen. Statt dessen erfolgt die Realisierung von Assistenzsystemen separiert voneinander (Schneider [Sch09] bezeichnet die Systeme als „Insellösungen“). Dies hat zur Folge, dass komplexe Situationen nicht vollständig erkannt werden und eine Handlungsempfehlung des Systems daher fehlerhaft sein kann.

Für eine korrekte Handlungsempfehlung ist es nötig, dass sich die entwickelten Systeme zunehmend intelligenter verhalten. Sie müssen aus Sensor-Rohdaten Objektwissen bilden, aus Objekten die Zusammenhänge in der Szene erkennen und die Entwicklung der Szene prognostizieren können. Nur dann können Assistenzsysteme auch in komplexen Situationen Risiken für den Menschen sichtbar machen und Entscheidungshilfen zur richtigen Zeit zur Verfügung stellen.

1.1 Motivation

Eine große Herausforderung für FAS ist nach wie vor der Aufbau eines Situationsverständnisses. Verfahren zur Umfelderkennung extrahieren dazu die Informationen aus den Signalen der Umfeldsensoren und bauen ein akkurates Umfeldmodell auf. Mit Hilfe von Hintergrundwissen wird das Wissen über die Objekte vervollständigt. Werden diese Informationen sinnvoll miteinander in Relation gesetzt, können daraus komplexe Zusammenhänge zwischen Objekten und deren Bedeutung für das eigene Fahrzeug abgeleitet werden – hierbei spricht man von der Interpretation von Fahrsituationen.

Zur Erzeugung eines Situationsverständnisses reicht ein alleiniges Verknüpfen der Informationen nicht aus, da die verschiedenen Informationsquellen teils widersprüchliche Informationen liefern oder die Informationen durch Fehler der Wahrnehmung beziehungsweise der zugrundeliegenden Modellannahmen verfälscht sind. Dies macht Verfahren notwendig, die zusätzlich Sicherheit und Relevanz von Informationen berücksichtigen können. Durch sogenannte probabilistische Verfahren ist es möglich, Informationen durch eine Wahrscheinlichkeitsaussage zu ergänzen und diese während der weiteren Verarbeitung zu berücksichtigen. Bisher werden probabilistische Verfahren maßgeblich für die Fehler- und Einflussanalyse bei trägen oder statischen Systemen eingesetzt. Der hohe Aufwand zur Berechnung der zugrundeliegenden mathematischen Ausdrücke und die Beschreibung der Systeme durch starre, vordefinierte probabilistische Netze erschwert den Einsatz in dynamischen Domänen. Für FAS

müssen probabilistische Verfahren eine Berechnung wahrscheinlichkeitsbehafteter Aussagen in Echtzeit gewährleisten. Ein fest definiertes Netz lässt sich nicht oder nur unzureichend an veränderliche Situationen anpassen. Für einen automatischen Aufbau von probabilistischen Netzen auf Basis von Umfeldinformationen fehlen bislang geeignete Methoden.

Die Erfassung der Fahrsituation durch maschinelle Methoden kann als Meilenstein bei der Weiterentwicklung von FAS gesehen werden.

1.2 Problemstellung und Zielsetzung

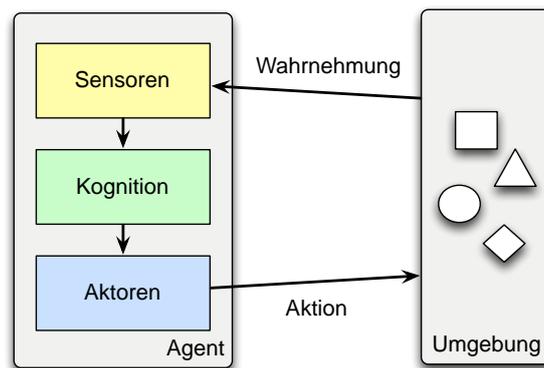


Bild 1.1: In einem kognitiven Regelkreis werden geeignete Aktionen aufgrund wahrgenommener Informationen berechnet [RN09].

Ein kognitives Assistenzsystem verhält sich wie ein kognitiver Agent in einem Regelzyklus mit seiner Umgebung. Dieser Zyklus beinhaltet die *Wahrnehmung*, die *Wissensverarbeitung* und das *Agieren* (siehe Abbildung 1.1). An diese Definition angelehnt, besteht die Wissensverarbeitung eines Assistenzsystems zumeist aus vier aufeinander aufbauenden Komponenten, die je nach Anforderung verschieden komplex sein können:

- Wahrnehmungskomponenten zur Verarbeitung von Signalen der Sensoren, sodass Objektwissen gebildet wird; dies beinhaltet üblicherweise auch die Fusion von Informationen über die Zeit;
- Analysekomponenten zum Aufbau eines Situationsverständnisses aus wahrgenommenen Informationen und zur Prädiktion der Entwicklung der Situation;
- Entscheidungskomponenten zur Bestimmung und Bewertung verschiedener Handlungsmöglichkeiten, welche den Systemzustand durch Veränderung direkt oder indirekt beeinflussbarer Parameter in einen optimalen Zustand überführen; sowie
- Aktionskomponenten, welche über entsprechende Schnittstellen auf einen Benutzer oder eine technische Komponente einwirken um schließlich eine Veränderung des Systemzustandes zu erreichen.

Für alle Komponenten gilt, dass die Verarbeitungsgeschwindigkeit neuer Informationen so hoch sein muss, dass Informationen nicht verloren oder verzögert werden. Die Verarbeitungs-

kette muss somit in Echtzeit ablaufen. Bei der Entwicklung dieser Komponenten ist ein gegenseitiger Einfluss zu berücksichtigen (siehe Abbildung 1.2).

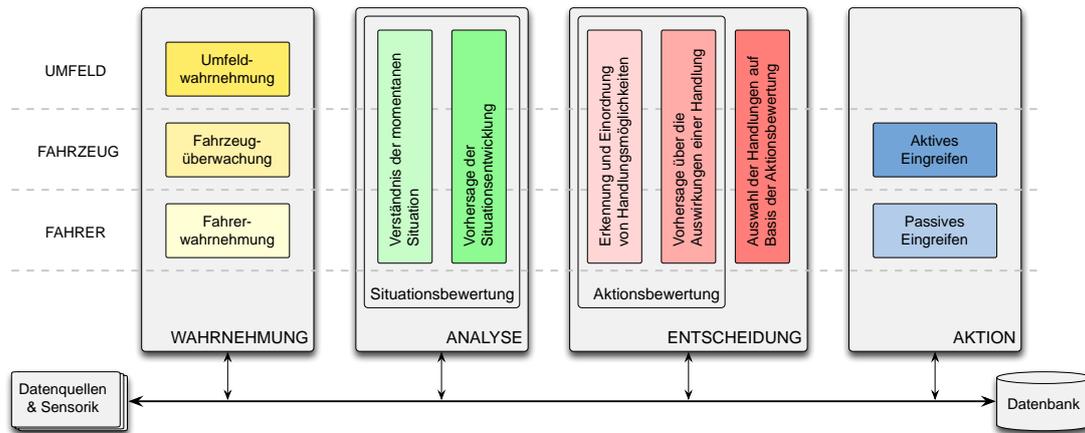


Bild 1.2: Adaptive FAS setzen sich aus Wahrnehmungs-, Analyse-, Entscheidungs- und Aktionskomponenten zur Informationsverarbeitung zusammen; nach [RHO09].

Für fortschrittliche Assistenzsysteme ist die Erkennung der vorherrschenden Fahrsituation notwendig. Bisherige Ansätze beziehen lediglich einzelne Objekte für die Situationserkennung ein. Eine genaue Situationserkennung kann allerdings nur durch die Erkennung aller für die jeweilige Situation relevanter Objekte ermöglicht werden. In dieser Arbeit wird eine Methodik zur Analyse von Fahrsituationen vorgestellt, die auf der Beschreibung der Elemente von Fahrsituationen in modellbasierter Form beruht. Zugrundeliegende Erkennungskomponenten stellen unscharfe Informationen bereit. Daher muss das Verfahren in der Lage sein, diese Unsicherheiten in die Situationsanalyse mit einfließen zu lassen.

Die Entwicklung eines Systems zur Interpretation der vorherrschenden Fahrsituation erfordert es, Methoden zur Detektion, Klassifikation und Verfolgung von Objekten adäquat mit Methoden zur probabilistischen Analyse von Informationen zu verknüpfen. Das entwickelte System muss in der Lage sein, Objekte und deren Attribute, sowie Relationen zwischen Objekten zu bestimmen.

Die aus dem Stand der Technik bekannten Methoden zur Objekterkennung liefern für bestimmte Objektarten gute Ergebnisse. Bei unbekanntem oder verschiedenen Objektarten kann eine stabile Objekterkennung aber nur durch Fusion verschiedener Sensoren erfolgen. Videosensoren werden heute in Fahrzeugen häufig eingesetzt und liefern präzise Informationen über die statische Verkehrsinfrastruktur. Auf dem Radarprinzip arbeitende Sensoren können über eine große Reichweite wichtige Informationen wie zum Beispiel Abstand und Relativgeschwindigkeit zu dynamischen Objekten erfassen. Tiefenbildgebende Sensoren können im näheren Umfeld des Fahrzeugs sowohl die Entfernung, die Ausrichtung wie auch die laterale Dimension von unbekanntem Objekten erfassen. Der Einsatz von tiefenbildgebenden Sensoren zusammen mit Videosensoren kann eine gleichzeitige Wahrnehmung, Klassifikation und Verfolgung von Objekten im Fahrzeugumfeld ermöglichen. Dafür müssen Methoden entwickelt werden, die für verschiedene Objektklassen die Vorteile der einzelnen Sensoren miteinander verknüpfen und eine robuste Umfelderkennung gewährleisten.

Auf Grundlage der Objektinformationen lassen sich Umfeldmodelle aufbauen, welche das Fahrzeugumfeld möglichst akkurat repräsentieren. Die Modelle werden dabei kontinuierlich an die Fahrsituation angepasst. Eine formale Repräsentation eines Weltmodells erlaubt es, die Grundkomponenten und -eigenschaften von Fahrsituationen, nämlich die Objekte, die Veränderungen der Objektzustände sowie Veränderungen der Objektrelationen generisch darzustellen. Die von der Umfelderkennung gemessenen Objektinformationen müssen in das Weltmodell einfließen können, ebenso müssen Relationen zwischen mehreren Objekten während der Laufzeit aufgebaut oder wieder entfernt werden können. Zusätzlich benötigtes Hintergrundwissen muss in diesem Weltmodell abzubilden sein. Das Weltmodell bildet die Grundlage für eine erfolgreiche Situationsanalyse.

Zur Analyse der Fahrsituation lassen sich aus dem Weltmodell mittels Schlussfolgerung weitere Informationen berechnen. Hierfür wird kontinuierlich ein graphisches Modell aus dem Weltmodell abgeleitet. Die im Weltmodell definierten probabilistischen Unsicherheiten und Abhängigkeiten von Variablen müssen vollständig in das graphische Modell übertragen werden. Die aus dem Stand der Technik bekannten Verfahren des statistischen relationalen Lernens (SRL) zeigen hierfür Verfahren auf, welche zur Lösung dieser Problemstellung beitragen können.

1.3 Begriffsbildung

Diese Arbeit befasst sich mit einem modellbasierten Ansatz zur probabilistischen Interpretation von Fahrsituationen. Die genaue Bedeutung der im Rahmen dieser Arbeit verwendeten Begriffe wird im Folgenden definiert:

1.3.1 Wissensmodellierung

Relevante Objekte, welche maßgeblichen Einfluss auf die Fahrsituation haben, werden als Modelle in einer formalen Sprache beschrieben. Ein Modell entspricht der Repräsentation des Wissens über ein Objekt. Es enthält sowohl dessen relevante Attribute als auch die Verknüpfungen zu weiteren Modellen, welche Einfluss auf das Objekt ausüben können (beispielsweise durch Interaktion). Durch die Wissensmodellierung lässt sich Objekt- und Hintergrundwissen derart speichern, sodass eine anschließende probabilistische Analyse der Informationen möglich ist. Zusätzlich lassen sich entwickelte Modellkonfigurationen anhand von Expertenwissen leicht verifizieren und modifizieren.

1.3.2 Analyse unter Berücksichtigung von Unsicherheiten

In einem System gespeichertes Wissen muss mit einem Glaubensgrad versehen werden, da es für ein System nicht möglich ist, eine reale Umgebung vollständig fehlerfrei zu messen. Ebenfalls existieren Bedingungen, welche von einem System nicht hergeleitet werden können (zum Beispiel ob ein Streckenabschnitt einer geplanten Route in naher Zukunft wegen eines Unfalls gesperrt ist). Diese Art der Informationsverarbeitung wird als probabilistische Interpretation von Informationen bezeichnet. Dennoch lässt sich aus probabilistischen Informationen eine

richtige Aktionsfolge ableiten, wenn unter Berücksichtigung des bekannten Wissens über die Umgebung eine rationale Entscheidung getroffen wird [RN09].

1.3.3 Interpretation der Fahrsituation

Die Fahrsituation wird im Allgemeinen als die Kombination des Zustands des Eigenfahrzeugs, der Zustände relevanter Objekte im Fahrzeugumfeld sowie die Veränderung dieser Zustände bezeichnet. Eine spezifische Anfrage an die Fahrsituation kann lauten: „Wie hoch ist das Risiko einer Kollision?“, oder „Hat mein Fahrzeug gegenüber einem anderen fremden Fahrzeug Vorfahrt?“. An diesen Beispielen wird deutlich, dass die Interpretation einer Fahrsituation gegenüber einer jeweils spezifischen Anfrage (Query) eine Aussage leistet. Die Aussagen müssen aus vorhandenen Informationen mit Hilfe der Wissensmodellierung schließbar sein. Im Gegensatz zur Interpretation auf Signal- oder Objektebene wird durch die Interpretation auf Situationsebene Wissen von unterschiedlichen Quellen genutzt. Im Kontext der Interpretation der Fahrsituation bedeutet dies die Einbeziehung von Informationen verschiedener Sensoren sowie verschiedenen Abstraktionsgraden.

1.4 Einordnung und wissenschaftlicher Beitrag

Im Stand der Forschung in Kapitel 3 werden Arbeiten, welche dieses Forschungsgebiet maßgeblich beeinflussen, detailliert vorgestellt. Bisherige Forschungsarbeiten legen den Schwerpunkt in die Verbesserung der Wahrnehmung, da eine gute Qualität der Informationen als Schlüssel für nachfolgende Prozesse gesehen wird. Für die Verknüpfung von Informationen zur Durchführung einer umfassenden Situationsanalyse werden entweder numerische Verfahren oder graphische Modelle eingesetzt. Numerische Verfahren bestimmen analytisch eine Situationsaussage, lassen sich aber nur in sehr exakt definierbaren Fahrsituationen berechnen. Außerdem führt die Wahl der Parameter häufig zu einer hohen Komplexität. Graphische Modelle berücksichtigen dagegen Unsicherheiten der Informationen, besitzen aber eine starre Modellstruktur. Dies führt zu Einschränkungen in der Situationsmodellierung und wirkt sich auf die Laufzeit der Verfahren aus.

Davon ausgehend wurde im Rahmen dieser Arbeit ein modellbasierter, objektorientierter und probabilistischer Ansatz entwickelt. Hintergrundwissen und aktuelle Umweltinformationen sind durch Prädikatenlogik beschrieben, wodurch Objekte, Attribute sowie Relationen zwischen Objekten formal repräsentiert werden können. Dies erlaubt die Beschreibung von Fahrsituationen und ermöglicht die Berücksichtigung der dynamischen Veränderungen des Fahrzeugumfelds schon bei der Modellerstellung. Die erzeugten graphischen Modelle werden flexibel an die vorherrschende Situation anpasst. Für die Erkennung relevanter Objekte und Attribute sowie die anschließende Analyse der Fahrsituation wurde eine durchgängige Verarbeitungskette herausgearbeitet. Die Erkennung relevanter Objekte und Attribute erfolgt auf Basis komplementärer Sensoren, da diese wesentlich robuster gegenüber der Verwendung von Einzelsensoren sind. Die Beschreibung der Prozesskette beginnt mit der sensorspezifischen Segmentierung von Objektdaten, anschließend erfolgt eine objektspezifische Klassifikation und schließlich erfolgt eine Zustandsschätzung der Objekte über die Zeit. Der entwickelte

Ansatz erstellt aus gemessenen Objektinformationen sowie einer formalen Beschreibung der Modellbasis ein graphisches Modell, aus welchem Informationen über die Fahrsituation abgeleitet werden können.

Die Beiträge dieser Arbeit stellen sich somit wie folgt dar:

1. Sensordatenverarbeitung: Entwicklung einer Verarbeitungskette zur Erkennung relevanter Objekte und Attribute mittels geeigneter komplementärer Sensoren;
2. Situationsmodellierung: Repräsentation von Fahrsituationen in flexibler, erweiterbarer Struktur durch wahrscheinlichkeitsbasierte Logik erster Ordnung (FOPL);
3. Situationsanalyse: Automatische Konstruktion von graphischen Modellen mit Bezug auf relevante Objekte und Relationen, sowie Situationsinterpretation durch probabilistische Inferenz;
4. Bewertung des Funktionsnachweises der Situationsanalyse unter realen Bedingungen und Umsetzung der Prozesskette auf Versuchsfahrzeugen.

1.5 Aufbau der Arbeit

Kapitel 2 beschreibt die Grundlagen dieser Arbeit. Die Sensorsysteme zur Wahrnehmung der Umgebung des Fahrzeugs werden erläutert. Daran schließt sich die Beschreibung stochastischer und logischer Verfahren an, auf denen die folgenden Kapitel aufbauen. Das Kapitel 2 schließt mit der Beschreibung zur Behandlung von unsicheren Informationen mittels graphischen Modellen und den Verfahren des statistischen relationalen Lernens.

In Kapitel 3 wird der Stand der Forschung präsentiert, welcher zur Erstellung dieser Arbeit herangezogen wurde und auf dem diese Arbeit aufbaut. Es werden die Methoden vorgestellt, welche Teilaspekte der hier adressierten Herausforderungen in vergleichbarer oder abweichender Weise zu lösen versuchen. Die Methoden werden analysiert und der Forschungsbedarf dieser Arbeit abgeleitet.

Kapitel 4 und 5 beschreiben die in dieser Arbeit umgesetzten Verfahren, welche zur Lösung der genannten Problemstellung beitragen. Die Kapitel sind getrennt nach den Verfahren der Wahrnehmung des Fahrzeugumfelds und den Verfahren der modellbasierten Datenfusion zur Situationsinterpretation. Wie in Abschnitt 1.2 beschrieben, sind beide Kapitel eng miteinander verknüpft.

Die umgesetzten Verfahren dieser Arbeit wurden umfangreich auf realen Versuchsfahrzeugen evaluiert. Eine Beschreibung dieser Evaluation ist in Kapitel 6 gegeben. Kapitel 7 schließt diese Arbeit mit der Zusammenfassung ab. Es werden die erreichten Beiträge dieser Arbeit zum Stand der Forschung beschrieben und eine Analyse des weiteren Forschungsbedarfs abgeleitet.

2 Grundlagen

In den nachfolgenden Abschnitten werden die Grundlagen dieser Arbeit vorgestellt. Der Abschnitt 2.1 beschreibt Sensorsysteme und deren Einsatzmöglichkeit für die Domäne der Fahrerassistenzsysteme (FAS). Besonderes Augenmerk liegt dabei auf komplementären Sensoren zur Messung von 2D- und 3D-Daten. In Abschnitt 2.2 werden die Grundlagen der formalen Wissensrepräsentation durch Logik beschrieben. Der Abschnitt behandelt zudem Grundlagen der eingesetzten probabilistischen Verfahren und Inferenzverfahren zum Schließen von Informationen aus unsicherem Wissen. Daran anschließend werden in Abschnitt 2.3 graphische Modelle zur Beschreibung von probabilistischen Netzen eingeführt. Abschnitt 2.4 schließt das Kapitel mit der Zusammenführung von Probabilistik und Logischen Systemen durch Einführung in die Grundlagen der wahrscheinlichkeitsbasierten Logik erster Ordnung (FOPL) ab.

2.1 Sensorsysteme zur Umfelderkennung

Der Mensch nimmt seine Umgebung mit Hilfe seiner fünf Sinnesorgane¹ wahr. Diese ermöglichen es, Stoffe zu schmecken, Düfte zu riechen, Oberflächen zu ertasten, Geräusche zu hören, die Umwelt zu sehen. Über spezielle Rezeptoren werden physikalische oder chemische Eigenschaften in elektrische Impulse umgewandelt und über Nervenbahnen im Körper weitergeleitet. Die komplementären Eigenschaften der Sinneswahrnehmung ermöglichen es Lebewesen, sich in bekannten wie unbekanntem Umgebungen zu orientieren und zu überleben. Für ein technisches System, welches in einer unbekanntem Umgebung agiert, ist die Erfassung des Umfelds von großer Bedeutung. Werden relevante Informationen nicht (rechtzeitig) erfasst, kann ein System ein gesetztes Ziel unter gewissen Rahmenbedingungen nicht optimal erreichen. Zur Wahrnehmung der Szene werden elektromagnetische Wellen oder Schallwellen durch einen Messfühler erfasst. Die Wellen sind entweder aufgrund anderer Effekte in der Szene bereits vorhanden oder werden durch das Messsystem selbst erzeugt. Man spricht hierbei von passiven respektive aktiven Systemen.

Passive Systeme messen elektromagnetische Strahlung (oder Schallwellen), welche von der Szene selbst durch physikalische Effekte erzeugt, gestreut oder reflektiert wird. Dabei wird die Intensität oder die Frequenz der Strahlung von einem Sensor erfasst und in ein elektrisches Signal umgewandelt. Im Gegensatz dazu wird bei aktiven Systemen elektromagnetische Strahlung (oder Schallwellen) erzeugt, die mit der Szene physikalisch interagiert. Es wird die Strahlung gemessen, die in Richtung des Sensors zurück reflektiert wird. Sind die Eigenschaften der erzeugten Wellen bekannt, können aus der Intensität oder der Laufzeit der gemessenen Strahlung Rückschlüsse auf die Szene gezogen werden.

Zur Wahrnehmung des Fahrzeugumfelds kommen in Fahrzeugen verschiedene Sensorsysteme zum Einsatz (siehe Abbildung 2.1). Unterschieden wird dabei zwischen bildgebenden

¹Allgemeiner Sprachgebrauch; die moderne Physiologie definiert neben den fünf Sinnen noch weitere Sinne.

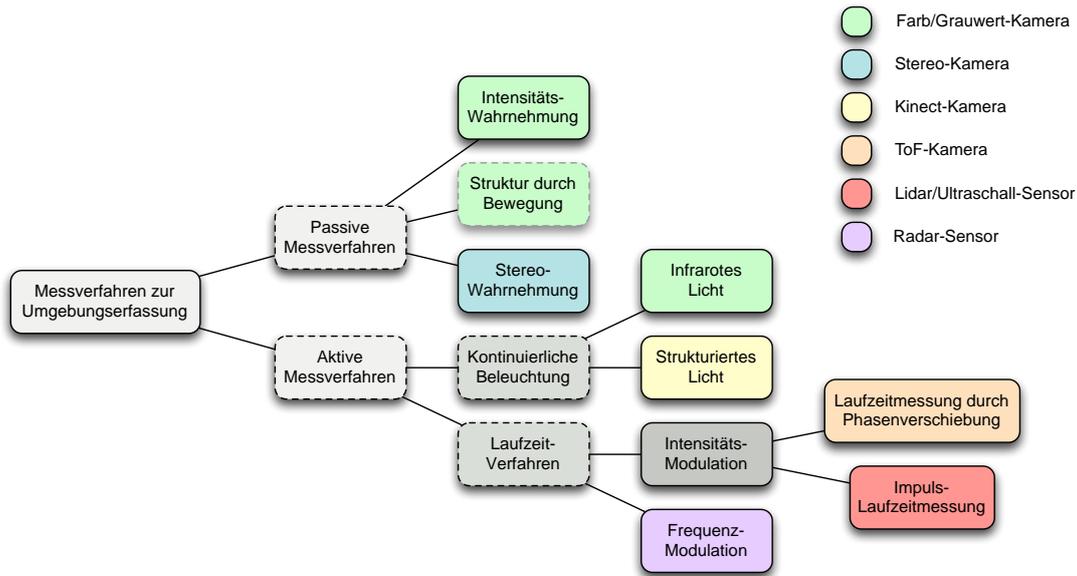


Bild 2.1: Taxonomie von Messverfahren zur Erfassung fahrrelevanter Informationen.

Systemen und Systemen auf Basis gebündelter Wellen. Bildgebende Sensorsysteme sind meist passive Systeme und ermöglichen eine hohe laterale Trennung von Informationen, also die Bestimmung flächiger Eigenschaften der Szene. Zu den passiven Verfahren gehört die Farb- oder Grauwert-Kamera, die häufig auch als High-Dynamic Range (HDR)-Kamera ausgelegt ist und damit einen besonders großen Helligkeitsbereich abdeckt. Bei passiven Systemen kann die Entfernung der Szene durch Triangulation zwischen korrespondierenden Bildpunkten zweier Bilder gemessen werden. Systeme auf Basis gebündelten Lichts oder Radiowellen erzielen eine hohe Reichweite und eine hohe Messgenauigkeit entlang der longitudinalen Achse. Aktive Systeme werden bevorzugt zur Bestimmung der Entfernungen in der Szene eingesetzt. Hierzu zählen Radar-, Lidar- und Lichtlaufzeit (ToF)-Sensoren.

Sensoren besitzen unterschiedliche charakteristische Auflösungseigenschaften. Die geometrische Auflösung beschreibt die räumliche Ausdehnung, die von einem Sensorbildpunkt erfasst wird. Die Empfindlichkeit beschreibt die kleinste Menge an Strahlung, für die der Sensor ein zuverlässiges elektrisches Signal liefert. Die temporale Auflösung oder Wiederholfrequenz bezeichnet den zeitlichen Abstand, zwischen welchem zwei aufeinanderfolgende Aufnahmen erfolgen können. Weiterhin ist relevant, welche Zusatzinformation durch einen Sensor messbar sind. Unter Ausnutzung des Doppler-Effekts können Radarsensoren zusätzlich zur Entfernung eines Objektes dessen Relativgeschwindigkeit bestimmen. Dies erhöht die Genauigkeit gegenüber Sensoren, welche diese Information durch temporale Analyse gewinnen, also durch Ableitung der Position nach der Geschwindigkeit über eine Sequenz von Messdaten.

2.1.1 Bildgebende Systeme im Automobil

Optische Kamerasysteme werden zur Umfeldwahrnehmung von FAS eingesetzt, beispielsweise zur Erkennung von Infrastruktur-Objekten. Diese Kamerasysteme lassen sich hinreichend genau durch das Lochkamera-Prinzip beschreiben (siehe Abbildung 2.2). Großer Vorteil der bildgebenden Systeme sind potentiell günstige Fertigungskosten, sowie ein großer horizon-

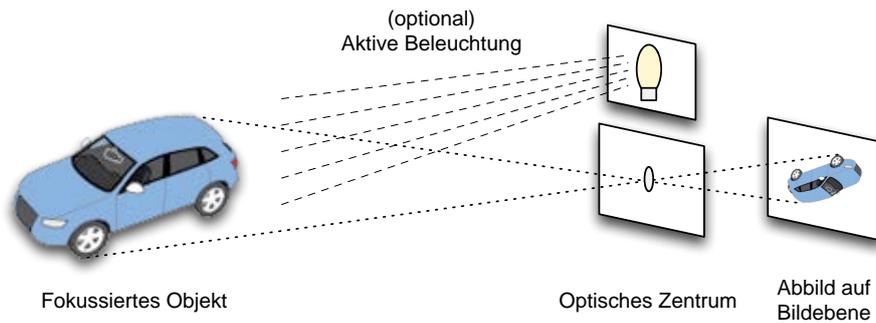


Bild 2.2: Messprinzip bildgebender Systeme; ein fokussiertes Objekt bildet sich durch das optische Zentrum in eine 2D-Bildebene ab.

taler und vertikaler Sichtbereich. Die Linsen zur Lichtbrechung können zum Teil direkt im Fahrzeug verbaut sein, zum Beispiel durch spezielle Schlitze in der Windschutzscheibe. Das optische System aus Chip und Linse wird durch seine intrinsischen Parameter definiert. Aktive Systeme benötigen zusätzliche Beleuchtungseinheiten, deren Licht durch die Szene zur Kamera zurückreflektiert wird.

Videokamera

Videokameras nutzen ein passives Messverfahren, senden also selbst kein Licht aus. Relevante Eigenschaften der Videokamera sind die horizontale und vertikale Auflösung, die maximal mögliche Bildrate und der Dynamikbereich des gemessenen Signals. Die Videokamera kann Intensitäten der Szene entweder als Grauwert- oder Farbwert-Information aufzeichnen. In der Regel unterstützen Grauwert-Kameras einen höheren Dynamikbereich, eine wichtige Eigenschaft für den Einsatz im Fahrzeug. Dort treten oftmals sowohl sehr dunkle wie auch sehr helle Lichtbedingungen innerhalb eines Bildes auf.

Die von der Videokamera gemessenen Informationen eignen sich besonders zur Analyse von Struktur- und Textureigenschaften in der Szene. Durch die Abbildung auf die Bildfläche geht jedoch die Tiefeninformation der Szene verloren. Die Bestimmung der Entfernung zu Objekten kann durch Modellwissen erfolgen. Lassen sich Modellannahmen mit ausreichender Zuverlässigkeit festlegen, so kann die tatsächliche Entfernung durch Rückprojektion des Modells aus der Bildebene in den 3D-Raum geschätzt werden.

Lichtlaufzeit (ToF)-Kamera

Die ToF Kamera gehört zur Klasse der aktiven Sensorsysteme. Im Unterschied zur Videokamera wird pro Bildpunkt die Entfernung der Szene anstelle des Helligkeitswertes gemessen und als sogenanntes Tiefenbild gespeichert. Durch einen ausgesendeten Infrarotlicht (IR)-Impuls wird die Szene ausgeleuchtet. Aus dem reflektierten Signal wird für jeden Bildpunkt mittels einer elektrischen Schaltung die benötigte Zeit des Impulses zurück zur Kamera gemessen [Sch+97]. Unter Kenntnis der Modulationsfrequenz des Signals und der Ausbreitungsgeschwindigkeit des IR-Lichtes wird die zurückgelegte Entfernung bestimmt.

Mit Hilfe eines sinusförmigen Signals mit bekannter Modulationsfrequenz kann die Entfernung zur Szene durch Phasenverschiebung berechnet werden. Diese berechnet sich aus der

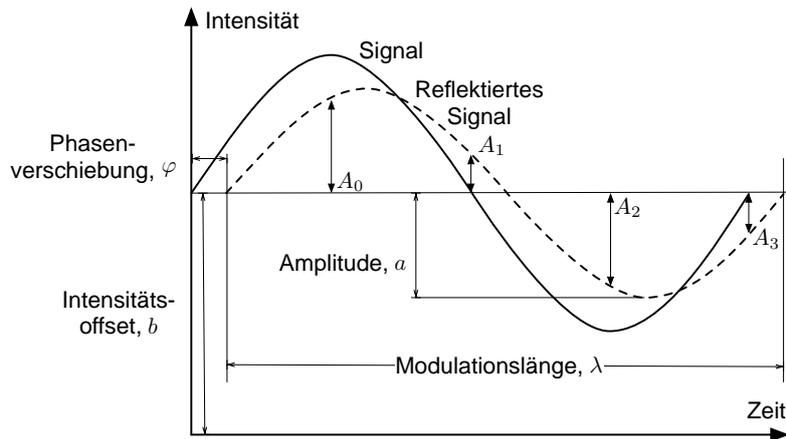


Bild 2.3: Durch vier Abtastungen lässt sich die Phasenverschiebung zwischen ausgesendetem und reflektiertem Signal bestimmen.

Kreuzkorrelation zwischen emittiertem und reflektiertem Signal. Der Zusammenhang zwischen der Distanz und der Phasenverschiebung ist durch

$$d = \frac{c}{4\pi \cdot f_{mod}} \cdot \varphi \quad (2.1)$$

gegeben, wobei d die Distanz zum Sensor, c die konstante Lichtgeschwindigkeit, φ die Phasenverschiebung und f_{mod} die Modulationsfrequenz ist. Die Phasenverschiebung lässt sich durch vier Abtastungen A_0, \dots, A_3 anhand Gleichung (2.2) bestimmen, wie in Abbildung 2.3 dargestellt ist. Zusätzlich können Rückschlüsse auf die Reflektivität der Szene (2.3) und den Gleichanteil des empfangenen Lichts (2.4) getroffen werden:

$$\varphi = \arctan \left(\frac{A_3 - A_1}{A_0 - A_2} \right), \quad (2.2)$$

$$a = \frac{\sqrt{(A_3 - A_1)^2 + (A_0 - A_2)^2}}{2}, \quad (2.3)$$

$$b = \frac{A_0 + A_1 + A_2 + A_3}{4}. \quad (2.4)$$

Die Daten des Tiefenbildes lassen sich unter Kenntnis der intrinsischen Parameter der Kamera in eine 3D-Punktwolke überführen. Hieraus kann ohne spezifisches Wissen über die Objektklasse eine Segmentierung der Punktdaten in Gruppen von Punkten erfolgen (beispielsweise durch Bereichswachstums-Verfahren).

Für den Einsatz im Fahrzeug wird die Kamera ähnlich wie eine Videokamera hinter der Windschutzscheibe verbaut. Die Beleuchtungseinheiten können getrennt von dem Sensor beispielsweise im Kühlergrill verbaut werden, wie in Abbildung 2.4 dargestellt. Der Einbau im Kühlergrill eignet sich, da die Beleuchtungseinheiten Abwärme produzieren und Wärme hier effizienter abzuführen ist.



Bild 2.4: Tiefenbild-Kamera in Testfahrzeug „CoCar“: Das System besteht aus zwei Beleuchtungseinheiten im Kühlergrill und einer Kamera hinter der Windschutzscheibe.

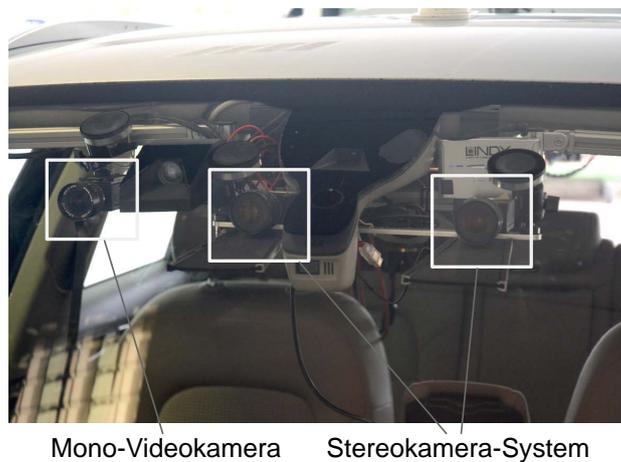


Bild 2.5: Zwei Kameras eines Stereokamera-Systems mit paralleler optischer Achse und eine Mono-Videokamera in „CoCar“.

Stereokamera

Ein weiterer geeigneter Sensor zur Messung von Entfernungsinformationen einer Szene ist die Stereokamera. Das Prinzip der Stereokamera beruht darauf, in zwei zeitsynchron bestimmten Bildern korrespondierende Objektpunkte zu bestimmen. Die Bilder werden durch zwei Kameras aufgezeichnet, welche ähnliche Bildeigenschaften haben und deren optische Achsen nahezu parallel sind (siehe Abbildung 2.5). Durch Kenntnis über den Positionsunterschied beider Kameras und den optischen Versatz des Objektpunktes (der Disparität) in den Kamerabildern kann die Entfernung des Objektpunktes zur Kamera bestimmt werden. Das Kernproblem der Stereokamera ist die effiziente Suche nach korrespondierenden Bildpunkten. Durch Verwendung der Epipolar-Geometrie kann die Abhängigkeit der Abbildung eines Objektpunktes in beiden Kamerakoordinatensystemen zueinander ausgedrückt werden. Heute existieren optimierte Methoden zur Bestimmung der Disparität für zwei zueinander ausgerichtete Bildpaare. Durch Parallelverarbeitung wird eine Verarbeitungszeit ähnlich der Bildwiederholraten der einzelnen Kameras erreicht [GEM09].

Aus der Disparität aller Bildpunkte zwischen beiden Kamerabildern lässt sich ein Tiefenbild berechnen. Somit kann die 3D-Position von Objekten bestimmt werden. Die Tiefenauflösung ist abhängig von der Basisbreite der beiden Kameras, also dem Abstand zwischen den Ka-

meras. Dabei sinkt die Tiefenauflösung reziprok mit steigendem Abstand des Objektes zum Kamerasystem.

Die Stereokamera vereint die Vorteile von tiefenbildgebenden Kameras und Videokameras. Entsprechend der Möglichkeiten der Videokameras kann ein hoher Dynamikbereich gewährleistet werden. Außerdem können die Eigenschaften der Kameras sehr spezifisch abgestimmt werden. Eine Projektion der Tiefenbild-Daten zu den Videodaten ist aufgrund des gleichen Ursprungsbildes nicht notwendig, weshalb die Segmentierung unbekannter Objekte aus 3D-Punktwolken und eine anschließende Klassifikation auf 2D-Bilddaten sehr günstig durchgeführt werden kann.

2.1.2 Radar- und Lidar-Systeme

Die Systeme auf Basis gebündelter Wellen werden in Radar (*Radio Detection and Ranging*) und Lidar (*Light Detection and Ranging*) Messsysteme unterteilt. Der Vorteil dieser Systeme liegt in der besonders hohen Messgenauigkeit der Entfernung, sowie insbesondere bei Radar in der Bestimmung der Geschwindigkeit von Objekten durch Ausnutzung des Doppler-Effekts.

Radar-Sensoren besitzen für sicherheitsorientierte FAS momentan die höchste Relevanz und werden für abstandsregelnde Tempomaten (ACC) und autonome Notbremsysteme (AEB) in Serie eingesetzt. Abhängig von der Messfrequenz eignen sich die Sensoren für die Nah- oder Fernbereichsüberwachung des Fahrzeugs. Im Gegensatz zu Lidar oder bildgebenden Sensoren ist das Messsystem robust gegenüber Umwelteinflüssen wie Schnee, starker Gischt oder Nebel. Eine gute Beschreibung verschiedener marktüblicher Sensoren und Verfahren zum Entwurf von Fahrfunktionen auf Basis von Radar-Sensoren findet sich in [Jor06].



ibeo LUX Lidar-Sensor

Bild 2.6: ibeo LUX Sensoren lassen sich durch ihre Bauform kompakt in der Stoßstange von Fahrzeugen integrieren.

Der Lidar-Sensor nutzt gebündeltes Licht zur Entfernungsmessung der Szene. Dabei wird der ausgesendete Lichtstrahl über einen rotierenden Spiegel gelenkt um über einen größeren Messbereich Informationen zu erhalten. Eine Abtastung des Umfelds erfolgt also nicht flächig, sondern punktuell. Neben der Entfernungsinformation wird durch das Lidar-System ebenfalls die Reflektanzeigenschaft in der Szene gemessen.

Zur weiteren Vergrößerung des Messbereichs werden mehrzeilige Lidar-Systeme eingesetzt. Im Gegensatz zu Kamerasystemen können Lidar-Systeme einen großen Öffnungswinkel von

über 180° besitzen, ohne dass dieser zu Verzerrungseffekten führt. Für Fahrzeuge kommen im Wesentlichen zwei Lidar-Systeme in Frage.

Der ibeo Lux Sensor ist ein vierzeiliger Lidar-Sensor mit einer Reichweite von bis zu 200 m [ibe13]. Er besitzt einen Winkelbereich von bis zu 110° und erreicht abhängig von der Konfiguration unterschiedlich hohe Winkelauflösungen bis zu $0,125^\circ$. Durch die hohe räumliche Auflösung können Objekte wie beispielsweise Fahrzeuge auch in großen Entfernungen gemessen werden. Die vier mit je $0,8^\circ$ vertikal aufgefächerten Messebenen erlauben die Erfassung von Objekten auch bei leichten Nickbewegungen.



Velodyne HDL-32E Lidar-Sensor

Bild 2.7: Rotierender Velodyne HDL-32E Lidar-Scanner auf dem Fahrzeugdach des Testfahrzeugs „CoCar“.

Der Velodyne Lidar-Scanner ist ein rotierender Laserscanner [Vel13]. Im Gegensatz zu Lidar-Scannern, welche das Licht durch einen rotierenden Spiegel ablenken, wird der gesamte Scanner durch Rotation bewegt. Je nach Bauweise kann der Scanner gleichzeitig 64 (HDL-64E) oder 32 (HDL-32E) Messebenen aufzeichnen. Diese sind mit je $0,4^\circ$ respektive $1,33^\circ$ vertikal aufgefächert. Durch seine Funktionsweise muss der Scanner, wie in Abbildung 2.7 gezeigt, auf dem Fahrzeugdach befestigt werden. Dafür liefert er ein sehr detailliertes 360° -Bild der Umgebung. Die hohe Reichweite und Genauigkeit befähigt dem Sensor die Eignung als Referenzsensor zur Bewertung der Qualität und Messgenauigkeit alternativer FAS-Sensoren. Die Rotationsgeschwindigkeit des Sensors liegt bei bis zu 15 Hz respektive 10 Hz.

2.1.3 Eignung zur Umfeldwahrnehmung

Die oben beschriebenen Sensoren sind zur Umfelderkennung für verschiedene FAS-Anwendungen unterschiedlich gut geeignet. Sie unterscheiden sich insbesondere in folgenden wichtigen Eigenschaften:

- Messung bestimmter physikalischer Eigenschaften wie Textur, Entfernung, Ausdehnung oder Geschwindigkeit;

- Signal-zu-Rauschverhältnis und Bestimmung der Sicherheit von Messungen;
- Maximale Reichweite bei der Erkennung von Objekten und entfernungsabhängige Tiefenauflösung;
- Geometrische Auflösung zur Trennfähigkeit zwischen Objekten und zur Texturerfassung;
- Feste oder zur geometrischen Auflösung abhängige Bildwiederholrate;
- Baugröße und mögliche Einbauposition im Fahrzeug;
- Fertigungskosten und -komplexität des Sensors sowie Zuverlässigkeit im Dauerbetrieb.

Aufgrund der unterschiedlichen Messprinzipien muss eine Auswahl von Sensoren abhängig von den jeweiligen FAS-Anwendungen getroffen werden. Können Objekte anhand ihrer Textur oder Struktur beschrieben werden, eignen sich bildgebende Sensoren wie Videokameras. Für die Erkennung von Hindernissen eignen sich tiefenbildgebende Kameras, Lidar- oder Radar-Sensoren, wobei sowohl die maximale Reichweite als auch die Tiefenauflösung zu berücksichtigen ist. Bei einem Stereokamera-System mit 12 cm Basisbreite ist der Messbereich aufgrund der nichtlinear fallenden Tiefenauflösung auf etwa 50 m eingeschränkt. Bei einer ToF-Kamera ist die Messreichweite abhängig von der Lichtstärke der eingesetzten Beleuchtungseinheit und von der Reflektanzeigenschaft von Objekten. Fehldetektionen können insbesondere durch ein schlechtes Signal-zu-Rauschverhältnis auftreten.

Für jeden Sensor muss dessen Einbauposition im Fahrzeug bekannt sein. Diese ist durch die extrinsischen Parameter des Sensors beschrieben. Bei optischen Systemen wird die Abbildungseigenschaft durch die intrinsischen Parameter spezifiziert. Die Umrechnung von einem Bildpunkt aus dem Sensorkoordinatensystem in einen Raumpunkt im Fahrzeugkoordinatensystem ist in Abschnitt 3.1.1 beschrieben.

2.2 Stochastik, Logik und Inferenz

Als Stochastik wird die gemeinsame Betrachtung mathematischer Teilgebiete der Wahrscheinlichkeitstheorie und der Statistik bezeichnet. Die Statistik erlaubt es, anhand von Beobachtungen zufälliger oder nicht spezifizierbarer Ereignisse, Aussagen über das zugrundeliegende Modell zu treffen. Die Wahrscheinlichkeitstheorie ordnet dem Eintreffen eines Ereignisses eine Wahrscheinlichkeit zu, deren Zahlenbereich zwischen Null und Eins definiert ist. Das Eintreffen eines Ereignisses gilt als sicher, wenn die Wahrscheinlichkeit genau Eins beträgt, und als nicht möglich, wenn die Wahrscheinlichkeit Null beträgt. Durch die gemeinsame Betrachtung lassen sich für Prozesse Glaubensgrade definieren, also Aussagen über den Ausgang eines Experiments.

Für diese Arbeit ist der Einsatz der Methoden der Stochastik geeignet. Da der Ausgang oder die Auswirkung einer Situation im Allgemeinen nicht exakt beschreibbar ist, werden zu einer Situation statistische Werte bestimmt und eine Wahrscheinlichkeit für das Eintreffen der einzelnen Auswirkungen berechnet.

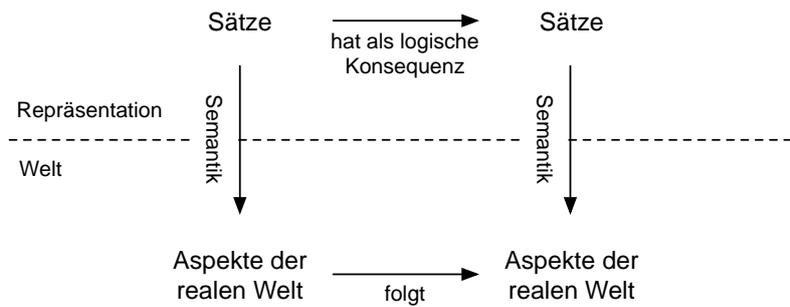


Bild 2.8: Logisches System: Sätze entsprechen formalisierten Aspekten der realen Welt; durch Inferenz lassen sich Folgerungen aus Sätzen ableiten [RN09].

Auf Grundlage einer formalen Beschreibung von Wissen in einer Wissensbasis ist es möglich, durch Logik (also der „Theorie des Folgerns“) auf weitere Informationen zu schließen [RN09]. Ein logisches System besteht dabei aus zwei Komponenten. Die Korrektheit einer Sprache wird durch ihre Syntax definiert, die Bedeutung einer Sprache durch ihre Semantik. Ein Satz ist eine nach den Regeln der Syntax korrekt definierte Formulierung. Mit Hilfe der Semantik lässt sich eine korrekt definierte Formulierung auf Aspekte der realen Welt abbilden.

Eine wahre Aussage über einen Satz wird als Welt und häufig auch als *Modell* bezeichnet. Auf dieser Grundlage lässt sich logisch schließen, also einen Satz logisch aus einem anderen Satz folgern. Dies wird auch als logische Inferenz bezeichnet. Die Inferenz leitet aus einem Modell Sätze ab, die wahr sind, wenn das Modell wahr ist. Die Zusammenhänge zwischen Semantik und Inferenz sind in Abbildung 2.8 dargestellt.

Da für Wahrscheinlichkeiten und Logik anstelle der genauen quantitativen Werte die qualitativen Beobachtungen eine höhere Relevanz besitzen, werden nach Pearl [Pea88] vier Beobachtungen definiert:

- Die Likelihood bewertet den Glaubensgrad über eine Eigenschaft,
- die Bedingung kann zur Darstellung von Sachverhalten herangezogen werden,
- die Relevanz bezeichnet den Einfluss einer Eigenschaft auf weitere Eigenschaften, und
- die Kausalität schließlich vermittelt den Grund des (Nicht-)Eintreffens einer Eigenschaft.

Die numerischen Größen leisten dabei lediglich eine Gewichtung.

2.2.1 Bedingte Wahrscheinlichkeit

Der nach dem englischen Mathematiker Thomas Bayes (1702–1761) benannte Bayes'sche Wahrscheinlichkeitsbegriff (engl. Bayesianism) interpretiert Wahrscheinlichkeit als Grad persönlicher Überzeugung (engl. degree of belief). Er unterscheidet sich damit von den objektivistischen Wahrscheinlichkeitsauffassungen wie dem frequentistischen Wahrscheinlichkeitsbegriff, der Wahrscheinlichkeit als relative Häufigkeit interpretiert.

Pearl [Pea88] führt die zugrundeliegenden Konzepte der Wahrscheinlichkeit P auf den Bayes'schen Formalismus zurück. Den drei Axiomen der Wahrscheinlichkeitstheorie genügt darin eine *Zufallsvariable*:

$$\begin{aligned} 0 &\leq P(A) \leq 1 \\ P(\text{sicher}) &= 1 \\ P(A \text{ oder } B) &= P(A) + P(B), \quad \text{wenn } A \text{ und } B \text{ sich gegenseitig ausschließen.} \end{aligned}$$

A und B sind Ereignisse. Diese Axiome besagen, dass ein Ereignis immer bestimmt werden kann, indem man es in sich gegenseitig ausschließende Unterereignisse zerlegt. Also lässt sich jedes Ereignis A als Verknüpfung der Ereignisse (A und B) sowie (A und $\neg B$) schreiben

$$P(A) = P(A, B) + P(A, \neg B). \quad (2.5)$$

Es folgt direkt, dass sich die Wahrscheinlichkeiten eines Ereignisses und seines negierten Ereignisses zu 1 summieren

$$P(A) + P(\neg A) = 1. \quad (2.6)$$

Durch den Bayes'schen Formalismus sind Aussagen über bedingte Wahrscheinlichkeiten möglich, also $P(A|B)$ – des Glaubens in A unter der Annahme dass B gilt. Wenn $P(A|B) = P(A)$ gilt, werden A und B als *unabhängig* bezeichnet. Gilt $P(A|B, C) = P(A|C)$, wird A und B als *bedingt unabhängig* gegeben C bezeichnet. Die bedingte Wahrscheinlichkeit lässt sich beschreiben durch

$$P(A|B) = \frac{P(A, B)}{P(B)}, \quad (2.7)$$

was die bedingte Beziehung von Ereignissen natürlicher beschreibbar macht als die Angabe der bedingten Wahrscheinlichkeit anhand gemeinsamer Ereignisse. So lassen sich beispielsweise kausale Zusammenhänge mit Hilfe der bedingten Wahrscheinlichkeit formulieren.

Satz von Bayes Grundlage für Bayes'sche Netze (BNs) (siehe Abschnitt 2.3.1) ist die oben erwähnte *bedingte Unabhängigkeit* von Zufallsvariablen. Gegenüber der vollständigen Beschreibung der Wahrscheinlichkeitsverteilung des Modells lassen sich die Modelle mit BNs mit deutlich geringerer Komplexität beschreiben. Bayes formuliert den Zusammenhang zwischen bedingten Abhängigkeiten von Zufallsvariablen wie folgt: Die Wahrscheinlichkeit $P(Y|X)$ einer Zufallsvariable Y gegeben einer Zufallsvariable X kann bestimmt werden, wenn die umgekehrte Wahrscheinlichkeit $P(X|Y)$ sowie die Wahrscheinlichkeit des Eintreffens von X und Y , $P(X)$ respektive $P(Y)$, bekannt sind.

$$P(Y|X) = \frac{P(X|Y) \cdot P(Y)}{P(X)}. \quad (2.8)$$

Nach dem Satz von Bayes ist also eine Umkehrung von Schlussfolgerungen möglich. $P(Y|X)$ wird auch als A-posteriori Wahrscheinlichkeit bezeichnet, $P(Y)$ als A-priori Wahrscheinlichkeit.

2.2.2 Aussagenlogik und Logik erster Ordnung

Als ein klassisches logisches System wird ein System bezeichnet, für dessen Semantik gilt, dass jede Aussage genau einen von zwei Wahrheitswerten hat und der Wahrheitswert einer zusammengesetzten Aussage eindeutig über die Wahrheitswerte ihrer Teilaussagen bestimmt ist.

Die Aussagenlogik (AL) ist ein klassisches logisches System, deren Syntax aus atomaren Sätzen besteht, welche *wahr* oder *falsch* sein können. Durch logische Verknüpfungen lassen sich atomare Sätze zu komplexen Sätzen verknüpfen. Eine gegebene Aussage kann somit geprüft werden, indem sie in ihre einzelnen atomaren Sätze zerlegt wird. Die Menge der logischen Verknüpfungen ist *nicht*, *und*, *oder*, *impliziert* und *genau dann, wenn*.

Die Mächtigkeit der AL ist beschränkt auf die logische Verknüpfung der atomaren Sätze. Da in der AL Fakten in der Welt also als *wahr* oder *falsch* bezeichnet werden, reicht die AL im Allgemeinen nicht aus, um den Großteil von Hintergrundwissen zu formalisieren. Die Logik erster Ordnung (FOL) (auch als Prädikatenlogik bezeichnet) erweitert die AL dahingehend, dass auch innerhalb atomarer Sätze Fakten über einige oder alle Objekte im Modell ausgedrückt werden können. Dies eignet sich insbesondere für die Darstellung von Fahrsituationen, da mit Hilfe der FOL Domänen durch Objekte und Relationen beschrieben werden können, wie in Kapitel 5 gezeigt wird.

Die FOL setzt sich aus vier Symbolen zusammen:

- Konstanten repräsentieren Objekte in der Domäne,
- Variablen stellen einen Bereich von Objekten der Domäne dar,
- Funktionen repräsentieren Abbildungen von Tupeln, und
- Prädikate stellen Attributeigenschaften sowie Relationen zwischen Objekten dar.

Das wesentliche Element der FOL ist das Prädikat. Außerdem wird die Menge logischer Symbole der AL in der FOL um die Quantoren *Allquantor* sowie *Existenzquantor* erweitert. Somit kann spezifiziert werden, wie ein Prädikat auf eine Menge von Objekten zutrifft. Die Menge der Objekte eines Modells wird als Domäne bezeichnet. Aussagen der FOL müssen jedoch *wahr* oder *falsch* sein.

2.2.3 Wahrscheinlichkeitsbehaftete Logik

Die Anwendung der Logik führt bei vielen realitätsnahen Beispielen dazu, dass die getroffenen Annahmen und gestellten Regeln nicht absolut eingehalten werden können. Dies lässt sich durch Hinzufügen von Ausnahmen kompensieren, wobei aber die Komplexität des logischen Systems erhöht wird.

Die wahrscheinlichkeitsbasierte Logik (PL) stellt ein System zur Verfügung, welches den Umgang mit unsicheren Regeln ermöglicht. Dadurch wird die Menge an notwendigen Regeln kompakt gehalten. Ähnlich wie Ereignisse einen Glaubensgrad besitzen, können logische Sätze eine Wahrscheinlichkeit besitzen. Mit Hilfe von PL lassen sich Wahrscheinlichkeiten über

Sätze schließen, welche logisch aus anderen Sätzen folgen. Der Wahrheitswert einer Aussage ist somit nicht mehr nur auf *wahr* oder *falsch* beschränkt, sondern wird um die Wahrscheinlichkeit im Intervall zwischen Null und Eins erweitert. Auch hier gelten die Axiome der Wahrscheinlichkeitstheorie.

Man spricht von wahrscheinlichkeitsbasierter Logik erster Ordnung (FOPL), wenn die Logik erster Ordnung mit der Wahrscheinlichkeitstheorie kombiniert wird. Dadurch lässt sich zusätzlich jedem atomaren Satz eine Wahrscheinlichkeit im Intervall zwischen Null und Eins zuordnen. Somit wird die gewünschte formale Repräsentation der Objekte und Relationen um die Beschreibung der Unsicherheit erweitert. Da die Domäne der Fahrsituationen inhärent unsicher ist, eignet sich hierfür die Beschreibung durch FOPL. Beispiele für FOPL sind in Abschnitt 2.4 beschrieben.

2.2.4 Inferenz

Wird aus einer Beobachtung auf eine Folgerung geschlossen, so spricht man von Inferenz. Dadurch kann anhand logischer Schlussfolgerungen aus bekanntem Wissen neue Information gewonnen werden. Die Entwicklung automatischer Inferenzsysteme gehört zum Forschungsfeld der „Künstlichen Intelligenz“. Ziel dieser Systeme ist das selbstständige Schließen auf neues Wissen aus einer gegebenen Wissensbasis. Dies erfolgt zum Beispiel durch das Stellen einer Anfrage (Query) an das Inferenzsystem, worauf durch logische Inferenz (zum Beispiel Rückwärtsverkettung) eine Antwort gesucht wird. Erfolgt die Bestimmung der Inferenz auf einer wahrscheinlichkeitsbehafteten Beobachtung oder Wissensbasis, so spricht man von probabilistischer Inferenz.

In Abschnitt 2.3.1 ist beschrieben, dass ein BN ein Modell für die Wahrscheinlichkeitsverteilung einer Welt darstellt. Bekannte Variablen werden darin als Evidenzen bezeichnet. Durch Inferenz lässt sich die Bestimmung der A-posteriori Wahrscheinlichkeit von Variablen (vergleiche Abschnitt 2.2.1) gegeben ihrer Evidenzen berechnen. Die Struktur von BNs ermöglicht es, durch die Kausalzusammenhänge die Verteilung einer Variablen gegeben der Evidenz im Netz zu bestimmen.

Die Inferenz von Wahrscheinlichkeitsverteilungen kann entweder exakt oder approximativ erfolgen. Verfahren zur Bestimmung der exakten Inferenz sind der *Belief Propagation (BP)* (Message Passing) sowie der *Junction-Tree* Algorithmus. Diese Verfahren können vom Aufwand her exponentiell ansteigen. Da die exakte Bestimmung der Inferenz häufig nicht notwendig ist, kann durch approximative Inferenz deutlich schneller eine ausreichend genaue Lösung gefunden werden. Die gängigen Verfahren nutzen dazu Markov chain Monte Carlo (MCMC) Methoden, wie beispielsweise *Gibbs-Sampling*. Eine gute Übersicht zu Methoden der probabilistischen Inferenz kann in [GT07] nachgelesen werden.

2.3 Graphische Modelle

Für die Darstellung von statistischen oder probabilistischen Zusammenhängen eignen sich insbesondere sogenannte graphische Modelle. Ein graphisches Modell besteht aus einer Menge von Knoten sowie einer Menge von Knoten verbindende Kanten. Während die Knoten des Gra-

phen eine Zufallsvariable beschreiben, wird die Abhängigkeit von Zufallsvariablen durch die Kanten des Graphen dargestellt. Die übliche Bezeichnung für graphische Modelle im Kontext der Stochastik lautet *probabilistisches Netz*.

Hierbei wird zwischen zwei Arten von Graphen unterschieden; den gerichteten und den ungerichteten Graphen. Gerichtete Graphen beschreiben einseitige Abhängigkeiten zwischen Zufallsvariablen, wie dies beispielsweise in Bayes'schen Netzen der Fall ist. In ungerichteten Graphen beschreiben die Kanten Wechselwirkungen zwischen Zufallsvariablen. Beispiele für ungerichtete Graphen sind Markov-Netze oder Faktorgraphen. Durch die Eigenschaft der ungerichteten Kanten können zyklische Abhängigkeiten dargestellt werden, dagegen lassen sich kausale Abhängigkeiten nicht direkt ablesen.

2.3.1 Bayes'sche Netze

Ein Bayes'sches Netz (BN) ist ein gerichteter Graph. Jeder Knoten des Graphen beschreibt eine Zufallsvariable sowie deren Eintrittswahrscheinlichkeit. Die Kanten des Graphen beschreiben die kausalen Zusammenhänge zwischen den Zufallsvariablen und haben eine Richtung. Ist eine Zufallsvariable abhängig von einer zweiten Zufallsvariablen, existiert eine Kante zwischen den Zufallsvariablen. Die Richtung der Kante beschreibt die Abhängigkeit. Das Bayes'sche Netz kann *unabhängige* und *bedingt unabhängige* Variablen direkt darstellen.

Für jede Variable ist die Wahrscheinlichkeit durch seine bedingte Wahrscheinlichkeitsverteilung (CPD) gegeben. Ein BN repräsentiert damit die vollständige Wahrscheinlichkeitsverteilung über eine Menge von Zufallsvariablen. Dadurch lässt sich für jede Variable ihre Wahrscheinlichkeit gegeben jeder Evidenz bestimmen. Ein BN ermöglicht es, eine komplexe Verteilung anhand der lokalen Abhängigkeit von Zufallsvariablen kompakt darzustellen.

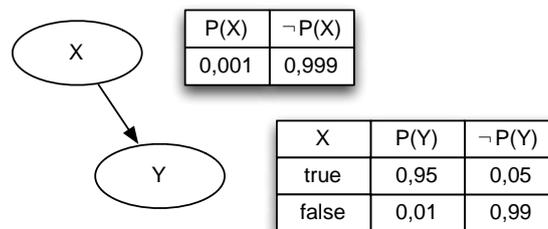


Bild 2.9: Ein einfaches BN sowie die bedingten Wahrscheinlichkeitstabellen; Y ist von X bedingt abhängig.

Die vollständige Spezifikation eines Bayes'schen Netzes $G = (X, E)$ lautet nach [RN09] wie folgt:

1. Eine Menge an Zufallsvariablen $X = X_1, X_2, \dots, X_n$ bildet die Knoten des Netzes. Variablen können kontinuierlich sein oder diskret, wenn sie nur endlich viele oder abzählbar unendlich viele Werte annehmen.
2. Eine Menge gerichteter Verknüpfungen oder Kanten E verbindet jeweils zwei Knoten. Wenn es eine Kante von Knoten X_i nach Knoten X_j gibt, sagt man, X_i ist ein Elternknoten von X_j : $X_i \in \mathbf{Pa}(X_j)$.

3. Jeder Knoten X_i hat eine bedingte Wahrscheinlichkeitsverteilung $P(X_i|\mathbf{Pa}(X_i))$, die den Effekt quantifiziert, den die Elternknoten für den Knoten verursachen.
4. Der Graph hat keine gerichteten Zyklen und ist damit ein gerichteter azyklischer Graph.

Da das BN durch seine Struktur die vollständige gemeinsame Wahrscheinlichkeitsverteilung repräsentiert, stellt es eine anschauliche Art zur Beschreibung probabilistischer Zusammenhänge dar. Das Netz ist vollständig definiert, wenn zu jedem Knoten die CPD gegeben ist (siehe Abbildung 2.9). Die vollständige Wahrscheinlichkeitsverteilung ist dann:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i|\mathbf{Pa}(X_i)). \quad (2.9)$$

2.3.2 Lernen in graphischen Modellen

Der Aufbau von graphischen Modellen erfolgt im Regelfall durch Expertenwissen. Der Experte muss dabei einerseits die genauen Knoten sowie die Kanten des Modells identifizieren, aber auch die genauen bedingten Wahrscheinlichkeitsverteilungen für die Knoten spezifizieren. Für ein vollständig beschriebenes graphisches Modell muss die vollständige Wahrscheinlichkeitsverteilung aus Gleichung (2.9) bestimmt werden. Dies bedeutet, dass für jede Variable X die CPD gegeben der Elternknoten $\mathbf{Pa}(X)$ definiert werden muss. Aufgrund der Komplexität der Modelle ist dies oftmals nicht möglich. Außerdem können für Variablen CPDs unbekannt sein. In vielen Fällen ist daher der Aufbau graphischer Modelle aufgrund von fehlendem oder ungenauem Wissen nur begrenzt möglich und sinnvoll.

Parameterlernen bezeichnet die Schätzung der CPD für Knoten. Anhand großer Stichprobenmengen kann mit Hilfe von Parameterlernen die CPD für unbekannte Variablen geschätzt werden. Ein möglicher Ansatz zum Parameterlernen ist das *Expectation Maximization* Verfahren.

Sind für einen Graphen die Knoten bekannt, nicht aber die Kanten welche die Knoten verbinden, so kann mit Hilfe von Strukturlernen die Struktur des graphischen Modells bestimmt werden. Anhand einer großen Stichprobenmenge wird das sogenannte Skelett des Graphen bestimmt. Zur Lösung von Strukturlernen werden bisher MCMC Verfahren eingesetzt.

2.3.3 Erweiterungen von graphischen Modellen

Das BN bildet die Grundlage für Bayes'sche Inferenz und erlaubt es, die Anzahl der notwendigen CPDs stark zu reduzieren und dennoch die vollständige Wahrscheinlichkeitsverteilung einer Welt zu spezifizieren. Dafür müssen die Zufallsvariablen, Abhängigkeiten und die Übergangswahrscheinlichkeiten der Knoten durch Experten entweder anhand von Vorwissen oder durch Parameterlernen definiert werden. Die resultierenden Modelle sind je nach Komplexität der darzustellenden Welt meist sehr groß und schwer zu erweitern.

Für spezielle Anwendungsfälle haben sich Erweiterungen des klassischen Modells etabliert, welche Strukturwissen über die Domäne zur Wahrung der Übersichtlichkeit und Verständlichkeit der Modelle ausnutzen. Die Modellierung zeitlicher Abhängigkeiten zwischen Variablen kann mittels Dynamischer Bayes'scher Netze (DBN) erfolgen. Für hierarchische Abhängigkeiten eignen sich Objektorientierte Bayes'sche Netze (OBN).

Dynamische Bayes'sche Netze

Nach Murphy [Mur02] modelliert ein DBN eine Zufallsvariable X_t gegenüber ihren zeitlichen Vorgängervariablen X_{t-1} und kann damit eine Dichteverteilung für zeitliche Entwicklungen darstellen

$$P(X_t|X_{t-1}) = \prod_{i=1}^N P(X_t^i|\text{Pa}(X_t^i)). \quad (2.10)$$

Dabei wird ein BN $G = (X, E)$ definiert durch das Paar (G_1, G_{\rightarrow}) (auch als 2-zeitiges temporales Bayes'sches Netz (2TBN) bezeichnet), wobei G_1 die initiale Wahrscheinlichkeitsverteilung $P(X_1)$ und G_{\rightarrow} den Übergang von $t - 1$ nach t modelliert. Die Übergangswahrscheinlichkeit $P(X_t^i|\text{Pa}(X_t^i))$ für jeden Knoten i des Netzes ist für alle $t > 1$ definiert. Einem DBN liegt die Annahme einer Markov-Kette erster Ordnung zu Grunde. Die Elternknoten $\text{Pa}(X_t^i)$ dürfen also entweder im gleichen Zeitabschnitt t oder im vorherigen Zeitabschnitt $t - 1$ liegen. Durch Ausrollen des 2TBN lassen sich die auf BN definierten Verfahren anwenden. Beispiele für DBNs sind das *Hidden Markov Model* sowie das *Kalman Filter* (siehe Abschnitt 3.3.3).

Objektorientierte Bayes'sche Netze

Ein OOBN [KP97] wendet die Mechanismen der Objektorientierung (d.h. Klassen und Instanzen) auf BN an. Eine OOBN Klasse ist ein BN Fragment, welches zusammengesetzt ist aus Eingangs-, Ausgangs- und internen Knoten:

- O : Menge von Ausgangsknoten, welche ihrerseits Elternknoten von Knoten außerhalb der Instanzen der Klasse sein können.
- I : Menge von Eingangsknoten, welche nicht in der Klasse liegen, allerdings als Elternknoten der Knoten innerhalb der Instanzen der Klasse verwendet werden. Eingangsknoten dürfen ihrerseits keine Elternknoten innerhalb der Klasse besitzen.
- P : Menge interner geschützter Knoten, deren Eltern- und Kindknoten innerhalb der Klasse liegen.

Ein OOBN ist somit ein gerichteter, azyklischer Graph $G = ((O, I, P), E)$. Die Eingangs- und Ausgangsknoten werden dabei als die Schnittstelle der Klasse bezeichnet. Gegeben den Schnittstellenknoten sind die internen Knoten bedingt unabhängig vom Rest des Graphen. Eine Klasse kann damit als stochastische Funktion betrachtet werden, welche für jeden Wert der Eingangsknoten eine Wahrscheinlichkeitsverteilung über die Werte ihrer Ausgangsknoten liefert. Dies ermöglicht einerseits die Wiederverwendung von Klassen (auch in verschiedenen Domänen) und erlaubt andererseits auch eine effiziente Berechnung der Inferenz, da die internen Knoten bedingt unabhängig vom Rest des Netzes gegeben I sind. Weiter gilt, dass dieselben zugrundeliegenden Parameter für alle Instanzen einer Klasse verwendet werden. Dies ist besonders für Parameterlernen eine wichtige Bedingung.

Für die Verknüpfung von Variablen zwischen OOBN-Klassen existieren zwei unterschiedliche Ansätze. Der Ansatz von Koller und Pfeffer [KP97] erzeugt die Kanten zwischen Knoten außerhalb der Klasse und Eingangsknoten der Klasse dynamisch. Erst bei der Instanziierung werden diese zwischen den jeweiligen Objektinstanzen erzeugt. Das führt dazu, dass die

Schnittstelle des OOBN abhängig von der Instanziierung unterschiedlich definiert ist. Auch die Vererbung von Klassen ist möglich, allerdings mit der Einschränkung, dass eine Basis-Klasse C bereits alle Eingabeknoten der vererbten Klassen C' beinhalten muss, also $I_C \supseteq I_{C'}$, wohingegen $O_C \subseteq O_{C'}$ gilt.

Bangsø [Ban04] stellt die Verknüpfung von Knoten außerhalb der Klasse zu Eingangsknoten mit Hilfe von Referenzknoten und Referenzkanten her. Alle Eingangsknoten entsprechen dabei also Referenzknoten und verweisen (ähnlich einer Referenz) auf einen entsprechenden Knoten außerhalb der Klasse. Diese kopierten Knoten erzeugen zwar eine höhere Komplexität des Modells, erlauben aber dafür, dass die Schnittstelle (O, I) des OOBN stets vollständig definiert ist. Dadurch lassen sich schon während der Klassendefinition sämtliche CPDs zuordnen. Des Weiteren erlaubt die Darstellung von Bangsø eine natürliche Vererbung von Klassen, nämlich, dass eine Klasse C' eine vererbte Klasse von C ist, wenn $I_C \subseteq I_{C'}$ und $O_C \subseteq O_{C'}$ gilt. Dies erlaubt die Erstellung von Oberklassen, welche später durch vererbte Klassen spezialisiert werden können. In dieser Arbeit werden die durch Bangsø definierten OOBNs eingesetzt.

2.4 Statistisches Relationales Lernen

Die starre Struktur von BNs ist oftmals nicht dafür geeignet, Domänen der echten Welt vollständig darzustellen. Typischerweise unterliegen die Modelle zur Beschreibung echter Domänen sowohl Unsicherheit als auch einer komplexen relationalen Struktur. Die Ansätze des statistischen Lernens liefern eine Methodik zum Umgang mit Unsicherheit, die Ansätze des relationalen Lernens sind für relationale Strukturen geeignet. Statistisches Relationales Lernen (SRL) vereint die Stärken beider Ansätze. Angelehnt an das relationale Schema von Datenbanken, können Attribute und Relationen mit Unsicherheit versehen werden. Das logische Fundament beruht auf Prädikatenlogik erster Ordnung, das probabilistische Fundament auf probabilistischen graphischen Modellen. Ein typisches Anwendungsgebiet von SRL ist das Lernen in Sozialen Netzen.

In den folgenden Abschnitten werden die für diese Arbeit wichtigsten Ansätze und deren Unterschiede nach Getoor und Taskar [GT07] vorgestellt.

2.4.1 Taxonomie der Verfahren

Die Ansätze des SRL lassen sich in drei Kategorien unterteilen: Die regelbasierten, die Entity-Relationship (ER)-basierten und die prozeduralen Verfahren.

Regelbasierte Verfahren Regelbasierte Verfahren des SRL basieren darauf, das Wissen über die Domäne in Form von Regeln zu spezifizieren. Diese sind mit einem Glaubensgrad versehen. Damit kann sehr leicht ausgedrückt werden, welche Fakten in einer Welt wahr sind und wie Fakten auf anderen Fakten beruhen. Prominenter Vertreter der regelbasierten Verfahren sind die sogenannten Markov Logik Netze (MLNs) [DR07], welche auf Markov Netzen basieren. Das logische Grundgerüst verwendet FOL und entspricht damit einer harten Bedingung für die Welt: Ist in einer möglichen Welt eine Regel nicht erfüllt, so ist die gesamte Welt nicht erfüllt; die Welt kann nach der Logik also nicht existieren. Durch MLNs werden die durch

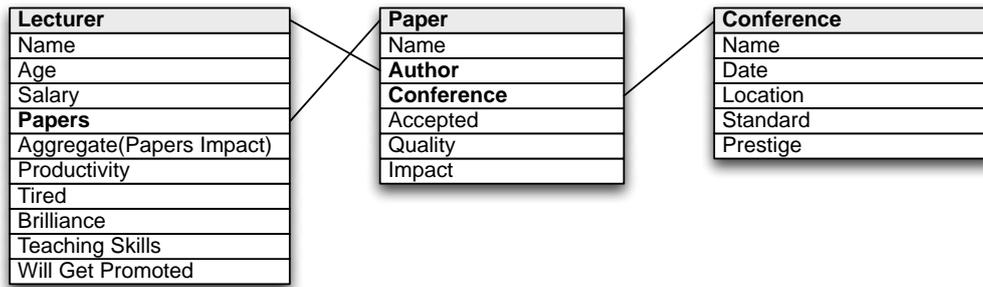


Bild 2.10: Beispiel der Universitäts-Domäne in relationaler ER-Darstellung; Klassen besitzen Attribute und sind durch Relationen miteinander verbunden [How10].

die FOL spezifizierten Regeln entschärft. Ist eine Regel durch eine Welt nicht erfüllt, so ist die Welt lediglich weniger wahrscheinlich, aber nicht unmöglich. Je weniger Regeln verletzt sind, desto wahrscheinlicher ist die Welt. Die Eigenschaft, die Wahrscheinlichkeit in Markov Netzen mit Hilfe von Gewichten zu repräsentieren, gilt ebenso für MLNs. Daher wird jede Regel mit einem Gewicht versehen, welches die Relevanz der Regel widerspiegelt. Der Einfluss der Regel auf die Gesamtwahrscheinlichkeit einer Welt wird dadurch definiert.

Entity-Relationship Verfahren ER-basierte Verfahren berücksichtigen Entitäten, deren Attribute und Relationen. Das Wissen über die Domäne wird dabei über die vorkommenden Entitäten und deren Relation untereinander modelliert, sowie welche Eigenschaften einer Entität auf Eigenschaften anderer Entitäten beruhen. Die in Abschnitt 2.4.2 beschriebenen Probabilistischen Relationalen Modelle (PRM) sowie die in Abschnitt 2.4.3 beschriebenen Objektorientierten Probabilistischen Relationalen Modelle (OPRM) gehören zur Klasse der ER-basierten Verfahren. Ein häufig verwendetes Anwendungsbeispiel für SRL-Verfahren ist durch die Universitäts-Domäne gegeben. Das relationale Modell dieser Domäne ist in Abbildung 2.10 gezeigt.

Prozedurale Verfahren Prozedurale Verfahren beschreiben das Wissen über eine Domäne in Form von Prozessen. Bei der Modellierung steht also der Ablauf im Vordergrund, wie eine Welt aufgebaut wird und wie ein Ereignis ein anderes Ereignis beeinflusst. BLOG (Bayes'sche Logik) gehört zur Sprachenklasse der Prozeduralen Verfahren. Der Vorteil dieser Sprache liegt darin, dass kein vollständiges Wissen über die Welt vorhanden sein muss, da die Welt iterativ aufgebaut wird. Dadurch lässt sich Wissen auch über unbekannte Objekte und über Domänen schließen, welche eine unbekannte oder veränderliche Anzahl von Objekten besitzen [MR07].

2.4.2 Probabilistische Relationale Modelle

Nach Getoor und Taskar [GT07] erweitern Probabilistische Relationale Modelle (PRM) die BNs durch die Konzepte von Objekten, ihrer Eigenschaften sowie die Relationen zwischen diesen. PRMs nutzen Klassen zur Darstellung von Objekten und Relationen. Jede Klasse entspricht dabei einem BN-Fragment, wobei alle Klassen gemeinsam einer wahrscheinlichkeits-behafteten Ontologie der Domäne entsprechen. PRMs integrieren also relationale Strukturen in OOBNs.

Ein PRM besteht aus zwei Komponenten, der *relationalen Komponente*, durch welche das logische Schema der Domäne widergespiegelt wird, und der *probabilistischen Komponente*, welche die probabilistischen Abhängigkeiten in der Domäne beschreibt.

Die relationale Komponente besteht aus einer Menge von Klassen $\mathbf{C} = \{C_1, C_2, \dots, C_n\}$. Jede Klasse besitzt eine Menge einfacher (oder auch beschreibender) Attribute und eine Menge komplexer (oder auch referenzierender) Attribute. Die Menge der einfachen Attribute einer Klasse C ist $\Delta_C = \{\delta_1, \delta_2, \dots, \delta_n\}$, die Menge der komplexen Attribute $\Phi_C = \{\phi_1, \phi_2, \dots, \phi_n\}$.

Ein einfaches Attribut hat einen vordefinierten, abzählbaren Wertebereich $\text{Val}(C_1.\delta_x)$. Ein komplexes Attribut hat einen vordefinierten Bereich $\text{Range}(C_1.\phi_x) = C_y$, wobei $C_y \in \mathbf{C}$ gilt. Während die einfachen Attribute Objekteigenschaften beschreiben, dienen die komplexen Attribute dazu, Relationen zwischen den Klassen herzustellen. Für jedes komplexe Attribut existiert außerdem ein inverses Attribut ϕ_x^{-1} , welches als inverse Funktion von ϕ_x dient. Die relationale Komponente orientiert sich somit wesentlich an dem Konzept relationaler Datenbanken.

Die probabilistische Komponente ist für die relationale Komponente eines PRM bestimmt durch eine Menge von Elternknoten $\mathbf{Pa}(\delta_x) = \{Pa_1, Pa_2, \dots, Pa_n\}$ für jede Klasse $C \in \mathbf{C}$ und jedes einfache Attribut $\delta_x \in \Delta_C$ und der CPD $P(\delta_x | \mathbf{Pa}(\delta_x))$. Die PRMs basieren somit auf BNs. Die vollständige Wahrscheinlichkeitsverteilung des Modells für alle möglichen Welten ist über die CPD der Attribute spezifiziert.

Durch die Instanz eines PRM wird dessen Semantik definiert. Dabei wird für jede Klasse $C \in \mathbf{C}$ eine Menge von Objekten instanziiert und alle einfachen und komplexen Attribute dieser Objekte spezifiziert. Ein PRM definiert eine Wahrscheinlichkeitsverteilung einer Welt über die Menge seiner Instanzen.

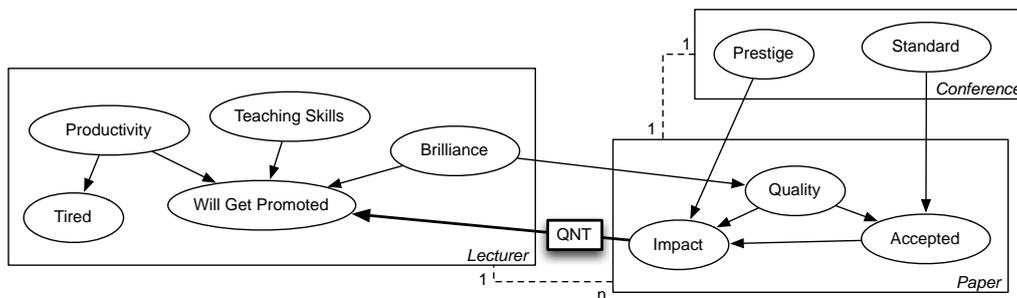


Bild 2.11: Probabilistisches Modell der Universitäts-Domäne in PRM-Form; Durchgezogene Pfeile stellen probabilistische Abhängigkeiten zwischen Attributen dar, gestrichelte Linien Relationen zwischen Klassen; der fette Pfeil entspricht einer Zählfunktion (QNT) zur Aggregation der *Impact*-Attribute aller *Paper*-Klassen [How10].

PRMs erlauben es außerdem, verschiedene Arten von Unsicherheiten zu modellieren. Dies sind die Unsicherheit über die Anzahl von Objekten, die Art der Referenz von Objekten sowie die Existenz von Objekten. Letztere erlaubt die Modellierung von Unsicherheiten mit Bezug auf die Erkennungsrate von Algorithmen.

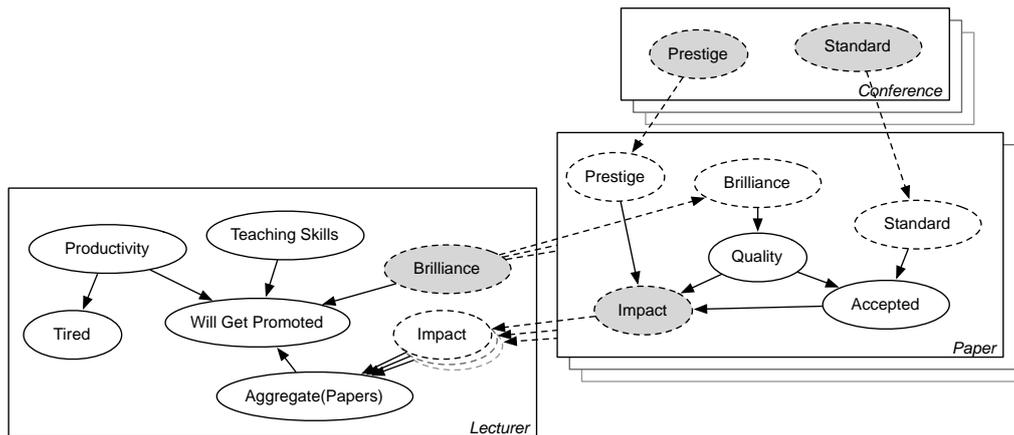


Bild 2.12: Probabilistisches Modell der Universitäts-Domäne in OPRM-Form; Durchgezogene Pfeile stellen probabilistische Abhängigkeiten zwischen Attributen dar, gestrichelte Attribute entsprechen Referenzknoten, grau gefärbte Attribute entsprechen Ausgangsknoten; gestrichelte Pfeile bilden Referenzkanten zwischen Ausgangsknoten und erzeugten Referenzknoten [How10].

2.4.3 Objektorientierte Probabilistische Relationale Modelle

Die Objektorientierten Probabilistischen Relationalen Modelle (OPRM) sind zu den im vorherigen Abschnitt vorgestellten PRMs eng verwandt. Allerdings besitzen sie einen stärkeren Bezug zum Konzept der Objektorientierung. Sie gehören ebenfalls zur Klasse der ER-basierten Verfahren der FOPL. Im Gegensatz zu BNs, welche attributbasierte Modelle sind, entsprechen die OPRMs klassenbasierten Modellen. Dadurch lassen sich probabilistische Modelle für Klassen von Objekten spezifizieren und probabilistische Abhängigkeiten des Modells auf eigene Attribute oder auf Attribute verknüpfter Klassen bilden. Die formale Sprachdefinition der OPRMs ist in [HS09] beschrieben und besteht aus zwei Komponenten. Einer relationalen Komponente und einer probabilistischen Komponente. Sie wird als Objektorientierte Probabilistische Relationale Modellierungssprache (OPRML) bezeichnet. Während die relationale Komponente die Beziehungen zwischen Klassen der Domäne beschreibt, dient der probabilistische Teil dazu, die probabilistischen Abhängigkeiten zwischen den Attributen in der Domäne zu formalisieren.

Die OPRML definiert also eine formale Grundlage zur Beschreibung von Domänen. Die so definierte Sprache erlaubt es, Domänen mit Attributunsicherheit, Strukturunsicherheit sowie Existenzunsicherheit zu modellieren. Ein aus dieser Sprache und einer Wissensbasis aufgebautes probabilistisches Modell ist ein OPRM und entspricht einer möglichen Welt. Ein OPRM spezifiziert die vollständige Wahrscheinlichkeitsverteilung über die Menge aller möglichen Welten.

Sowohl in PRMs als auch in OPRMs werden die Modelle einer Domäne aus einer Wissensbasis erzeugt. Die graphischen Modelle der Sprache entsprechen OOBNs (siehe Abschnitt 2.3.3). Die Unterscheidung zu den oben spezifizierten PRMs liegt im Wesentlichen in der Art der verwendeten OOBNs, welche zur Berechnung der Wahrscheinlichkeitsverteilung aus dem Modell erzeugt werden. Das OPRM verwendet das von Bangsø [Ban04] definierte OOBN-Konzept und nutzt die Vorteile der stets definierten Schnittstelle der OOBN-Klassen sowie der Vererbung von Klassen aus.

Das durch Howard und Stumptner [HS09] vorgestellte Konzept ermöglicht die formale Repräsentation von dynamischen Situationen unter Berücksichtigung von Unsicherheiten. Die vorgestellte Methodik eignet sich zur Beschreibung von Fahrsituationen. OPRMs haben bislang Anwendung in der Situationsanalyse im taktisch-militärischen Umfeld gefunden, worauf im Stand der Forschung in Abschnitt 3.4.3 eingegangen wird.

3 Stand der Forschung

Im folgenden Kapitel wird der Stand der Forschung aufgearbeitet und darauf aufbauend eine Aussage über den Forschungsbedarf dieser Arbeit getroffen. Die Analyse der Fahrsituation erfordert es, die relevanten Objekte im Fahrzeugumfeld mit Hilfe von Sensoren zu erfassen. In Abschnitt 2.1 wurde bereits ein Überblick zu geeigneten Sensorsystemen dargestellt. Darauf aufbauend werden zunächst in Abschnitt 3.1 Verfahren analysiert, welche die spezifischen Besonderheiten bei der Verarbeitung von Sensordaten und die Grundkonzepte zur Erkennung relevanter Objekte vorstellen.

Anschließend folgt in Abschnitt 3.2 die Analyse von Ansätzen zur Umfeldwahrnehmung mit geeigneten Sensorsystemen, wobei der Schwerpunkt auf der Erkennung dynamischer Objekte liegt. Die Eigenschaften der Sensoren und Erkennungsverfahren bringen mit sich, dass häufig auch Geisterobjekte durch das System erkannt werden. Eine etablierte Methode, diese Objekte zu unterdrücken und gleichzeitig die Erkennung gültiger Objekte zu stabilisieren, liegt im Einsatz von temporalen Filtern. Geeignete Methoden werden dazu in Abschnitt 3.3 vorgestellt. Abschnitt 3.4 beschreibt bestehende Methoden zur Repräsentation von Hintergrundwissen auf Basis von Logik und Probabilistik sowie zur Interpretation von Fahrsituationen aus der Domäne der Fahrerassistenzsysteme (FAS). Das Kapitel wird in Abschnitt 3.5 durch eine Diskussion des Standes der Forschung sowie des Forschungsbedarfs dieser Arbeit abgeschlossen.

Dieses Kapitel stellt damit konzeptuelle Ansätze vor, die Lösungen zu den folgenden Problemfeldern aufzeigen:

- Abschnitt 3.1: Methoden zur Wahrnehmung des Fahrzeugumfelds,
- Abschnitt 3.2: Verfahren zur Erkennung dynamischer Objekte,
- Abschnitt 3.3: Methoden zur temporalen Informationsverarbeitung und Zustandsschätzung, sowie
- Abschnitt 3.4: Ansätze zur Wissensrepräsentation und formalen Situationsbeschreibung.

3.1 Sensordatenverarbeitung

Aufgrund der sehr heterogenen Sensorsysteme und der daraus gewonnenen Informationen werden an dieser Stelle Konzepte erklärt, die übergreifend bei der Entwicklung von FAS berücksichtigt werden müssen. Zunächst wird die Konvention der Koordinatensysteme für FAS beschrieben sowie die jeweiligen Transformationen in die Sensorkoordinatensysteme definiert. Anschließend wird die grundlegende Datenverarbeitungskette beschrieben, welche aus Informationsvorverarbeitung, Berechnung von Objekthypothesen und temporaler Verifikation von Objekthypothesen besteht. Die Bildung von Objekthypothesen erfolgt zunächst durch eine

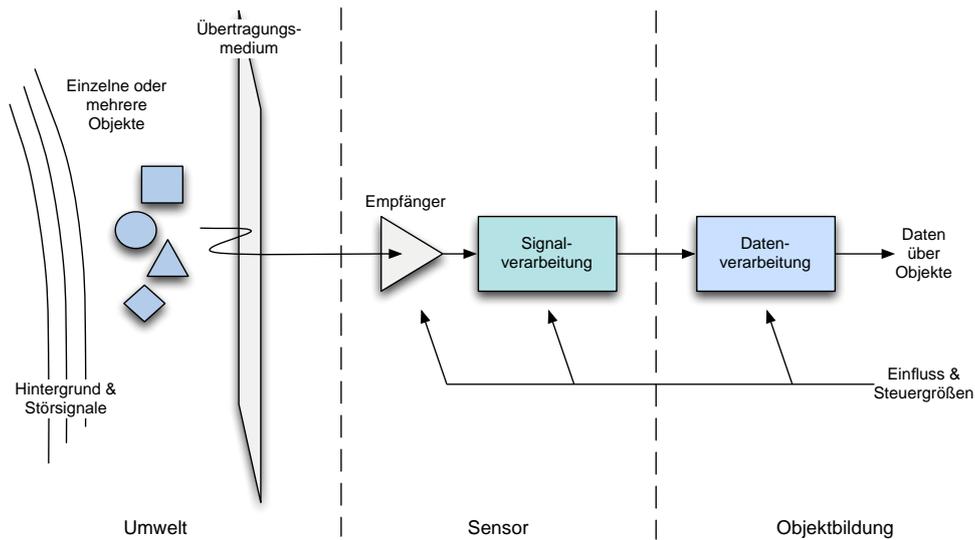


Bild 3.1: Komponenten eines Sensorsystems: Sensoren erfassen die Umwelt und bilden daraus Objektwissen; nach Blackman und Popull [BP99].

Objektsegmentierung und einer anschließenden (optionalen) Objektklassifikation. Die temporale Verifikation ermöglicht es, den Zustand eines Objektes über einen längeren Zeitraum, also über mehrere Messzyklen, zu filtern. Daraus lassen sich eine genauere Zustandsbeschreibung sowie zusätzliche Informationen gewinnen. Ist das verarbeitende System in der Lage, die durch die Sensoren gelieferten Daten ausreichend schnell zu verarbeiten, so spricht man von einer echtzeitfähigen Verarbeitung. Diese Anforderung ist für FAS sehr wichtig.

Ein allgemeines Sensorsystem besteht aus einem oder mehreren Sensoren, welche die Umwelt durch geeignete Signalmessungen mit Hilfe von Empfängern wahrnehmen. Wie in Abbildung 3.1 gezeigt, besteht die Umwelt immer sowohl aus relevanten Objekten sowie aus Hintergrund- und Störobjekten. Ziel der Signalverarbeitung ist es, die Signale des Empfängers so zu verarbeiten, dass eine möglichst gute Trennung zwischen relevanten Objekten und Hintergrund erfolgt. An die Signalverarbeitung schließt sich die Objektbildung an, welche die gemessenen Signale schließlich zu Objekten zusammenfasst und daraus ein virtuelles Abbild der in der Umwelt tatsächlich vorhandenen Objekten erstellt.

3.1.1 Koordinatensystem-Konvention

Die Konvention der zu verwendenden Koordinatensysteme für FAS ist in der DIN 70000 sowie international in der ISO 8855 definiert. Diese schreiben für ein Fahrzeug ein rechtwinkliges Achsensystem vor, bei welchem in Fahrtrichtung gesehen die x -Achse nach vorne, die y -Achse nach links und die z -Achse nach oben zeigt. Hierdurch wird ein rechtshändiges Koordinatensystem beschrieben, in dem der Drehsinn mathematisch positiv definiert ist. Der Referenzpunkt des fahrzeugfesten Inertialsystems mit $(x, y, z)^T = (0, 0, 0)^T$ liegt am vordersten Punkt in der Mitte des Fahrzeuges auf der Bodenebene (siehe Abbildung 3.2). Dieses Referenzsystem wird als Fahrzeugkoordinatensystem bezeichnet.

Bei der Transformation von Sensorinformationen in das Fahrzeugkoordinatensystem muss unterschieden werden, ob die verfügbaren Informationen in 2D oder in 3D vorliegen. Bei

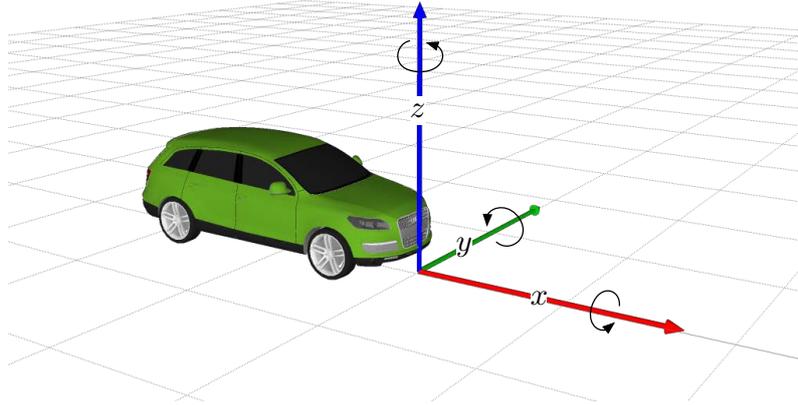


Bild 3.2: Definition des fahrzeugfesten Inertialsystems.

2D-Daten müssen zur Transformation sowohl die extrinsischen wie auch die intrinsischen Kameraparameter bekannt sein, sowie ein zusätzlicher Wert, welcher dem Tiefenwert des Bildpunktes in der Szene entspricht. Bei 3D-Daten reichen dagegen die extrinsischen Kameraparameter aus, um eine Transformation der Information aus dem Kamerakoordinatensystem in das Fahrzeugkoordinatensystem durchzuführen.

Sind die Kameraparameter bekannt, lässt sich die Transformation für einen gegebenen Objektpunkt $\mathbf{o} = (o_x, o_y, o_z)^T$ in $\mathbf{p} = (p_x, p_y, p_z)^T$ von Kamerakoordinaten in Fahrzeugkoordinaten ausdrücken mit:

$$\begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} = \left(YPR \begin{bmatrix} o_x \\ o_y \\ o_z \end{bmatrix} \right) + \begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix}, \quad (3.1)$$

mit

$$YPR = \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix}, \quad (3.2)$$

wobei $c_x, c_y, c_z, \theta_x, \theta_y, \theta_z$ die extrinsischen Parameter der Kamera in das Fahrzeugkoordinatensystem bezeichnen. $(c_x, c_y, c_z)^T$ bezeichnet den Translationsvektor, θ_x den Roll-Winkel, θ_y den Nick-Winkel und θ_z den Gier-Winkel.

Soll aus den 2D-Daten eines Objektpunktes der 3D-Punkt in Fahrzeugkoordinaten bestimmt werden, muss eine Schätzung des kartesischen Abstands \hat{d} des Messpunktes zur Kamera durchgeführt werden. Der in Bildkoordinaten definierte Punkt $\mathbf{i} = (u, v)^T$ lässt sich in den Objektpunkt \mathbf{o} unter Berücksichtigung der intrinsischen Parameter überführen:

$$\begin{bmatrix} o_x \\ o_y \\ o_z \end{bmatrix} = \begin{bmatrix} d \\ \frac{(u_0 - u)d}{f_u} \\ \frac{(v_0 - v)d}{f_v} \end{bmatrix}, \quad (3.3)$$

wobei durch (u_0, v_0) der Versatz der optischen Bildachse zum Mittelpunkt des Sensors und durch (f_u, f_v) die Fokusslänge der Linse beschrieben ist.

3.1.2 Vorverarbeitung

Die Sensordatenverarbeitung besteht zunächst darin, eine möglichst hohe und zugleich stabile Datenqualität den anschließenden Verarbeitungsalgorithmen bereitzustellen. Den Algorithmen der Sensordatenverarbeitung liegen Annahmen über die Daten zu Grunde, insbesondere in Form von Schwellwerten, Filtergrößen und -stärken. In der Vorverarbeitung werden einerseits die aufgezeichneten Messdaten normalisiert und andererseits das vorhandene Messrauschen in den Sensordaten reduziert. Eine Normalisierung ist notwendig, wenn Umwelteinflüsse einen systematischen Fehler auf die Messdaten bewirken. Bei einer Videokamera führen beispielsweise unterschiedliche Lichtverhältnisse dazu, dass die Helligkeit der aufgezeichneten Bilder stark variiert. Dem lässt sich durch eine Helligkeitsnormalisierung entgegenwirken. Ebenfalls kann das Messrauschen systematisch gefiltert werden, indem für die Objekterkennung wenig oder nicht relevante Daten frühzeitig unterdrückt werden. Dabei können zusätzliche, durch den Sensor aufgezeichnete Informationen wertvolle Hinweise darüber liefern, welche Messinformation als zuverlässig und welche als weniger zuverlässig klassifiziert wird. Bei aktiven Kamerasystemen korreliert beispielsweise die Intensität des reflektierten Lichts (also die gemessene Signalamplitude) mit der Zuverlässigkeit der gemessenen Information.

3.1.3 Bildung von Objekthypothesen

Nach der Vorverarbeitung erfolgt die Erkennung einer Menge von Objekten. Aus Gründen der Systemlaufzeit wird oftmals auf komplexe Algorithmen verzichtet und statt dessen eine höhere Falsch-Positiv-Rate in Kauf genommen. Daher spricht man auch von Objekthypothesen, welche möglicherweise nicht mit tatsächlich im Umfeld vorhandenen Objekten korrespondieren. Durch die Erstellung von Objekthypothesen wird der Suchraum für die wesentlich komplexeren Algorithmen zur Verifikation von Hypothesen eingeschränkt, sodass Anforderungen an die Laufzeit der Algorithmen auch bei rechenintensiven Operationen erreicht werden können.

Für die Bildung von Objekthypothesen wird zuerst eine Segmentierung in den Sensordaten durchgeführt. Dabei werden Bereiche in den Sensorinformationen markiert, welche einem Objekt zugeordnet werden. Diese Bereiche werden als Region of Interest (ROI) bezeichnet und beschreiben die Eingabedaten für die anschließende Objektklassifikation sowie für die Bestimmung der Objektgröße. Die Objektklassifikation bestimmt für jede ROI eine Zuordnungswahrscheinlichkeit zu bekannten Objektklassen. Dabei werden in der Regel mit Hilfe von maschinellen Lernverfahren Klassifikatoren für die Objektklassen im Voraus trainiert. Die Klassifikation kann für flach oder hierarchisch strukturierte Klassen erfolgen.

Objektsegmentierung

Die Unterscheidung von Objekten und Hintergrund wird durch die Objektsegmentierung geleistet. Die während einem Messzyklus aufgezeichneten Sensorinformationen werden objektspezifisch analysiert. Je nach Messprinzip kommen zur Objektsegmentierung unterschiedliche Ansätze und Merkmale zur Geltung.

In Videodaten lässt sich mittels ansichtsbasierter Merkmale eine Segmentierung erreichen, zum Beispiel mittels Hough-Raum Analyse oder Symmetriemerkmale, während cluster-

ziehungsweise modellbasierte Verfahren für die Segmentierung von Punktwolken geeignet sind. Zum Beispiel Cluster anhand der Euklidischen Distanz, Regionenwachstums-Verfahren oder der Random Sample Consensus (RANSAC) Algorithmus (siehe [Jäh05; FB81]). Das zugrundeliegende Prinzip besteht darin, dass ein Objekt sich durch konstante Merkmale beziehungsweise durch starke Differenz gegenüber seiner Umgebung identifizieren lässt und sich dadurch vom Hintergrund klar abhebt. Bei modellbasierten Verfahren muss außerdem ein wesentlicher Teil des Objektes beobachtbar und darf nicht verdeckt sein.

Typischerweise auftretende Fehler bei der Objektsegmentierung sind Untersegmentierung, Übersegmentierung und Verschmelzung von Objekten. Untersegmentierung bezeichnet dabei Segmentierungsfehler, bei denen mehrere Objekte zu einem Segment fälschlicherweise zusammengefasst werden. Bei Übersegmentierungsfehlern werden Objekte durch zu kleine Segmente beschrieben. Besonders die clusterbasierten Ansätze neigen zu Untersegmentierung von Objekten, da nahe beieinander platzierte Objekte zu einem Segment zusammengefasst werden.

Objektklassifikation

Die Objektklassifikation vollführt die Trennung eines Merkmalsraumes in einzelne Klassen. Sie wird auch als Mustererkennung bezeichnet. Damit Objektklassen spezifisch voneinander unterschieden werden können, werden in einem Klassifikationsschritt für eine Objektklasse spezifische Merkmale so miteinander kombiniert, dass eine Aussage darüber getroffen werden kann, ob das untersuchte Objekt einer Klasse zugehörig ist oder nicht. Dies ist möglich, wenn durch Merkmalerkennung die beschreibenden Merkmale (der Merkmalsvektor) einer Objektklasse bestimmt werden können. Der Merkmalsvektor beschreibt voneinander disjunkte Objekteigenschaften, welche über die Objektklasse generalisieren und eine optimale Beschreibung der Objektklasse im Merkmalsraum liefert. Für alle Objekte einer Objektklasse sollte der Merkmalsvektor identisch sein.

Da die Erstellung der Merkmalsvektoren in der Regel nicht trivial ist, wird mit Hilfe einer statistisch relevanten Grundmenge einer Objektklasse ein Merkmalsvektor mit Hilfe maschineller Lernverfahren trainiert. Zur anschließenden Trennung des Merkmalsraumes eignen sich gängige Klassifikationsverfahren wie Boosting Ansätze [VJ01] oder Support Vektor Maschinen (SVM) [CS00]. Die Objektklassifikation kann in der Regel überwacht oder unüberwacht stattfinden. Für die Anwendung in dieser Arbeit eignet sich die unüberwachte Klassifizierung, da das System Entscheidungen autark treffen muss. Je nach eingesetztem Sensorsystem können zur Objektklassifikation die gleichen Sensorinformationen wie zur Objektsegmentierung verwendet oder komplementäre Sensordaten eingesetzt werden.

3.1.4 Verifikation von Objekthypothesen

Die erstellten Objekthypothesen enthalten in der Regel nicht ausschließlich gültige Objekte sondern auch Hintergrundobjekte, die fälschlicherweise als gesuchte Objekte interpretiert wurden. Dies wird in Kauf genommen, wenn das Verfahren zur Erstellung der Objekthypothesen dafür weniger Ressourcen benötigt, weniger komplex ist oder mit höherer Wahrscheinlichkeit eine größere Menge gültiger Objekte detektiert - dennoch sollte eine Objektliste nur

eine minimale Zahl ungültiger Objekte enthalten. Die Verifikation von Objekthypothesen dient dazu, unsichere oder falsche Hypothesen zu markieren und gültige Hypothesen zu bestätigen. Dies kann unterschiedlich erreicht werden:

- Regelbasierte Verkettung von modellbasierten Merkmalen,
- Klassifikation von ansichtsbasierten Merkmalen,
- Verfolgung von temporalen Merkmalen.

Sowohl die regelbasierte Verkettung modellbasierter Merkmale als auch die Klassifikation ansichtsbasierter Merkmale stellen rechenaufwändige Entscheidungsprozesse dar, besonders wenn diese auf alle Eingangsdaten angewendet werden müssen. Ist das Verhältnis von auftretenden Fehldetektionen zu gültigen Detektionen gering, lässt sich eine Verifikation von Objekthypothesen durch temporale Fusion erreichen. Hierbei wird jede Objekthypothese über einen längeren Zeitraum beobachtet und während dessen als Objektsequenz (auch als Track bezeichnet) gespeichert. Die Sequenz wird fortlaufend überprüft, ob die Objekthypothese kontinuierlich beobachtet wird und sich gemäß der erwarteten Modellannahmen verhält. Ist dies der Fall, wird die Objekthypothese als verifiziert markiert, anderenfalls gelöscht. Hierdurch wird die Erkennungsrate gültiger Objekte trotz Vorhandensein von einzelnen Fehldetektionen stabilisiert.

3.1.5 Weiterführende Datenverarbeitung

Mit der Erstellung der verifizierten Objekthypothesen ist der Prozess der Objektbildung abgeschlossen. Auf Grundlage der gebildeten Objekte lässt sich mit Verfahren zur Situationsanalyse und -bewertung nun eine Aussage über die vorhandene Situation erstellen und der Einfluss der Objekte auf das eigene System messen. Diese Verarbeitungsschritte sind geprägt durch Inferenzverfahren, also dem Schließen weiterer Informationen durch Logik anhand des gebildeten Wissens und einer Wissensbasis.

3.2 Umfeldwahrnehmung

Die Erkennung der relevanten Objekte im Fahrzeugumfeld umfasst die Erkennung dynamischer und semistatischer Objekte, also Objekte, welche nicht zur Verkehrsinfrastruktur gehören sondern beweglich sind und somit für das eigene Fahrzeug ein plötzliches Hindernis darstellen können. Die Erkennung von statischen Objekten ist in den meisten Fällen gleichzusetzen mit der Erkennung von Infrastruktur-Objekten, worauf im zweiten Teil dieses Abschnittes kurz eingegangen werden soll.

3.2.1 Erkennung dynamischer Objekte

Der folgende Abschnitt zeigt, wie die Erkennung von dynamischen Objekten erfolgen kann. Zunächst werden Verfahren zur Erkennung von Fahrzeugen vorgestellt. Anschließend werden Arbeiten vorgestellt, die sich mit der Erkennung von Objekten unabhängig von deren Objektklasse beschäftigen.

Fußgänger gelten im Verkehrsgeschehen als die schwächsten Teilnehmer, weshalb Unfälle zwischen Fahrzeugen und Fußgängern oftmals schwerwiegend sind. Die Erkennung von Fußgängern stellt ein eigenständiges, komplexes Forschungsgebiet dar und steht im Rahmen dieser Arbeit nicht im Vordergrund. Die Arbeit von Enzweiler und Gavrilu [EG08] liefert eine umfassende Übersicht zur Fußgängererkennung auf Videodaten. Eine Klassifikation von Fußgängern ausschließlich auf Lichtlaufzeit (ToF)-Daten ist in [EM07] vorgestellt. Erwähnenswert ist die Arbeit von Ess u. a. [Ess+09]; diese erreicht eine deutliche Verbesserung in der Erkennung von Fußgängern durch Kombination von Stereo-Tiefenbildern mit hochauflösenden Videodaten.

Erkennung von Fahrzeugen

Die Erkennung vorausfahrender Fahrzeuge ist eine Grundvoraussetzung für das inzwischen in vielen Fahrzeugen verbaute ACC. Aufgrund des Einsatzgebietes dieses Systems auf Autobahnen ist eine große Reichweite des Sensors und die Bestimmbarkeit der Objektgeschwindigkeit von großer Bedeutung, weshalb sich besonders der Radar-Sensor für diesen Einsatzzweck eignet. In [Jor06] wurden Verfahren vorgestellt, mit welchen sich anhand der Signal-Information des Radar-Sensors Fahrzeughypothesen erstellen lassen. Dabei wird die Ausdehnung der Objekte anhand der Reflexionscharakteristik geschätzt, sowie Scheinbewegungen durch Spiegelungen an Leitplanken unterdrückt. In der Arbeit wird besonders auf die Herausforderung einer stabilen und zuverlässigen Datenassoziation hingewiesen, welche in der Arbeit zu einer Erweiterung des Probabilistic Data Association (PDA)-Verfahrens geführt hat. Dieses sogenannte *EPDA*-Verfahren berechnet zugleich eine Existenzwahrscheinlichkeit für jedes verfolgte Objekt, mit der eine Aussage über die Zuverlässigkeit der Objekthypothese ermöglicht wird. Dabei werden Modellannahmen für die Klasse *Fahrzeug* getroffen. Die Verfahren sollten jedoch auch für andere Objektklassen anwendbar sein.

Neben dem Radar-Sensor hat sich mittlerweile auch der Video-Sensor im Fahrzeug als quasi-Standard etabliert. Dieser hat als reiner 2D-Sensor allerdings den Nachteil, weder die Entfernung eines Objektes noch die genaue Geschwindigkeit direkt aus den Sensordaten zu erfassen. Um zunächst die Distanz eines Objektes korrekt zu schätzen, wird auf Modellannahmen zurückgegriffen.

Hoffmann [Hof06] stellt in seiner Arbeit Bildoperatoren zur Erzeugung von Fahrzeughypothesen vor. Die Schattenkante des Fahrzeuges auf der Fahrbahnoberfläche ermöglicht eine gute initiale Hypothese und liefert gleichzeitig eine Schätzung der Entfernung des Fahrzeuges. Dies ist möglich, wenn sowohl die genaue Einbauposition des Sensors bekannt ist als auch der Fahrbahnverlauf vor dem Fahrzeug eben ist. Da das Schattenmerkmal zur Fahrzeugerkennung alleine nicht spezifisch genug ist, werden zusätzlich das Symmetrie-Merkmal sowie eine Kombination verschiedener Kantenmerkmale eingesetzt. Die vorgestellten Merkmale eignen sich zur Erkennung der Fahrzeugfront sowie des -hecks, allerdings nicht für schräg positionierte Fahrzeuge. Mit Hilfe des Joint Probabilistic Data Association (JPDA)-Verfahrens und des Interacting Multiple Model (IMM)-Filters werden anschließend fehlerhafte Hypothesen gefiltert und Fahrzeughypothesen verifiziert, sowie die Objektgeschwindigkeit und die Bewegungsrichtung ermittelt. Weitere Arbeiten verwenden zur Suche nach Fahrzeughypothesen schnelle

Klassifikatoren, welche mittels dem Sliding-Window Verfahren die Bilddaten in verschiedenen Skalierungen nach ansichtsbasierten Merkmalen durchsuchen. Hierzu eignen sich besonders Boosting-Klassifikatoren mit Haar-Wavelets. Aufgrund der Unsicherheit in der Schätzung der Entfernung und der schlechteren Zuverlässigkeit bei der Erkennung von Fahrzeugen aus Videodaten werden häufig zusätzliche Sensoren in die Verarbeitungskette integriert oder durch das System ausschließlich Warnhinweise gegeben.

Neben der Erkennung von Fahrzeughypothesen stellt auch die Klassifikation von Bildausschnitten zur Fahrzeugerkennung eine Herausforderung dar. Sun, Bebis und Miller [SBM06b] beschreiben hierfür Merkmale, welche zur videobasierten Erkennung von Fahrzeugen geeignet sind. Neben den schon oben genannten Schatten-, Symmetrie- und Kantenmerkmalen werden auch Farb- sowie Texturmerkmale vorgeschlagen. Allerdings ist zu erwarten, dass diese Merkmale in inhomogenen Szenarien wie beispielsweise in der Innenstadt, schwer vom Hintergrund unterscheidbar sind. Zur Hypothesenverifikation werden templatebasierte Methoden wie die Erkennung des Kennzeichens oder die Erkennung der U-Form von Fahrzeugkanten erwähnt. Als wesentlich relevanter werden die ansichtsbasierten Methoden in [SBM06a] eingeschätzt, welche eine Entscheidung *Fahrzeug/nicht Fahrzeug* mit Hilfe von Klassifikationsverfahren treffen. Hierbei erzielen besonders die Haar- und Gabor-Merkmale in Kombination mit Support Vektor Maschinen (SVMs) vielversprechende Ergebnisse.

Erkennung unbekannter Objekte

Unabhängig von der Güte bei der Erkennung bekannter Objektklassen wie *Fahrzeuge* oder *Fußgänger* lässt sich das Fahrzeugumfeld dennoch nicht ausreichend wahrnehmen. Dies liegt daran, dass Objekte entweder nur teilweise sichtbar sind oder es sich um unerwartete Objektarten handelt (beispielsweise ein auf der Fahrspur liegender Ast). Die Erkennung unbekannter Objekte fällt mit monoskopischen Videodaten schwer, da die Unterscheidung zwischen Hintergrund-Objekten und unbekanntem Vordergrundobjekt meist nur durch Modellwissen möglich ist. Der Einsatz tiefenbildgebender Sensorik, also ToF-Kamera, Stereokamera oder mehrzeiliger Laserscanner erleichtert die Erkennung von Objekten auf der Fahrfläche, da Annahmen über die Position von Objekten getroffen werden, jedoch nicht über die Objektklasse. Gemeinsam bei den hier vorgestellten Ansätzen ist, dass die ermittelte Objektkontur und die Position von Objekten die relevante Information für nachfolgende Systeme zur Situationsbewertung darstellt.

Fardi u. a. [Far+06] stellen auf Grundlage der Daten einer Photomischdetektor (PMD)-Kamera ein Verfahren vor, welches durch Segmentierung der PMD-Daten Objekthypothesen bildet. Zunächst werden in einem Vorverarbeitungsschritt ungültige Bildpunkte mit Hilfe des Amplitudenbildes und einem Homogenitätskriterium gefiltert. Mit Hilfe der Reflektanzinformation des Amplitudenbildes werden zunächst die Bildpunkte ausgeblendet, für welche eine gültige Reflexion nicht messbar ist. Das Homogenitätskriterium führt dazu, Bereiche mit einer starken Tiefenvarianz in der Umgebung zu unterdrücken. Zur Segmentierung von Objekten wird die Annahme getroffen, dass ein Objekt im Tiefenhistogramm eine Histogrammklasse belegt und die zum Objekt gehörenden Bildpunkte räumlich nebeneinander liegen. Die getroffene Annahme ist allerdings nur für Objekte gültig, die koplanar parallel zur Sensorebene



Bild 3.3: Stereodaten werden durch die Stixel-Repräsentation zu Stiften zusammengefasst, welche auf der freien Fahrfläche aufsetzen [PF10].

stehen und keine longitudinale Ausdehnung besitzen. Zur Segmentierung lassen sich dann entweder Regionenwachstums-, Kontursuche- oder Pyramidenverkettungs-Verfahren anwenden.

Von Hu, Lamosa und Uchimura [HLU05] wird ein Ansatz für die Segmentierung von Stereodaten vorgeschlagen, bei welchem die Informationen des Tiefenbildes (in diesem Fall genauer der Disparitätskarte) einmal entlang der v -Bildkoordinate und einmal entlang der u -Bildkoordinate in ein Histogramm projiziert werden. Das Ergebnis wird als U/V -Disparitätenkarte bezeichnet. Ersteres entspricht dann einer Projektion der Tiefendaten auf die x/y -Ebene im Kamerakoordinatensystem, letzteres entspricht der Projektion der Tiefendaten auf die x/z -Ebene. In diesen U/V -Disparitätenkarten zeichnen sich Objekte, welche senkrecht auf der Fahrbahnoberfläche stehen, als Linien in der U -Disparitätenkarte ab und können durch Anwendung einer Linienextraktion detektiert werden. Gleichzeitig lässt sich aus der V -Disparitätenkarte die Neigung der Bodenebene extrahieren, was auch bei gekrümmten Fahrbahnoberflächen eine Filterung der zur Fahrbahnfläche gehörenden Bildpunkte erlaubt.

Ein anderer Ansatz bei der Erkennung unbekannter Objekte wird durch die Stixel-Repräsentation verfolgt. Pfeiffer und Franke [PF10] beschreiben ein Verfahren, das für die Verarbeitung dichter Stereodaten geeignet ist. Dazu werden Bereiche im Tiefenbild zu Stiften zusammengefasst, die auf der Fahrbahn aufsetzen (siehe Abbildung 3.3). Zunächst wird die vor dem Fahrzeug liegende freie Fläche ermittelt und in jeder Bildspalte die untere Kante eines möglichen Objektes markiert. Anschließend wird die obere Begrenzung des Objektes durch die Modellannahme bestimmt, dass ein Objekt auf der Bodenebene steht und sich dessen Tiefendaten vom Hintergrund absetzen. Die resultierenden Stifte entsprechen den sogenannten statischen Stixeln und beschreiben Teile von Objekten. Ein Fahrzeug ist dabei aus mehreren Stixeln zusammengesetzt, ein Fußgänger aus entsprechend weniger Stixeln. Bevor die Stixel aber zu Objekten zusammengefasst werden, wird für jedes Stixel dessen räumliche Bewegung mittels Kalman Filter über die Zeit geschätzt. Erst daran anschließend werden räumlich nahe beieinander liegende Stixel mit gleicher oder ähnlicher Bewegung zu Objekten zusammengefasst. Dadurch ist es möglich, auch Objekte mit nicht rigider Form zu detektieren, wie zum

Beispiel LKWs mit Anhängern.

Weitere Arbeiten beruhen auf der Segmentierung von Velodyne-Punktwolken oder der Objektbeschreibung mittels Belegtheitsgitter [MS13; BGD10]. Auf eine nähere Beschreibung wird an dieser Stelle verzichtet. Zusammenfassend wird deutlich, dass die Erkennung unbekannter Objekte auf Basis von Tiefenbildern eine wichtige Erweiterung zur klassischen Detektion bekannter Objektarten darstellt. Die gewonnene Information über Objektausdehnung und -schwerpunkt erleichtert es, auf Basis eines komplementären Sensors eine Objektklassifikation und eine robuste temporale Objektverfolgung an die Erkennung anzuschließen.

3.2.2 Erkennung von Infrastruktur-Objekten

Die Erkennung von Infrastruktur-Objekten kann mit Hilfe von Videokameras oder tiefenbildgebenden Kameras durchgeführt werden. Zu den relevanten Infrastruktur-Objekten zählen auf Autobahnen im Wesentlichen Fahrspur-Informationen sowie Verkehrszeichen.

Die aktuellen Fahrspur-Begrenzungen lassen sich beispielsweise mit Hilfe des Canny-Bildoperators robust extrahieren [AZ03]. Diese Begrenzung liefert für den Spurhalteassistent (LKS) die notwendige Datengrundlage. Da Fahrspuren als Klothoiden modelliert werden, erzeugen die meisten Fahrspurschätzer entweder Geradenstücke oder Klothoidenmodelle. Zur temporalen Fusion wird dabei entweder das Kalman Filter oder Partikelfilter eingesetzt [Vac08; Gum+09].

Verkehrszeichen besitzen genormte Formen und verwenden Piktogramme, was deren Erkennung und Klassifikation begünstigt. Runde Geschwindigkeitszeichen lassen sich in Videodaten sehr gut mittels der Radialen Symmetrie detektieren, einer Variante der Hough-Transformation [BZF08].

Im innerstädtischen Bereich zählen des Weiteren Kreuzungen und Ampelanlagen zu relevanten Infrastruktur-Objekten. Als ergänzende Informationsquelle bietet sich der elektronische Horizont an, welcher aus digitalen Kartendaten und einem Ortungssystem wie beispielsweise dem Globalen Positionierungssystem (GPS) aufgebaut wird.

Die Verwendung der Information über Schildart und Position aus digitalen Karten als zusätzliche Informationsquelle ermöglicht eine weitere Steigerung der Zuverlässigkeit des Systems. Allerdings müssen Widersprüche zwischen der sensoruellen Erkennung und der digitalen Karte berücksichtigt werden. Diese aufzulösen kann beispielsweise mit Hilfe der Dempster-Shafer Fusion erfolgen [NGZ09].

3.3 Temporale Informationsfusion

Die temporale Informationsfusion ermöglicht es, für den Zustand eines Objektes eine Glaubensaussage zu treffen, indem das Objekt über einen längeren Zeitraum verfolgt wird und mit den Erwartungen eines Modells verglichen wird. Einflussfaktoren für die Glaubensaussage sind die im Modell angenommene Sicherheit über die Zustandsfortsetzung und der erwartete Fehler bei der Transformation des Systemzustandes in den Messzustand. Aus dem Vergleich lässt sich sowohl eine Prädiktion des Systemzustandes erreichen, sowie eine Aussage über die

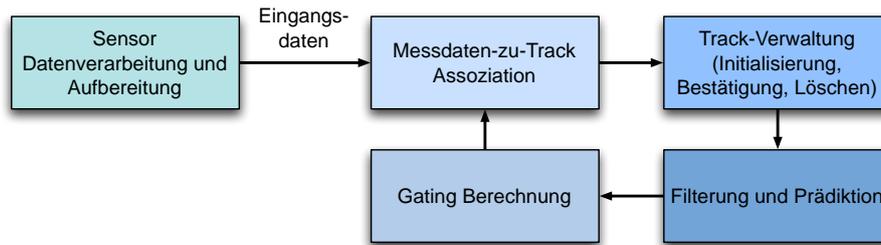


Bild 3.4: Komponenten und Ablauf eines *Multi-Target-Tracking* (MTT) Systems [BP99].

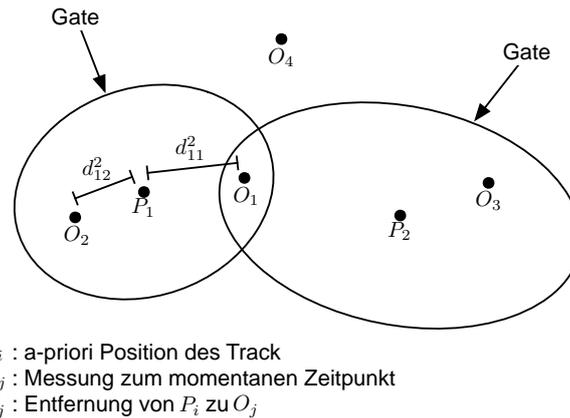


Bild 3.5: Gating zur Assoziation zwischen Messdatum und Track.

Wahrscheinlichkeit der Existenz des Systems treffen. Wie wichtig eine exakte Zustandsfortführung für ein Assistenzsystem ist, wird in [BF08] beschrieben.

Der Ablauf der temporalen Fusion folgt dem Prinzip eines konventionellen Multi-Target Tracking (MTT) Systems und ist in Abbildung 3.4 dargestellt. Nach Blackman und Popull [BP99] liegt diesem ein rekursives System zugrunde. In zurückliegenden Zeitschritten wurden bereits aus beobachteten Objekten Tracks erstellt. Im aktuellen Zeitschritt werden neue Objektmessungen als Eingangsdaten vom Sensor empfangen und der Tracking-Zyklus startet. Zunächst erfolgt eine Assoziation zwischen bereits vorhandenen Tracks und den aktuellen Messdaten. Wird für ein Messdatum keine gültige Zuordnung gefunden, kann daraus ein neuer Track erzeugt werden. Tracks mit einer zugeordneten Messung werden durch den Messwert aktualisiert. Die Häufigkeit dieser Aktualisierung bestimmt die Zuverlässigkeit des Tracks. Tracks mit geringer Zuverlässigkeit werden gelöscht, sobald sich den Tracks keine Messungen mehr zuordnen lassen. Schließlich wird der A-priori Zustand des Tracks für den Zeitpunkt der nächsten zu erwartenden Messung berechnet und für diesen ein sogenanntes Gate bestimmt.

3.3.1 Datenassoziation

Im Datenassoziationsschritt werden die aktuellen Messdaten den Tracks zugewiesen und gegebenenfalls neue Tracks erstellt beziehungsweise alte Tracks gelöscht. Die Datenassoziation umfasst:

- die Berechnung der Gates,

- die Bestimmung einer Zuordnung von Messdaten und Tracks, sowie
- die Track-Verwaltung.

Durch die Verwendung von Gates lässt sich effizient bestimmen, welche Messdaten für die Berechnung der Assoziation zu einem Track in Frage kommen, da sich durch Gates die Anzahl der möglichen Kandidaten stark reduziert lässt. Die Gates werden in der Regel im Zustandsraum des Objektes erstellt, aber auch symbolische Merkmale wie die Klasse eines Objektes können verwendet werden. Die Differenz zwischen jeder Messung und der korrespondierenden Schätzung wird mit Hilfe eines normalisierten Abstandsmaßes d_{ij}^2 bestimmt, wobei i den geschätzten Zustand des Tracks bei Eintreffen der Messung und j die Messung kennzeichnet. Für das Gate wird ein maximal zulässiger Fehler definiert, der zwischen Messung und Track auftreten darf. Ist der Wert des normalisierten Abstandes kleiner als der zulässige Fehler, entspricht die Messung einem gültigen Kandidaten für ein Track-Update.

Die genaue Zuordnung zwischen Messdaten und Tracks kann durch unterschiedliche Strategien erreicht werden. Die Global Nearest Neighbor (GNN)-Zuordnungsstrategie erreicht, dass für jeden Track möglichst eine Messung für das Track-Update zugeordnet wird, gleichzeitig darf eine Messung allerdings auch maximal einen Track aktualisieren.

Andere Strategien behandeln alle im Gate des Tracks liegenden Messungen als gültige Zuordnung. Im Update-Schritt wird dann auf Grundlage eines probabilistischen Ansatzes der Einfluss jeder Messung gewichtet. Geeignete Verfahren sind das PDA und das JPDA Verfahren. Beide Verfahren berechnen für jeden Kandidaten über alle Tracks diejenige Zuordnung, welche statistisch die höchste Güte besitzt. Während das PDA Verfahren die Annahme trifft, dass genau eine Messung einem Track zugehört und alle anderen Messungen Falsch-Positive sind, können durch das JPDA Verfahren mehrere Messungen zu einem Track-Update beitragen. Auch für nahe beieinander liegende Objekte können dadurch gültige Zuordnungen erstellt werden [BDH09].

3.3.2 Filterung des Systemzustandes

Nach Abschluss der Datenassoziation muss jeder Track durch die ihm zugeordnete Messung aktualisiert werden. Zunächst wird der letzte bekannte Systemzustand in den A-priori Zustand mit Hilfe eines Zustandsübergangsmodells überführt. Anschließend wird dieser Zustand mit Hilfe eines Messmodells in den Messraum transformiert und das Residual zwischen erwarteter Messung und tatsächlicher Messung bestimmt. Ungenauigkeiten im Zustandsübergangsmodell sowie im Messmodell lassen sich durch Rauschprozesse beschreiben.

Wenn für ein Objekt mehrere alternative Zustandshypothesen gleichzeitig verfolgt werden, spricht man von Multi-Hypothesen Tracking (MHT). Bei unsicheren oder widersprüchlichen Zuordnungskombinationen zwischen Messungen und Tracks werden für alle wahrscheinlichen Assoziationen entsprechende Hypothesen erzeugt. Im weiteren Verlauf wird kontinuierlich bewertet, ob eine Hypothese ein Objekt korrekt beschreibt oder ob die Hypothese falsch ist. Beim MHT müssen mittels Pruning unwahrscheinliche Hypothesen entfernt oder ähnliche Hypothesen fusioniert werden, damit die Komplexität und damit der Rechenaufwand beschränkt bleibt [Rei79].

3.3.3 Kalman Filter

Das Kalman Filter liefert eine optimale Prozessschätzung für Systeme, deren Prozesse und Messprinzip gaußschem Rauschen unterliegen [WB95]. Dabei ist es ausreichend, den mittleren Prozesszustand und die zugehörige Kovarianzmatrix zu schätzen um die Wahrscheinlichkeitsdichtefunktion des Prozesszustandes ermitteln zu können.

Ziel des Kalman Filters ist der Rückschluss des Prozesszustandes aus mit Fehlern behafteten Messungen. Durch das Kalman Filter kann ein Prozess rekursiv geschätzt werden, welcher durch eine stochastisch lineare Differenzfunktion beschrieben ist. Sind das Prozessrauschen sowie das Messrauschen korrekt modelliert, wird der *mittlere quadratische Fehler* (MSE) durch das Filter minimiert. Allgemein gilt dabei, dass ein Prozess durch ein Zustandsübergangsmodell, eine optionale Steuereingabe sowie ein Messmodell beschrieben ist:

- \mathbf{x}_k = Prozesszustand; n -dimensional zum Zeitpunkt k
- $\hat{\mathbf{x}}_k$ = Schätzung des Prozesszustands zum Zeitpunkt k
- \mathbf{y}_k = Messung; m -dimensional zum Zeitpunkt k
- \mathbf{u}_k = Steuereingabe; l -dimensional zum Zeitpunkt k
- P_k = Kovarianzmatrix; $n \times n$ zum Zeitpunkt k
- A = Zustandsübergangsmodell; $n \times n$ Matrix überführt den Zustand \mathbf{x}_{k-1} des vorherigen Zeitschrittes $k - 1$ in den Zustand zum aktuellen Zeitpunkt k
- H = Messübergangsmodell; $m \times n$ Matrix überführt den Zustand \mathbf{x}_k in die Messung \mathbf{y}_k
- B = Steuermatrix; $n \times l$ Matrix, welche \mathbf{u}_{k-1} in \mathbf{x}_k überführt
- Q = Kovarianz des Prozessrauschens
- R = Kovarianz des Messrauschens

Der zeitdiskrete Prozess des Zustandsübergangs ist gegeben durch die Systemgleichung

$$\mathbf{x}_k = A \cdot \mathbf{x}_{k-1} + B \cdot \mathbf{u}_{k-1} + \mathbf{w}_{k-1}, \quad (3.4)$$

die Messgleichung ist gegeben durch

$$\mathbf{y}_k = H \cdot \mathbf{x}_k + \mathbf{v}_k, \quad (3.5)$$

wobei die Unsicherheiten im Prozess durch zwei voneinander unabhängigen Zufallsvariablen \mathbf{w}_k und \mathbf{v}_k beschrieben werden. Sie sind über die folgende Normalverteilungen definiert:

$$p(\mathbf{w}) \sim \mathcal{N}(0, Q), \quad (3.6)$$

$$p(\mathbf{v}) \sim \mathcal{N}(0, R). \quad (3.7)$$

Auf Grundlage des Zustandsübergangsmodells A lässt sich der Zustand zum Zeitpunkt \mathbf{x}_{k-1} in den aktuellen Zustand \mathbf{x}_k überführen. Die aus dem Prozesszustand erwartete Messung lässt sich durch das Messübergangsmodell H beschreiben. Aus dem ermittelten Fehler zwischen erwarteter und tatsächlicher Messung lässt sich folglich eine Aussage über den Prozesszustand ableiten.

Mit Hilfe zweier sich zyklisch wiederholender Schritte wird nun der Prozesszustand geschätzt. Der erste Schritt beschreibt zunächst die Vorhersage des Prozesszustandes zum Zeit-

punkt der Messung:

$$\hat{\mathbf{x}}_{k|k-1} = A \cdot \hat{\mathbf{x}}_{k-1} + B \cdot \mathbf{u}_{k-1} \quad (3.8)$$

$$P_{k|k-1} = A \cdot P_{k-1} \cdot A^T + Q. \quad (3.9)$$

Der daran anschließende zweite Schritt korrigiert den geschätzten Prozesszustand durch die Messung:

$$K_k = P_{k|k-1} \cdot H^T (H \cdot P_{k|k-1} \cdot H^T + R)^{-1} \quad (3.10)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + K_k (\mathbf{y}_k - H \cdot \hat{\mathbf{x}}_{k|k-1}) \quad (3.11)$$

$$P_k = (I - K_k \cdot H) P_{k|k-1}, \quad (3.12)$$

wobei K_k das Kalman Gain bezeichnet, welches den Einfluss der Messung auf den geschätzten Prozesszustand bestimmt. Der Vorteil des Kalman Filters beim Einsatz als MTT ist die automatische Bestimmung des Kalman Gain, da dieses sich in Abhängigkeit des Prozess- und Messrauschens selbst einstellt.

3.3.4 Interacting Multiple Model Filter

Zur Berücksichtigung unterschiedlicher Objekt-Bewegungen lässt sich der Systemzustand eines Objektes auch über mehrere Modelle bewerten, wobei zu jedem Zeitschritt alle Modelle untereinander in Bezug auf den Prozess bewertet werden. Durch Mischung der Modellergebnisse lassen sich auch schwierig zu beschreibende Prozesse oder Prozesse mit wechselnden Systemzuständen beschreiben (beispielsweise ein Fahrzeug, welches bei zunächst konstanter Geschwindigkeit ein Bremsmanöver durchführt). Eine gängige Erweiterung ist die Schätzung eines Prozesses mit mehreren parallel arbeitenden Filtern. Eine solche Erweiterung ist das Interacting Multiple Model (IMM) Filter.

Das IMM Filter erweitert den klassischen Kalman Filter Ansatz dahingehend, dass Prozesszustände und Kovarianzmatrizen für mehrere Modelle entsprechend einem Markov Ansatz kombiniert werden. Die Anzahl der gleichzeitig zu verwendenden Modelle ist dabei nicht beschränkt, sondern kann problemspezifisch gewählt werden [BP99].

Es wird angenommen, dass zu jedem Zeitschritt eine Wahrscheinlichkeit p_{ij} existiert, zu dem ein Prozess von einem Modellzustand i in einen anderen Zustand j wechselt. Die Übergangswahrscheinlichkeit P_T wird anhand von Prozesswissen definiert. Mit den Filterschritten (3.8) bis (3.12) des Kalman Filters gilt für jedes Modell $i \in 1 \dots N$:

$\hat{\mathbf{x}}_{i,k-1}$ = Zustandsschätzung zum Zeitpunkt $k - 1$ des Filters i

$P_{i,k-1}$ = Kovarianzmatrix zum Zeitpunkt $k - 1$ des Filters i

$\mu_{i,k-1}$ = Wahrscheinlichkeit, dass sich der Prozess in Modellzustand i befindet, gegeben $k - 1$

$\mu_{ij,k-1}$ = Bedingte Wahrscheinlichkeit, dass der Prozess von Zustand i in Zustand j zum Zeitpunkt $k - 1$ gewechselt hat

Abbildung 3.6 zeigt den Ablauf des IMM Filters. Im ersten Schritt erfolgt für alle Modelle eine Modell-Interaktion mit den aktuellen Modellwahrscheinlichkeiten $\mu_{i,k-1}$. Die Zustandsschätzung $\hat{\mathbf{x}}_{i,k-1}$ und die Kovarianzmatrix $P_{i,k-1}$ sind für jedes Modell durch dessen vorherigen

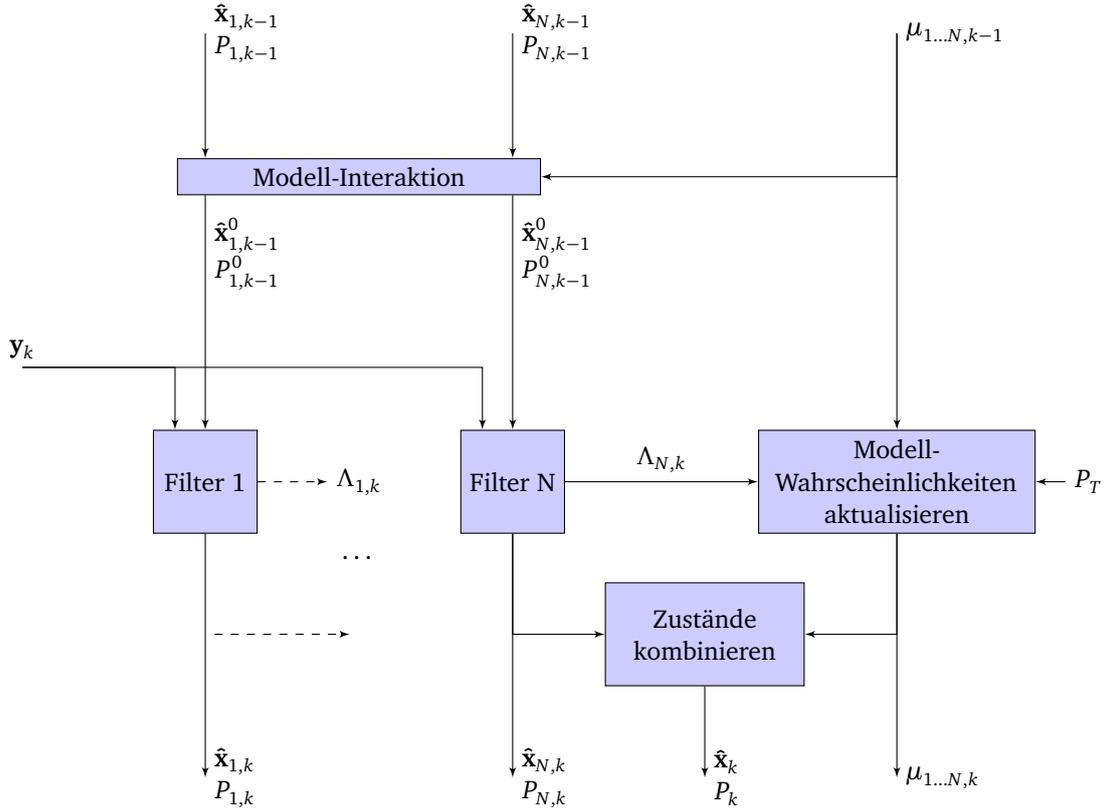


Bild 3.6: Blockdiagramm des IMM Filters, nach Jordan [Jor06].

Zustand gegeben. Die Modell-Interaktion erzeugt nun auf der Annahme eines Prozessübergangs in den Zustand j neue Filterzustände $\hat{\mathbf{x}}_{j,k-1}^0$ und Kovarianzmatrizen $P_{j,k-1}^0$.

Für jedes Modell $j = 1 \dots N$ wird ein leicht modifiziertes Kalman Filter ausgeführt. Zunächst wird dabei der Zustand zum Zeitpunkt der nächsten Messung k auf Basis der Mischzustände und Kovarianzen gebildet. Darauf folgt die Messdaten-Assoziation. Damit diese gültige Ergebnisse für alle Modelle erreicht, wird eine fusionierte Prozessprädiktion $\hat{\mathbf{x}}_{k|k-1}$ und Kovarianzmatrix $P_{k|k-1}$ berechnet. Anschließend wird im herkömmlichen Sinne die Korrektur der modellspezifischen Prädiktion anhand der zugeordneten Messdaten vorgenommen, sowie das Modell mit der höchsten Wahrscheinlichkeit ausgewählt. Die Zustandsschätzung $\hat{\mathbf{x}}_k$ wird schließlich aus der Kombination der einzelnen Modellzustände und der entsprechenden Modellwahrscheinlichkeiten gebildet. Die einzelnen Filterschritte des IMM Filters sind wie folgt gegeben:

1. Zunächst werden die Mischwahrscheinlichkeiten $\mu_{ij,k-1}$ für alle Modelle $i, j = 1 \dots N$ berechnet, basierend auf der Übergangsmatrix P_T und den Modellwahrscheinlichkeiten $\mu_{i,k-1}$:

$$\mu_{ij,k-1} = \frac{1}{\bar{c}_j} \cdot p_{ij} \cdot \mu_{i,k-1} \quad (3.13)$$

$$\bar{c}_j = \sum_{i=1}^N p_{ij} \cdot \mu_{i,k-1}. \quad (3.14)$$

2. Anschließend werden die Modelle anhand der berechneten bedingten Wahrscheinlichkeiten gemischt. Dies erfolgt auf der Annahme, dass der Prozess von Zustand i in Zustand j zum Zeitpunkt $k - 1$ gewechselt ist, für alle Modelle $j = 1 \dots N$:

$$\hat{\mathbf{x}}_{j,k-1}^0 = \sum_{i=1}^N \mu_{ij,k-1} \cdot \hat{\mathbf{x}}_{i,k-1} \quad (3.15)$$

$$P_{j,k-1}^0 = \sum_{i=1}^N \mu_{ij,k-1} \cdot \left\{ P_{i,k-1} + \left[\hat{\mathbf{x}}_{i,k-1} - \hat{\mathbf{x}}_{j,k-1}^0 \right] \cdot \left[\hat{\mathbf{x}}_{i,k-1} - \hat{\mathbf{x}}_{j,k-1}^0 \right]^T \right\}. \quad (3.16)$$

3. Für die Berechnung der Gates und zur Assoziation von Messungen wird eine gemischte Zustandsschätzung und Kovarianzmatrix berechnet:

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{j=1}^N \bar{c}_j \cdot \hat{\mathbf{x}}_{j,k|k-1} \quad (3.17)$$

$$P_{k|k-1} = \sum_{j=1}^N \bar{c}_j \cdot P_{j,k|k-1}. \quad (3.18)$$

4. Mittels den bekannten Kalman Gleichungen wird für jedes Modell $j = 1 \dots N$ die Zustandsschätzung $\hat{\mathbf{x}}_{j,k}$ sowie die Kovarianzmatrix $P_{j,k}$ berechnet. Die gemeinsame Messung \mathbf{y}_k wird für alle Modelle im Korrekturschritt des Kalman Filters verwendet.
5. Anschließend erfolgt die Bestimmung der Likelihood-Funktion für jedes Modell $j = 1 \dots N$. Für jedes Modell j wird die Residualkovarianzmatrix $S_{j,k}$ und das Distanzmaß $d_{j,k}^2$ bestimmt. Unter Annahme der Dimension des Messvektors M ist die Likelihood-Funktion für die Messung:

$$\Lambda_{j,k} = \frac{1}{\sqrt{(2\pi)^M \cdot \det(S_{j,k})}} \cdot e^{-d_{j,k}^2/2}. \quad (3.19)$$

Der statistische Abstand $d_{j,k}^2$ zwischen Messung und Prozesszustand ist gegeben durch:

$$d_{j,k}^2 = \tilde{\mathbf{y}}_{j,k}^T \cdot S_{j,k}^{-1} \cdot \tilde{\mathbf{y}}_{j,k}, \quad (3.20)$$

wobei $\tilde{\mathbf{y}}_{j,k}$ den Differenzvektor zwischen aktueller und erwarteter Beobachtung für das Modell j und $S_{j,k}$ die entsprechende Kovarianzmatrix bezeichnet:

$$\tilde{\mathbf{y}}_{j,k} = \mathbf{y}_k - H \cdot \hat{\mathbf{x}}_{j,k|k-1} \quad (3.21)$$

$$S_{j,k} = H \cdot P_{j,k|k-1} \cdot H^T + R. \quad (3.22)$$

6. Anhand der Likelihood-Funktionen werden die Modellwahrscheinlichkeiten aktualisiert:

$$\mu_{j,k} = \frac{1}{c} \cdot \Lambda_{j,k} \cdot \bar{c}_j \quad (3.23)$$

$$c = \sum_{j=1}^N \Lambda_{j,k} \cdot \bar{c}_j. \quad (3.24)$$

7. Im letzten Schritt werden die unterschiedlichen Zustände der Modelle miteinander kombiniert, um einen resultierenden Zustand für den Zeitpunkt k zu bilden. Dieser stellt den Ausgang des IMM Filters dar:

$$\hat{\mathbf{x}}_k = \sum_{j=1}^N \hat{\mathbf{x}}_{j,k} \cdot \mu_{j,k} \quad (3.25)$$

$$P_k = \sum_{j=1}^N \mu_{j,k} \cdot \left\{ P_{j,k} + [\hat{\mathbf{x}}_{j,k} - \hat{\mathbf{x}}_k] \cdot [\hat{\mathbf{x}}_{j,k} - \hat{\mathbf{x}}_k]^T \right\}. \quad (3.26)$$

3.4 Situationsanalyse

Im Forschungsfeld der *Intelligenten Fahrzeuge* wird der Bestimmung der Fahrsituation eine hohe Bedeutung zugemessen, da dies ausschlaggebend für die Angemessenheit einer Aktion ist, also beispielsweise dem Setzen einer Meldung an den Fahrer oder die automatische Betätigung der Bremse zur Einleitung einer Notbremsung. Bei konventionellen Systemen ist die Verarbeitungskette zwischen Wahrnehmung und Situationsanalyse (SA) üblicherweise sehr kurz. In den meisten Fällen besteht diese aus einem Sensor und einer direkt auf dessen Signalverarbeitung aufbauenden Analyse zur Umsetzung einer Assistenzfunktion. Dabei werden gezielt weitere Informationen ausgeschlossen, um die Systemkomplexität und den mit der Entwicklung einer Assistenzfunktion entstehenden Testaufwand (und somit die Entwicklungskosten) gering zu halten. Bei diesen Systemen kann auch ohne den Einsatz von komplexen Inferenzmechanismen eine spezifische SA erfolgen. In Abschnitt 3.4.1 wird dies näher beschrieben.

Eine umfassende, vollständige Betrachtung der aktuellen Situation kann dann gelingen, wenn das Wissen der verschiedenen Komponenten zur Umfelderkennung, also der Wahrnehmung der Infrastruktur, der dynamischen Objekte und des elektronischen Horizonts, in einer gemeinsamen Wissensbasis zusammengetragen wird und die Wissensbasis in einer maschinell verständlichen Form aufgebaut ist. Dazu muss das Wissen also einerseits formal repräsentiert werden, damit es maschinell verarbeitet werden kann. Andererseits muss eine Anbindung an die verschiedenen Wahrnehmungskomponenten existieren, damit die formale Repräsentation mit der Veränderung der Situation mitgeführt werden kann. Dann kann aus der formalen Repräsentation ein graphisches Modell abgeleitet und zur Analyse der Fahrsituation herangezogen werden. Die Art der Repräsentation von Wissen und der Mechanismus zum Schließen von Information hängt eng miteinander zusammen, weshalb bei der Analyse der Verfahren im folgenden Abschnitt stets Repräsentation und SA gemeinsam betrachtet werden. Eine Beschreibung solcher Systeme ist in den Abschnitten 3.4.2 und 3.4.3 gegeben.

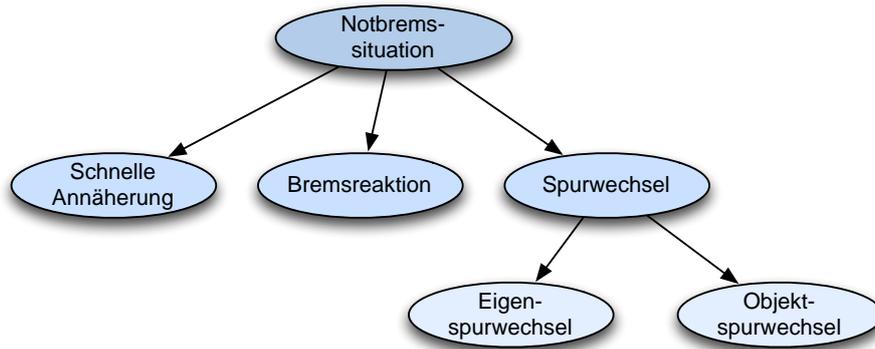


Bild 3.7: Situationsteilbaum zur Identifikation einer Notbrems-situation [Sch09].

3.4.1 Situationsanalyse ohne logische Ausdrücke

Auch in konventionellen Assistenzsystemen ist die Durchführung einer SA von großer Bedeutung. Diese erfolgt direkt gekoppelt an eine Umfeldmessung und kann somit als Überprüfung der Relevanz einer Messung verstanden werden.

Situationserkennung mit probabilistischen Netzen

Eine Analyse von Verfahren zur Situationsanalyse für FAS erfolgt durch Schneider [Sch09]. In der Arbeit wird ein generischer und probabilistischer Ansatz zur Erkennung relevanter Fahr-situationen und Fahrmanöver vorgestellt. Schwerpunkt der Arbeit ist die Aggregation von Informationen verschiedener Sensoren im Fahrzeug in einer einheitlichen Situationserkennung, um daran anschließend unterschiedliche FAS mit Informationen zu versorgen.

Hierzu werden aus den Informationen der Fahrumgebung und der Eigenfahrzeugbeschreibung wahrscheinlichkeitsbehaftete Situationen und Manöver extrahiert. Die Situationsmerkmale sind mittels Regeln formuliert, welche allerdings keiner speziellen formalen Sprache folgen, sondern umgangssprachlich formuliert sind, wie beispielsweise „Befindet sich im Verkehrsraum vor mir ein Fahrzeug“. Diese Regeln werden als Situationsknoten in einem graphischen Modell verknüpft, ähnlich der in [DBB02] vorgestellten Modelle (siehe Abbildung 3.7). Die Sensorinformationen werden durch eine unscharfe Zugehörigkeitsfunktion¹ in eine symbolische Repräsentation übertragen und dienen als Evidenzen im graphischen Modell. Vorteil des Verfahrens ist, dass die Unsicherheiten der Umfelderkennung durch das graphische Modell mittels der Gauss'schen Fehlerfortpflanzung bis zur Situationsentscheidung fortgeführt werden, sodass Assistenzfunktionen auf diese Information zurückgreifen können. Zusätzlich lässt sich die Relevanz von Merkmalen spezifisch beeinflussen, indem jedes Merkmal individuell gewichtet werden kann. Der gewählte Ansatz ist dadurch robust bezüglich den Attributen eines Objektes in der Fahrumgebung.

Das graphische Modell dient zur Fusion der Merkmale und zur Bestimmung der Situation. Mit Hilfe von Inferenzmechanismen lässt sich eine Situations- und Manöverentscheidung herleiten. Zur Behandlung komplexer Situationen lässt sich das graphische Modell aus bestehenden Situationsteilbäumen kaskadieren, der Netzaufbau ist dadurch modular orientiert.

¹sogenannte Fuzzy-Funktionen

Bei der Erzeugung der Situationsteilbäume durch den Domänenexperten wird die kausale Eigenschaft von Bayes'schen Netzen (BNs) ausgenutzt, sodass natürlich verständliche Netze entstehen. Das probabilistische Modell wird durch bedingte Wahrscheinlichkeitsverteilungen (CPDs) dargestellt. Der gewählte methodische Ansatz basiert auf starren BNs. Diese Modellierung verhindert ein robustes Verhalten gegenüber Veränderungen der Anzahl von Objekten und Relationen in der Fahrumgebung.

Numerische Situationsanalyse anhand der Time-To-React

Die Bestimmung der Kollisionszeit (TTC) ist ein wichtiges Merkmal zur Bestimmung der Kritizität einer Situation. Wird eine gewisse Zeit bis zu einer erwarteten Kollision unterschritten, so lösen aktuelle Assistenzsysteme zunächst eine Warnung an den Fahrer aus, gefolgt von einer Vorkonditionierung des Fahrzeuges auf eine mögliche Notbremsung. Unterbleibt eine Aktion des Fahrers, kann das Fahrzeug im Extremfall selbst eine Notbremsung einleiten, um die Folgen bei einem Unfall zu mindern. Man spricht dabei von einem Kollisionsminderungssystem (CMS).

In der Arbeit von Hillenbrand, Spieker und Kroschel [HSK06] wird dargestellt, wie auf Grundlage der TTC und ähnlicher Merkmale wie der Reaktionsreservezeit (TTR) eine SA zur Entscheidungsfindung unter Berücksichtigung von Sensor- und Prädiktionsunsicherheiten in CMS erreicht werden kann. Die SA erfolgt dabei analytisch gegenüber den Steuermöglichkeiten des Fahrers, also des Lenkradeinschlags, der Brems- und Gaspedalposition, woraus die minimalen Ausweichtrajektorien nach links beziehungsweise rechts sowie die maximale Verzögerung respektive die maximale Beschleunigung resultieren. Ein vergleichbarer Ansatz wird in [BCS10] verfolgt.

Das eigene Fahrzeug sowie ein Hindernis auf offener Fläche werden in einem kontinuierlichen Zustandsraum beobachtet. Durch numerische Analyse lässt sich daraus die maximale Reservezeit TTR entweder zum Ausweichen oder zum Bremsen bestimmen, welche als Kritizitätsmaß in die Situationsbewertung einfließt. Dabei wird die Schätzunsicherheit der Objektverfolgung bei der Bestimmung einer minimalen und maximalen TTR berücksichtigt. In Abhängigkeit der mittleren TTR wird eine mögliche Risikostufe durch Schwellwerte bestimmt, wobei jeder Stufe eine feste maximale Bremsverzögerung zugewiesen ist. Die Zuverlässigkeit des Systems wurde über Monte Carlo Simulationsläufe anhand verschiedener Situationsbeispiele nachgewiesen. Aufgrund der numerischen Modellierung der Methode zeigt der Ansatz ein robustes Verhalten gegenüber Attributveränderungen des eigenen Fahrzeuges oder des Hindernisses. Eine Übertragbarkeit auf Situationen mit mehreren Fahrzeugen oder Einschränkungen durch die Verkehrsinfrastruktur im Sinne eines generischen Ansatzes erscheint nicht-trivial.

Durchgängig probabilistische Objekt- und Situationsanalyse

Eine Herausforderung bei Systemen mit Berücksichtigung von Unsicherheit ist die durchgängige Verwendung der probabilistischen Attribute. Ein methodischer Ansatz dafür wird in der Arbeit von Schubert und Wanielik [SW11] dargestellt. Darin wird das Konzept der ad-

aptiven Wahrscheinlichkeitsknoten präsentiert, durch welche sich die Unsicherheit bei der Zustandsschätzung an die SA durchreichen lässt.

Sowohl die Objektverfolgung wie auch die Fahrspurerkennung stellt in jedem Zeitschritt jeweils die momentane Zustandsschätzung als Normalverteilung des Systemzustandes sowie der Schätzfehlerkovarianzmatrix zur Verfügung. Gegenstand der Arbeit ist die Formulierung einer gemeinsamen Beschreibung der Unsicherheit der Objektbildung und der SA. Dies ist durch die Annahme motiviert, dass die meisten Verfahren zur Bildung von Objekten Wahrscheinlichkeitsdichtefunktionen aufstellen, diese aber bei der Überführung in die probabilistischen Modelle zur SA häufig nicht vollständig berücksichtigt werden.

Die Überführung der kontinuierlichen Systemzustände in diskrete Zufallsvariablen unter Berücksichtigung der Wahrscheinlichkeitsdichtefunktionen erfolgt durch sogenannte „soft evidences“. Kernaspekt der Arbeit ist, dass die Wahrscheinlichkeitsverteilung einer Variablen aus einem Intervall des Systemzustandes abgeleitet werden kann. Die Berücksichtigung der Unsicherheit erfordert, dass der Wert der Zufallsvariable außerhalb der Intervallgrenzen nicht Null sein darf. Unterliegt ein Wert einer Varianz und liegt der Zustand außerhalb des Intervalls, so ist dessen Wahrscheinlichkeit durch Zuweisung einer Gaussfunktion gegeben, die nach dem Abstand zur Intervallgrenze sowie der Varianz gewichtet ist. Bei unendlich hoher Varianz der Systemzustände ergibt sich eine Gleichverteilung der diskreten Zufallsvariablen.

Zunächst wird die *Normalisierte Laterale Position* (NLP) jedes Objektes gegenüber den erkannten Fahrspuren gebildet. In der Varianz der NLP wird dabei sowohl die Kovarianz des Objektzustandes sowie die Kovarianz des Fahrspurzustandes berücksichtigt. Daraus lässt sich die Wahrscheinlichkeit bestimmen, mit welcher sich ein Objekt auf einer Fahrspur befindet. Auf ähnliche Weise wird die Kollisionsgefahr zwischen einem Objekt und dem eigenen Fahrzeug mittels der *Deceleration To Safety Time* (DST) bestimmt. Auch hier werden die Kovarianzen der Objektverfolgung berücksichtigt. Somit lässt sich die Wahrscheinlichkeit bestimmen, ob eine gefährliche Situation gegenüber der DST zu einem Objekt vorliegt.

In der Arbeit wird gezeigt, wie die Ausgaben einer Objektverfolgung und eines Fahrspurerkennungssystems bei vollständiger Berücksichtigung der Unsicherheit an ein BN angebunden werden können. Ausgabe des Systems ist die Überwachung der lateralen Manöver, also ob ein Spurwechsel in der aktuellen Situation empfehlenswert ist. Ein vergleichbarer Ansatz wird durch Kasper [Kas13] aufgezeigt. Darin werden Objektorientierte Bayes'sche Netze (OOBNs) zur Modellierung der *Fahrzeuge* und *Fahrspuren* sowie der Relationen *Fahrzeug-Fahrzeug* und *Fahrzeug-Fahrspur* eingesetzt.

Zur SA werden die Werte der Erkennungskomponenten in diskrete Zufallsvariablen der Modelle überführt, wobei die Modelle starr strukturiert sind. Die gewählten Methoden sind für die spezifisch erstellten Relationen der Fahrsituation robust. Die Arbeiten lassen aber offen, wie eine generische Beschreibung von Situationen und eine automatische Konstruktion der Modelle erfolgen kann (vergleiche Abschnitt 3.5).

3.4.2 Ontologiebasierte Ansätze

Ontologien stellen eine gemeinsame Repräsentationsform für Wissen aus verschiedenen Anwendungen dar. Die Verwendung von Ontologie zur Beschreibung von Hintergrundwissen ist

ein häufig verwendetes Werkzeug im Bereich der semantischen Datenanalyse, zum Beispiel im Forschungsfeld des „Semantic Web“. Im Kontext des Forschungsfelds der „Künstlichen Intelligenz“ ist eine Ontologie definiert als eine explizite, formale Repräsentation des Wissens über eine Domäne. Eine Ontologie repräsentiert dazu die Objekte einer Domäne in Form von Konzepten, sowie zusätzlich die Eigenschaften, die diese Objekte besitzen, die Inter-Objekt Relationen, die Rollen und die Ereignisse und Prozesse, in welchen Objekte untereinander agieren. Durch eine Ontologie wird ein gemeinsames Verständnis über eine Menge von Daten spezifiziert.

Eine gemeinsame Sprachbeschreibung ist durch die Web Ontology Language (OWL) gegeben. Die formale Repräsentation beruht dabei auf Beschreibungslogik (DL), welche eine Untermenge der Logik erster Ordnung (FOL) ist. DL beschreibt die relevanten Konzepte einer Domäne (die sogenannte Terminologie) und verwendet diese Konzepte zur Spezifikation der Eigenschaften von Objekten, welche in der Domäne vorkommen (sogenannte Zusicherungen). Für DL existieren effiziente Mechanismen zum Prüfen und Schließen der spezifizierten Wissensbasis:

- Subsumption, das Schließen von konkretem Wissen auf allgemeines Wissen;
- Satisfiability, das Prüfen auf Konsistenz der definierten Konzepte;
- Equivalence, das Prüfen auf Gleichheit verschiedener Konzepte derselben Instanzen.

In DL ist es außerdem möglich, Wissen über nicht spezifizierte Objekte als *unbekannt* zu modellieren.

Für DL-Systeme existieren bislang keine effizienten Methoden zur Berücksichtigung von Unsicherheiten. Die Inferenzmechanismen für Ansätze wie PR-OWL auf Basis von Multi Entity Bayesian Networks (MEBN) besitzen einen exponentiellen Aufwand, wodurch sich probabilistische Ontologien bislang nicht für die Verwendung zur Situationsbeschreibung in FAS eignen [LCJ08].

Fallbasiertes Schließen zur Situationsanalyse

Regelbasierte Systeme zur Beschreibung von Fahrsituationen codieren Wissen in den verwendeten Regeln. Diese Regeln lassen sich nicht allgemein schließen oder automatisch erweitern, sondern müssen statt dessen vom Entwickler eines Systems spezifiziert werden. Die Interpretation von Fahrsituationen mittels fallbasiertem Schließen (CBR) stellt eine günstige Möglichkeit dar, automatisch Wissen aus Beobachtungen zu gewinnen und neues Wissen aus ähnlichen Situationen abzuleiten [Vac08].

Objekte und Relationen einer Fahrsituation werden zunächst mittels OWL-Konzepten beschrieben. Die Situationsbeschreibung entspricht einem Szenengraphen, welcher die Szene durch Instanzen der in der Ontologie definierten Konzepte und Attribute repräsentiert. Aus der Situationsbeschreibung und einer Situationsbewertung wird ein Fall gebildet. Ein ähnlicher Ansatz wird in [Gra+13] verfolgt.

Das Hintergrundwissen enthält eine Menge bekannter Fälle, welche durch den Entwickler vorgegeben oder im Laufe der Zeit durch das System eingelernt werden. Der CBR-Zyklus besteht aus vier Schritten, wie in Abbildung 3.8 gezeigt ist. Zunächst wird die aktuelle Situation

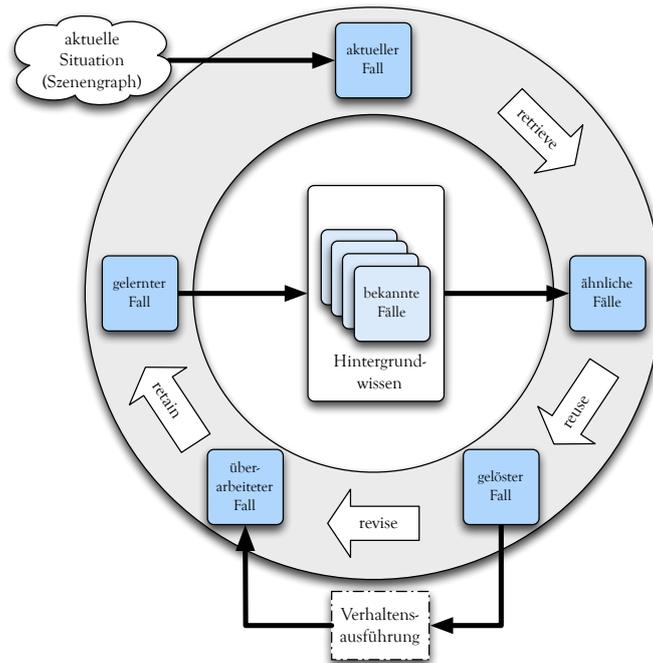


Bild 3.8: Vacek [Vac08] setzt fallbasiertes Schließen zur Situationsanalyse ein; der Ablauf des Verfahrens besteht aus vier Schritten: *retrieve*, *reuse*, *revise* und *retain*.

anhand der Ontologie in einen Fall konvertiert und verwandte Fälle aus dem Hintergrundwissen mit Hilfe einer Ähnlichkeitsfunktion ausgewählt. Die Lösung des ähnlichsten Falles wird als initialer Lösungsversuch zur Verhaltensausführung herangezogen. Nach Bedarf wird der ähnliche Fall an den vorgegebenen Fall angepasst und schließlich wieder als Hintergrundwissen abgespeichert. Durch die Mechanismen der DL ist es außerdem möglich, aus mehreren spezialisierten Fällen allgemeine Fälle abzuleiten.

Die Herausforderung besteht in der Beschreibung der Ähnlichkeitsfunktion, welche Fälle miteinander vergleicht. Die Beschreibung der Situation in Form des Szenengraphs erfolgt deterministisch. Unsicherheiten über die Attribute der Objekte werden nicht verwendet. Dies führt im Laufe der Zeit zu einer großen Anzahl gespeicherter Fälle. Eine mögliche Lösungsmethodik bietet das Case-Based Approximate Reasoning Verfahren.

Ontologiebasiertes Schließen von Verkehrsregeln

Forschungsgegenstand der Arbeit von Hülsen, Zöllner und Weiss [HZW11] ist der Aufbau eines Situationsverständnisses im Straßenverkehr. Hierfür werden Ontologien zur Beschreibung von Straßenkreuzungen genutzt und ein Mechanismus zum Schließen von Verkehrsregeln für Kreuzungen implementiert. Die Ontologie ist in Form logischer Regeln spezifiziert. Inkonsistenzen in der definierten Wissensbasis können durch die Methodik automatisch festgestellt werden. Der Aufbau der Ontologie erfolgt modulbasiert, sodass diese bei Bedarf erweiterbar ist. Die Arbeit zeigt anschaulich, dass dadurch auch komplexe Straßennetze dargestellt werden können.

Durch die Inferenzmechanismen der DL lassen sich insbesondere spezifizierte Kreuzungstopologien und darauf befindliche Verkehrsteilnehmer auf Regelkonformität prüfen beziehungs-

weise Regelverletzungen feststellen. Gleichzeitig können Relationen zwischen Elementen der Szene automatisch abgeleitet werden, wie zum Beispiel Vorfahrt-gewähren oder Ist-auf-Relationen.

Ontologien erlauben es, Fragen über konkrete Instanzen zu stellen um explizite Informationen zu generieren. So ist es beispielsweise möglich, eine Liste aller Fahrzeug-Tupel anzufragen, für welche die Vorfahrt-gewähren Relation gilt. Das Schließen von Information ist allerdings sehr aufwändig, da die Zusicherungsaspekte der Ontologie gegenüber allen Objekten ausgewertet werden müssen. Durch die Einschränkung, Unsicherheiten in der Ontologie nicht berücksichtigen zu können, muss die Ontologie durch Sonderfälle erweitert werden. Anderenfalls treten Inkonsistenzen auf, welche keine Aussage über die Situation zulassen.

3.4.3 Wissensbasierte Ansätze

Auch die wissensbasierten Ansätze setzen Ontologien zur Beschreibung der relevanten Objekte der Domäne ein. Allerdings erfolgt das Schließen von Information nicht direkt auf der definierten Ontologie (zum Beispiel durch DL), sondern, statt dessen auf logischen Konstrukten oder erzeugten probabilistischen Modellen, welche nicht direkt mit der Ontologie verknüpft sein müssen.

Die SA mit Logik erster Ordnung (FOL) erlaubt es, logische Schlüsse aus formalen Beschreibung von Hintergrundwissen zu ziehen. Im Gegensatz zur DL, welche gemeinsam in den oben vorgestellten Ansätzen verwendet wird, kann Wissen durch FOL in allgemeinerer Form spezifiziert werden. Gängig sind dabei regelbasierte oder Entity-Relationship (ER)-Verfahren, wie diese in Abschnitt 2.4.1 vorgestellt worden sind.

Im Folgenden sind Beispiele genannt, welche die FOL zur SA nutzen. Die Methodik der Risiko-Schablonen beschreibt Fahrsituationen in Form prädikatenlogischer Regeln. Die vorherrschende Situation ist in Form einer Wissensbasis abgelegt und wird durch Inferenz mit den Risiko-Schablonen in eine Risiko-Bewertung überführt. Die Methodik der wissensgetriebenen Situationsanalyse überführt Objekte und Relationen in prädikatenlogische ER-Strukturen, die zum Aufbau graphischer Modelle genutzt werden. Mittels Inferenz erfolgt auf den erzeugten Modellen eine Situationsanalyse.

Risiko-Muster zur Situationsanalyse

Die in [Lat+05] beschriebene Arbeit führt eine Analyse von Fahrsituationen durch und ordnet Objekten einen Risiko-Wert zu. Die eingesetzte Methode verwendet hierfür eine Wissensbasis und sogenannte Risiko-Muster zur Verknüpfung von Objektkonstellationen mit Risiko-Werten.

Die Wissensbasis enthält eine qualitative Repräsentation der Objekte und Attribute. Die symbolischen Werte sind aus Sensorinformationen abgeleitet und entsprechen Merkmalen, welche zur Beschreibung von Fahrsituationen geeignet sind.

Risiko-Muster stellen spezifische Objektkonstellationen dar, welche mit einem Risiko-Wert verknüpft sind. Die Muster werden durch Regeln der FOL formalisiert. Zur Bestimmung der SA wird die aufgebaute Wissensbasis in einem Inferenzschritt mit den im Hintergrundwissen gespeicherten Objektkonstellationen verglichen (siehe Abbildung 3.9). Wird ein Muster in einer entsprechenden Situation erkannt, so wird jedem beteiligten Objekt der Risiko-Wert des

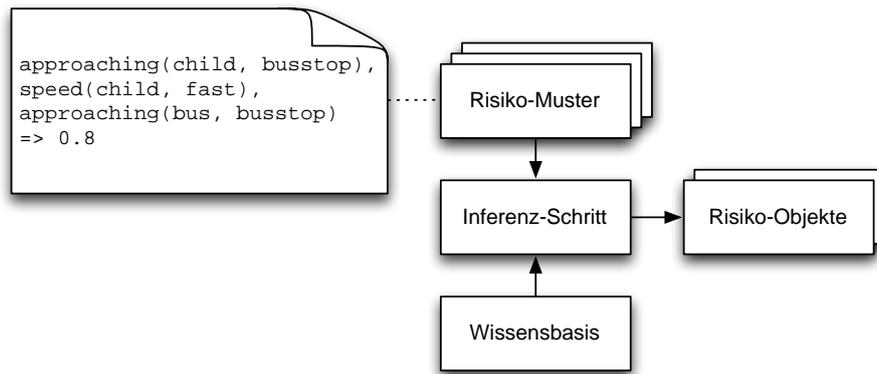


Bild 3.9: Risiko-Muster entsprechen Objektkonstellationen mit einem Risiko-Wert; In einem Inferenz-Schritt wird die Wissensbasis mit den Mustern verglichen und Risiko-Werte den Objekten zugeordnet.

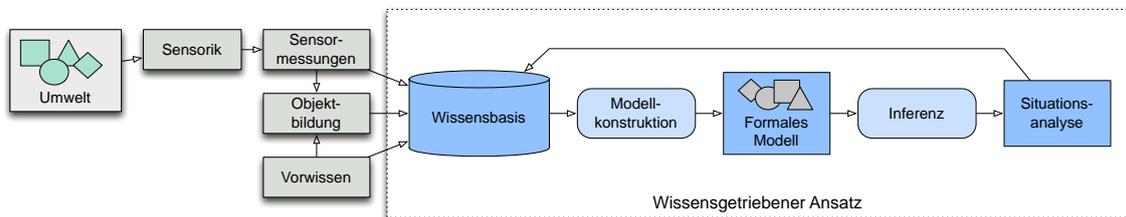


Bild 3.10: Wissensgetriebene Interpretation mittels Objekt-Attribut Schablonen [How10].

Musters zugewiesen. Objekte mit dem höchsten Risiko werden anschließend für den Benutzer durch das System hervorgehoben.

Die als Hintergrundwissen durch Regeln formalisierten Fahrsituationen berücksichtigen keine Unsicherheit. Dies kann insbesondere dazu führen, dass die Wissensbasis auf falsche oder gar keine Fahrsituationen abgebildet wird, falls spezifizierte Regeln nicht exakt zutreffen. Eine Erweiterung der spezifischen Objektkonstellationen um weitere Regeln ist nicht generisch möglich.

Wissensgetriebene Situationsanalyse

Eine Methode auf Basis der wahrscheinlichkeitsbasierten Logik erster Ordnung (FOPL) zur Situationsanalyse in taktisch-militärischen Domänen wird in der Arbeit von Howard [How10] beschrieben. Der Ansatz verwendet als formale Repräsentation die in Abschnitt 2.4.3 vorgestellte Objektorientierte Probabilistische Relationale Modellierungssprache (OPRML). Der konzeptionelle Ablauf der wissensgetriebenen Interpretation ist in Abbildung 3.10 gezeigt.

Zunächst werden mit Hilfe der Modellierungssprache durch einen Domänenexperten die für die Beschreibung der Domäne notwendigen Objektklassen und deren Attribute spezifiziert. Die Attribute umfassen sowohl mit Unsicherheit behaftete Attribute sowie komplexe Attribute, welche zur Darstellung von Relationen eingesetzt werden. Diese Repräsentation ist in der Wissensbasis gespeichert.

Sobald durch Sensormessungen neue Objektinformationen erfasst werden, erfolgt eine Instanziierung der entsprechenden Objektklasse aus der Wissensbasis, indem eine Objektinstanz

in ein formales Modell hinzugefügt und gegebenenfalls mit bestehenden Instanzen verbunden wird. Die Instanz verwendet dabei alle Informationen der Klasse aus der Wissensbasis, insbesondere das probabilistische Modell der Klasse. Gleichzeitig lassen sich Attribute durch Sensorinformationen verfeinern.

Das formale Modell stellt damit einerseits eine äquivalente Repräsentation der Umwelt dar, andererseits besitzt es ein vollständiges probabilistisches Modell der mit Unsicherheit behafteten Attribute. Auf diesem probabilistischen Modell wird durch Inferenz eine SA berechnet.

Zur formalen Repräsentation der Klassen und Instanzen wird die von Minsky [Min74] eingeführte Frame-Repräsentation verwendet, eine objektzentrierte Repräsentationsform. Dabei werden Objekte als sogenannte Frames modelliert, die Attribute und Relationen von Objekten als Slots des jeweiligen Frame.

Die beschriebene Methodik zeigt auf, wie Hintergrundwissen einer Domäne mit Konzepten der FOPL in einer Wissensbasis abgelegt werden kann. Diese Wissensbasis beinhaltet das probabilistische Modell der Domäne auf Klassenebene. Die vorherrschende Situation kann durch den Aufbau eines formalen Modells in der Repräsentation abgebildet und aktualisiert werden. Des Weiteren ist es möglich, Domänen mit Attributunsicherheit, Strukturunsicherheit sowie Existenzunsicherheit zu modellieren.

3.4.4 Diskussion

Die vorgestellten Repräsentationsmethoden stellen Konzepte zur Fahrsituationsanalyse vor, die entweder auf logikbasierter Beschreibung (DL) oder Logik erster Ordnung (FOL) aufsetzen. Spezifisch ausgewählte Situationen werden auch mittels Bayes'scher Netze (BNs) beziehungsweise Markov Ketten beschrieben.

Die Arbeiten zur SA durch BNs zeigen eine gute Leistung bei der Erkennung von Fahrsituationen, besonders da mit Unsicherheit behaftete Informationen ausgewertet werden können. Allerdings sind die Methoden stets auf spezifische Anwendungen limitiert. Eine einfache Anpassung an veränderte Situationen sowie die Wiederverwendung schon modellierter Teilbäume erfolgt nicht mit methodischen Konzepten. Die Situationsbäume von Schneider [Sch09] lassen sich zwar für verschiedene Situationen miteinander kombinieren, eine umfeldabhängige automatische Konstruktion der Situationsbäume ist jedoch nicht möglich.

Durch ontologiebasierte Ansätze kann eine Beschreibung von Situationen in generischer Weise erfolgen und eine Inferenz auf Situationen ist möglich. Die Anpassung der Modelle an veränderte Situationen erfolgt dabei mit Hilfe der gewählten Repräsentation. Allerdings lassen sich Unsicherheiten in den Attributen der Objekte nicht berücksichtigen. Dies führt zu einem starken Anstieg der Komplexität. Im Ansatz von Vacek [Vac08] kann die Fallbasis selbst in einfachen Anwendungen sehr groß werden, da unterschiedliche Objektpositionen in diskrete Fälle unterteilt werden und Relationen nicht modelliert werden.

Die Verwendung von Situationsklassen zusammen mit einer Wissensbasis ermöglicht es, Rückschlüsse aus bekannten Situationen auf Objekte zu ziehen. In [Lat+05] ist dies allerdings nur für wenige modellierte Situationen möglich. Es ist nicht möglich, durch die Zuweisung eines Risiko-Wertes auf eine spezifische Objektkonstellation ein generisches Modell zu erhalten. Außerdem ist das System nicht robust gegenüber Attributveränderungen der Objekte. Rück-

schlüsse auf die Situation sind nicht möglich, sofern für eine vorherrschende Situation keine entsprechende Objektkonstellation existiert. Die Repräsentation ist durch Verwendung von FOL-Regeln zwar erweiterbar, allerdings ist zu erwarten, dass die Regelbasis für unterschiedliche Objektkonstellationen schnell anwächst.

Ein großer Nachteil bei der Modellierung von Situationen durch graphische Modelle ist deren starre Struktur. Dies führt zu Einschränkungen, welche sich entweder auf die Laufzeit der Verfahren oder auf die Quantität der betrachteten Information niederschlagen. Aus diesen Gründen werden bisher analytische Methoden zur Bewertung der Fahrsituation bevorzugt. Diese führen aber zu einem hohen Aufwand bei der Wahl der Parameter und garantieren nur dann zuverlässige Ergebnisse, wenn die zugrundeliegenden Modelle das Problem exakt beschreiben.

Die Modellierung durch die OPRML nach Howard [How10] greift auf FOPL zurück und verwendet hierfür eine ER-basierte Beschreibung für Situationen. Die eingesetzte Frame-Repräsentation ist gut geeignet, um Objekte und Relationen als Klassen in einer Wissensbasis beschreiben zu können. Die Konstruktion von formalen Modellen erlaubt es, flexibel angepasste graphische Modelle für die Inferenz von Informationen zu berechnen.

Der Großteil bisheriger Forschungsarbeiten an FAS legt den Schwerpunkt in die Verbesserung der Wahrnehmung, da eine gute Qualität der Informationen als Schlüssel für nachfolgende Prozesse gesehen wird. Der Stand der Forschung zeigt auf, dass bei der Verknüpfung verschiedener Umfeldinformationen zum Aufbau einer SA geeignete Methoden fehlen, die abhängig von den wahrgenommenen Informationen Modelle zur Inferenz erstellen.

3.5 Forschungsbedarf

Im folgenden Abschnitt wird der Forschungsbedarf zur Situationsanalyse (SA) für FAS unter Berücksichtigung der vorgestellten Arbeiten beschrieben. Bisherige Arbeiten haben sich mit der Entwicklung spezifischer, eigenständiger FAS beschäftigt. Eine über einzelne Funktionen hinausgehende Betrachtung der Forschungsfelder Umfelderkennung und Situationsanalyse wurde bislang durch Forschungsarbeiten nicht hinreichend untersucht. Besonders eine durchgängige Situationsbetrachtung aus Sicht der Einflussfaktoren Fahrer, Fahrzeug und Umfeld wurde noch nicht durchgeführt. Dies wird auch durch die Arbeit von Rendon-Velez, Horváth und Opiyo [RHO09] festgestellt. Die Betrachtung der Fahrsituation wurde in bisherigen Forschungsarbeiten nur ansatzweise erforscht.

Anforderungen Nach [How10] sollte ein Ansatz zur Situationsanalyse drei Anforderungen erfüllen, nämlich *Generizität*, *Robustheit* und *Modularität*:

Generisch deshalb, weil eine Methode losgelöst von prototypischen Systemen und Speziallösungen einer Domäne sein sollte und bezüglich unterschiedlicher Aspekte generalisieren sollte. Bezogen auf Fahrsituationen soll der Ansatz für unterschiedliche Situationen verwendbar und insbesondere erweiterbar sein. Gleichwohl sollte eine Wiederverwendbarkeit von bereits modellierten Klassen und Relationen ermöglicht werden (Abbildung 3.11).

Robust, da sich der Ansatz konsistent bezüglich der Variation innerhalb von Situationen verhalten sollte und damit eine einheitliche Funktionsweise unabhängig von konkreten Objekten

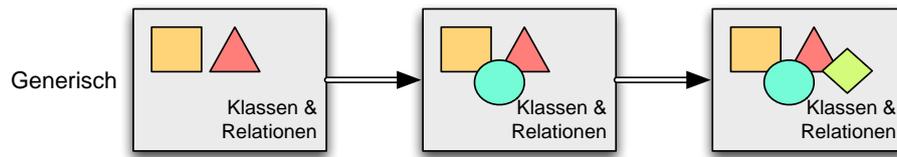


Bild 3.11: Veranschaulichung der Anforderung *Generisch*: Die Domäne sollte bezüglich Klassen und Relationen erweiterbar sein.

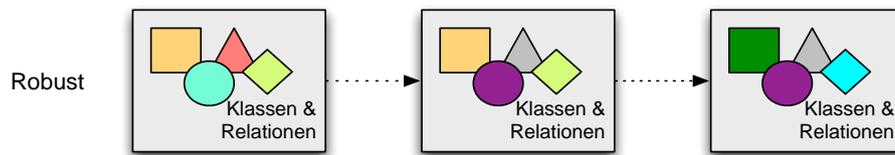


Bild 3.12: Veranschaulichung der Anforderungen *Robustheit*: Gewährleistung einer einheitlichen Funktionsweise auch bei Veränderungen in der Domäne.

oder Objekt-Attributwerten garantiert. In dieser Arbeit müssen dynamische Veränderungen von Fahrsituationen berücksichtigt werden, beispielsweise die variable Anzahl von Objekten im Umfeld des Fahrzeugs. Dies sollte nur minimale Auswirkungen im formalen Modell bewirken. Das graphische Modell sollte sich ohne manuelle Eingriffe an die Fahrsituation anpassen (Abbildung 3.12).

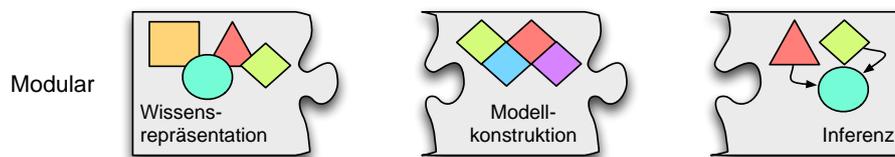


Bild 3.13: Veranschaulichung der Anforderungen *Modularität*: Wissensrepräsentation, Modellkonstruktion und Inferenzverfahren stellen separate Komponenten dar.

Modular, um Wiederverwendbarkeit, Erweiterbarkeit und Validierbarkeit zu fördern. Daraus folgt, dass Einzelkomponenten separat weiterentwickelt und verbessert werden können. Die Komponenten zur SA sollten es ermöglichen, die Wissensrepräsentation getrennt von der Konstruktion des formalen Modells und dem Mechanismus zur Inferenz von Informationen zu erstellen (Abbildung 3.13).

Umfeldererkennung Die Verbesserung der Umfeldwahrnehmung ist fortwährend Gegenstand der Forschung. Die für FAS zur Verfügung stehende Sensorik ist inzwischen weit ausgereift. Für viele der relevanten Umgebungsobjekte existieren hochentwickelte Verfahren zur Informationsextraktion. Beispiele dafür sind in den Abschnitten 3.2.1 und 3.2.2 genannt. Dennoch erreichen die Verfahren oftmals nicht ausreichend hohe Erkennungsraten. Besonders in ungünstigen Situationen, wenn Objekte teilweise verdeckt sind (ein Fußgänger hinter einem Fahrzeug), die gewohnte Ansicht nicht messbar ist (Seitenansicht eines Fahrzeuges) oder gar mehrere widersprüchliche Informationen gemessen werden (Fahrspurmarkierung in Baustellen). Unter Berücksichtigung der eingeschränkt vorhandenen Rechenleistung und der

geforderten Geschwindigkeit der Verarbeitungskette sind die Grenzen der Umfelderkennung schnell erreicht.

Eine Lösung liegt in der verstärkten Nutzung von entfernungsmessenden Sensoren, welche Tiefenbilder der Umgebung aufzeichnen und modellunabhängige Verfahren zur Objekterkennung ermöglichen. Mit effizienten Verfahren lassen sich Objekthypothesen aus den Tiefenbildern erstellen. In Kombination mit Videosensoren zur Klassifikation der Objekte wird daraus eine große Menge an Informationen extrahiert. Die Verfahren zur Informationsfusion müssen aufeinander abgestimmt werden, sodass die jeweiligen Stärken der Sensoren spezifisch ausgenutzt werden können. Eine Fusion erzielt dann einen deutlichen Mehrwert gegenüber dem separaten Betrieb der Sensoren. Unsicherheiten in der sensorischen Erkennungsleistung können dabei berücksichtigt werden.

Eine konzeptionelle Verarbeitungskette, welche die Objekterkennung und Fusion für FAS auf geeigneten Sensoren in der geforderten Weise realisiert, wird in Kapitel 4 in dieser Arbeit präsentiert.

Modellierung Eine Repräsentation von Fahrsituationen soll es erlauben, sowohl einfache als auch komplexe Situationen gleichermaßen darzustellen und generisch zu beschreiben. Die Herausforderung liegt darin, mit einem geeigneten Mechanismus die Fahrsituation formal und damit maschinell verständlich zu repräsentieren. Zur Beschreibung der momentanen Fahrsituation müssen der Zustand des Eigenfahrzeugs, die Zustände relevanter Objekte im Fahrzeugumfeld sowie die Veränderung dieser Zustände berücksichtigt werden. Wie im vorherigen Abschnitt festgestellt, erfolgt die Modellierung von Situationen meist durch starre Strukturen. Dies erschwert ein robustes Verhalten gegenüber einer variablen Anzahl von Objekten und verringert eine Erweiterbarkeit gegenüber neuer Situationen.

Die Verfahren der wissensgetriebenen Situationsanalyse zeigen aus methodischen Aspekten die größte Eignung für diese Arbeit. Die formale Repräsentation von Entitäten und Relationen erfolgt mittels FOPL. Dies ermöglicht die Berücksichtigung von Attribut-, Struktur- und Existenzunsicherheiten. Daher wird im Rahmen dieser Arbeit die Objektorientierte Probabilistische Relationale Modellierungssprache (OPRML) zur wissensgetriebenen SA aufgegriffen. Es wird erörtert, inwiefern sich Fahrsituationen darin repräsentieren lassen und ob eine SA im gewünschten Sinne darauf möglich ist.

Situationsanalyse Die Situationsanalyse (SA) sollte es erlauben, Fragen über die Fahrsituation maschinell zu beantworten, wie beispielsweise „Wie hoch ist das Risiko einer Kollision mit irgendeinem Objekt?“, „Welche Fahrspuren sind durch andere Fahrzeuge belegt?“ oder „Wie hoch ist das Risiko einer Kollision mit einem Objekt auf meiner Fahrspur?“. SA muss also im Wesentlichen die Bedeutung der Konzepte *Objekt* und *Relation* verstehen können. Unter Objektverständnis ist dabei gemeint, eine Kenntnis über die Objekte in der Umgebung zu besitzen. Unter dem Verständnis über Relationen wird die Kenntnis über aktuell bestehende Objektzusammenhänge sowie die Änderungen der Relationen im Kontext des Fahrens verstanden. Relationen umfassen dabei sowohl Objekte einer Klasse (*Fahrzeug*) als auch Objekte verschiedener Klassen (*Fahrzeug* und *Fahrspur*).

Im Kontext von FAS sollte durch SA ein Verständnis über die Fahrsituation aufgebaut werden. Dies geschieht in Hinblick auf die beobachteten Zustände (Position, Geschwindigkeit, Manöver, . . .) anderer Objekte sowie des eigenen Fahrzeugs. Sichere Aussagen über Fahrsituationen sind häufig nicht möglich, da sie mit Unsicherheit behaftet sind. Kann diese Eigenschaft von Seiten der SA nicht berücksichtigt werden, müssen Ausnahmen für Regeln explizit modelliert werden oder stochastische Modelle in viele Einzelfälle diskretisiert werden. Die Vorteile bei der Berücksichtigung von Probabilistik liegen damit auf der Hand. Es erleichtert die natürliche Modellierung von Situationen und erlaubt es, Glaubensgrade über Situationsaspekte zu formulieren. Treffen alle Systemkomponenten jeweils eine Aussage über die Zuverlässigkeit ihrer berechneten Ergebnisse, dann kann die Wahrscheinlichkeit des Eintreffens von Ereignissen bestimmt werden. Durch die Modellierung der Kausalität von Ereignissen kann zudem auch auf die Ursache von Ereignissen gefolgert werden, sodass gezielt auf eine Ursache reagiert werden kann.

Eine Analyse von Fahrsituationen sollte also auf einer logischen probabilistischen Repräsentation beruhen und darüber Inferenz betreiben. Zur Berechnung der Inferenz müssen die probabilistischen Formulierung in graphische Modelle überführt werden, was als Modellkonstruktion bezeichnet wird. Eine automatische Konstruktion erlaubt es, graphische Modelle adaptiv an die vorherrschende Situation anzupassen. Die Modellkonstruktion, sowie die darauf aufbauenden Inferenzmechanismen, müssen in den geforderten Zykluszeiten der Assistenzsysteme erfolgen, damit eine Situationsaussage rechtzeitig zur Aktionsplanung gewährleistet werden kann. Kapitel 5 beschreibt die Umsetzung dieser Anforderungen in dieser Arbeit.

4 Wahrnehmung des Fahrzeugumfelds

Zur Interpretation der Situation müssen Objekte und Attribute, welche Einfluss auf die Fahr-situation haben, zunächst durch Sensoren erfasst werden, mit denen Fahrzeuge ausgerüstet sind. Die Sensoren überwachen dazu das Fahrzeug selbst sowie das Umfeld und zeichnen In-formationen kontinuierlich auf. In dieser Arbeit wird ein komplementäres Sensorsystem beste-hend aus Lichtlaufzeit (ToF)-Sensoren sowie Video-Sensoren eingesetzt und für die Eignung zur Umfeldwahrnehmung bewertet.

Für die Detektion von Fahrzeugen, Fußgängern und Hindernissen wurde dafür ein An-satz entwickelt, durch welchen zuerst aus Tiefen- oder Videodaten Objekthypothesen er-stellt werden und anschließend durch Zustandsschätzung eine Verifikation der Hypothesen erfolgt. Erlangte Objektinformationen bestehen aus der Position, Dimension und Orientie-rung der Objekte. Für Tiefendaten kommen abhängig von der Auflösung des ToF-Sensors Regionenwachstums- oder Histogrammbasierte Verfahren zum Einsatz. Für Videodaten wer-den abhängig von der Objektklasse ansichtsbasierte Merkmale verwendet. Anschließend wer-den die Objekthypothesen anhand der Videodaten klassifiziert. Die Klassifizierungsverfahren für Fahrzeuge sowie für Fußgänger nutzen dabei Support Vektor Maschinen oder Klassifikator-Kaskaden. Klassifizierte Objekthypothesen werden durch temporale Zustandsschätzung verifi-ziert und dynamische Attribute wie die Bewegungsrichtung oder -geschwindigkeit abgeleitet. Dazu wird ein Multi-Target Tracking (MTT) System auf Basis des Kalman Filters eingesetzt. Unter Berücksichtigung der fahrzeugeigenen Bewegung werden durch das Interacting Multi-ple Model (IMM) Filter statische und dynamische Objekte voneinander unterschieden.

Für die Bestimmung der Fahrsituation sind Umfeldinformationen nicht nur im Bereich vor dem Fahrzeug, sondern auch neben oder hinter dem Fahrzeug von Relevanz. Im Rahmen dieser Arbeit wurden die Verfahren daher zur Überwachung des Fahrzeugumfelds neben dem Fahrzeug untersucht.

4.1 Konzept der Umfelderkennung

Der Ablauf der Detektion und die Bestimmung der Attribute von Objekten im Fahrzeugumfeld mittels komplementären Sensoren wird im folgenden Abschnitt erklärt. Das Konzept sieht dabei vor, dass Objekthypothesen sowohl anhand der Tiefendaten der ToF-Kamera als auch anhand der Videodaten der Videokamera äquivalent erzeugt werden können.

Bei der Konzeption der Umfelderkennung wurde besonders darauf geachtet, eine vollstän-dige Trennung zwischen den Schichten Segmentierung, Klassifikation und Zustandsschät-zung zu gewährleisten. Dies erlaubt die Ergänzung um weitere Komponenten der jeweili-gen Schicht, ohne die jeweils anderen Schichten anpassen zu müssen. Die zwischen diesen drei Schichten ausgetauschten Daten müssen dabei alle Informationen zur Projektion der Da-ten in komplementäre Sensordaten mitführen. Zusätzlich werden die Daten in jeder Schicht

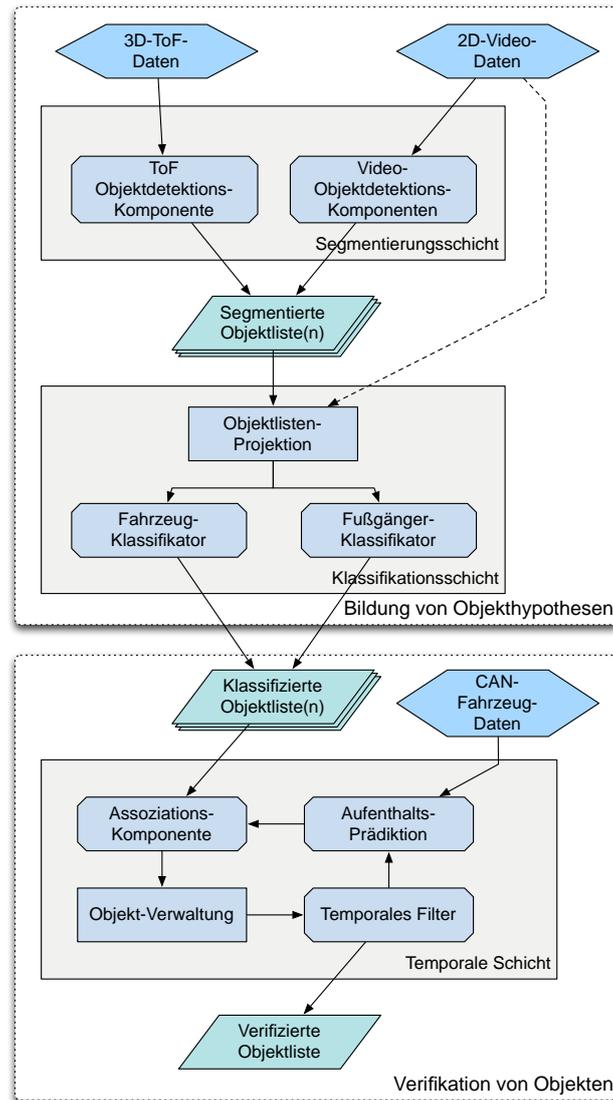


Bild 4.1: Konzeptioneller Ablauf der Umfelderkennung mit komplementären Kameras und temporaler Fusion zur Bildung von Objektwissen.

durch die hinzugewonnene Information erweitert. Die ursprüngliche Information wird dabei nicht modifiziert. Die resultierende Objektliste der temporalen Zustandsschätzung enthält für jedes Objekt alle Informationen der gesamten Historie des Objektes. Dadurch können die temporalen Schätzer besonders flexibel das Objektwissen interpretieren und beispielsweise eine temporale Klassifikation beziehungsweise einen Konfliktwert zwischen Klassen bestimmen.

Abbildung 4.1 zeigt das Konzept der Umfelderkennung. In der Segmentierungsschicht erfolgt je nach den zur Verfügung stehenden Sensordaten und abhängig von den verwendeten Erkennungsalgorithmen eine Hypothesenbildung. Pro Sensor und Algorithmus wird eine Objektliste erstellt. Die Objektlisten werden an die Klassifikationsschicht übergeben. In Abhängigkeit des eingesetzten Klassifizierungsalgorithmus wird geprüft, ob dieser die Sensordaten verarbeiten kann. Sind die Sensordaten identisch zur Segmentierungsschicht, so kann die Klassifikationskomponente auf die segmentierte Region of Interest (ROI) zurückgreifen. Werden unterschiedliche Sensordaten eingesetzt, wird die segmentierte ROI in die Sensordaten des Klassifizierungsalgorithmus projiziert. Pro Klassifikationsverfahren wird die Objektliste

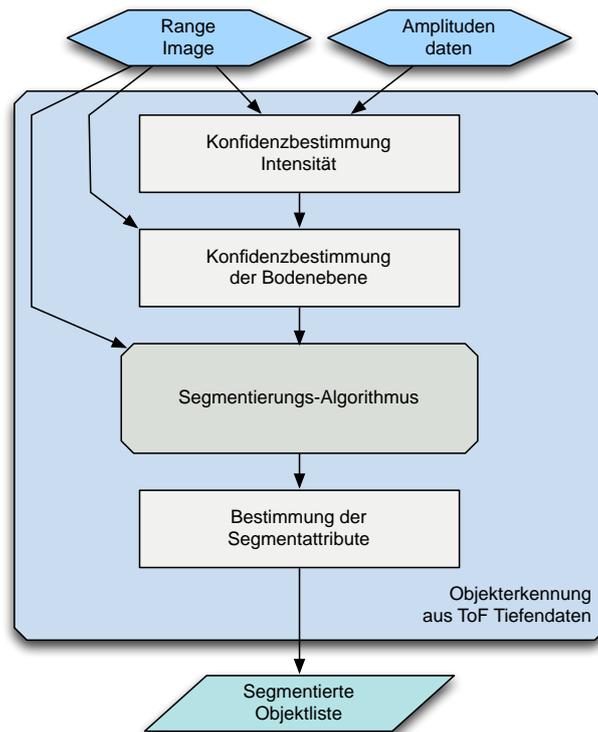


Bild 4.2: Konzeptioneller Ablauf zur Bildung von Objekthypothesen aus ToF-Tiefendaten durch verschiedene Segmentierungsverfahren.

mit Klasseninformationen angereichert. Die Objektlisten werden an die temporale Schicht übergeben. Diese führt die Objektlisten zusammen. Hierfür berechnet eine Assoziationskomponente die Zuordnung zwischen den Hypothesen der aktuellen Objektlisten und den Hypothesen der Objektliste der vorherigen Zyklusschritte. Ein temporales Filter aktualisiert die Zustandsschätzung für jedes Objekt. Das Resultat der Verarbeitungskette ist eine verifizierte Objektliste, welche der Situationsanalyse (SA) zur Verfügung gestellt wird.

4.2 Objekterkennung aus Tiefendaten

Die geometrische Auflösung der ToF-Sensordaten unterscheidet sich stark voneinander. Der Informationsgehalt pro Bildpunkt hängt damit von dem eingesetzten Sensor ab. Bei Sensoren mit einer hohen Auflösung (> 3000 Bildpunkte) und niedriger Reichweite kann die Zuverlässigkeit einer Signalinformation aus einer Menge zusammengehöriger (meist benachbarter) Bildpunkte abgeleitet werden. Beispielsweise besitzt ein einzelner Bildpunkt des Photomischdetektor (PMD) 41k-S2 Sensors (200×200 Bildpunkte) in $7m$ eine Messfläche von $\approx 6cm^2$ und trägt somit nur einen kleinen Teil der Information, welche zur Segmentierung eines Objektes in der Größe einer Person notwendig ist. Bei Sensoren mit sehr geringer Auflösung (≈ 1000 Bildpunkte) und hoher Reichweite dagegen fällt jedem Bildpunkt ein relevanter Informationsgehalt zu, beispielsweise nimmt die Messfläche eines Bildpunktes in ungünstigen Fällen mehrere Quadratmeter ein. Bei dem PMD 1k-S Sensor (64×16 Bildpunkte) besitzt ein Bildpunkt in $7m \approx 136cm^2$ und in $25m \approx 0,17m^2$ Messfläche. Eine Person mit einer Fläche von $\approx 0,85m^2$ bildet sich somit in $> 25m$ Entfernung auf etwa 5 Bildpunkte ab.

Aus diesem Grund werden zur Objektsegmentierung, abhängig von der eingesetzten Sensorik, verschiedene Segmentierungsverfahren im Rahmen dieser Arbeit eingesetzt. Für Sensoren mit geringer Auflösung wird ein Regionenwachstums-Verfahren, wie in Abschnitt 4.2.1 beschrieben, genutzt. Dieses wendet eine Homogenitätsbewertung auf jeden Bildpunkt gegenüber dessen direkter Nachbarschaft an und nutzt diese Bewertung zur Clusterbildung. Für Sensoren mit hoher Auflösung wird ein Histogrammbasiertes Verfahren eingesetzt, welches Punktemengen anhand einer stochastischen Funktion bewertet und daraus Objektausdehnung und -lage bestimmt. Dieses ist in Abschnitt 4.2.2 beschrieben.

4.2.1 Objektsegmentierung durch Regionenwachstum

Für die Bestimmung von Objekt-ROI aus gering aufgelösten PMD-Tiefendaten wird ein Regionenwachstums-Verfahren eingesetzt, welches für jeden Bildpunkt mittels einer Bewertungsfunktion eine Gruppenzugehörigkeit bestimmt. Trotz der höheren Komplexität dieses Verfahrens ist aufgrund der geringeren Auflösung der Kamera der Rechenaufwand des Verfahrens akzeptabel. Die Anforderung, eine Segmentierung in der geforderten Zykluszeit durchzuführen, kann somit eingehalten werden. Abbildung 4.2 zeigt den konzeptionellen Ablauf der Objektsegmentierung.

Der Eingangsdatensatz der PMD-Kamera besteht aus zwei Teilbildern: den Tiefendaten sowie den Amplitudendaten. Die Tiefendaten entsprechen der radialen Entfernungsmessung vom optischen Zentrum der Kamera zur Szene. Die Signalamplitude entspricht dabei der Intensität des von der Szene zurückreflektierten Signals.

Die Tiefendaten enthalten sowohl nutzbare Signal-Information als auch Rauschen. Eine Konfidenzbestimmung jedes Bildpunktes ist anhand der korrespondierenden Amplitudendaten möglich. Die Zuverlässigkeit der Tiefeninformation sinkt für einen Bildpunkt, wenn eine zu geringe Intensität des Signals durch die Szene reflektiert wird. Bildpunkte werden als ungültig markiert, wenn die gemessene Intensität einen Schwellwert h_{intens} unterschreitet, also

$$g(x) = \text{invalid, falls } h(x) < h_{intens}, \quad (4.1)$$

wobei $g(x)$ dem Tiefenwert und $h(x)$ der Amplitude des Bildpunktes x entspricht.

Anschließend an die Intensitätskompensation findet eine Filterung der zur Bodenebene gehörenden Bildpunkte statt. Unter Berücksichtigung der intrinsischen und extrinsischen Kameraparameter lässt sich das Tiefenbild in eine Punktwolke umwandeln, welche die Messdaten als 3D-Punkte im Fahrzeugkoordinatensystem repräsentiert. Messdaten, welche von der Bodenebene reflektiert werden, treten in relevanter Menge nur im Nahbereich des Fahrzeugs auf, da die Reflektanzeigenschaft der Bodenebene bei sehr flachem Einfallswinkel gegen Null strebt. Ein Bildpunkt wird als ungültig markiert, wenn dessen Position \mathbf{p} im Fahrzeugkoordinatensystem unter der angenommenen Bodenebene mit $z \approx 0$ und innerhalb des Nahbereichs $x_{nearfield}$ des Fahrzeug liegt:

$$g(x) = \text{invalid, falls } p_x < x_{nearfield} \wedge p_z < 0. \quad (4.2)$$

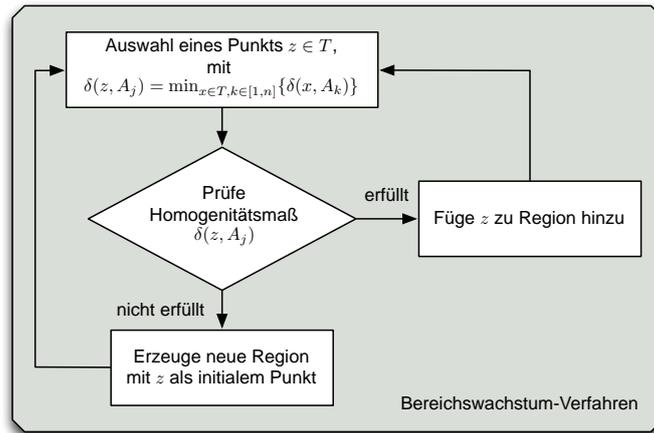


Bild 4.3: Ablauf der Objektsegmentierung durch Regionenwachstum; mithilfe eines Homogenitätsmaßes wird für Bildpunkte geprüft, ob diese Teil einer Region sind.

Die ungültigen Bildpunkte des Tiefenbildes werden im nachfolgenden Verfahren nicht berücksichtigt. Das Bereichswachstums-Verfahren ist ein iterativer Algorithmus, welcher ein Bild in einzelne Segmente zerlegt (siehe Abbildung 4.3). In jedem Schritt wird ein Bildpunkt ausgewählt und zu einem Bereich hinzugefügt, für welches ein Homogenitätsmaß erfüllt ist, oder ein neues Segment erzeugt, falls für keinen schon existierenden Bereich das Kriterium erfüllt ist. Dies wird wiederholt, bis alle Bildpunkte des Bildes Regionen zugeordnet sind. Beispiele dafür sind das Seeded Region Growing (SRG)-Verfahren von Adams und Bischof [AB94] oder das von Lin, Jin und Talbot [LJT01] vorgestellte Unseeded Region Growing (URG). Letzteres Verfahren zur unüberwachten Segmentierung von Graustufenbildern eignet sich zur angestrebten Segmentierung von Tiefendaten, da initiale Bereiche nicht manuell vorgegeben werden müssen, sondern durch den Algorithmus gewählt werden.

Für eine Menge an Regionen A_1, A_2, \dots, A_n ist T die Menge aller noch nicht klassifizierten Punkte, die an mindestens einer Regionengrenze liegen:

$$T = \left\{ x \notin \bigcup_{i=1}^n A_i \wedge \exists k : N(x) \cap A_k \neq \emptyset \right\}, \quad (4.3)$$

wobei $N(x)$ die direkt benachbarten Bildpunkte des Punktes x sind. Weiter ist ein Abstandsmaß definiert mit

$$\delta(x, A_i) = |g(x) - \text{mean}[A_i]|, \quad (4.4)$$

mit $g(x)$ als der Tiefenwert des Bildpunktes x und $\text{mean}[A_i]$ als der mittlere Tiefenwert der Region A_i . Der Index i muss dabei das Kriterium $N(x) \cap A_i$ erfüllen. Ist $T \neq \emptyset$, so beträgt der minimale Abstand eines Punktes z zu einer Region A_j , $j \in [1, n]$:

$$\delta(z, A_j) = \min_{x \in T, k \in [1, n]} \{\delta(x, A_k)\}. \quad (4.5)$$

Ist $\delta(z, A_j)$ kleiner als die maximal zulässige Tiefendifferenz δ_{max} zwischen zwei Regionen, wird der Bildpunkt z zu der Region A_j hingefügt und der mittlere Tiefenwert $\text{mean}[A_j]$ der Region aktualisiert. Ansonsten wird eine neue Region A_{n+1} mit z als initialem Punkt hinzuge-

fügt.

Das dargestellte Homogenitätskriterium $\delta(x, A_i)$ ermöglicht es, die Segmentierung auf einfache Art zu variieren. Durch Adaption des Homogenitätskriteriums kann die Segmentierung beispielsweise auch anhand der Ähnlichkeit der Amplitude zwischen dem gewählten Bildpunkt und der Region erfolgen oder ein unterschiedliches Ähnlichkeitsmaß zwischen vertikal und horizontal benachbarten Bildpunkten definiert werden. Auch eine Adaption des Schwellwertes δ_{max} abhängig von der mittleren Entfernung der Region $\text{mean}[A_i]$ kann die Segmentierungsergebnisse verbessern.

Nach Abschluss der Bereichsegmentierung folgt eine Segmentverifikation, welche den Informationsgehalt eines Bildpunktes berücksichtigt. Ein Segment ist gültig, wenn es eine Objektfläche mit minimaler Fläche λ_{min} beschreibt. Da die betrachtete Messfläche pro Bildpunkt mit zunehmender Entfernung quadratisch steigt, führt dies dazu, dass unter einer gewissen Entfernung Segmente mit wenigen Bildpunkten als ungültig aussortiert werden. Gleichzeitig werden Segmente, welche eine Messung in großer Entfernung beschreiben aber nur aus einem Bildpunkt bestehen, als gültig markiert.

Die gültigen Segmente beschreiben Objekthypothesen, welche durch die PMD-Kamera in diesem Messzyklus beobachtet wurden. Zur Weiterverarbeitung in der Umfelderkennung, insbesondere für die Durchführung der Klassifikation und der Zustandsschätzung, werden für jede Objekthypothese die Form beschreibenden Attribute bestimmt. Dies sind der Schwerpunkt des Objektes sowie die umschließende Box (BBox).

Das hier beschriebene Verfahren wird auf einem autonomen Fahrzeug zur Hindernisbestimmung und zur Realisierung von ACC-Funktionen eingesetzt [Vac+07; Sch+07; Sch+08; Koh+13]. Des Weiteren wird es in Arbeiten zur Informationsfusion zwischen PMD- und Videokamera zur Klassifikation von Fußgängern genutzt. Der Ansatz in [Nat+08] zeigt auf, wie der Ablauf einer komplementären Objektsegmentierung und -klassifikation umgesetzt werden kann.

4.2.2 Histogrammbasierte Objektsegmentierung

Histogrammbasierte Verfahren zur Objektsegmentierung eignen sich, wenn ein Objekt durch eine größere Menge an Bildpunkten beschrieben ist und der Einfluss eines einzelnen Bildpunktes für die Segmentierung eines Objektes daher nicht entscheidend ist. Mit Hilfe von Histogrammen wird eine stochastische Zählung der Bildpunkte vorgenommen, indem ähnliche Merkmale zusammengefasst und die Häufigkeit jedes Merkmals gezählt wird. Anhand des erstellten Histogramms lassen sich Merkmalsgrenzen bestimmen, über welche auf Objekte geschlossen werden kann. Durch Rückzuordnung aus dem Merkmalsraum auf die Bildpunkte lassen sich schließlich die Objekt-ROI folgern. Histogrammbasierte Verfahren sind aufgrund der Dimensionsreduktion sehr effizient und damit für große Bilddaten gut geeignet.

Eine Vorverarbeitung der Daten gemäß der in Abschnitt 4.2.1 beschriebenen Algorithmen ist nur zu Teilen notwendig, da durch die stochastische Eigenschaft der Segmentierung ein robustes Verhalten erreicht wird. Die Filterung nach der Konfidenz jedes Bildpunktes anhand dessen Amplitudenwert führt zu einer verbesserten Segmentierung, weshalb dieser Vorverarbeitungsschritt ebenfalls auf die Tiefendaten angewendet wird. Eine feste Extraktion der

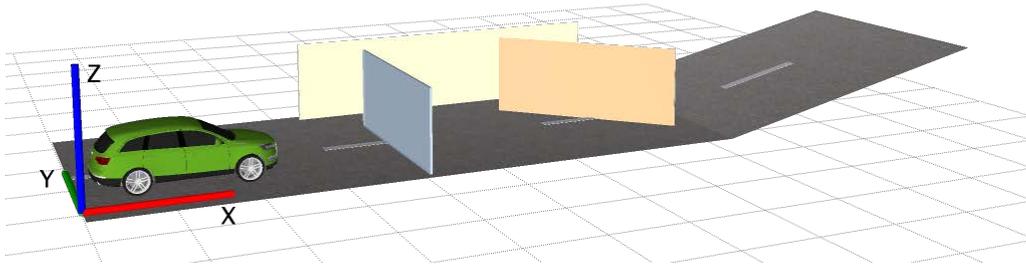


Bild 4.4: Messbare Flächen-Konfigurationen sind senkrecht auf der x - y und x - z -Ebene stehende Flächen.

Bodenebene ist allerdings nicht notwendig, da die Bestimmung der Bodenebene durch den Segmentierungsalgorithmus geleistet wird.

Basierend auf einer Modellannahme, dass sich jedes Objekt im Fahrzeugumfeld durch eine eingrenzende, rechteckige Fläche beschreiben lässt, werden Flächen durch das Verfahren extrahiert, wie in Abbildung 4.4 dargestellt. Senkrecht auf der x - y -Ebene stehende Strukturen sind zum Beispiel die Heck- oder Frontansicht von Fahrzeugen, welche sich vor dem eigenen Fahrzeug befinden, Leitplanken, neben dem eigenen Fahrzeug fahrende Objekte und seitlich einscherende Fahrzeuge sowie quer stehende Begrenzungen. Diese entsprechenden extrahierten Flächen sind in der Abbildung als hellblaue, hellgelbe und orangene Flächen gekennzeichnet. Auch Flächen senkrecht zur x - z -Ebene, also horizontale Flächen wie die Fahrbahn oder Steigungen können durch den Algorithmus segmentiert werden.

Der Ansatz nach Hu, Lamosa und Uchimura [HLU05] beschreibt, wie schräge Flächen aus Stereo-Disparitätsmessungen extrahiert werden. Dabei bilden sich senkrecht auf der x - y -Ebene stehende Flächen als Linien in dem Raum ab, der zwischen den Bildspalten u und den Disparitätswerten d_{Disp} aufgespannt wird. Flächen senkrecht zur x - z -Ebene bilden sich als Linien in dem durch die Bildzeilen v und den Disparitätswerten d_{Disp} aufgespanntem Raum ab. Die resultierenden Räume werden als U - respektive V -Disparitätsräume bezeichnet. Durch den Ansatz wird die Dimension der Daten informationserhaltend reduziert, um anschließend mittels einer Liniensuche effizient markante Strukturen zu finden. Die Extraktion der Linien erfolgt durch eine Hough-Transformation mit anschließender Liniensegment-Extraktion.

Zur Segmentierung von ToF-Tiefendaten und zur Verbesserung der Laufzeit wird der Algorithmus wie in [SRZ10] und [TS11] beschrieben eingesetzt und verbessert. Das Verfahren kann in vier Schritte unterteilt werden. Zunächst werden die Räume entlang der horizontalen (u) und vertikalen (v) Bildachse berechnet. Anschließend werden diese Räume in den Hough-Raum transformiert. Hier wird eine modifizierte Line Pattern Hough Transformation (LPHT) zur Suche nach Geradensegmenten in den Räumen angewendet. Jedes gefundene Liniensegment entspricht einer Fläche in der Szene. Zuletzt erfolgt die Bestimmung der vertikalen beziehungsweise horizontalen Ausdehnung jeder Fläche. Der Segmentierungsansatz ist in Abbildung 4.5 konzeptionell gezeigt.

Berechnung der U/V -Räume Die Berechnung der U/V -Räume erfolgt anhand der in [HLU05] vorgestellten Gleichungen. Der U -Raum wird ermittelt, indem das Tiefenbild spaltenweise analysiert wird und jeder auftretende Entfernungswert an seiner Position im U -Raum

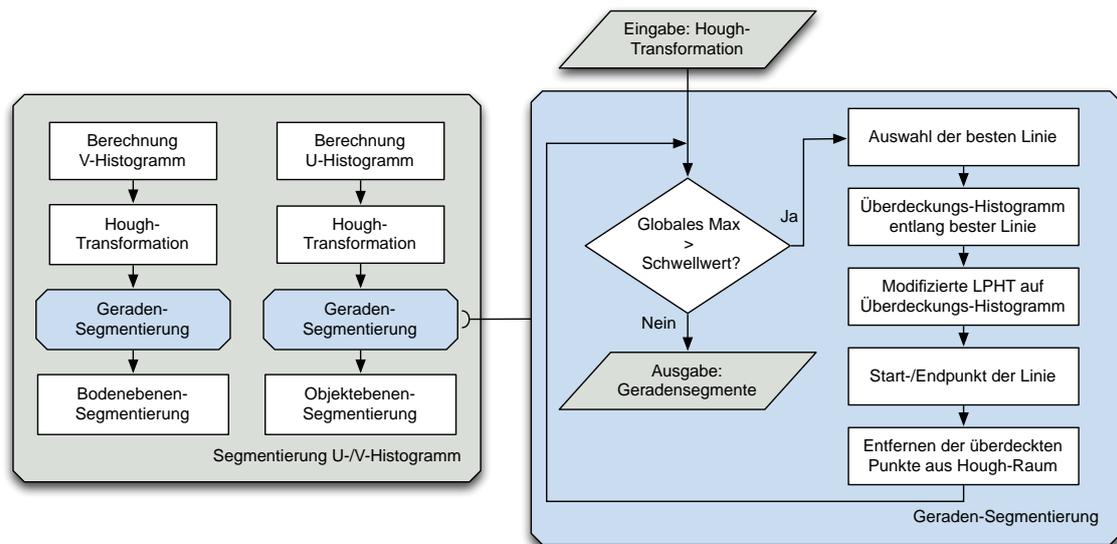


Bild 4.5: Ablauf der Histogrammbasierten Objektsegmentierung (links) und der modifizierte Liniensuche (rechts).

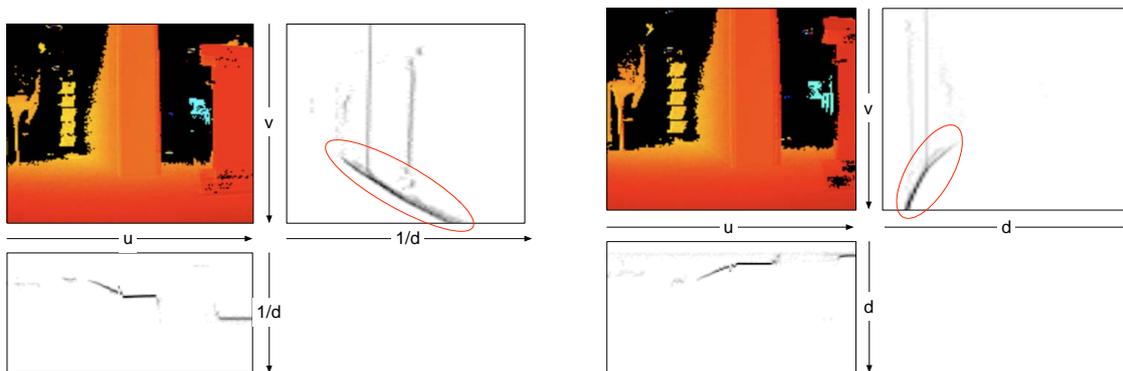


Bild 4.6: Tiefenbild in Falschfarbendarstellung mit entsprechenden U/V -Räumen. Die Bodenebene im V -Raum ist rot umrandet; links: Reziproke Erstellung des V -Raums; rechts: Erstellung des V -Raums mit äquidistanten Tiefenwerten.

inkrementiert wird. Der V -Raum wird analog berechnet, indem das Tiefenbild zeilenweise analysiert wird.

Die Dimension der U/V -Räume ist definiert durch die Auflösung der Eingangsdaten sowie der Anzahl der berücksichtigten Tiefenstufen des Tiefenbildes. Die Klassen der Histogramme bilden die Tiefenstufen ab. Die Auflösung des U -Raumes ist damit durch die Anzahl der Bildspalten und der Anzahl an Klassen des Histogramms gegeben.

Wie in [SRZ10] gezeigt wird, bilden sich parallel zur y -Achse liegende Flächen als Linien ab, wenn sie reziprok zur Entfernung in den V -Raum eingetragen werden. Analog bilden sich parallel zur z -Achse liegende Flächen als Linien in den U -Raum ab. Solange der Nickwinkel der Kamera gegenüber der Fahrbahnebene klein ist, können die sichtbaren Flächen von Objekten als Flächen parallel zur z -Achse approximiert werden.

Ein resultierendes U/V -Raum-Paar ist in der Abbildung 4.6 links dargestellt. Abbildung 4.6 rechts zeigt das Ergebnis der Berechnung mit äquidistanter Eintragung der Tiefendaten. Dort bilden sich Flächen nicht als gerade Linien ab. Die Bodenebene im V -Raum ist in beiden Dar-

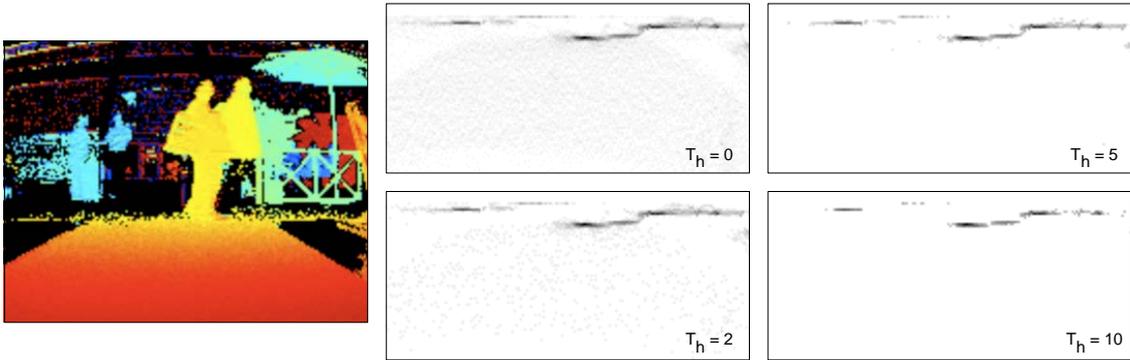


Bild 4.7: Unterdrückung des Grauschleiers im U -Raum durch unterschiedliche Akkumulator-Schwellwerte T_h .

stellungen rot umrandet. Je höher der Akkumulatorwert für eine Histogrammklasse ist, desto dunkler bildet sich dieser Punkt in der Abbildung ab. Deutlich zu erkennen sind die geraden Linien, welche durch die im Bild erkennbaren Hindernisse beziehungsweise die Bodenebene resultieren. Senkrecht auf der Bodenebene stehende Objekte führen an den entsprechenden Stellen im U -Raum also zu deutlich erkennbaren Linien. Flächen, welche nicht normal zur jeweiligen Projektionsebene stehen, führen allerdings zu Rauschen in dem jeweiligen Raum. Die Bodenebene erzeugt im U -Raum einen Grauschleier. Dieser lässt sich jedoch nach Analyse der Histogrammverteilung filtern, indem ein unterer Schwellwert T_h eingeführt wird, welcher dem minimal zulässigen Akkumulatorwert für einen Bildpunkt entspricht. Abbildung 4.7 zeigt die resultierenden U -Räume für unterschiedliche Schwellwerte $T_h = \{0, 2, 5, 10\}$.

Linienextraktion Im folgenden Abschnitt ist beschrieben, wie sich aus gefilterten Räumen effizient Linien extrahieren lassen. Abbildung 4.5 rechts zeigt den konzeptionellen Ablauf der Linienextraktion. Diese wird separat für den U - und V -Raum durchgeführt. Zunächst wird der jeweilige Raum in dessen korrespondierenden Hough-Raum mittels der Standard Hough Transformation (SHT) überführt. Das beste Liniensegment einer Geraden wird durch die LPHT ermittelt und anschließend im Hough-Raum entfernt. Dies wird für die jeweils beste Linie solange wiederholt, bis ein Abbruchkriterium erreicht wird. Das Abbruchkriterium ist erreicht, wenn eine maximale Anzahl extrahierter Linien im Hough-Raum gefunden wurde oder kein deutliches Maximum mehr zu erkennen ist.

Die SHT kann effizient berechnet werden, da per Definition weder vertikale Linien im U -Raum noch horizontale Linien im V -Raum existieren dürfen. Die Dimension des Hough-Raums für den U -Raum lässt sich auf den Winkelbereich $\omega_u = [0 + \nu, \pi - \nu]$ einschränken, wobei ν dem Winkel entspricht, mit welchem Hindernisflächen um die z -Achse rotiert sein können. Entsprechend lässt sich die Dimension des V -Raums einschränken auf $\omega_v = \left[-\frac{\pi}{2} + \nu, \frac{\pi}{2} - \nu\right]$.

Da die Dimension des U/V -Raums durch das Tiefenbild bekannt ist, lassen sich zur Optimierung der Laufzeit die trigonometrischen Funktionen für den Hough-Raum approximativ vorberechnen. Eine Faltung der Bilddaten mit einer Filtermaske ist nicht notwendig. Die Räume können direkt in die Hough-Räume überführt werden. Außerdem werden die Hough-

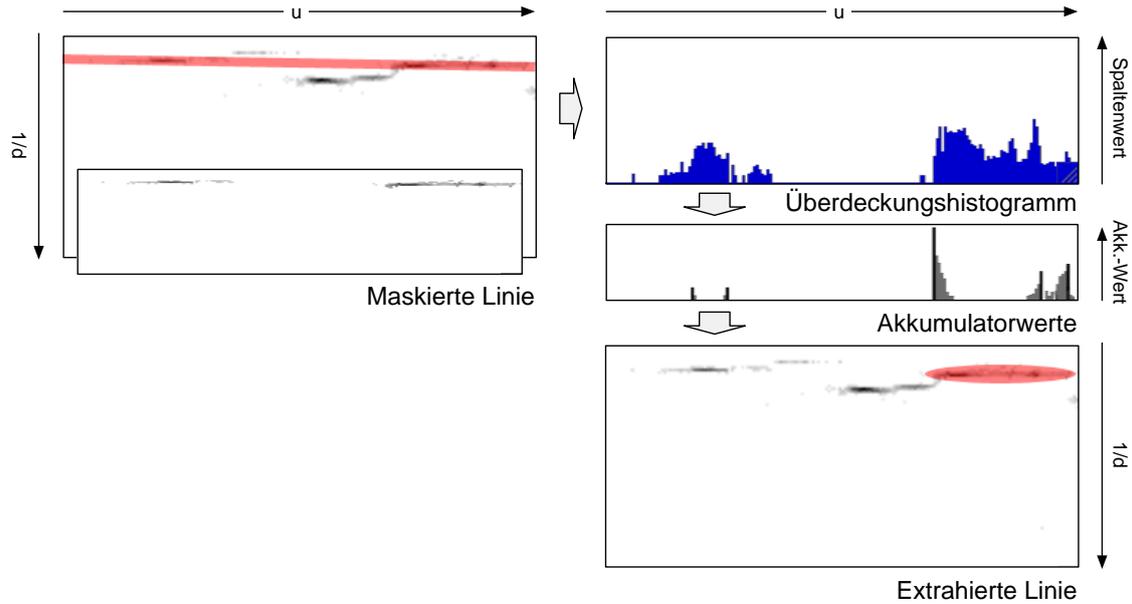


Bild 4.8: Linienextraktion; 1. Bild: Maskierte Linie mit ϵ -Bereich; 2. Bild: Überdeckungshistogramm entlang der Maskierung; 3. Bild: Akkumulatorwerte der Start- und Endpunkte von Linien; 4. Bild: Extrahierte Linie.

Daten durch den jeweiligen Akkumulatorwert Akk_{hist} des Raumes gewichtet, sodass deutliche Punkte im Raum einen stärkeren Einfluss im Hough-Raum besitzen:

$$r_u(\omega_u) = \left(u \cos \omega_u + \frac{1}{d} \sin \omega_u \right) \text{Akk}_{\text{hist}} \left[u, \frac{1}{d} \right] \quad (4.6)$$

$$r_v(\omega_v) = \left(\frac{1}{d} \cos \omega_v + v \sin \omega_v \right) \text{Akk}_{\text{hist}} \left[\frac{1}{d}, v \right]. \quad (4.7)$$

Hohe Werte im Hough-Raum entsprechen deutlichen Geraden im Histogrammbild. Eine Gerade ist durch die Hough-Parameter (r, ω) beschrieben. Das entsprechende Liniensegment der Geraden, also der Start- und Endpunkt, ist zunächst unbekannt. Zur Bestimmung dieser Linie wird ein Überdeckungshistogramm aus den Punkten unter der Geraden (r, ω) und einem ϵ -Bereich des Histogrammbildes berechnet (siehe Abbildung 4.8; die rot markierte Gerade entspricht dem stärksten Punkt im Hough-Raum (r, ω)).

Innerhalb dieses ϵ -Bereichs werden für den U -Raum alle Histogrammwerte spaltenweise zu einem Überdeckungswert akkumuliert. In dem Überdeckungshistogramm wird anschließend der Start- und Endpunkt der Linie bestimmt. Der ursprüngliche LPHT-Algorithmus nach Yip [Yip94] berechnet für ein 2D-Bild die relative Konnektivität jedes Punktes zur Bestimmung der Start- und Endpunkte einer Linie. Der Algorithmus wurde dahingehend verändert, dass die relative Konnektivität nun entlang eines 1D Linienhistogramms bestimmt wird, was die Komplexität des LPHT-Algorithmus von $O(MN \log(M) \log(N))$ auf $O(N \log(N))$ senkt. Ein weiterführender Vergleich beider Algorithmen kann in [TS11] nachgelesen werden. Die Modifikation nutzt die jeweiligen Vorteile der SHT und des LPHT. Die Position von Geraden lässt sich zuverlässig durch die SHT ermitteln, während die Bestimmung der Start-/Endpunkte durch die LPHT effizient erfolgt.

Der modifizierte Algorithmus zur Bestimmung der relativen Konnektivität ist als Pseudo-

Algorithmus 1 Modifizierte Berechnung der relativen Konnektivität entlang eines 1D Histogramms nach [TS11].

```

/* Sei  $H$  das Überdeckungshistogramm */
/* Sei  $N$  die Anzahl der Histogrammklassen */
/* Sei  $t$  der Schwellwert der Histogrammklasse */
/* Sei  $n$  die Anzahl kollinearere Punkte */
/* Sei  $\Delta_{max}$  der maximale relative Versatz */
/* Sei  $Akk_{LPHT}$  der Line Patterns Hough Akkumulator */
/* Sei  $m$  der Akkumulator-Schwellwert */
for  $x_1 = 1$  to  $N$  do
  if  $H(x_1) \geq t$  then
    for  $x_2 = x_1 + 1$  to  $x_1 + \Delta_{max}$  do
      if  $H(x_2) \geq t$  then
         $\Delta = x_2 - x_1$ 
        if  $H(x_1 - \Delta) < t$  then
          /* Startpunkt gefunden */
           $x_n = x_2$ 
           $n = 2$ 
           $ready = 0$ 
          while  $ready = 0$  do
            if  $H(x_n + \Delta) \geq t$  then
              /* weiteren Linienpunkt gefunden */
               $n = n + 1$ 
               $x_n = x_n + \Delta$ 
            else
              /* Endpunkt gefunden */
               $ready = 1$ 
            end if
          end while
          if  $n \geq m$  then
            /* Genügend Linienpunkte gefunden */
            /* Startpunkt zu Akkumulator hinzufügen */
             $Akk_{LPHT}[x_1] = Akk_{LPHT}[x_1] + n$ 
            /* Endpunkt zu Akkumulator hinzufügen */
             $Akk_{LPHT}[x_n] = Akk_{LPHT}[x_n] + n$ 
          end if
        end if
      end if
    end for
  end if
end for

```

code in Algorithmus 1 gegeben. Eine Menge aufeinanderfolgender, äquidistanter und kollinearere Punkte führt nach der Definition der relativen Konnektivität zu hohen Akkumulatorwerten Akk_{LPHT} . Da der berücksichtigte relative Versatz Δ zwischen zwei Punkten inkrementell erhöht wird, ist das Verfahren robust gegen kleine Lücken entlang einer Linie. Die wesentlichen Einflussparameter des Algorithmus sind:

1. Schwellwert t : Dieser Parameter entscheidet, welchen Überdeckungswert ein Punkt mindestens besitzen muss, damit er als möglicher Linienpunkt berücksichtigt wird;

2. Maximaler relativer Versatz Δ_{max} : Dieser Parameter bestimmt, welcher relative Versatz innerhalb von Linien erlaubt ist. Wird er groß gewählt, verschmelzen Segmente zu einer gemeinsamen Linie. Wird er klein gewählt, zerfallen Linien in unterbrochene Segmente. Der Parameter stellt daher eine Abhängigkeit zum Rauschanteil der Daten dar.
3. Schwellwert m : Filtert das Rauschen im LPHT-Akkumulator, welches durch schwache Konnektivität zwischen Punkten zustande kommt.

Nach Berechnung der relativen Konnektivität entsprechen die zwei höchsten Akkumulatorwerte dem Start- und Endpunkt der Linie. Anschließend werden die zur Linie gehörenden Punkte aus dem Histogrammbild und dem dazugehörigen Hough-Raum entfernt, sodass weitere, schwächere Linien in den folgenden Zyklen gefunden werden können. Nach Erreichen des Abbruchkriteriums endet die Liniensextraktion. Abbildung 4.9 zeigt beispielhaft alle extrahierten Liniensegmente eines Tiefenbildes.

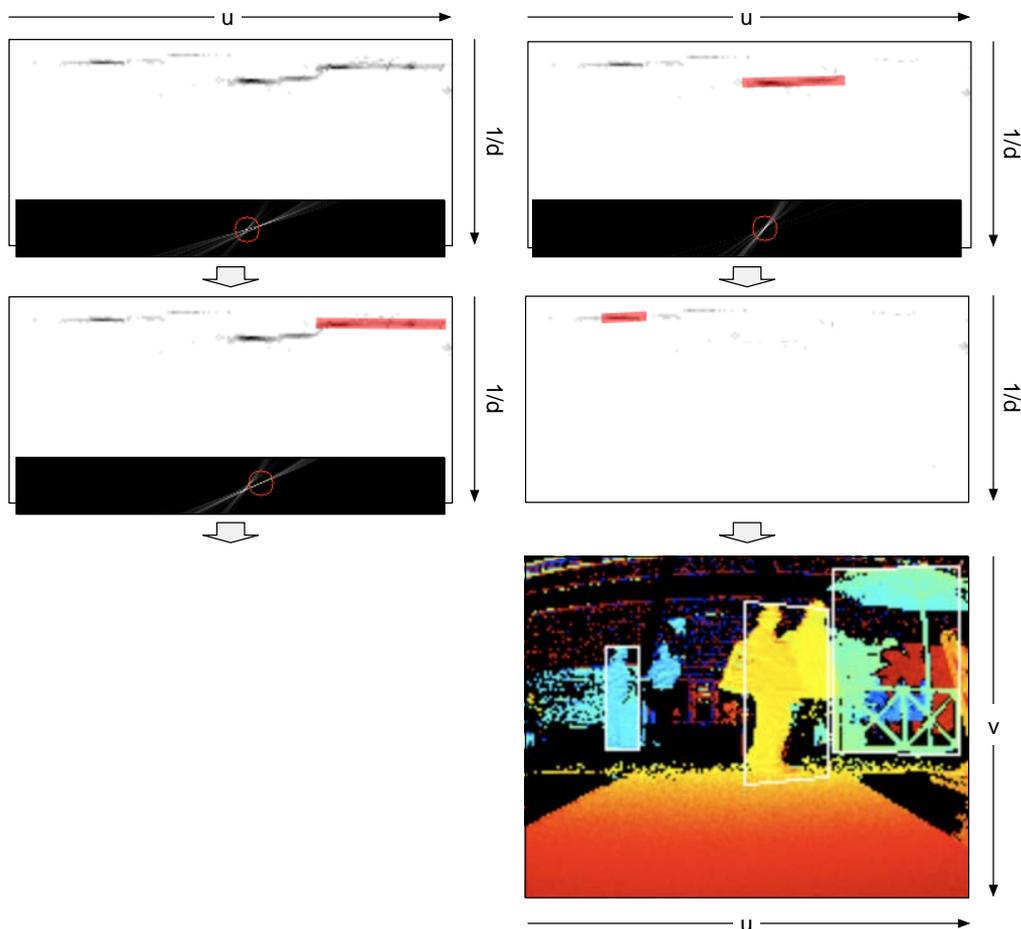


Bild 4.9: Fortlaufend werden extrahierte Linien aus dem Hough-Raum und dem U -Raum entfernt; nach dem dritten Durchlauf wurden alle deutlichen Linien im U -Raum segmentiert.

Objektflächen und befahrbare Flächen Der Start- und Endpunkt einer Linie im U -Raum entspricht der linken und rechten Kante einer Objektfläche im Tiefenbild. Zur Bestimmung der oberen und unteren Kante der Fläche wird für jede Linie ein vertikales Überdeckungshistogramm berechnet, indem für jeden Tiefenwerte entlang der Bildzeile ein Überdeckungswert

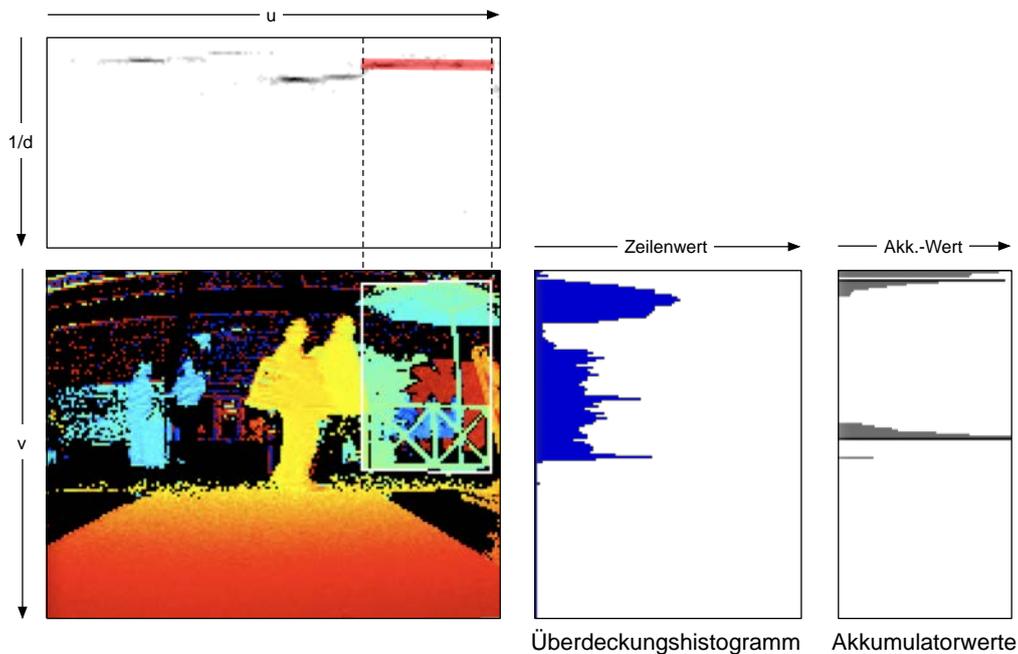


Bild 4.10: LPHT zur Bestimmung der vertikalen Ausdehnung von Objekten; oben: Betrachtetes Liniensegment; rechts: Vertikales Überdeckungshistogramm und berechnete Akkumulatorwerte der vertikalen Start-/Endpunktsuche; unten links: resultierende Fläche als weiße Box umrandet.

akkumuliert wird, falls dieser mit dem berechneten Tiefenwert der Linie übereinstimmt. Anschließend wird der LPHT-Algorithmus auf dem vertikalen Überdeckungshistogramm ausgeführt, wobei die resultierenden Maxima der oberen und unteren Flächengrenze entsprechen. Damit ist die Fläche eines Objektes vollständig bestimmt. Eine beispielhafte Flächensegmentierung ist in Abbildung 4.10 dargestellt.

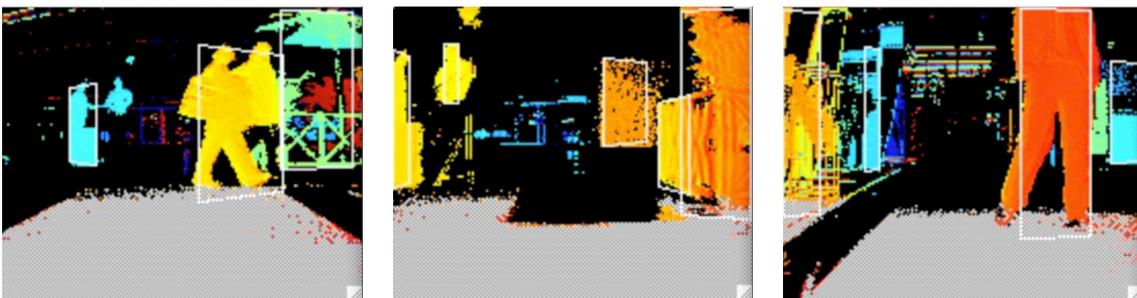


Bild 4.11: Ergebnisse der segmentierten Objektflächen und der befahrbaren Bodenfläche; Objektflächen sind mit weißen Boxen umrandet, die befahrbare Fläche ist grau texturiert dargestellt.

Die Bestimmung befahrbarer Ebenen erfolgt sehr ähnlich. Die extrahierten Linien im V -Raum entsprechen den Flächen, welche parallel zur y -Achse liegen. Der Nickwinkel θ zwischen den Flächen und der Kamera sowie der vertikale Versatz z_{offset} folgen direkt aus dem Parameter (ω, r) der Liniengleichung im Hough-Raum des V -Raumes. Flächen, welche einen Nickwinkel kleiner θ_{max} und einen Versatz kleiner $z_{\text{offset max}}$ zur Kamera besitzen, werden als befahrbare Flächen angenommen. Anschließend werden zeilenweise die Punkte des Tiefenbildes bestimmt, deren Tiefenwert dem Tiefenwert der befahrbaren Fläche entspricht. In Ab-

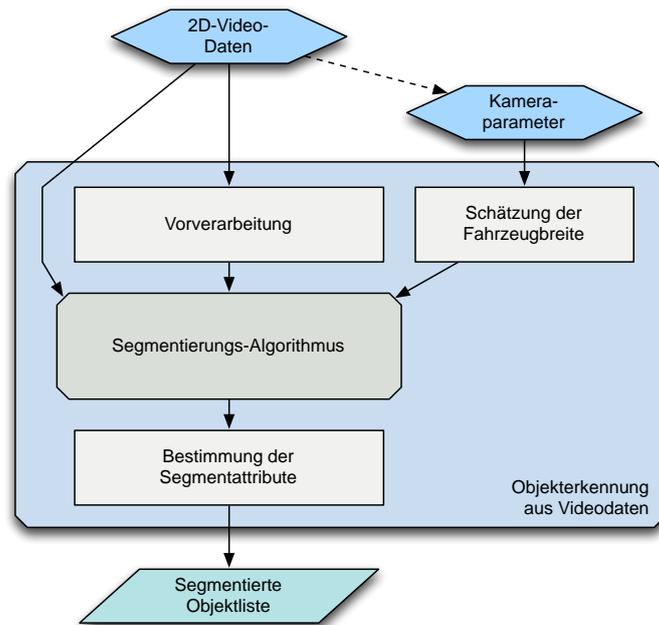


Bild 4.12: Konzeptioneller Ablauf der videobasierten Objekterkennung, das Resultat ist eine Liste segmentierter Objekte.

Abbildung 4.11 sind die Objektflächen mit weißen Boxen markiert. Die befahrbare Bodenfläche ist grau texturiert dargestellt.

4.3 Videobasierte Objekterkennung

Die Verfahren der videobasierten Objekterkennung nutzen monoskopisch gemessene Videodaten zur Erzeugung von Objekthypothesen. Dazu werden Modellannahmen über mögliche Objektarten (*Fahrzeuge, Fußgänger*) getroffen und ansichtsbasierte Merkmale gewählt, welche eine diskriminative Beschreibung der Objektart in den Bilddaten erlauben. Für Objekthypothesen lässt sich eine sehr genaue laterale Position in den Videodaten bestimmen. Außerdem können die Videodaten gleichzeitig für mehrere Assistenzfunktionen wie zur Geschwindigkeitszeichenerkennung (TSR) oder für Spurhalteassistenten (LKS) eingesetzt werden.

Aufgrund der höheren Komplexität bei der Extraktion ansichtsbasierter Merkmale ist mit einer höheren Laufzeit der Verfahren zu rechnen. Eine Schätzung der Entfernung von Objekten ist nur anhand von Modellwissen möglich. Dies kann anhand der Relation zwischen gemessener Objektausdehnung in Bildpunkten und erwarteter Objektgröße stattfinden. Alternativ kann die Entfernung anhand des Kontaktpunktes eines Objektes auf der Fahrbahnoberfläche bei Kenntnis der 3D-Lage der Oberfläche bestimmt werden. Der Ablauf der videobasierten Objekterkennung ist in Abbildung 4.12 dargestellt.

4.3.1 Fahrzeugerkennung bei Tag

Der videobasierten Fahrzeugerkennung bei Tag liegt die Annahme zugrunde, dass Fahrzeuge im nahen Fahrzeugumfeld mittels ansichtsbasierter Merkmale in Videodaten erkennbar sind. Die Schätzung initialer Objekthypothesen folgt dem in [Hof06] vorgeschlagenen Verfahren,

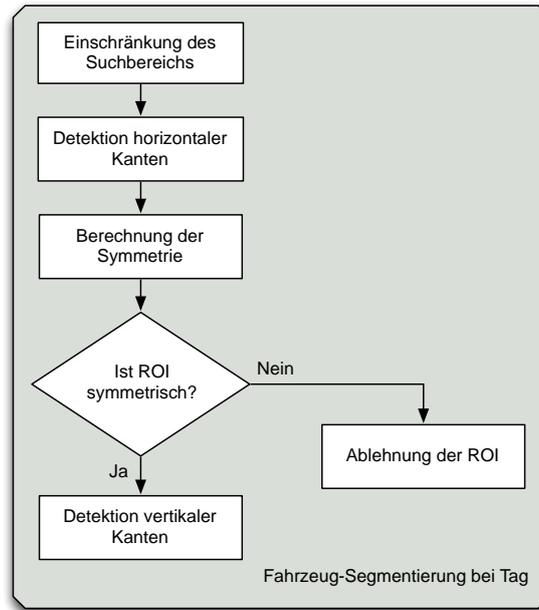


Bild 4.13: Konzept der videobasierten Fahrzeugerkennung bei Tag.

unterscheidet sich aber durch anschließende Verifikation von Hypothesen durch Klassifikation mit ansichtsbasierten Merkmalen (vergleiche Abschnitt 4.4.2).

Die Erkennung ist auf Front- beziehungsweise Heckansichten von Fahrzeugen beschränkt. Ansätze zur Erkennung querender Fahrzeuge mittels ansichtsbasierter Merkmale werden in dieser Arbeit nicht aufgegriffen. Das Konzept zur Erkennung von Fahrzeugen bei Tag ist in Abbildung 4.13 gezeigt. Anhand der zugrundeliegenden Kameraparameter und Modellannahmen werden Schätzwerte initialisiert, welche die Abhängigkeit zwischen der erwarteten Ansicht eines Fahrzeugs und der Position im Videobild herleiten. Aufgrund der perspektivischen Betrachtung der Szene nimmt mit steigender Entfernung eines Fahrzeugs die beobachtete Breite $width_{2D}$ in Bildpunkten ab (siehe Abbildung 4.14). Diese Beziehung kann aus den bekannten Eigenschaften der Kamera abgeleitet werden:

$$width_{2D} = \frac{width_{3D} \cdot f_u}{\hat{d}}, \quad (4.8)$$

wobei \hat{d} der geschätzten Entfernung des Objektes, $width_{3D}$ der angenommenen Fahrzeugbreite in Metern und f_u der Fokallänge der Kamera entspricht. Die Fahrzeugbreite wird als restriktive Bedingung für die Merkmalsextraktion eingesetzt. Innerhalb der Fahrzeugklasse wird die Breite mit 1,8 m als konstant angenommen, was aufgrund der geringen Variation der tatsächlichen Fahrzeugbreiten zulässig ist.

Die korrekte Schätzung der Fahrzeugbreite in Bildpunkten beeinflusst die Erkennungsgüte der vertikalen Kanten- und Symmetriemerkmale erheblich. Die Schätzung der Entfernung des Objektes \hat{d} basiert auf der Annahme, dass die Lage der Fahrbahnebene vor dem Fahrzeug bekannt ist. Zunächst wird die Entfernung eines Fahrzeugs zur Kamera anhand der Position auf der Fahrbahnebene unter Berücksichtigung der Kameraparameter aus Abschnitt 3.1.1 ge-

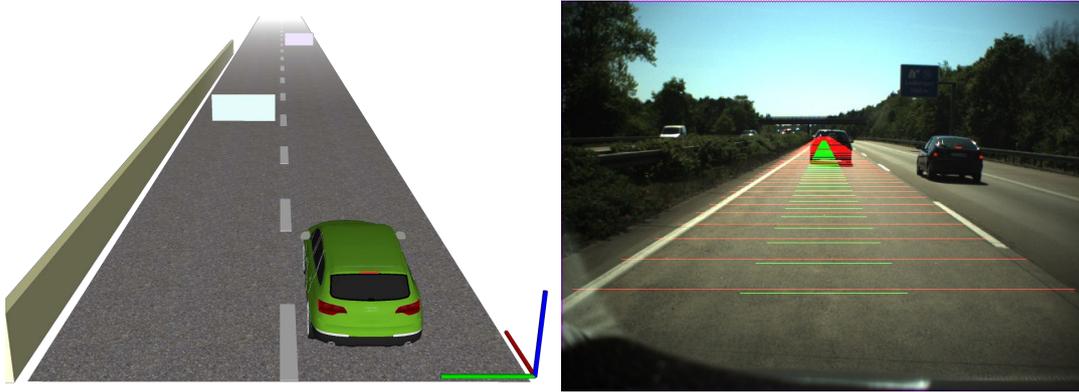


Bild 4.14: Links: Objekte gleicher Breite besitzen mit steigender Entfernung eine geringere Breite in Bildpunkten; rechts: Erwartete minimale und maximale Breite der Unterkante von Fahrzeugen, ermittelt anhand intrinsischer und extrinsischer Kameraparameter.

schätzt. Die Bildzeile v steht mit der Entfernung des Objektes in folgendem Zusammenhang:

$$\hat{d} = c_x + \frac{c_z}{c_z - p_z(v)} \cdot (p_x(v) - c_x), \quad (4.9)$$

wobei sich $\mathbf{p}(v)$ aus Gleichung (3.1) und (3.3) für den Bildpunkt $\mathbf{i} = (0, v)^T$ und $d = 1,0$ m ergibt.

Als Merkmale zur Initialisierung von Fahrzeughypothesen kommen horizontale und vertikale Kanten sowie das Symmetriemerkmal zum Einsatz, da sich Front- und Heckansichten von Fahrzeugen durch diese Kanten und eine symmetrische Struktur hervorheben. Der Ablauf der Segmentierung erfolgt vom Groben zum Feinen: Zunächst wird nach horizontalen Kanten, insbesondere der Schattenkante eines Fahrzeugs auf der Fahrbahnoberfläche gesucht. Dies erlaubt eine präzise initiale Detektion der Fahrzeughypothese, solange die Länge der Kante der erwarteten Fahrzeugbreite entspricht. Anschließend wird mit der Suche nach der Symmetrieachse fortgefahren, sodass die Mitte des Fahrzeugkandidaten präzise bestimmt werden kann. Schließlich folgt die Suche nach vertikalen Kanten zur Bestimmung der linken und rechten äußeren Kante der Hypothese. Das Ergebnis der Hypothesengenerierung ist die Region of Interest (ROI) potentieller Fahrzeuge im Videobild.

Detektion von horizontalen Kanten Lässt sich für ein Fahrzeug mindestens eine starke horizontale Kante beobachten, beispielsweise der Schlagschatten des Fahrzeugs auf der Fahrbahn, dann kann darauf aufbauend eine Hypothese über die Position des Fahrzeugs im Fahrzeugkoordinatensystem abgeleitet werden. Die Kantendetektion erfolgt in dem festgelegten Suchfenster, wobei der Gradient der Kante in horizontaler und vertikaler Richtung der Orientierung der Kante entspricht. Eine Kante, deren Orientierung innerhalb eines eingestellten Orientierungsintervalls liegt, deren Betrag überhalb eines festgelegten Schwellwertes liegt und deren Länge der zulässigen Fahrzeugbreite in Bildpunkten entspricht, wird als zulässige Kante bezeichnet. Jede zulässige Kante erzeugt eine ROI, für welche nachfolgend der Symmetriewert bestimmt wird.

Berechnung der Symmetrie Ausgehend von der resultierenden ROI der horizontalen Kantensuche erfolgt eine lokale Symmetrieberechnung. Anhand der erwarteten Breite und Höhe eines Fahrzeugs lässt sich ein Symmetriefilter initialisieren, welcher bei ebener Straße für achsensymmetrische Objektansichten die höchste Filterantwort erzeugt, wenn die Objektansicht in Bildpunkten mit der berechneten Breite des Filters übereinstimmt. Weicht die Breite eines Objektes in Bildpunkten von der erwarteten Breite ab, sinkt der Wert der Filterantwort. Zunächst wird entlang jeder Bildzeile der ROI deren Symmetriewert gebildet und anschließend werden die Werte spaltenweise akkumuliert. Der Symmetriewert einer Bildzeile, ausgehend von einem Bildpunkt, wird bestimmt indem die Intensitätswerte der Bildpunkte links und rechts verglichen werden. Die Symmetrieberechnung entspricht also der Bestimmung der Achsensymmetrie.

Der symmetrische Bereich kann, wie in [Hof06] beschrieben, durch einen geraden und einen ungeraden Funktionsanteil beschrieben werden. Der gerade Anteil ist für den zu untersuchenden Punkt (u, v) und den horizontalen Versatz Δ_u gegeben durch:

$$E(\Delta_u, u, v) = \frac{1}{2} (g(u + \Delta_u, v) + g(u - \Delta_u, v)), \quad (4.10)$$

der ungerade Anteil durch:

$$O(\Delta_u, u, v) = \frac{1}{2} (g(u + \Delta_u, v) - g(u - \Delta_u, v)). \quad (4.11)$$

Der ungerade Anteil $O(\Delta_u, u, v)$ sollte also für symmetrische Strukturen im Gegensatz zum geraden Anteil mittelwertfrei sein. Über einen Bereich der Breite $2w$ ist der mittelwertbereinigte gerade Anteil gegeben durch:

$$E_n(\Delta_u, u, v, w) = E(\Delta_u, u, v) - \frac{1}{2w} \int_{-w}^w E(\Delta_u, u, v) d\Delta_u. \quad (4.12)$$

Zur Bewertung der Signifikanz des geraden und des ungeraden Anteils in dem zu untersuchenden Bereich, wird eine entsprechende Signalenergie als Symmetriemaß S berechnet:

$$S(u, v, w) = \frac{\int_{-w}^w E_n(\Delta_u, u, v, w)^2 d\Delta_u - \int_{-w}^w O(\Delta_u, u, v)^2 d\Delta_u}{\int_{-w}^w E_n(\Delta_u, u, v, w)^2 d\Delta_u + \int_{-w}^w O(\Delta_u, u, v)^2 d\Delta_u}. \quad (4.13)$$

Befindet sich die vertikale Symmetrieachse in der Mitte der umschließenden Box (BBox), wird die Hypothese akzeptiert.

Detektion vertikaler Kanten Abschließend werden die seitlichen Kanten der Fahrzeughypothese gesucht und die BBox verbessert. Eine Fahrzeughypothese sollte möglichst vollständig in der BBox erscheinen und möglichst mittig durch die BBox beschrieben werden. Die untere horizontale Kante der BBox bleibt unverändert.

Entlang den Bildspalten der BBox wird der lokale vertikale Kantenbetrag aufsummiert. Die äußeren Begrenzungen entsprechen den zwei Maxima der Spaltensumme. Eine Fahrzeughy-

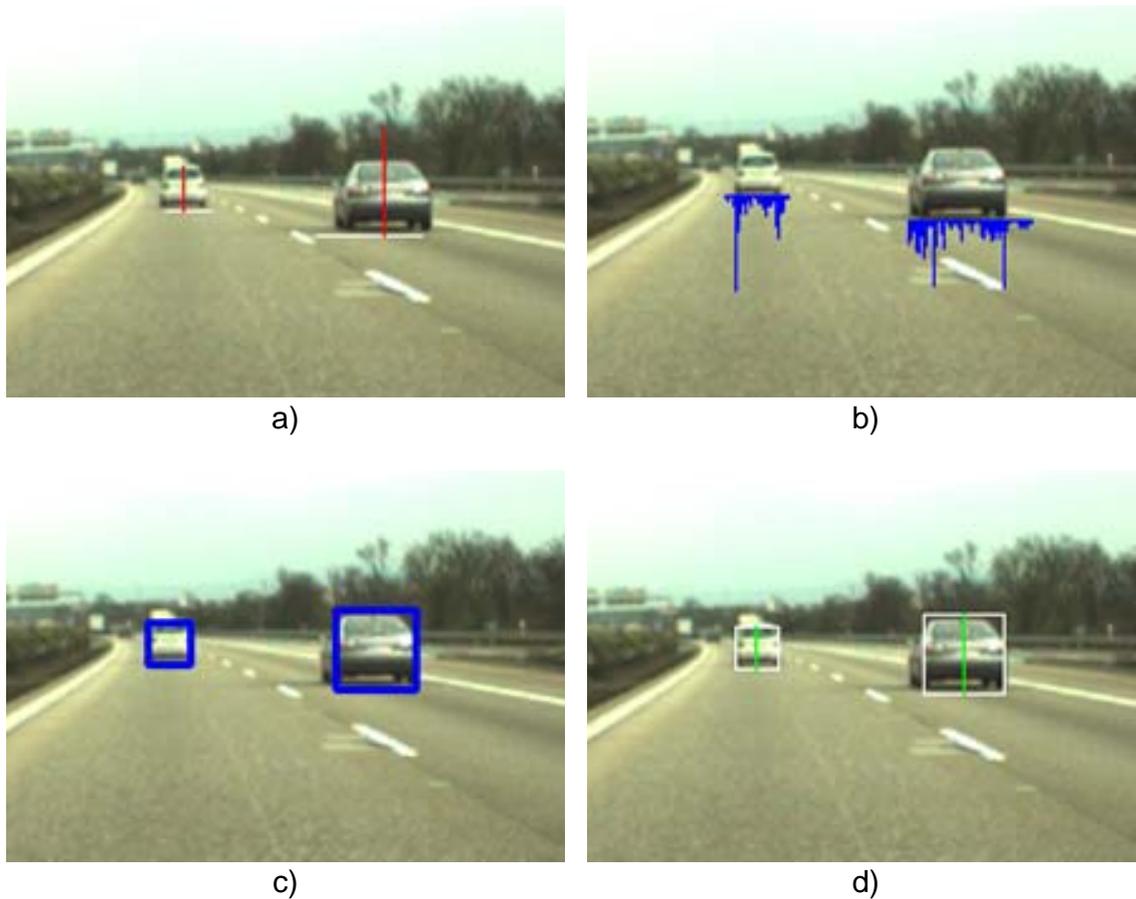


Bild 4.15: Verarbeitungsschritte der Fahrzeugerkennung bei Tag: a) Bestimmung der horizontalen Kante (weiß) und vertikalen Symmetrieachse (rot); b) Bestimmung der vertikalen Kanten; c) Korrektur der ROI; d) Erstellte Fahrzeughypothese.

pothese ist vollständig beschrieben durch die untere horizontale Kante, die Symmetrieachse sowie die beiden seitlich begrenzenden vertikalen Kanten.

Die Verarbeitungsschritte der videobasierten Fahrzeugerkennung sind in Abbildung 4.15 gezeigt.

4.3.2 Fahrzeugerkennung bei Nacht

Bei Nacht sind die oben vorgestellten Merkmale, falls überhaupt, nur schwer zu erkennen (vergleiche Abbildung 4.16). Dennoch ist eine videobasierte Erkennung von Fahrzeugen bei Dunkelheit erreichbar. Dafür werden die Lichter als Merkmale genutzt. Dieser Abschnitt stellt die in [Car09; SCZ10] beschriebene Methode vor.

Das Konzept zur Segmentierung von Fahrzeugen bei Dämmerung und Dunkelheit lässt sich in drei Abschnitte gliedern, wie in Abbildung 4.17 dargestellt ist. Für die Segmentierung werden in den Videodaten Fahrzeuglichter gesucht, welche durch hell im Videobild auftretende Lichtflecken charakterisiert sind. Die Videodaten werden dazu mit einer Filtermaske gefaltet. Diese basiert auf einer 2D-Gaussfunktion. Die Breite der Maske wird in Abhängigkeit zur erwarteten Größe des Lichtflecks gewählt. Im zweiten Schritt werden Lichtflecken entsprechend ihrer Position im Videobild durch einen regelbasierten Ansatz miteinander zu Lichtpaaren kombiniert. Zuletzt wird die Güte der ROI jedes Lichtpaares durch eine Symmetrieprüfung

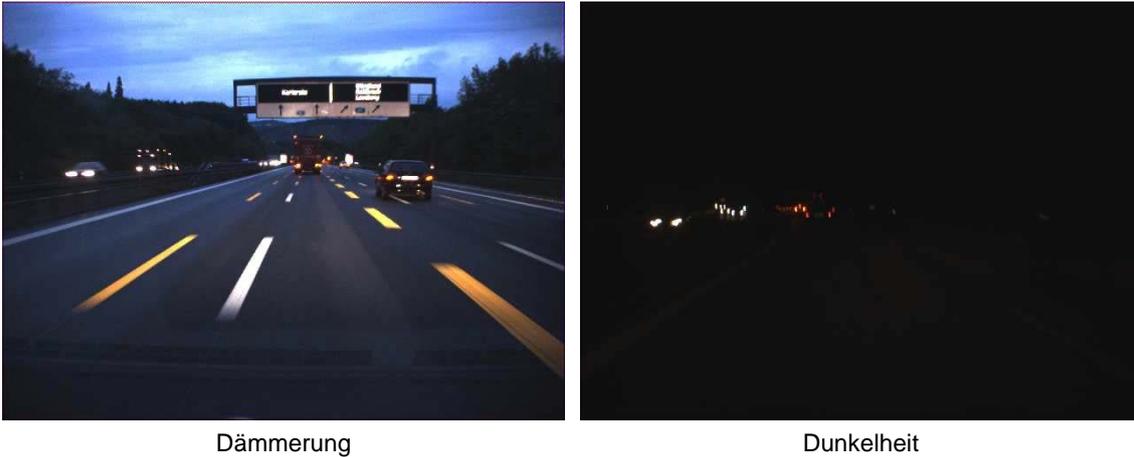


Bild 4.16: Aufgezeichnete Videobilder einer Farbkamera bei Dämmerung und bei Dunkelheit.

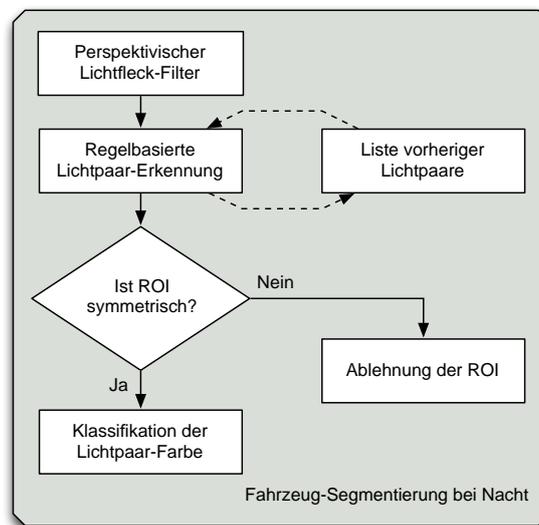


Bild 4.17: Konzeptioneller Ablauf der videobasierten Fahrzeugerkennung bei Nacht.

bewertet. Abhängig von den verwendeten Kameradaten stehen Farbinformationen zur Verfügung. Hiermit ist eine Differenzierung zwischen den verschiedenen auftretenden Lichttypen, also Front-, Heck- oder Bremsleuchten, möglich.

Bildvorverarbeitung Die gründliche Bildvorverarbeitung ist besonders wichtig, da die wechselnden Lichtbedingungen bei Dämmerung und Dunkelheit nicht ausreichend durch kamera-interne Funktionen ausgeglichen werden können. Es ist notwendig, die Videodaten zunächst zu normalisieren. Dies führt zu stabileren Segmentierungsergebnissen.

Zunächst wird eine Farbraum-Konvertierung des Bildes in Graustufen vorgenommen. Anschließend wird durch ein iteratives Verfahren der Bildkontrast so lange verringert, bis ein mittlerer Kontrastwert unterschritten wird. Die Anpassung erfolgt jeweils durch Subtraktion der mittleren Bildintensität von dem Graustufenbild. Helle Bilder werden hierdurch stärker beeinflusst als dunkle Bilder.

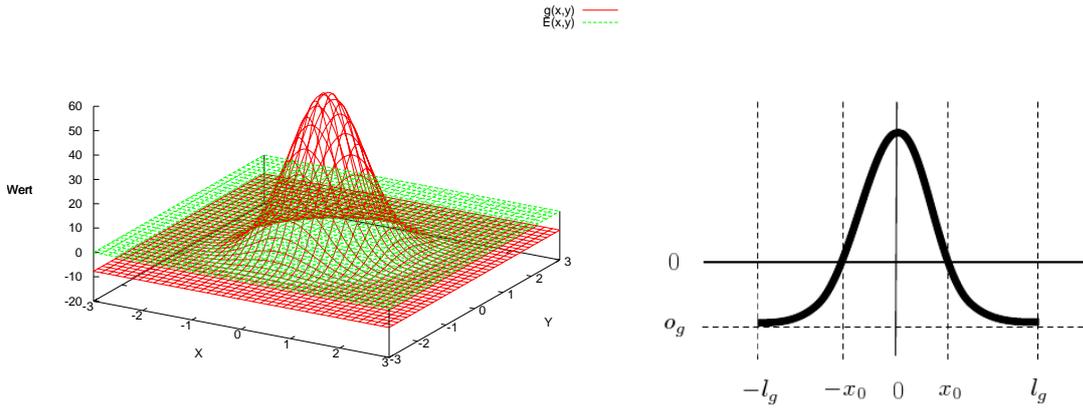


Bild 4.18: Links: 3×3 Maske für 2D Gauss-Filterkernel; rechts: Schnitt durch den Filterkernel $g(x) - o_g$ im Intervall $[-l_g, l_g]$.

Segmentierung heller Lichtflecken Als Filterfunktion wird eine 2D-Gaussfunktion verwendet, deren Breite an die erwartete Breite von Fahrzeuglichtern im Kamerabild angepasst wird (siehe Abbildung 4.18 links). Die Filterfunktion begünstigt die Suche nach hellen, kreisrunden Bereichen. Die Größe eines Lichtflecks ist abhängig von der Entfernung \hat{d} des Fahrzeugs. Unter Kenntnis der Fahrbahnebene, der Kameraparameter und der erwarteten Höhe der Lichter \hat{h}_{light} über der Fahrbahn beträgt \hat{d} in Abhängigkeit von der Bildzeile v :

$$\hat{d} = c_x + \frac{c_z - \hat{h}_{\text{light}}}{c_z - p_z(v)} \cdot (p_x(v) - c_x). \quad (4.14)$$

Vergleiche hierzu Gleichung (4.9). Wird \hat{d} in Gleichung (4.8) eingesetzt, kann die Breite des Filterkerns für nahe Fahrzeuge bestimmt werden. Für weit entfernte Fahrzeuge wird eine minimale Filtergröße definiert.

Der 2D Gauss-Filterkernel ist angelehnt an eine 1D Filtermaske, wie diese beispielsweise zur Segmentierung von Fahrbahnmarkierungen in [Gum+09] beschrieben ist. Gleichung (4.15) definiert eine Gaussverteilung $g(x)$ mit Erwartungswert μ und Varianz σ und die Verteilungsfunktion $c_g(x)$:

$$g(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4.15)$$

$$c_g(x) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x - \mu}{\sigma \sqrt{2}} \right) \right) \quad (4.16)$$

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (4.17)$$

Von der Gaussglocke wird ein Versatz o_g abgezogen, sodass das Integral über der Glockenfunktion Null ergibt. Die negativen und positiven Flächen zwischen der Glocke und der x -Achse im Intervall $[-l_g, l_g]$ sind also identisch (siehe Abbildung 4.18 rechts). Mit Hilfe der

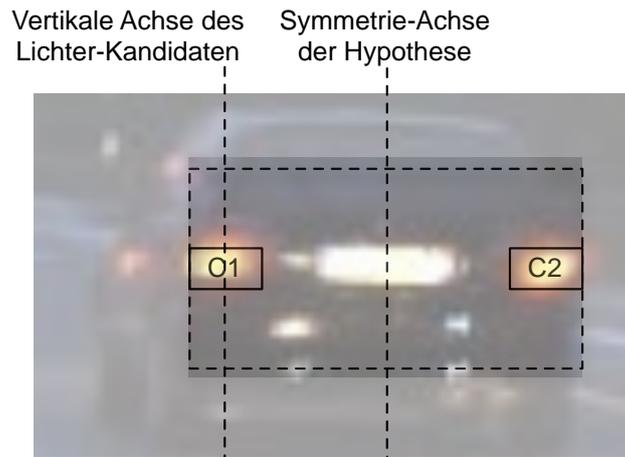


Bild 4.19: Schätzung der Ähnlichkeit zwischen zwei Lichter-Kandidaten C_1, C_2 zur Bestimmung der besten Fahrzeughypothese.

Verteilungsfunktion $c_g(x)$ lassen sich die Nulldurchgänge $\pm x_0$ wie folgt bestimmen:

$$0 = (l_g - x_0) \cdot g(x_0) - [c_g(-x_0) - c_g(-l_g)] - [c_g(0) - c_g(-x_0) - (x_0 \cdot g(x_0))], \quad (4.18)$$

$$o_g = g(x_0). \quad (4.19)$$

Das Filter unterdrückt homogene Intensitäten, während Intensitätsverteilungen in der Art von Lichtern hervorgehoben werden. Die Filtermasken werden aus Laufzeitgründen für die einzelnen Entfernungsstufen vorberechnet. Anhand der Filterantwort wird ein Binärbild durch das Schwellwertverfahren erstellt.

Regelbasierte Kombination von Lichterpaaren Die im Binärbild vorhandenen Einträge entsprechen einzelnen Lichthypothesen. Zunächst werden für die Einträge deren Kontur sowie deren BBox berechnet. Jede BBox wird dann als Lichter-Kandidat betrachtet, die Menge aller Lichter-Kandidaten ist $\mathbf{C} = \{C_1, \dots, C_n\}$. Jeder Lichter-Kandidat C_i besitzt eine Farbeigenschaft L_{rgb} , die Position des Kandidaten im Bild (u, v) , sowie die Größe des Lichtes (w, h) .

Für die Aktualisierung von Fahrzeughypothesen werden Lichter-Kandidaten \mathbf{C}' der vorhergehenden Verarbeitungsschritte hinzugezogen, was die Robustheit des Systems steigert. Jede Fahrzeughypothese besitzt die folgenden Zustandseigenschaften:

- Ein Minimum von zwei erkannten Lichtflecken $L_1, L_2 \in \mathbf{C}'$,
- optional ein dritter Lichtfleck $L_3 \in \mathbf{C}'$, welcher dem Bremslicht entspricht.

Nun wird die Ähnlichkeit der aktuellen Lichter-Kandidaten zu den vorhergehenden Lichter-Kandidaten bestimmt. Dabei muss eine Ähnlichkeit sowohl in der erwarteten Position der Kandidaten vorliegen, als auch der Größen-Unterschied zwischen den Kandidaten gering sein. Ist eine Zuordnung gültig, wird die Fahrzeughypothese weitergeführt. Dabei ist es möglich, dass Fahrzeughypothesen mit nur einem korrekt zugeordneten Licht weitergeführt werden.

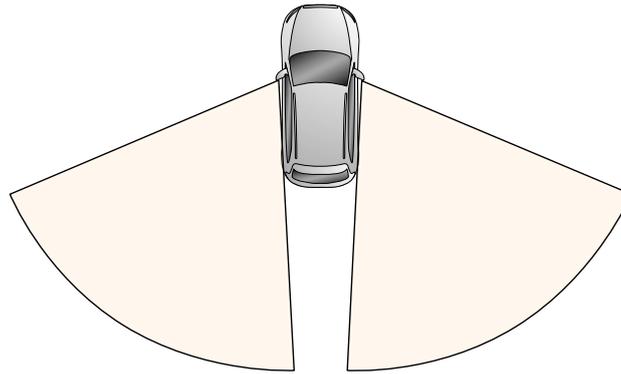


Bild 4.20: Sensorische Abdeckung des Fahrzeugumfelds durch Videokameras im Toten-Winkel Bereich.

Neue Fahrzeughypothesen werden aus Lichter-Kandidaten gebildet, welche nicht zu vorherigen Fahrzeugkandidaten zugeordnet werden. Dafür werden anhand von Regeln gemeinsame Paare $C_1, C_2 \in \mathbf{C}, C_1 \neq C_2$ gesucht:

- Für zwei Lichter-Kandidaten wird zunächst deren vertikale Position verglichen;
- anschließend wird das Größenverhältnis zwischen beiden Lichter-Kandidaten verglichen.

Sowohl der vertikale Positionsunterschied als auch das Größenverhältnis dürfen bestimmte Schwellwerte nicht verletzen. Oftmals werden mehrere mögliche Lichter-Paare gefunden. Mit Hilfe der normalisierten Summe quadratischer Differenzen wird die Ähnlichkeit der beschriebenen Fläche zwischen zwei Lichter-Kandidaten C_1 und C_2 berechnet. Abbildung 4.19 zeigt dazu beispielhaft zwei Lichter-Kandidaten.

Anhand des Farbwertes L_{rgb} der Licht-Kandidaten lassen sich falsche Fahrzeughypothesen unterdrücken. Auf Basis des HSV-Farbraums¹ werden die Lichter anhand Sättigung und Farbwert in die Klassen *weiß*, *rot* und *undefiniert* getrennt. Zur Klassifikation wird der Mittelwert der Region des Lichter-Kandidaten bestimmt. Stimmen die Lichter-Farben überein, so kann zwischen entgegenkommenden Fahrzeugen (weißes Licht) und vorausfahrenden Fahrzeugen (rotes Licht) unterschieden werden. Die Suche nach einem dritten Bremslicht erfolgt nur, wenn die Lichter einer Fahrzeughypothese übereinstimmend als *rot* klassifiziert werden und somit eine Fahrzeug-Rückseite beobachtet wird.

4.3.3 Fahrzeugerkennung auf Nebenspuren

Die videobasierte Objekterkennung kann durch Anpassungen des Verfahrens zur Erkennung von Fahrzeugen neben oder hinter dem eigenen Fahrzeug verwendet werden. Somit lassen sich auch neben dem Fahrzeug befindliche Fahrbahnen überwachen. Es wird unterschieden, ob die Fahrspur frei oder durch ein anderes Fahrzeug belegt ist. In [Zhu10] ist der Ablauf der Erkennung von Fahrzeugen auf Nebenspuren mittels ansichtsbasierter Merkmale und einer flussbasierten Weiterverfolgung beschrieben. Die Anbringung der rückwärts gerichteten

¹Farbraum, welcher nach Farbwert, Sättigung und Intensität getrennt ist; im Gegensatz zu RGB-Farbraum, welcher nach den Farben Rot, Grün und Blau getrennt ist.

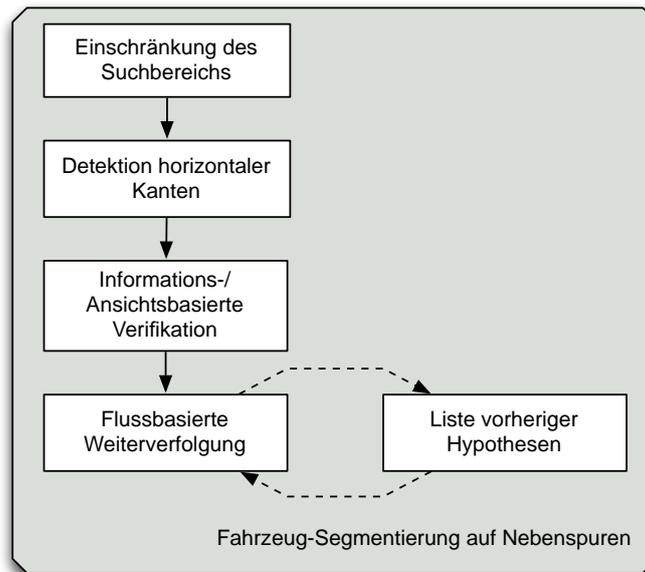


Bild 4.21: Konzeptioneller Ablauf der videobasierten Erkennung von Fahrzeughypothesen auf Nebenspur.

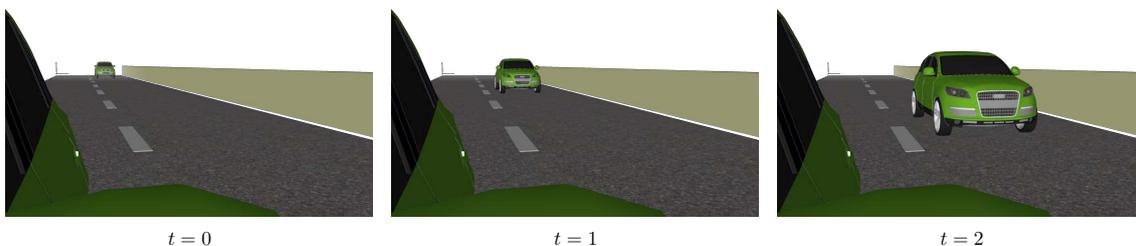


Bild 4.22: Perspektivische Veränderung während der Annäherung eines Fahrzeugs auf der Nachbarspur.

Kameras ist in Abbildung 4.20 gezeigt. Durch diese Position lässt sich der Tote-Winkel Bereich des Fahrzeugs optimal überwachen.

Der konzeptionelle Ablauf ist in Abbildung 4.21 dargestellt. Aufbauend auf der Erkennung von Fahrzeugen bei Tag (siehe Abschnitt 4.3.1) wird das Verfahren an die besonderen Eigenschaften der Überwachung von Nebenspur angepasst. Die Herausforderung der Objektsegmentierung liegt darin, dass durch die Bewegung eines Fahrzeugs auf einer Nachbarspur die Ansicht des Fahrzeugs im Kamerabild perspektivisch verzerrt wahrgenommen wird. Dieser Effekt ist in Abbildung 4.22 dargestellt. Während das Fahrzeug zum Zeitpunkt $t = 0$ noch sehr zuverlässig mittels Merkmalen wie Symmetrie und Kanten erkannt werden kann, ist eine Erkennung zum Zeitpunkt $t = 2$ mittels Symmetrie nicht mehr möglich.

Die Segmentierung erfolgt durch die Detektion der horizontalen Kante und einer daran anschließenden Verifikation der Hypothese. Für die Berücksichtigung der perspektivischen Veränderung wird auf die Hypothesen der vergangenen Zeitschritte zurückgegriffen. Sie werden durch eine deformierbare Gitterstruktur flussbasiert weiterverfolgt. So wird die perspektivische Veränderung der Ansicht von Fahrzeugen beispielsweise während eines Überholvorgangs berücksichtigt.

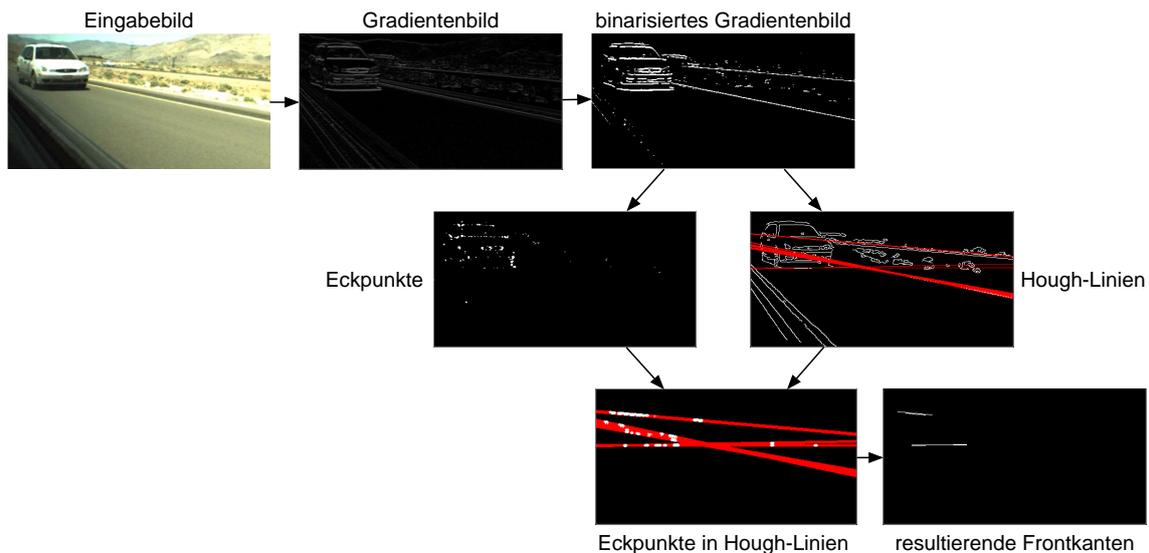


Bild 4.23: Suche nach horizontalen Fahrzeugkanten anhand von Gradientenmerkmalen.

Einschränkung des Suchbereiches Der Vorverarbeitungsschritt zur Fahrzeugsegmentierung auf Nebenspuren besteht aus einer Konvertierung des Eingangsbildes in ein Graustufenbild, einer Kontrastverbesserung und einer affinen Transformation des Bildes. Durch die Transformation entsprechen horizontale Fahrzeugkanten für den erwarteten Spurbereich im transformierten Bild horizontalen Bildkanten. Die affine Transformation lässt sich direkt aus den bekannten Kameraparametern bestimmen.

Zur Verbesserung der Laufzeit wird nicht der gesamte Bildbereich nach potentiellen Fahrzeugen abgesucht, sondern lediglich der Bereich, in welchem Fahrzeuge auf einer Nachbarspur erwartet werden. Die Position von benachbarten Fahrspuren kann durch Sensoren erfasst oder durch Modellwissen geschätzt werden. Eine sensorische Erkennung der Spuren bietet hierfür eine höhere Zuverlässigkeit. Bei Verwendung von Modellwissen kann anhand der bekannten Anbringung der Kamera im Fahrzeug die erwartete Position von Nachbarspuren im Bild berechnet werden.

Detektion der Fahrzeugunterkante Für Fahrzeuge auf Nebenspuren gelten vergleichbare Annahmen wie in Abschnitt 4.3.1. Der Schlagschatten eines Fahrzeuges bildet auf der Fahrbahn eine gut detektierbare Kante und wird zur Initialisierung einer Fahrzeughypothese genutzt (siehe Abbildung 4.23).

Eckpunkte weisen in u - und v -Richtung Intensitätsänderungen auf. Die Intensitätsänderungen lassen sich durch Suche mittels des Harris Detektors im binarisierten Gradientenbild des Eingabebildes extrahieren [HS88]. Hierfür wird ein Strukturtensor gebildet. Ein Eckpunkt ist durch zwei große Eigenwerte λ_1 und λ_2 des Tensors charakterisiert. Die gefundenen Eckpunkte werden durch morphologische Dilatation hervorgehoben.

Im binarisierten Gradientenbild werden starke Kanten mit Hilfe der Hough-Transformation markiert. Die resultierenden Hough-Linien werden dazu genutzt, die Liste der gefundenen Eckpunkte zu filtern. Zur Bestimmung von Kandidaten für die Suche der Frontunterkante werden die Eckpunkte, welche durch starke Kanten gestützt werden, auf ihre Kontinuität hin

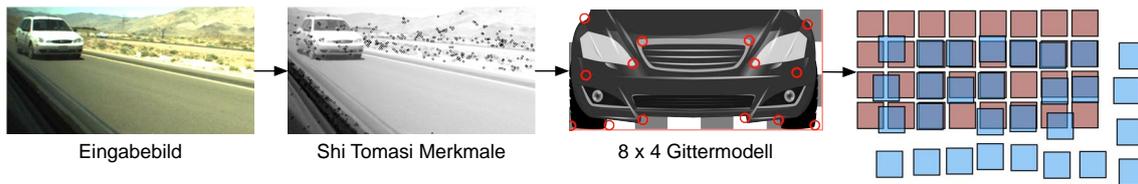


Bild 4.24: Gittermodell mit 8×4 Elementen, Kreise beschreiben im Gitter eingetragene Merkmale.

geprüft. Die resultierenden horizontalen Kanten werden dann als Frontunterkanten weitergeführt. Entsprechend einer Modellannahme wird ausgehend von der Unterkante eine BBox bestimmt, welche Kante die Fahrzeughypothese umschließt.

Modellbasierte Verifikation Das oben beschriebene Verfahren liefert eine Menge initialer Hypothesen. Für jede Hypothese wird überprüft, ob diese ein Fahrzeug beschreibt oder nicht. Die dafür eingesetzten Merkmale umfassen:

- Symmetrieeigenschaft der BBox,
- Varianz der Grauwerte innerhalb der BBox,
- Stärke der vertikalen Kanten der BBox, sowie
- Lage und Größe der Hypothese im Fahrzeugkoordinatensystem.

Die Berechnung der Symmetrieeigenschaft erfolgt für die oben bestimmte BBox wie in Abschnitt 4.3.1 beschrieben. Anschließend wird mittels der Grauwertvarianz geprüft, ob die durch die BBox beschriebene Fläche eine inhomogene Textur beschreibt. Ist dies nicht der Fall, liegt die Annahme nahe, dass durch die BBox kein Fahrzeug, sondern Teile der Fahrbahnoberfläche oder des Himmels beschrieben sind.

An den seitlichen Rändern der BBox wird angenommen, dass deutliche vertikale Kantenstrukturen im Bild sichtbar sind. Zur Überprüfung werden in dem Gradientenbild die Werte aller Kanten über die Bildspalten aufsummiert. Nach Glättung des Kantenprofils durch Filterung mit einem Gaußkernel werden die Ränder der BBox geprüft, ob diese den Maxima der vertikalen Kanten entsprechen.

Zuletzt wird für die Hypothese geprüft, wie deren Lage und Größe im Fahrzeugkoordinatensystem ist. Anhand der Lage der Fahrspur erfolgt eine Rückprojektion der 2D-BBox in die 3D-Fahrzeugkoordinaten. Dabei liegt die Annahme zugrunde, dass die detektierte Fahrzeugunterkante auf der Fahrbahnoberfläche liegt. Die Größe der BBox in Fahrzeugkoordinaten sollte der Breite einer typischen Fahrzeughypothese von etwa 1,8 m entsprechen.

Darauf aufbauend kann eine ansichtsbasierte Verifikation durchgeführt werden. Diese ist in Abschnitt 4.4.2 beschrieben.

Flussbasierte Weiterverfolgung Durch den seitlichen Versatz von Fahrzeugen können die oben beschriebenen Merkmale ab einem Abstand von weniger als 15m nicht mehr zuverlässig erkannt werden. Beispielsweise weist die Ansicht keine Symmetrie mehr auf, die oben genannten Verifikationsschritte würden zur Ablehnung der Hypothese führen. Aus diesem Grund wird

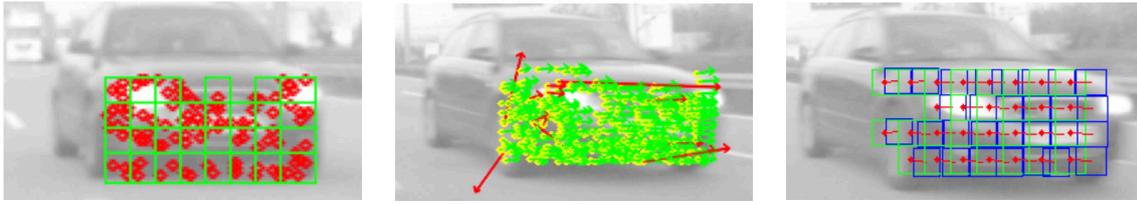


Bild 4.25: Flussbasierte Weiterverfolgung durch Gittermodell: Merkmale werden zunächst der Gitterstruktur zugeordnet, jede Gitterzelle wird anhand des mittleren Flusses individuell weitergeführt, was zu einer Verformung des Gittermodells führt.

für die Verfolgung der Fahrzeuge in geringer Entfernung eine flussbasierte Weiterverfolgung gewählt, welche auf einem deformierbaren Gittermodell beruht (siehe Abbildung 4.24).

Innerhalb dem durch die BBox beschriebenen Rechteck werden zunächst Merkmale gesucht, welche sich über mehrere Bilder gut verfolgen lassen. Zur Suche nach invarianten Merkmalen ist der Shi-Tomasi Operator geeignet. Dieser basiert auf dem oben genannten Struktur-Tensors des Harris Detektors und berechnet aus dem Minimum der Eigenwerte des Tensors ein nahezu eindeutiges Merkmal zur Verfolgung von Bildpunkten [ST94]. Zur Berechnung des optischen Flusses wird die Methode von Lukas und Kanade eingesetzt, bei der es sich um eine lokale Methode mit hoher Robustheit gegen Rauschen handelt. Es kann vorkommen, dass Merkmale bei der Verfolgung im darauffolgenden Bild außerhalb des Bildes liegen oder wegen Verdeckung verschwinden. Durch Template-Matching wird der Bereich um das Merkmal und der Bereich um den korrespondierenden nachfolgenden Merkmalspunkt zur Bestimmung der Güte des Flussvektors miteinander verglichen.

Die gültigen Merkmalspunkte werden in Gitterzellen des deformierbaren Gittermodells eingetragen, wie in Abbildung 4.25 gezeigt ist. Für das Gittermodell liegt die Annahme zugrunde, dass innerhalb einer Gitterzelle die flussbasierte Bewegungsschätzung rein translatorisch ist. Notwendigerweise muss nicht in jeder Gitterzelle ein Merkmalspunkt vorhanden sein. Werden mehrere Merkmalspunkte innerhalb einer Gitterzelle eingetragen, wird über diese der Mittelwert und die Varianz gebildet. Gitterzellen ohne Merkmalspunkte oder mit zu hoher Varianz werden als ungültig markiert. Das deformierbare Gittermodell wird weitergeführt, solange eine ausreichend hohe Anzahl an Gitterzellen als gültig markiert ist.

4.4 Klassifikation

Auf Basis der in Abschnitt 4.2 und 4.3 beschriebenen Verfahren existieren nun Listen von Hypothesen, welche Objekte mit ihrer Position und Ausdehnung im 3D-Raum beschreiben. Eine Klassifikation erlaubt mit Hilfe der Texturinformation der 2D-Kameradaten eine Unterscheidung zwischen den Klassen *Fahrzeugfront*, *Fahrzeugheck*, *Fußgänger* beziehungsweise *Hintergrund*. Die Klassifikationsverfahren besitzen üblicherweise einen hohen Rechenaufwand, was aber durch die geringe Anzahl der Hypothesen und durch Vorwissen über Position und Ausdehnung ausgeglichen wird. Die in Abschnitt 3.1.3 vorgestellten Methoden beschreiben Klassifikationsverfahren mit Support Vektor Maschinen (SVMs) oder Klassifikator-Kaskaden. Hierfür müssen zunächst Modelle anhand von Trainingsdaten berechnet werden.

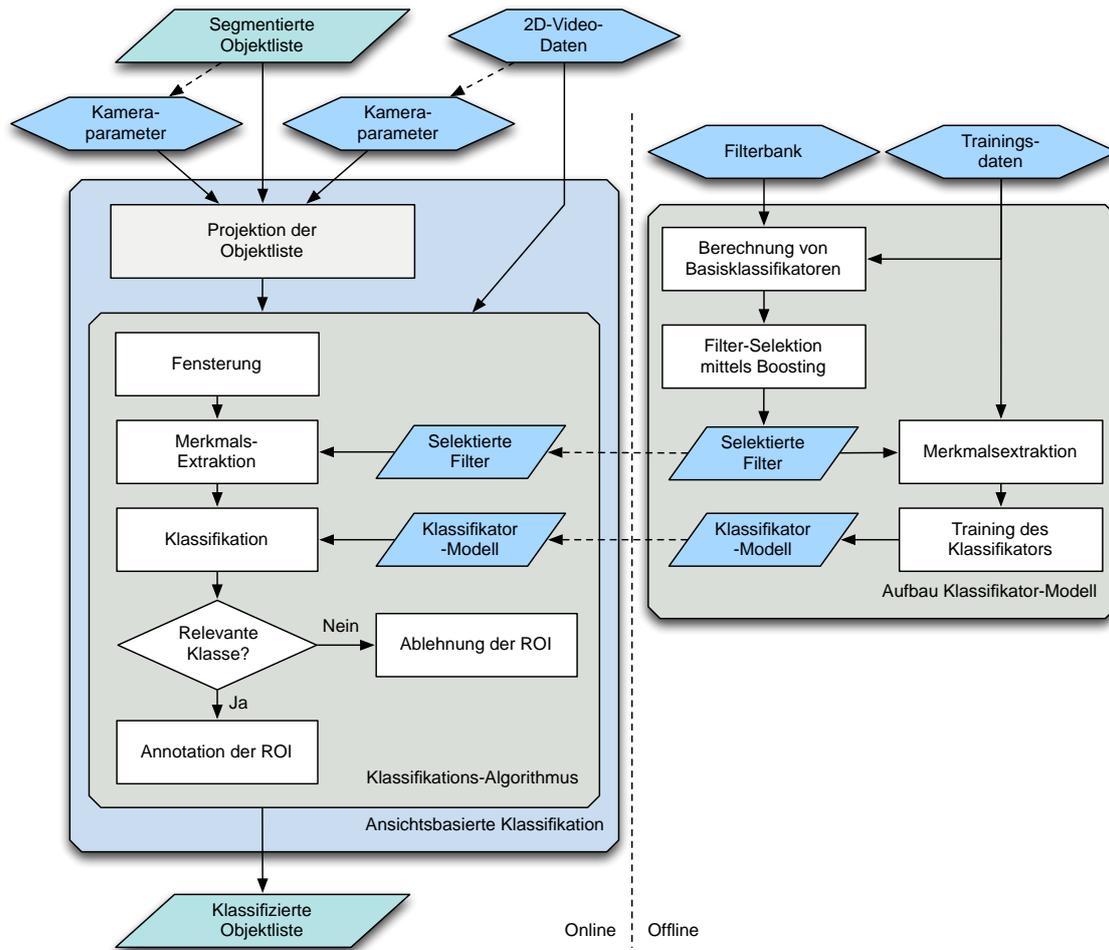


Bild 4.26: Konzeptioneller Ablauf der Klassifikation segmentierter Hypothesen durch ansichtsbasierte Merkmale.

Abbildung 4.26 beschreibt den konzeptionellen Ablauf der Klassifikation. Die Klassifikation erfolgt auf Basis einzelner Bildausschnitte der 2D-Daten, welche die zu untersuchende Hypothese beinhalten. Durch Projektion der BBox aus dem Koordinatensystem der Messsensorik in das 2D-Bild wird eine ROI zur Klassifikation erstellt. Auf der ROI werden ansichtsbasierte Merkmale berechnet und in einen Merkmalsvektor überführt. Abhängig von der Art der zu untersuchenden Klasse unterscheiden sich die Verfahren zur Merkmalsextraktion. Der Merkmalsvektor wird durch einen Klassifikator bewertet, welcher auf ein vorher trainiertes Modell zurückgreift. Das Ergebnis der Klassifikation ist eine Schätzung der Zugehörigkeit zu einzelnen Klassen des Modells. Die Klasseninformation der ROI wird mit der Zugehörigkeitsverteilung aktualisiert und in der Objektliste eingetragen.

4.4.1 Entzerrung flächiger Objekte

Die Projektion von Hypothesen in die 2D-Daten erfolgt auf Basis der gemessenen, sichtbaren Fläche der Hypothesen. Je nach Segmentierungsverfahren kann dies entweder die BBox eines Objektes sein oder eine im Raum stehende Fläche, welche die Hauptfläche eines Objektes repräsentiert. Wurde die Objekthypothese nicht durch den zur Klassifikation verwendeten Sensor generiert, muss die Hypothese in das Videobild des 2D-Sensors projiziert wer-

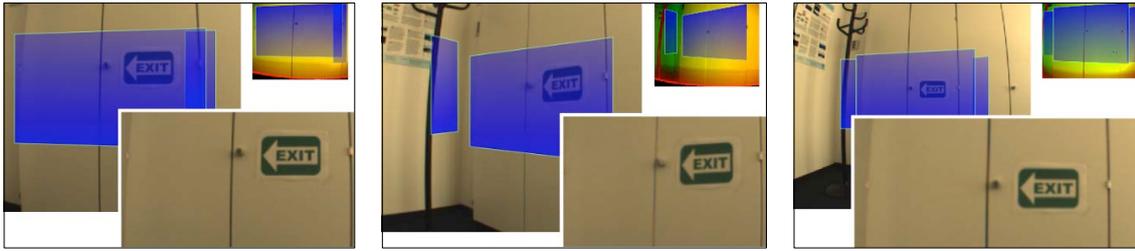


Bild 4.27: Entzerrung von Texturen durch isometrische Projektion anhand schräger Flächen.

den [Koh+10]. Für eine im Raum stehende Fläche lassen sich die Eckpunkte der Fläche in das 2D-Bild transformieren und durch eine ROI beschreiben.

Aus der Parameterdarstellung der 3D-Fläche kann der Sichtvektor zum betrachteten Objekt abgeleitet werden. Eine isometrische Textur-Extraktion ermöglicht eine Entzerrung der Ansicht von Objekten im 2D-Bild. Dies ist besonders dann hilfreich, wenn eine Klassifikation bei schräger Ansicht eines Objektes sehr wahrscheinlich fehlschlägt. Die Tiefeninformation der 3D-Fläche kann außerdem zur Größennormierung des Objektes verwendet werden (siehe Abbildung 4.27). Somit kann eine verbesserte ansichtsbasierte Klassifikation rechteckiger oder runder Formen erreicht werden.

4.4.2 Klassifikation von Fahrzeugen

Das Konzept der bildbasierten Klassifikation von Fahrzeugen folgt dem in [SBM06a] beschriebenen Verfahren. Der Rechenaufwand fällt gegenüber dem beschriebenen Ansatz geringer aus, da die Verfahren nur auf die Verifikation von Hypothesen, also zur Bestimmung der Klassenzugehörigkeit, eingesetzt werden.

Eingabe für das Verfahren ist ein normalisiertes Grauwert-Bild der ROI. Für die Berechnung des Merkmalsvektors wird das Eingabebild in Teilbilder zerlegt. Auf diesen werden jeweils die Textur-Merkmale berechnet. Durch die gegenseitige Überlappung der Teilbilder kann eine höhere Robustheit gegenüber kleinen Variationen der Position der Hypothese im Eingabebild erreicht werden. In Abbildung 4.28 ist eine Unterteilung eines 32×32 Bildpunkte großen Bildausschnittes in 9 Subfenster dargestellt.

Aufbau des Klassifikator-Modells mittels Gabor-Merkmalen Für die Merkmalsextraktion von Fahrzeugen wird das Gabor-Merkmal eingesetzt. Das Merkmal eignet sich zur Unterscheidung von Texturen. Ein zweidimensionales Gabor-Filter entspricht einer Gaussfunktion, die mit einer Sinusfunktion einer bestimmten Frequenz und Orientierung gefaltet wird. Durch die freie Wahl der Frequenz f und Orientierung θ der Sinusfunktion und der Skalierung σ_x, σ_y können Gabor-Filter spezifisch an die Eingabebilder angepasst generiert werden.

Die Herausforderung für das Erreichen einer optimalen Klassifikation liegt in der Auswahl von Gabor-Filtern aus einer sogenannten Filterbank, welche die Merkmale der Klasse möglichst stabil beschreiben. Eine Filterbank besteht aus einer Menge einzelner Gabor-Filter mit unterschiedlichen Parametersätzen $\{\theta, f, \sigma_x, \sigma_y\}$.

Unter der Annahme, dass sich die Klassen hinsichtlich ihrer Textur im Bild voneinander unterscheiden, werden die drei stochastischen Momenten eines Gabor-Filters der Filterbank

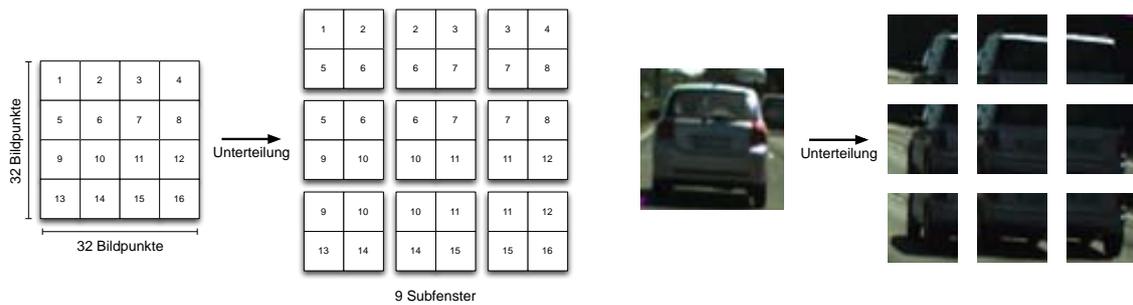


Bild 4.28: Unterteilung eines Eingabebildes in 9 überlappende Subfenster.

nach Faltung mit dem Eingabebild ausgewählt (der Mittelwert μ , die Standardabweichung σ und die Schiefe γ).

Die Auswahl geeigneter Gabor-Filter erfolgt in dieser Arbeit iterativ mittels des Boosting Ansatzes. Anschaulich kann der Boosting Ansatz dazu eingesetzt werden, einen starken Klassifikator aus einer Menge von Basisklassifikatoren zu bilden. Dies geschieht durch Linearkombination der Basisklassifikatoren, wobei der Gesamt-Klassifikationsfehler minimiert wird.

Ein Basisklassifikator besteht aus einer SVM, welche einen spezifischen Gabor-Filter zur Extraktion der Momente verwendet. Wie in [SJ09] beschrieben ist, wird für jeden Basisklassifikator auf einem Trainingsdatensatz dessen Klassifikationsfehler bestimmt. Durch Boosting werden die Basisklassifikatoren entsprechend ihres Fehlers geordnet. Die Gabor-Merkmale der t besten Klassifikatoren werden nach t Iterationen aus der Filterbank ausgewählt und gespeichert. Die SVM-Modelle der Basisklassifikatoren werden verworfen.

Unter Verwendung der t ausgewählten Gabor-Filter wird ein SVM-Modell offline erstellt. Für jedes Subfenster s der Eingabebilder des Trainingsdatensatzes werden die drei Momente mit t geeigneten Gabor-Filtern bestimmt. Die Größe des Merkmalsvektors \vec{m} beträgt damit $3 \times s \times t = 27t$:

$$\vec{m} = (\mu_{11}\sigma_{11}\gamma_{11}, \dots, \mu_{9t}\sigma_{9t}\gamma_{9t})^T. \quad (4.20)$$

Das Training des SVM-Modells erfolgt auf Basis ausgeschnittener und nach Klassen sortierter Bildausschnitte. In den Merkmalsvektor gehen keine zeitlichen Sequenzen ein, der Klassifikator bewertet damit nur Bildausschnitte. Das gelernte SVM-Modell sowie die Auswahl der Gabor-Filter zur Merkmalsextraktion werden für die Online-Klassifikation vorgehalten.

Klassifikation mittels Gabor-Merkmalen Die Klassifikation erfolgt unter Verwendung der ausgewählten Gabor-Filter und dem eingelernten SVM-Modell. Eingabebilder werden in ihrer Größe normalisiert und in ihre Subfenster zerlegt. Auf den Subfenstern wird der Merkmalsvektor \vec{m} wie oben beschrieben bestimmt und durch die SVM eine Zugehörigkeitsverteilung für die unterschiedlichen Klassen berechnet.

Alternative Klassifikationsansätze Anstelle der oben beschriebenen Gabor-Filter kann eine robuste Merkmalsextraktion auch durch Histogram of Oriented Gradients (HOG)-Merkmale erreicht werden [Zhu10]. Gleiche Objekte zeichnen sich durch eine ähnliche Intensität der Gradienten und der Kantenrichtungen aus. Die Extraktion des HOG-Merkmals berechnet eine Histogrammverteilung der Gradienten aller Kanten innerhalb eines Bildausschnittes. Das

Extraktionsverfahren hat sich in der Literatur insbesondere zur Erkennung von Fußgängern bewährt [DT05].

Der konzeptionelle Aufbau des Klassifikator-Modells ist ähnlich dem oben beschriebenen Ansatz. Das Eingabebild der ROI wird wie oben beschrieben in einzelne Bildausschnitte zerlegt. Für jeden Bildausschnitt wird die Histogrammverteilung bestimmt. Zur Unterdrückung des Gleichanteils in Bildausschnitten wird ein normalisiertes Histogramm aus den Histogrammen aller benachbarten Bildausschnitte gebildet. Die normalisierten Histogramme der Bildausschnitte bilden den Merkmalsvektor. Zur Klassifikation wird mit den Merkmalsvektoren des Trainingsdatensatzes ein SVM-Modell eingelernt.

4.4.3 Klassifikation von Fußgängern

Auf die Klassifikation von Fußgängern soll an dieser Stelle nicht detailliert eingegangen werden. Die Erkennung von Fußgängern in Videodaten wurde im Rahmen dieser Arbeit wie in [Göt11] beschrieben umgesetzt und ähnelt im Wesentlichen den oben beschriebenen Ansätzen zur Klassifikation von Fahrzeugen.

Die Klassifikation von Fußgängern kann mittels HOG-Merkmalen und SVM-Klassifikatoren oder alternativ dazu durch AdaBoost-Kaskaden und Haar-Wavelets erfolgen. In Abschnitt 6.4.2 der Evaluation sind Ergebnisse der videobasierten Fußgängerklassifikation anschaulich gezeigt.

4.5 Temporale Zustandsschätzung

Die temporale Zustandsschätzung hat zum Ziel, die Objekthypothesen über die Zeit zu stabilisieren und zusätzliche temporale Merkmale für die Hypothesen zu bestimmen. Der Tracking-Schritt nimmt eine Liste klassifizierter Objekthypothesen entgegen und bildet daraus eine gesammelte Objektliste aller detektierten Hypothesen in einem gewissen Zeitintervall. Somit kann über die Zeit ein Objektwissen gebildet werden, welches mit Informationen über Zustandsunsicherheiten und Existenzsicherheiten versehen ist. Fehlerhaft detektierte Hypothesen können abgelehnt und für ausbleibende Messungen zumindest eine Prädiktion der wahrscheinlichen Aufenthaltsposition bestimmt werden.

Basis des Trackings ist ein Multi-Target Tracking (MTT)-System nach Blackman und Popull [BP99] (vergleiche Abschnitt 3.3). In Abbildung 4.29 ist schematisch der Ablauf des Zustandsschätzers gezeigt. Dieser besteht aus einer Assoziations-Komponente, welche eine Zuordnung zwischen den bisherigen Objektlisten und einer neu eintreffenden Objektliste vornimmt und eine Existenzwahrscheinlichkeit jedes Objektes bestimmt. Auf Basis der Ergebnisse der Assoziation erfolgt die Aktualisierung der Objekte, wobei gültige Objekte bestätigt und ungültige Objekte entfernt werden. Ein temporales Filter auf Basis des Kalman Filters führt die Berechnung der Objektparameter durch und ermittelt die Schätzfehler-Kovarianzen für die Objekte. Für die Assoziation der nächsten Hypothesen wird die Objektliste prädiziert.

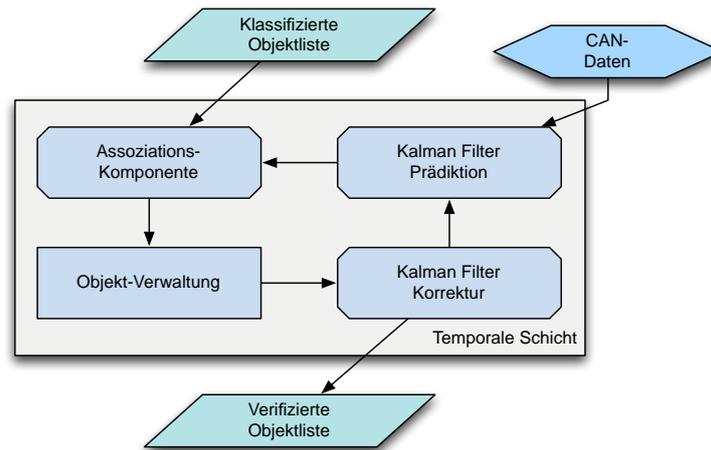


Bild 4.29: Ablauf der Zustandsschätzung von Objekthypothesen über die Zeit mit einem MTT-System.

Die Grundlagen zur Filterung mittels Kalman Filter sind in Abschnitt 3.3.3 beschrieben. Die Verfolgung einer Hypothese erfolgt auf Basis der 3D-Positionsschätzung. Ein Objekt wird beschrieben durch den Zustandsvektor \mathbf{x}_k zum Zeitpunkt k . Der Zustandsvektor besteht aus

$$\mathbf{x}_k = \left(x, y, z, v_x, v_y, v_z, a_x, a_y, a_z \right)^T, \quad (4.21)$$

wobei die Dimensionen x, y, z entkoppelt voneinander betrachtet werden.

Bewegungsmodelle Als Bewegungsmodelle eignen sich das Modell konstanter Geschwindigkeit (CV), konstanter Beschleunigung (CA) oder das Beschleunigungsmodell nach Singer [Sin70]. Letzteres schätzt Position, Geschwindigkeit und Beschleunigung eines Objektes abhängig von dessen angenommener Manöverdauer. Die diskretisierte Entwicklung der Beschleunigung entspricht einem Markov Prozess 1. Ordnung:

$$a(k+1) = \rho_m a(k) + \sqrt{1 - \rho_m^2} \sigma_m r(k), \quad (4.22)$$

wobei gilt

$$\rho_m = e^{-\beta T}, \beta = 1/\tau_m \quad (4.23)$$

$$\tau_m, \sigma_m = \text{Manöverdauer eines Objektes,} \quad (4.24)$$

Standardabweichung der Manöverdauer

$$r(k) = \text{Mittelwertfreie Zufallsvariable mit Gauss'scher Normalverteilung} \quad (4.25)$$

$$T = \text{Messwiederholdauer.} \quad (4.26)$$

Für die Annahme, dass die Messwiederholdauer sehr klein ist ($T \ll \tau_m$), entspricht das Modell dem CA-Modell. Zur besseren Darstellung sind im Folgenden lediglich die Matrizen zur Schätzung der Längsbewegung von Objekten dargestellt. Die Kalman-Matrix des Zustands-

übergangsmodells A_{dyn} lautet

$$A_{\text{dyn},x} = \begin{pmatrix} 1 & T & \frac{T^2}{2} \\ 0 & 1 & T \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.27)$$

Anhand der Manöverdauer τ_m lässt sich die Bewegungsdynamik des geschätzten Zustands adaptiv anpassen. σ_v^2 entspricht der Varianz der zulässigen Objektbeschleunigung. Damit ist die Kovarianz des Prozessrauschens Q_{dyn}

$$Q_{\text{dyn},x} = \frac{2\sigma_{v,x}^2}{\tau_m} \begin{pmatrix} \frac{T^5}{20} & \frac{T^4}{8} & \frac{T^3}{6} \\ \frac{T^4}{8} & \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^3}{6} & \frac{T^2}{2} & T \end{pmatrix}. \quad (4.28)$$

Das Messübergangsmodell H_{dyn} und die Kovarianz des Messrauschens R_{dyn} sind:

$$H_{\text{dyn},x} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \quad (4.29)$$

$$R_{\text{dyn},x} = \sigma_{w,x}^2. \quad (4.30)$$

Eine Unterscheidung zwischen statischen und dynamischen Objekten ist durch das Interacting Multiple Model (IMM)-Verfahren möglich, welches in Abschnitt 3.3.4 näher beschrieben ist. Für das statische Bewegungsmodell gelten die Zustandsübergangsmatrix $A_{\text{stat},x}$ und das Messübergangsmodell $H_{\text{stat},x}$:

$$A_{\text{stat},x} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (4.31)$$

$$H_{\text{stat},x} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}. \quad (4.32)$$

Da dem Modell die Annahme zugrund liegt, dass keine Objektbewegung auftritt, werden Änderungen in der Position von Objekten über ein mittelwertfreies Rauschen angenommen:

$$Q_{\text{stat},x} = \sigma_{v,x}^2 \begin{pmatrix} T^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (4.33)$$

Die Messunsicherheit wird über die Kovarianz des Messrauschens dargestellt:

$$R_{\text{stat},x} = \sigma_{w,x}^2. \quad (4.34)$$

Kompensation der Eigenbewegung Die Messungen von Objekten erfolgen aus Sicht des sich bewegenden Eigenfahrzeugs. Daher wird für statische Objekte die Relativgeschwindigkeit zwischen statischem Objekt und Eigenfahrzeug gemessen. Die Kompensation der Eigenbewegung des Fahrzeugs erfolgt über den optionalen Kontrollvektor \mathbf{u}_{k-1} im Kalman Filter

(vergleiche Abschnitt 3.3.3). Die Kompensation erfolgt durch Betrachtung der als konstant angenommenen momentanen Eigengeschwindigkeit $v_{ego,k}$:

$$B_x = \begin{pmatrix} T & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (4.35)$$

$$\mathbf{u}_{k-1,x} = \begin{pmatrix} -v_{ego,k} \\ 0 \\ 0 \end{pmatrix}. \quad (4.36)$$

Die Kompensation der Geschwindigkeit erfolgt ohne Berücksichtigung der Gierrate des Fahrzeugs. Dies ist ausreichend, solange die Gierrate des Fahrzeugs gering ist. Während eines Beschleunigungsmanövers kann die Kompensation nicht alle Bewegungseffekte ausgleichen, diese werden aber ausreichend über das Prozessrauschen ausgeglichen. Der Kontrollvektor \mathbf{u}_{k-1} wird aus den Daten der fahrzeuginternen Sensorik beim Eintreffen einer neuen Objektliste erzeugt.

Unterscheidung zwischen statischen und dynamischen Objekten Für die Unterscheidung zwischen statischen und dynamischen Objekten wird das IMM-Filter eingesetzt. Das Filter ist in Abschnitt 3.3.4 beschrieben. Die Initialisierung des Filters erfolgt mit zwei Modellen: eines zur Schätzung dynamischer Objekte und eines zur Schätzung statischer Objekte. Das IMM-Filter vergleicht in jedem Zeitschritt, welches Filter zur Beschreibung der Bewegung eines Objektes unter Berücksichtigung der Eigenbewegung besser geeignet ist. Dabei werden die Modellwahrscheinlichkeiten für alle Filter aktualisiert.

Für jedes Objekt wird dessen Zustand aus der Kombination der Zustände der beiden Modelle aktualisiert. Dies geschieht unter Berücksichtigung der Modellwahrscheinlichkeiten beider Modelle. Anhand der Modellwahrscheinlichkeiten kann entschieden werden, ob es sich um ein statisches oder ein dynamisches Ziel handelt. Das jeweilige Filter stellt außerdem die Schätzsicherheit des Objektzustandes sowie ein Existenzmaß zur Verfügung, welches aus der Messhistorie gebildet wird.

4.6 Diskussion

In diesem Kapitel wurde ein konzeptioneller Ablauf vorgestellt, welcher Methoden zur Erkennung, Klassifikation und zur zeitlichen Verfolgung in eine durchgängige Verarbeitungskette integriert. Das Konzept erlaubt insbesondere den Einsatz komplementärer Sensoren bei der Segmentierung und Klassifikation von Objekthypothesen. Die in diesem Kapitel vorgestellten Konzepte beschreiben Verfahren zur Erkennung von Objekten aus Tiefendaten sowie aus Videodaten.

Hierdurch ergeben sich mehrere Vorteile: Durch die Verwendung unterschiedlicher Sensor-konzepte zur Segmentierung und Klassifikation können Unsicherheiten in der Bestimmung von Objekt-Attributen verringert werden. Dies betrifft die Position sowie die Ausdehnung der Objekte, ebenso wie die Zugehörigkeit zu Objektklassen. Gleichzeitig lässt sich die Laufzeit der

eingesetzten Verfahren verbessern. Insbesondere videobasierte Erkennungsverfahren besitzen einen hohen Rechenaufwand, falls das komplette Videobild nach Hypothesen durchsucht werden muss. Die Funktion dieses durchgängigen Konzepts wird in Abschnitt 6.4.2 evaluiert.

Die unterschiedlichen Eigenschaften der tiefenbildgebenden Sensoren erfordert es, sensorspezifische Anpassungen an den Verfahren vorzunehmen oder unterschiedliche Verfahren zu wählen. Alle Erkennungsverfahren liefern resultierende Objekthypothesen, welche in den nachfolgenden Schritten der Verarbeitungskette ausgewertet werden. Das Verfahren zur zeitlichen Verfolgung leistet darauf aufbauend eine Bildung verifizierter Objekte. Somit lassen sich zusätzliche Attribute wie die Objektgeschwindigkeit sowie Schätzsicherheiten anhand der Messdaten ermitteln. Die Verfahren werden in Abschnitt 6.3 näher untersucht.

Einige der genannten Ansätze decken nur Teilaspekte der Forschung an Verfahren zur Umfelderkennung in Fahrzeugen ab. Für die Bestimmung einer auf Umfeldmessungen aufbauenden Situationsanalyse sind Erfahrungen mit realen, mit Unsicherheiten behaftete Informationen eminent wichtig. Aus diesen Informationen dennoch Schlüsse ziehen zu können stellt den wesentlichen Aspekt des nächsten Kapitels dar.

5 Modellbasierte Datenfusion zur Situationsinterpretation

Die Analyse von Fahrsituationen erfolgt durch Überführung von Informationen der Fahrzeugumgebung in ein Umgebungsmodell. Die Informationen werden aus Messungen der Sensoren gewonnen. Das Umgebungsmodell ist dazu initial aus Vorwissen aufgebaut und wird im Laufe der Situationsanalyse (SA) kontinuierlich durch weitere Informationen verfeinert und konkretisiert. Es können Informationen verschiedener Sensoren und informationsverarbeitender Algorithmen zusammen abgelegt werden.

In dieser Arbeit wird ein modellbasierter Ansatz zur Darstellung von Fahrsituationen umgesetzt. Zur Wissensrepräsentation liegt die Objektorientierte Probabilistische Relationale Modellierungssprache (OPRML) zugrunde, mit welcher Klassen und Attribute, sowie Relationen zwischen Klassen und Attributen beschrieben werden. Damit wird sowohl Hintergrundwissen wie auch aktuelle Umfeldinformationen abgebildet. Die Wissensrepräsentation liefert die Ausgangsbasis zur automatischen Konstruktion graphischer Modelle. In den graphischen Modellen werden probabilistische Abhängigkeiten zwischen Informationen dargestellt. Dadurch ist es möglich, Aussagen zu bestimmen, die nur aus Fusion mehrerer zugrundeliegender Informationen getroffen werden können. Schlussfolgerungen mittels Inferenzverfahren berechnen eine probabilistische Aussage über die vorherrschende Fahrsituation gegeben der Messungen der Umfeld-Sensoren.

5.1 Methodisches Konzept der wissensbasierten Interpretation

Im folgenden Abschnitt ist beschrieben, wie die formalen Sprachelemente definiert sind, der Aufbau des Weltmodells durch Frame-Slot-Konzepte umgesetzt wird und der logische Ablauf sowie die Anbindung der Umfeldwahrnehmung erfolgt. Diese Schritte sind in [Kuh11] detailliert ausgeführt.

Die schematische Darstellung des Ablaufes ist in Abbildung 5.1 dargestellt. Eingabe des Systems sind Objekthypothesen, welche durch die in Kapitel 4 beschriebenen Verfahren erstellt werden. Die Wissensbasis stellt die Konzepte bereit, welche durch die formale Sprache beschrieben sind. In Abhängigkeit von der Umfelderkennung wird aus den Konzepten das aktualisierte Weltmodell aufgebaut, welches die Instanzen der Konzepte enthält und schritt haltend zur Umfelderkennung eine Repräsentation der momentanen Fahrsituation darstellt. Die probabilistische Komponente des Weltmodells codiert ein probabilistisches Netz, aus welchem sich Situationsaussagen zur Fahrsituation ableiten lassen.

5.1.1 Formale Beschreibung

Der formalen Beschreibung liegen die Sprachelemente der OPRML zugrunde. Sie besteht aus zwei Komponenten. Der relationalen Komponente sowie der probabilistischen Komponente.

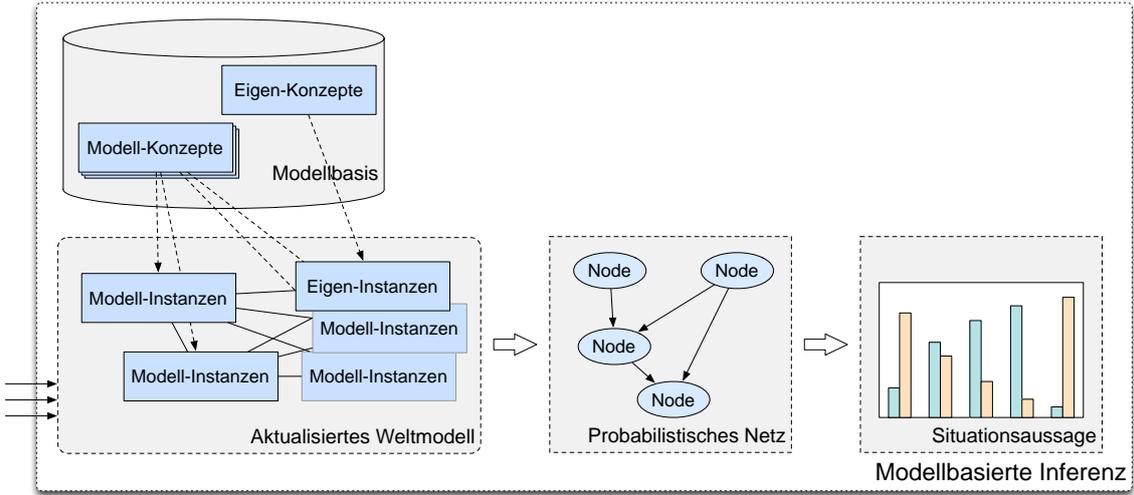


Bild 5.1: Konzeptioneller Aufbau des Weltmodells und des probabilistischen Netzes zur Interpretation der Fahrsituation.

Der relationale Teil ist gegeben durch:

- Eine Menge von Klassen, $\mathbf{C} = \{C_1, C_2, \dots, C_n\}$ und einer partiellen Ordnung \mathbf{C}, \sqsubseteq , welche die Klassenhierarchie bestimmt.
- Eine Menge einfacher Attribute, $\mathbf{\Lambda}_C = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$, für jede Klasse $C \in \mathbf{C}$. Dies sind nicht-probabilistische Attribute der Klasse.
- Eine Menge beschreibender Attribute, $\mathbf{\Delta}_C = \{\delta_1, \delta_2, \dots, \delta_n\}$, für jede Klasse $C \in \mathbf{C}$. Attribut δ_x der Klasse C_1 wird bezeichnet mit $C_1.\delta_x$. Jedes beschreibende Attribut hat einen Domänentyp $\text{Dom}(C_1.\delta_x) = C_1$ und einen Bereich $\text{Range}(C_1.\delta_x) = \text{Val}(C_1.\delta_x)$, wobei $\text{Val}(C_1.\delta_x)$ einer vordefinierten, abzählbaren Menge von Werten entspricht, also $\text{Val}(C_1.\delta_x) = \{\text{Val}_1, \text{Val}_2, \dots, \text{Val}_n\}$;
- Eine Menge komplexer Attribute, $\mathbf{\Phi}_C = \{\phi_1, \phi_2, \dots, \phi_n\}$, für jede Klasse $C \in \mathbf{C}$. Attribut ϕ_x der Klasse C_1 wird bezeichnet mit $C_1.\phi_x$. Komplexe Attribute beschreiben funktionale Zusammenhänge zwischen Klassen. Jedes komplexe Attribut hat einen Domänentyp $\text{Dom}(C_1.\phi_x) = C_1$ und einen Bereich $\text{Range}(C_1.\phi_x) = C_y$, wobei $C_y \in \mathbf{C}$ gilt. Des Weiteren existiert für jedes komplexe Attribut sein inverses Attribut, welches die inverse Funktion von ϕ darstellt. Wenn also $\text{Dom}(C_1.\phi) = C_1$ und $\text{Range}(C_1.\phi) = C_y$ gilt, dann ist $\text{Dom}(C_y.\phi^{-1}) = C_y$ und $\text{Range}(C_y.\phi^{-1}) = C_1$. Diese Attribute werden als komplex bezeichnet, da sie die Tatsache ausdrücken, dass ihr Wert einer weiteren Klasse entspricht. Komplexe Attribute können mehrwertig sein und somit auf eine Menge von Klassen Bezug nehmen.

Der probabilistische Teil ist wie folgt definiert:

- Für jede Klasse $C \in \mathbf{C}$ existiert ein bedingtes Wahrscheinlichkeitsmodell für jedes beschreibende Attribut $\delta_x \in \mathbf{\Delta}_C, P(\delta_x | \mathbf{Pa}(\delta_x))$, wobei $\mathbf{Pa}(\delta_x) = \{\text{Pa}_1, \text{Pa}_2, \dots, \text{Pa}_n\}$ die Menge an Elternknoten von δ_x ist. Attribute müssen kein Wahrscheinlichkeitsmodell besitzen, wenn sie nicht von anderen Attributen probabilistisch abhängen.

Die Klassenebene wird zur Beschreibung der Modellkonzepte genutzt. Für das Weltmodell werden durch Instanziierung der Klassen die Elemente der Instanzebene erzeugt. Die Umsetzung der OPRML erfolgt auf Basis von Frame-Slot Konzepten [Min74]. Die Klassenebene besteht aus:

- Einer Menge von Frame-Klassen $\mathbf{C} = \{C_1, C_2, \dots, C_n\}$,
- einer Ordnungsrelation auf \mathbf{C} , welche die Klassenhierarchie definiert,
- einer Menge von Slots unterschiedlicher Art:
 - Einfache Slots zur Beschreibung einfacher Attribute,
 - probabilistische Slots zur Darstellung beschreibender Attribute, sowie
 - referenzierende Slots zur Abbildung der komplexen Attribute.

Die Instanzebene besteht aus:

- Einer Menge von Frame-Instanzen $\mathbf{I} = \{I_1, I_2, \dots, I_n\}$,
wobei jede Frame-Instanz I_i eine Instanziierung genau einer Frame-Klasse $C_k \in \mathbf{C}$ ist.
- Einer Menge von Slot-Instanzen,
wobei die Slot-Instanzen eine Instanziierung aller Slots der Frame-Klasse C_k sind, von welcher die Frame-Instanz I_i instanziiert ist. Dabei wird die Klassenhierarchie umgesetzt, also auch Slots der übergeordneten Frame-Klassen berücksichtigt.

Das Frame-Modell ermöglicht die Umsetzung der formalen Sprachelemente unter Berücksichtigung der jeweiligen Eigenschaften spezifischer Attribute. Die Verknüpfung zwischen Klassen und Instanzen wird durch das Modell gewährleistet. Frame-Instanzen leiten sich also direkt aus Frame-Klassen ab, äquivalent dazu leiten sich Slot-Instanzen aus den Slot-Klassen ab. Ein Frame kann dabei beliebig viele Slots enthalten, jeder Slot besitzt jedoch eine eindeutige Frame-Zuordnung. Innerhalb der Slot-Instanzen können Werte abhängig von der Art des Slots gespeichert werden.

Die einfachen Slots können einfache Werte speichern und entsprechen damit bekannten Werten, ähnlich zu Einträgen in einer relationalen Datenbank. Die probabilistischen Slots können mit Unsicherheit behaftete Werte speichern und entsprechen damit Zufallsvariablen. Die probabilistischen Abhängigkeiten der Zufallsvariable sowie die bedingte Wahrscheinlichkeitsverteilung (CPD) werden in Form einer Liste gespeichert. Die referenzierenden Slots stellen Relationen zu anderen Frame-Klassen her, ähnlich den Fremdschlüsseln einer relationalen Datenbank. Dabei werden innerhalb des Slots referenzierte Slot-Instanzen anderer Frames abgelegt.

Die Speicherung von sicheren Daten erfolgt also durch einfache Slots. Der Wertebereich einfacher Slots ist ein einfacher Datentyp, eine Instanz des Slots kann genau einen Wert speichern. Erlaubt sind dabei Zeichenketten, Ganzzahlen oder Fließkommazahlen. Zusammen mit den referenzierenden Slots kann daraus zunächst mindestens die Mächtigkeit einer relationalen Datenbank dargestellt werden.

Zur Darstellung unsicherer Daten werden die probabilistischen Slots eingesetzt. Diese besitzen als Wertebereich ein vordefiniertes, abzählbares Tupel von benannten Werten. Dabei wird zwischen zwei unterschiedlichen Arten von probabilistischen Slots unterschieden:

- Die beschreibenden Slots (auch *Descriptive*-Slots genannt) besitzen auf Klassenebene eine bedingte Wahrscheinlichkeitsverteilung (CPD) und können mit anderen probabilistischen Slots verknüpft sein. Die CPDs und die Verknüpfungen zu Slots werden bei der Instanziierung übernommen.
- Die zusammenfassenden Slots (auch als *Aggregate*-Slots bezeichnet) sind auf Klassenebene mit einem anderen probabilistischen Slot verknüpft und durch eine Aggregationsfunktion definiert. Die Aggregationsfunktion beschreibt eine CPD für jede Anzahl an verknüpften Instanzen. Bei der Instanziierung werden die verknüpften Instanzen und die CPD durch die Anzahl verknüpfter Instanzen festgelegt. Auf Klassenebene beschreiben zusammenfassende Slots eine prinzipielle Beziehung zu verknüpften Slots.

Die Darstellung der Relationen einer Klasse zu Fremdklassen erfolgt durch referenzierende Slots. Diese Slots haben als Wertebereich eine Frame-Klasse und liefern eine Referenz zu dieser. Sie importieren dabei spezifische probabilistische Slots einer fremden Frame-Klasse in die eigene Frame-Klasse. Dadurch werden auf Klassenebene beschreibende Slots der referenzierten Klasse in die referenzierende Klasse übertragen. Hierdurch kann aus probabilistischer Sicht eine bedingte Abhängigkeit von Knoten einer Klasse zu Elternknoten einer Fremdklasse hergestellt werden. Es lassen sich zwei Typen referenzierender Slots unterscheiden:

- Einfache Referenz-Slots importieren während der Instanziierung genau eine Frame-Instanz (einer Frame-Klasse),
- mehrfache Referenz-Slots (*Multireferenz*-Slots) importieren mehrere Frame-Instanzen einer Frame-Klasse während der Instanziierung.

Die Multireferenz-Slots stehen in enger Beziehung mit den Aggregate-Slots. Werden durch einen Multireferenz-Slot mehrere Frame-Instanzen einer Klasse referenziert, so fasst ein dazu gehöriger Aggregate-Slot die Informationen der Frame-Instanzen zusammen. Da ein beschreibender Slot nur mit probabilistischen Slots verknüpft werden kann, müssen die Informationen von einem Multireferenz-Slot durch einen Aggregate-Slot gebündelt werden. Aus probabilistischer Sicht bedeutet dies, dass ein Knoten von einem zusammenfassenden Knoten abhängt, welcher durch eine Spezialfunktion wiederum von einer Vielzahl anderer Knoten probabilistisch abhängig ist.

5.1.2 Probabilistisches Modell

Die oben beschriebenen probabilistischen Slots bilden das probabilistische Modell zur Inferenz von Informationen. Zunächst gilt:

- auf Klassenebene:
 - \mathbf{C} die Menge aller Frame-Klassen
 - $\Pi_{\mathbf{C}} = \{\pi_1, \pi_2, \dots, \pi_n\} \forall C \in \mathbf{C}$ die Menge aller probabilistischen Slot-Klassen

$\Delta_C = \{\delta_1, \delta_2, \dots, \delta_m\} \subset \Pi_C \quad \forall C \in \mathbf{C}$ die Menge aller Descriptive-Slot-Klassen

$\Xi_C = \{\xi_1, \xi_2, \dots, \xi_l\} \subset \Pi_C \quad \forall C \in \mathbf{C}$ die Menge aller Aggregate-Slot-Klassen

- auf Instanzebene:

\mathbf{I} die Menge aller Frame-Instanzen

$\Phi_I = \{\phi_1, \phi_2, \dots, \phi_k\} \quad \forall I \in \mathbf{I}$ die Menge aller probabilistischer Slot-Instanzen.

Auf Klassenebene besteht das probabilistische Modell aus:

- einem bedingten Wahrscheinlichkeitsmodell $P(\delta_x | \mathbf{Pa}(\delta_x)) \quad \forall \delta_x \in \Delta_C \quad \forall C \in \mathbf{C}$, wobei $\mathbf{Pa}(\delta_x)$ die Menge aller Eltern von δ_x ist,
- einem funktionellen Wahrscheinlichkeitsmodell $f(\xi_x | \mathbf{F}(\xi_x)) \quad \forall \xi_x \in \Xi_C \quad \forall C \in \mathbf{C}$, wobei $\mathbf{F}(\xi_x)$ der Elternknoten von ξ_x ist, dessen Anzahl von Instanzen noch unbekannt ist; f ist eine Aggregationsfunktion.

Auf Instanzebene besteht das probabilistische Modell aus:

- einem bedingten Wahrscheinlichkeitsmodell $P(\phi_x | \mathbf{Pa}(\phi_x)) \quad \forall \phi_x \in \Phi_I \quad \forall I \in \mathbf{I}$.

Das bedingte Wahrscheinlichkeitsmodell und der aus den probabilistischen Slots bestimmte gerichtete Graph auf Instanzebene können direkt als Bayes'sches Netz (BN) betrachtet werden. Das bedingte Wahrscheinlichkeitsmodell kann anhand der Slot-Klassen erstellt werden und ist bei der Instanziierung von Klassen bereits vorhanden. Das funktionelle Wahrscheinlichkeitsmodell dient der Definition des probabilistischen Modells für Multireferenzen, da hier die Anzahl an Elternknoten erst auf Instanzebene bekannt ist. Durch die Beschreibung wird die Aggregationsfunktion bestimmt, welche die CPD auf Instanzebene spezifiziert. Zur vollständigen Bestimmung der CPD durch die Aggregationsfunktion wird als Eingabe die Anzahl der Zustände des Aggregate-Slots, die Anzahl der Elternslots sowie die Anzahl der Zustände jedes Elternslots benötigt. Die Unterteilung des Modells in Frames ermöglicht es zudem, das Frame als Fragment des vollständigen BN zu betrachten, was den Vorteilen eines Objektorientierten Bayes'schen Netzes (OOBN) entspricht.

5.1.3 Modellerstellung mittels Frame-Slot Konzepten

Zur Analyse der Fahrsituation wird das Frame-Slot-Konzept an die zu bestimmende Situation spezifisch angepasst und die Anzahl der Instanzen sowie der Werte an die Umfelderkennung adaptiert. Die Menge der Frame- und Slot-Klassen können zunächst beliebig und unabhängig voneinander spezifiziert werden. Allerdings sei hervorzuheben, dass eine Slot-Klasse nur dann verwendet werden kann, wenn diese einer Frame-Klasse zugehört.

Das zugrundeliegende Frame-Konzept ist in Abbildung 5.2 dargestellt und enthält im Wesentlichen die Information über das Vater-Konzept sowie Felder zur Beschreibung des Konzepts und eine Liste von Slot-Konzepten, welche dem Frame zugeordnet sind. Das zugrundeliegende Slot-Konzept ist eine abstrakte Beschreibung einer Slot-Klasse. Die konkreten Konzepte für einfache, referenzierende oder probabilistische Slot-Konzepte übernehmen die Attribute des abstrakten Slot-Konzepts. Je nach Art des Konzepts werden Felder vorgehalten, um

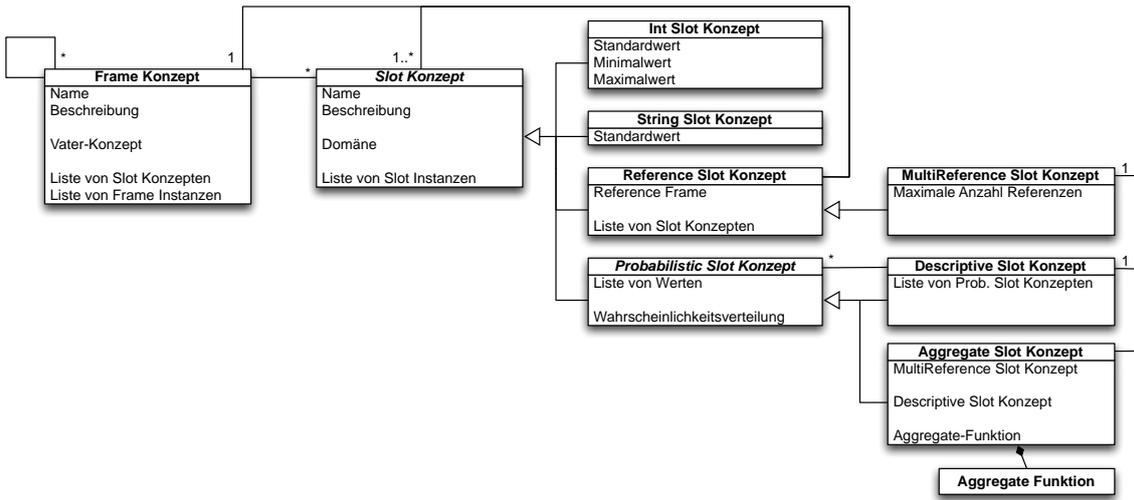


Bild 5.2: Spezifikation der Frame- und Slot-Konzepte in UML-Notation.

Verweise auf Frame-Konzepte, weitere Slot-Konzepte, zugrundeliegende CPDs oder anderes zu speichern.

Sobald alle notwendigen Konzepte erstellt wurden, können Relationen zwischen den Konzepten festgelegt werden. Hierfür werden die Felder verknüpft, welche auf andere Konzepte verweisen. Dies betrifft beispielsweise die Festlegung der Hierarchie der Frames mittels Vater-Konzepten, die Zuordnung von Slots zu Frames mittels der Liste von Slot-Konzepten oder auch die Verknüpfung von Slots zu einem referenzierenden Slot-Konzept.

Für Multireferenz-Slots müssen die entsprechenden Aggregationsfunktionen festgelegt werden. Der Aggregationsfunktion kann funktionell *Summe*, *Modus*, *Mittelwert*, *Minimum*, *Maximum*, *Anzahl* oder eine benutzerdefinierte Funktion zugeschrieben werden. Da auf Klassenebene die konkrete Anzahl von Elternknoten unbekannt ist (vergleiche funktionelles Wahrscheinlichkeitsmodell f), kann die CPD erst während der Instanziierung bestimmt werden.

Für beschreibende, probabilistische Slots müssen initiale CPDs festgelegt werden. Diese stellen Vorwissen dar, können aber bei der Instanziierung durch konkretes Wissen (beispielsweise durch Umfeldmessungen) verändert werden. Für jede CPD müssen dabei die Wertebereiche des Slots und dessen Elternslots berücksichtigt werden.

Bevor eine Instanziierung der Konzepte in das Weltmodell erfolgt, müssen alle Felder mit Werten gefüllt sein. Sobald diese Bedingung erfüllt ist, können Frames und die zugehörigen Slots instanziiert werden.

Durch die Instanziierung werden in vielen Fällen mehrfach Instanzen von einer Klasse abgeleitet. Eine eindeutige Namenskonvention erleichtert die Unterscheidung zwischen Klassen und Instanzen, wie in Abbildung 5.3 gezeigt. Dies erlaubt die Aktualisierung der Instanzen des Weltmodells in Abhängigkeit von der Umfeldwahrnehmung. Die Grafik zeigt die formale Spezifikation der Modelle in OPRML-Darstellung. Darin wird die enge Verknüpfung zwischen Frames und Slots besonders deutlich. Ist ein Slot mit einem Frame verknüpft, so verweist das *Domain*-Element auf den Frame. In dem Beispiel ist ein einfacher Frame dargestellt, welchem ein beschreibender Slot zugeordnet ist. Der beschreibende Slot besitzt im *Range*-Element die Information über dessen Wertebereich sowie im *Distribution*-Element die initial angenomme-

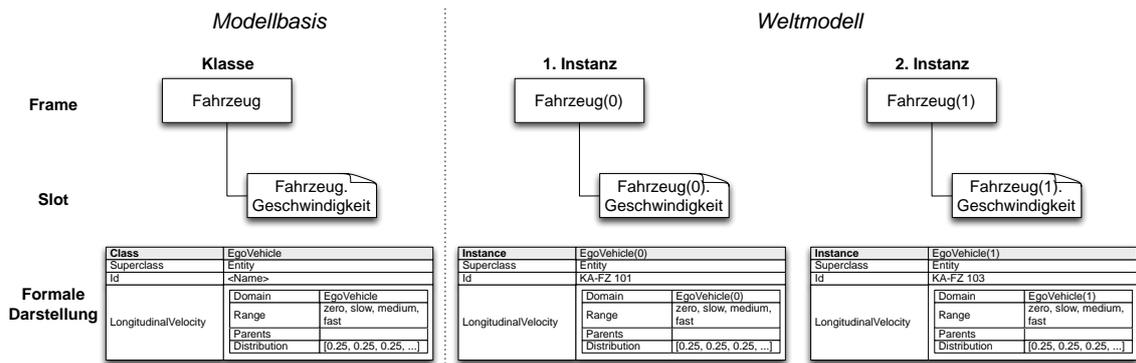


Bild 5.3: Die eindeutige Benennung von Klassen und Instanzen erlaubt die Unterscheidung und spätere Identifikation der Elemente in der Modellbasis und im Weltmodell; in der unteren Zeile ist die formale Darstellung von zusammengehörender Frame-Slot-Klassen sowie Frame-Slot-Instanzen gezeigt.

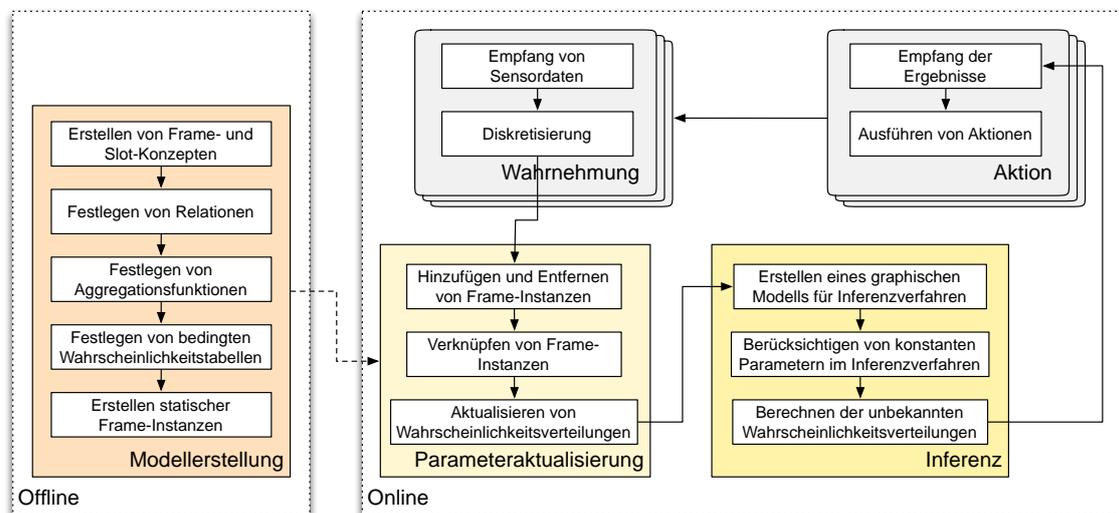


Bild 5.4: Ablauf der wissensbasierten Interpretation mit Trennung der Online- und Offline-Arbeitsschritte.

ne CPD über den Wertebereich. Durch das *Parents*-Element können probabilistische Abhängigkeiten spezifiziert werden. In dem gezeigten Fall existiert keine Abhängigkeit auf einen anderen beschreibenden Slot.

5.1.4 Logischer Ablauf und Anbindung der Umfeldwahrnehmung

Im Folgenden ist die Anbindung der Umfeldwahrnehmung an das Weltmodell beschrieben [Kuh11]. Die Parameteraktualisierung führt die Instanzen des Weltmodells schritthaltend mit den aktualisierten Umgebungsinformationen mit. Pro Zeitschritt werden die Instanzen anhand der Sensordaten aktualisiert. Dabei werden alle Frame-Instanzen des Weltmodells überprüft, ob das jeweils korrespondierende Objekt durch die Umfeldwahrnehmung im aktuellen Zeitschritt beobachtet wurde.

Nicht (mehr) benötigte Frame-Instanzen und die dazugehörigen Slot-Instanzen werden aus dem Weltmodell entfernt. Für neu detektierte Objekte durch die Umfeldwahrnehmung werden entsprechende Frame-Instanzen in das Weltmodell eingesetzt. Für auf Klassenebene ver-

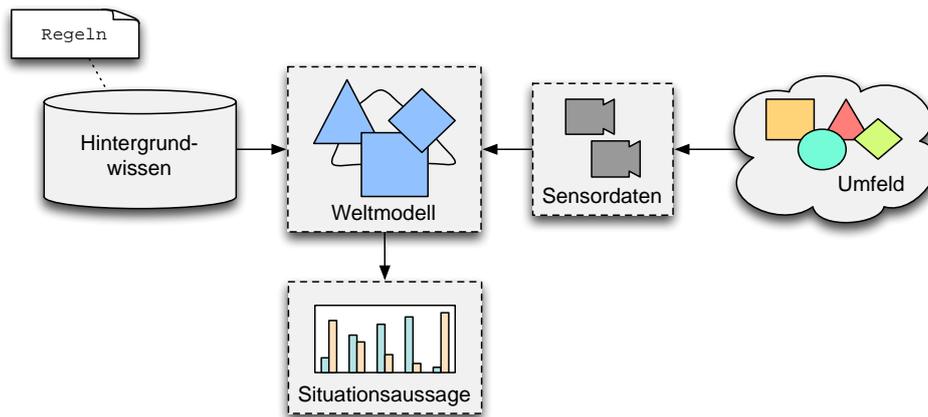


Bild 5.5: Schematische Beschreibung des Informationsfluss zum Aufbau des Weltmodells aus Hintergrundwissen und Umfeldinformationen und zur Bestimmung einer Situationsaussage.

knüpfte Slot-Konzepte werden ebenfalls Instanzen in das Weltmodell eingetragen und mit den Frame-Instanzen verknüpft.

Das Weltmodell entspricht damit einer aktuellen Repräsentation der durch die Umgebungswahrnehmung gemessenen Informationen in Form von Frame- und Slot-Instanzen. Im Weltmodell werden Frame-Instanzen, welche untereinander zum Beispiel durch referenzierende Slots in Relation stehen, miteinander verknüpft. Dies erfolgt auf Basis der festgelegten Klassendefinition der Modellbasis. Für Multireferenz-Slots werden alle existierenden Frame-Instanzen eines Konzepts mit dem Multireferenz-Slot verbunden.

Schließlich werden die CPDs der probabilistischen Slot-Instanzen anhand der Messungen aktualisiert. Die Aktualisierung erfolgt aus zweierlei Gründen. Entweder werden durch die Umgebungswahrnehmung konkrete Informationen zu Attributen gemessen, sodass direkt die CPD eines Attributs aktualisiert wird, oder die Anzahl der verknüpften Instanzen eines Multireferenz-Slots hat sich verändert, sodass hierfür eine an die neue Anzahl angepasste CPD bestimmt werden muss.

Im Inferenzschritt werden die unbekannt Parameter der Attribute im Weltmodell bestimmt und Aussagen über Attribute berechnet, die durch die kausalen Zusammenhänge beobachtbar sind. Hierfür wird ein probabilistisches Netz erstellt, welches aus den Frame-Slot-Instanzen abgeleitet wird. Existieren innerhalb des Netzes konstante Parameter, so werden diese markiert und gelten als Vorwissen während des Inferenzschrittes. Für die Berechnung der Inferenz wird das loopy Belief Propagation (BP)-Verfahren gewählt, welches eine approximative Schätzung der Wahrscheinlichkeiten liefert. Die Umsetzung der Inferenzberechnung erfolgt mit der freien Softwarebibliothek `libDAI` [Moo10]. Diese stellt effiziente Inferenzverfahren für Graphen mit diskreten Knoten bereit.

5.2 Ablauf der Modellierung von Fahrsituationen

In diesem Abschnitt wird dargestellt, dass das gewählte Konzept einen generischen, robusten und modularen Ansatz zur Analyse von Fahrsituationen leistet.

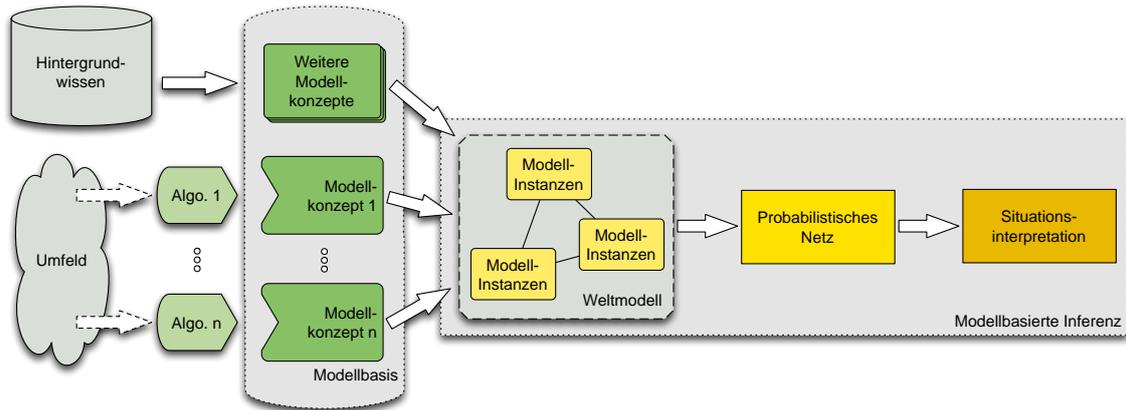


Bild 5.6: Konzeptioneller Zusammenhang von Hintergrundwissen, Umfeldinformationen, informationsverarbeitender Algorithmen, der Modellbasis und der modellbasierten Inferenz zur Situationsinterpretation für FAS.

Das in dieser Arbeit entwickelte Konzept basiert auf der Überführung von zwei verschiedenartigen Informationen in ein adäquates Weltmodell. Einerseits sind das Informationen die durch Hintergrundwissen vorhanden sind, andererseits aktuelle Umfeldinformationen, die mit Hilfe von Sensoren erfasst werden. Mit Hilfe des Weltmodells lassen sich aus der formalen Beschreibung der Objekte und der Relationen zwischen diesen Aussagen über die Fahrsituation ableiten. Dieser Zusammenhang ist in Abbildung 5.5 dargestellt. Für den Aufbau eines Weltmodells muss eine geeignete Repräsentation gewählt werden, mit der Hintergrundwissen und Umfeldinformationen gleichermaßen dargestellt und abgelegt werden können.

Die Repräsentation ist entscheidend für die Interpretation von Situationen. Eine ungeeignete Repräsentation führt entweder zu unnötig hoher Komplexität oder erschwert die Abbildung der durch die Umfelderkennung generierten Informationen. Die Wissensrepräsentation des hier präsentierten Verfahrens erfolgt durch Verwendung der Prädikatenlogik. Damit können logische Zusammenhänge in Form von Entitäten und Relationen dargestellt werden. Die konzeptuelle Idee ist dem Ansatz der wissensgetriebenen SA nachempfunden, welcher in Abschnitt 3.4.3 beschrieben wurde.

In [SZ11] sind die notwendigen Anforderung der Anpassung der Modellbasis an die im Fahrzeug zur Verfügung stehenden Erkennungssysteme beschrieben. Abbildung 5.6 zeigt den konzeptionellen Aufbau der Umsetzung mit verschiedenen Algorithmen zur Wahrnehmung der Umgebung. Die formale Repräsentation ist Grundlage der Modellbasis, in der Modellkonzepte beschrieben sind.

Dem Ansatz liegt die Annahme zugrunde, dass sich Fahrsituationen durch ihre Objekte, deren Attribute sowie die Relationen zwischen diesen beschreiben lassen. Eine Fahrsituation wird im Sinne ihrer Objekte und Relationen formal in einer Modellbasis erfasst. Die Messung der Objekte und Relationen erfolgt durch die im Fahrzeug vorhandenen Erkennungsverfahren. Den Algorithmen lassen sich spezifische Modellkonzepte zuordnen. Weitere Modellkonzepte folgen aus der Beschreibung von Hintergrundwissen. Aus der Modellbasis wird zur Interpretation der Fahrsituation ein Weltmodell abgeleitet und kontinuierlich aktualisiert. Abhängig von den Umfeldinformationen werden Instanzen in das Weltmodell dynamisch eingeführt oder gelöscht. Die Modellinstanzen definieren ein probabilistisches Netz, auf dem mittels Inferenz

Informationen über die Fahrsituation abgeleitet werden.

Spezifikation von Fahrsituationen In der Spezifikation werden die Informationen und Zusammenhänge definiert, auf welche durch SA geschlossen werden soll. Aus Modellierungssicht bedeutet dies, dass entsprechende Schnittstellen zu den Wahrnehmungsverfahren festzulegen sind und der Informationsfluss zwischen Modellen abgebildet wird. Hierbei ist Expertenwissen sehr nützlich, da beispielsweise Wissen über früher modellierte Situationen genutzt werden kann. Objekte und Relationen können für Hintergrundwissen oder Modellkonzepte übernommen werden. Während der Spezifikation wird festgelegt, welche Objekt-Attribute von Unsicherheit behaftet sind und welche Attribute als sicher gegeben sind. Außerdem werden Abhängigkeiten zwischen Objekt-Attributen und Objekten bestimmt.

Definition von Modellkonzepten An die Spezifikation schließt sich die Definition der Modellkonzepte an. Die Modellkonzepte werden formal als Objekte und Relationen in der Modellbasis beschrieben. Die Definition umfasst sowohl die allgemeinen Konzepte des Hintergrundwissens, spezifische Konzepte über Algorithmen sowie relationale Konzepte, die Zusammenhänge zwischen Konzepten herstellen. Zur formalen Beschreibung der Modellkonzepte wird die OPRML verwendet. Dies ermöglicht es, die probabilistischen und relationalen Spezifikationen getrennt und konsistent voneinander zu beschreiben.

Um eine vollständige Definition der Modellkonzepte zu erreichen, muss zunächst der Bereich der beschreibenden Attribute festgelegt werden. Da beschreibende Attribute der Modellkonzepte diskret sind, ist der Wertebereich eine vordefinierte, abzählbare Menge von Werten. Für die Anbindung der Umfelderkennung wird daher festgelegt, wie eine Überführung der kontinuierlichen Variablen in den diskreten Raum erfolgt. Äquivalent werden die probabilistischen Abhängigkeiten zwischen den Attributen und die Abhängigkeit zu anderen Modellkonzepten durch die Verwendung komplexer Attribute festgelegt.

Da die Definition der Modellkonzepte generisch ist, können einmal definierte Konzepte wiederverwendet werden. Dies führt langfristig zu einer Sammlung von Modellkonzepten, welche leicht für die Beschreibung verschiedener Situationen herangezogen werden können und die Beschreibung neuer Situationen vereinfacht.

Aufbau des Weltmodells Kern der modellbasierten Inferenz ist das Weltmodell, welches die Verknüpfung zwischen Modellbasis und Umfeldinformation herstellt. Das Weltmodell entspricht dem aktuellen Zustand der Umgebung, also den tatsächlich in der Umgebung vorhandenen Objekten. Das bedeutet, dass für jedes gemessene Objekt die formalen Modellkonzepte in das Weltmodell eingetragen werden. Dies erfolgt in Form von Modellinstanzen, die direkt aus den Modellkonzepten abgeleitet werden. Die Modellkonzepte sind also Schablonen für Modellinstanzen (siehe Abbildung 5.7).

Die Modellinstanzen werden automatisch, basierend auf Sensorinformationen und Vorwissen erzeugt und die Attribute der Instanzen kontinuierlich aktualisiert. Wird ein Objekt durch die Umfelderkennung nicht mehr wiedergefunden, wird die entsprechende Instanz aus dem Weltmodell gelöscht. Besitzen die Modellkonzepte probabilistische Abhängigkeiten zu anderen Konzepten, werden Verknüpfungen zwischen den konkreten Instanzen hergestellt. Da An-

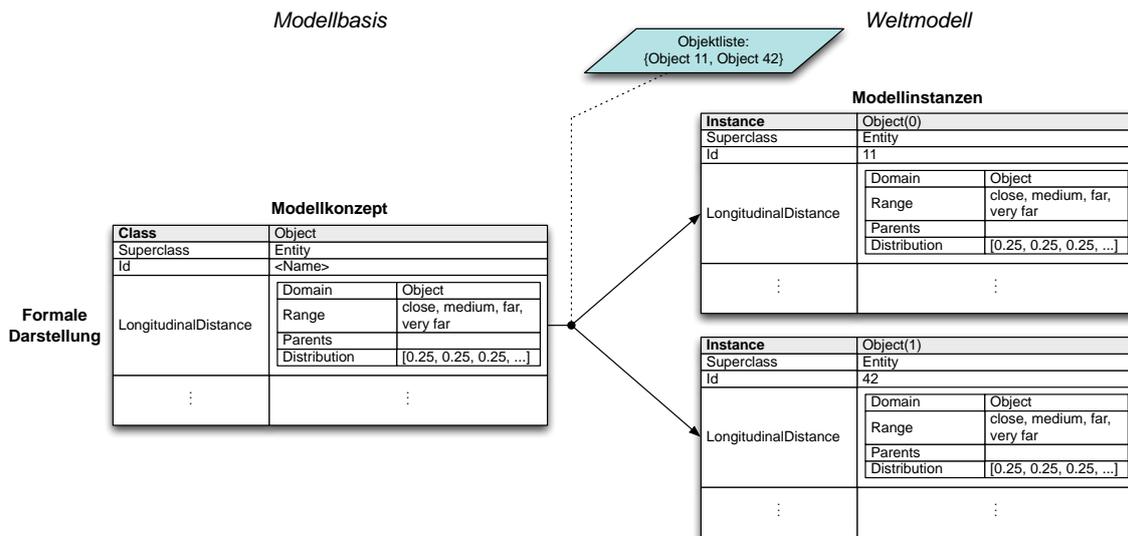


Bild 5.7: Aufbau des Weltmodells: Modellkonzepte der Modellbasis dienen als Schablonen für Modellinstanzen im Weltmodell; das Weltmodell ist abhängig von der Objektliste der Umfelderkennung.

derungen in der Umgebung durch das Umfeldmodell sofort an das Weltmodell weitergegeben werden, verhält sich das System robust. Die Abbildung des Umfeldmodells in das Weltmodell in Form der Attribute sowie der Relationen ermöglicht, dass die dynamischen Aspekte der Umwelt unter Berücksichtigung der Unsicherheiten korrekt durch das Weltmodell repräsentiert werden.

Aufbau des probabilistischen Netzes Das Weltmodell wird dazu genutzt, ein graphisches Modell der probabilistischen Attribute und den kausalen Zusammenhängen zu bilden. Das zugrundeliegende Netz kann als BN betrachtet werden. Es ist allerdings nicht notwendig, das vollständige BN aus dem Weltmodell abzuleiten. Häufig reicht es, Teilbäume des Netzes für einzelne Instanzen des Modells aufzubauen. Dies kann zu einer Reduktion des Aufwands in der Berechnung der Inferenz führen, da sich Änderungen in der Umgebung nur auf Teilkomponenten des Netzes auswirken und Wahrscheinlichkeitsverteilungen für interne Knoten vorberechnet werden können. Hierbei wird die Eigenschaft der Kapselung von Knoten in objektorientierten Bayes'schen Netzen (OOBNs) ausgenutzt, da interne Knoten der OOBN-Klasse d-separiert (*bedingt unabhängig*) gegenüber ihrer Schnittstelle sind.

Das probabilistische Netz und die Inferenzberechnung können getrennt vom Entwurf des Modells und der Instanziierung der Modellinstanzen betrachtet werden. Dies erlaubt eine modulare Entwicklung der Komponenten.

Ableitung der momentanen Fahrsituation Auf Basis des graphischen Modells erfolgt die Bestimmung der Fahrsituation. Aus der Kombination des Zustands des Eigenfahrzeugs, der Zustände relevanter Objekte im Fahrzeugumfeld sowie den Veränderungen dieser Zustände werden neue Informationen berechnet. Durch das probabilistische Netz sind die probabilistischen Attribute und Abhängigkeiten definiert. Umfeldmessungen fließen in das Netz als Evidenzen ein. Zur Bestimmung der Fahrsituation lassen sich Anfragen über die Attribute an

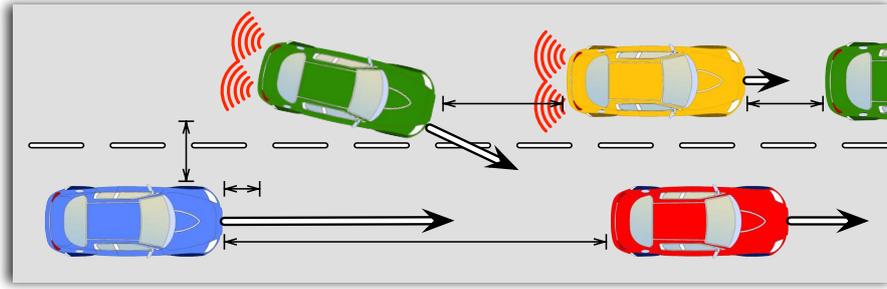


Bild 5.8: Annäherungssituation innerhalb einer Menge von Fahrzeugen.

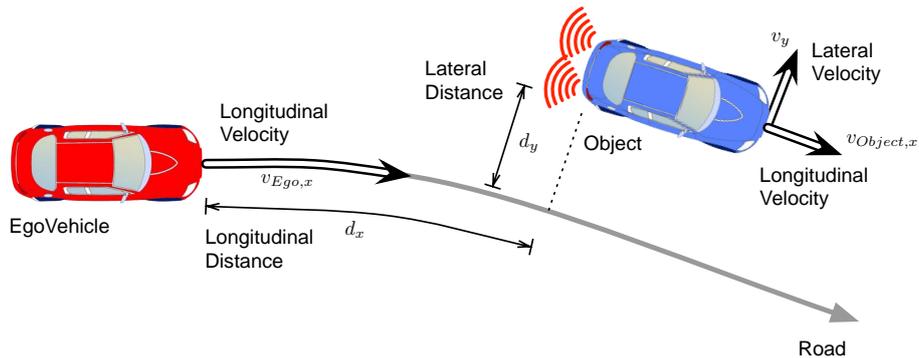


Bild 5.9: Relevante Attribute einer Annäherungssituation an ein vorausfahrendes Fahrzeug.

das Netz stellen. Durch Inferenz kann die exakte oder näherungsweise exakte Wahrscheinlichkeitsverteilung des Netzes bestimmt und daraus die momentane Fahrsituation abgeleitet werden.

5.3 Anwendungsbeispiel der Methodik

Eine beispielhafte Anwendung der oben vorgestellten Methodik verdeutlicht im Folgenden, wie Objektorientierte Probabilistische Relationale Modelle (OPRM) zur probabilistischen Interpretation von Fahrsituationen zum Einsatz kommen [SZ11]. Hierbei wird das Risiko einer Kollision des eigenen Fahrzeugs mit Fahrzeugen in der unmittelbaren Umgebung geschätzt. Das Risiko soll als *gering* bezeichnet werden, wenn der laterale Abstand zwischen zwei Fahrzeugen *groß* ist. Es soll als *hoch* bezeichnet werden, wenn sowohl der laterale als auch der longitudinale Abstand *gering* ist und eine Annäherung der Fahrzeuge stattfindet. Ein hohes Kollisionsrisiko muss dennoch nicht zwingend zu einer Kollision führen.

In Abbildung 5.8 ist eine Fahrsituation dargestellt, wie dies auf einer mehrspurigen Straßen bei dichtem Verkehr auftreten kann. Hier ist die Schätzung des Risikos einer möglichen Kollision mit den unmittelbar benachbarten Fahrzeugen von Bedeutung. Durch die gewählte Methodik ist es möglich, unter Einbeziehung aller Objekte in der Umgebung des Fahrzeugs eine gemeinsame Betrachtung des Risikos durchzuführen.

Bestimmung der relevanten Attribute Ausgehend von der skizzierten Fahrsituation werden die wesentlichen Attribute identifiziert, welche durch die Umgebungserkennung gemessen werden können. Zur Schätzung des Risikos besitzen die Attribute *Geschwindigkeit* sowie

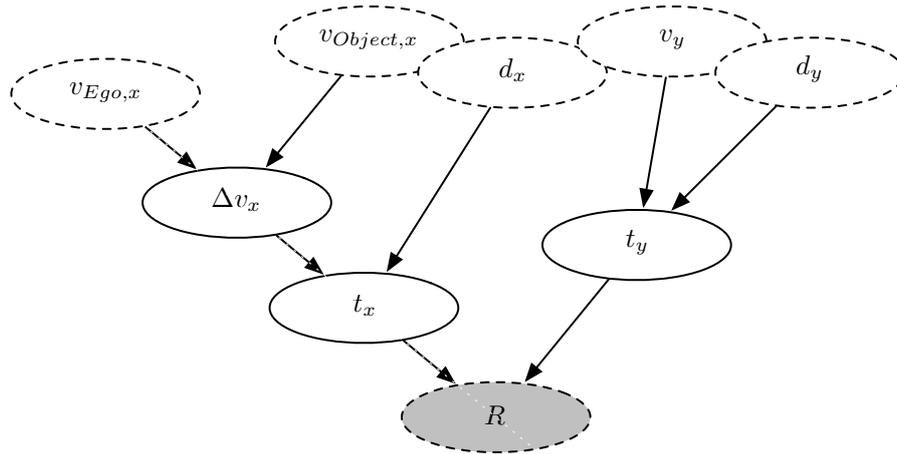


Bild 5.10: Graphisches Modell zur Bestimmung des Kollisionsrisikos gegenüber einem weiteren Objekt entsprechend zur Abbildung 5.9.

Abstand zwischen Fahrzeugen eine große Bedeutung. In Abbildung 5.9 sind die relevanten Attribute zwischen zwei Fahrzeugen schematisch dargestellt. Durch die Umfelderkennung erfolgt zunächst eine Schätzung der longitudinalen und lateralen Distanz d_x und d_y zum jeweiligen Objekt. Durch temporale Filterung können unabhängig von der eingesetzten Umgebungssensorik die relativen longitudinalen und lateralen Geschwindigkeiten Δv_x und v_y bestimmt werden.

Wie in [Kuh11] beschrieben, erfolgt die Bestimmung dieser Attribute entlang eines gekrümmten Koordinatensystems, welches an der Fahrspur des eigenen Fahrzeugs (Ego-Fahrzeug) ausgerichtet ist. Die relative Geschwindigkeit Δv_x ist die Differenz zwischen der absoluten Fahrzeuggeschwindigkeit $v_{Ego,x}$ und der Geschwindigkeit des Objektes $v_{Object,x}$. Das Risiko einer Kollision R ist als Funktion abhängig von den Kollisionszeiten t_x und t_y :

$$R(t_x, t_y). \quad (5.1)$$

Die oben beschriebenen Zusammenhänge lassen sich in Form bedingter Wahrscheinlichkeiten als BN, wie in Abbildung 5.10 gezeigt, darstellen:

$$\Delta v_x(v_{Ego,x}, v_{Object,x}) \Rightarrow P(\Delta v_x | v_{Ego,x}, v_{Object,x}) \quad (5.2)$$

$$t_x(d_x, \Delta v_x) \Rightarrow P(t_x | d_x, \Delta v_x) \quad (5.3)$$

$$t_y(d_y, v_y) \Rightarrow P(t_y | d_y, v_y) \quad (5.4)$$

$$R(t_x, t_y) \Rightarrow P(R | t_x, t_y). \quad (5.5)$$

Daraus lässt sich die Kritikalität R gegeben den Eingangsvariablen als Produkt der Einzelwahrscheinlichkeiten berechnen:

$$P(R | v_{Ego,x}, v_{Object,x}, d_x, v_y, d_y) = P(R | t_x, t_y) \cdot P(t_x | d_x, \Delta v_x) \cdot P(t_y | d_y, v_y) \cdot P(\Delta v_x | v_{Ego,x}, v_{Object,x}). \quad (5.6)$$

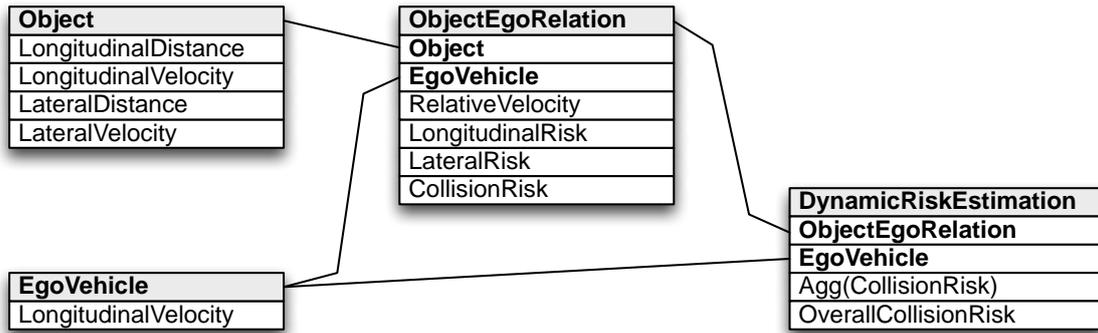


Bild 5.11: Relationale Darstellung der Klassen und Attribute zur Beschreibung der Fahrsituation; fett geschriebene Attribute stellen Relationen auf Fremdklassen her.

Modellierung der Fahrsituation In Abbildung 5.11 ist die relationale Darstellung der Zusammenhänge abgebildet. Die Modellierung der Fahrsituation erfolgt in Entity-Relationship (ER)-Darstellung durch Ausnutzung der Logik erster Ordnung (FOL). Die Klassen dieser Domäne sind $C = \{Object, EgoVehicle, ObjectEgoRelation, DynamicRiskEstimation\}$ und stehen untereinander in Relation. Die Klasse *ObjectEgoRelation* importiert Informationen der Klassen *Object* sowie *EgoVehicle*. Die Klasse *DynamicRiskEstimation* importiert Informationen der Klasse *ObjectEgoRelation* sowie *EgoVehicle*.

Attribute sind klassenspezifisch. Für die Klasse *Object* sind Attribute $\Delta_{Object} = \{LongitudinalDistance, LateralDistance, \dots\}$ mit fest definierten, abzählbaren Wertebereichen. Für das Attribut ist der Wertebereich $Val(Object.LongitudinalDistance) = \{close, medium, far, very\ far\}$.

Die Beschreibung der Relationen zwischen Objekten erfolgt durch referenzierende Attribute. In dieser Domäne sind dies $\Phi_{ObjectEgoRelation} = \{Object, EgoVehicle\}$ und $\Phi_{DynamicRiskEstimation} = \{ObjectEgoRelation, EgoVehicle\}$. Der Wertebereiche von $\Phi_{ObjectEgoRelation}$ ist gegeben durch $Range(ObjectEgoRelation.Object) = \{Object\}$ und $Range(ObjectEgoRelation.EgoVehicle) = \{EgoVehicle\}$. In Abbildung 5.11 sind die Attribute fett geschrieben, welche auf eine Fremdklasse verweisen.

Anhand der relationalen Darstellung kann an dieser Stelle bereits gezeigt werden, dass pro Objekt eine eigene Risiko-Berechnung gegenüber dem Ego-Fahrzeug erfolgt. In der Klasse *DynamicRiskEstimation* werden die Einzelrisiken zu einem Gesamtrisiko fusioniert. Die Aggregation der Einzelrisiken erfolgt über das Aggregationsattribut. Dieses ist definiert als $Agg(DynamicRiskEstimation.CollisionRisk) = \max(ObjectEgoRelation.CollisionRisk)$. Die Maximum-Aggregation bestimmt eine CPD, welche der Wahrscheinlichkeitsverteilung des Objektes mit dem höchsten Risiko entspricht.

Aufbauend auf der ER-Beschreibung erfolgt die Definition des probabilistischen Modells. Dabei werden die probabilistischen Abhängigkeiten zwischen Attributen sowie die CPDs festgelegt. Für jedes probabilistische Attribut sind dessen Wertebereich sowie die Wertebereiche der Elternknoten zu definieren. Für referenzierende Attribute ist festzulegen, welche probabilistischen Attribute der Fremdklasse in die eigene Klasse importiert werden. Die Bestimmung der jeweiligen CPD erfolgt durch Expertenwissen oder durch maschinelles Lernen. Das probabilistische Modell stellt die initiale Wahrscheinlichkeitsverteilung der Modellbasis dar. CPDs der einzelnen Attribute können nach Instanziierung der Klasse auf Instanzebene verfeinert

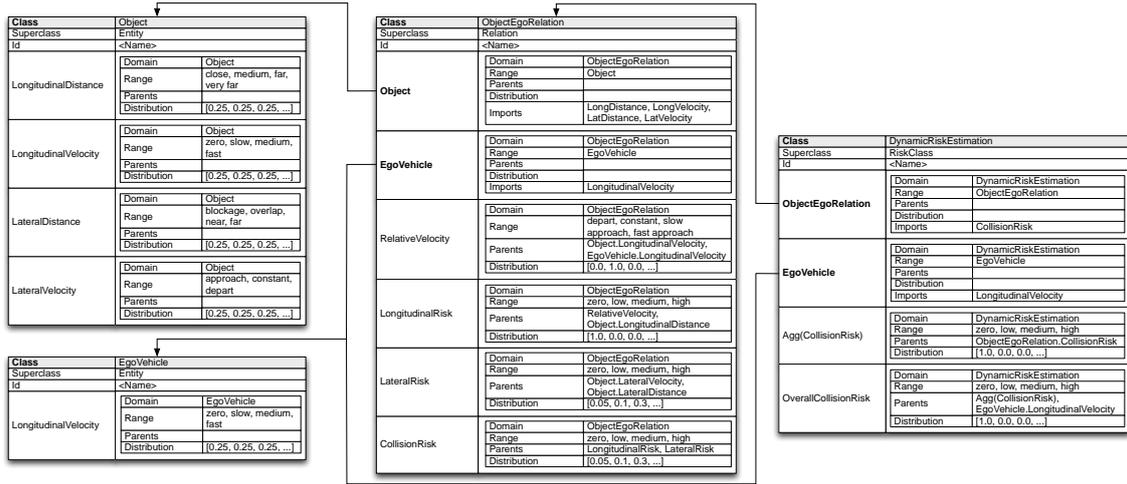


Bild 5.12: Formale Darstellung der Modellbasis als Klassenstruktur zur Beschreibung der Fahrsituation.

werden.

Das Einlernen der CPDs anhand von Trainingsdaten erfolgt auf aus dem Modell abgeleiteten OOBNs (vergleiche Abbildung 5.13). Die Verfahren zum Parameterlernen auf OOBNs zeigen hierfür die größte Eignung. Diese schätzen die CPD eines probabilistischen Attributs unter Berücksichtigung der Evidenz aller Instanzen der Klasse sowie der Klassenstruktur. Im Rahmen der vorliegenden Arbeit konnte dieser Ansatz nur konzeptuell untersucht werden.

Die OPRML-Darstellung beschreibt formal alle Fahrsituationen, wie sie in der modellierten Welt auftreten können. Abbildung 5.12 zeigt die Repräsentation der Fahrsituation in OPRML-Darstellung.

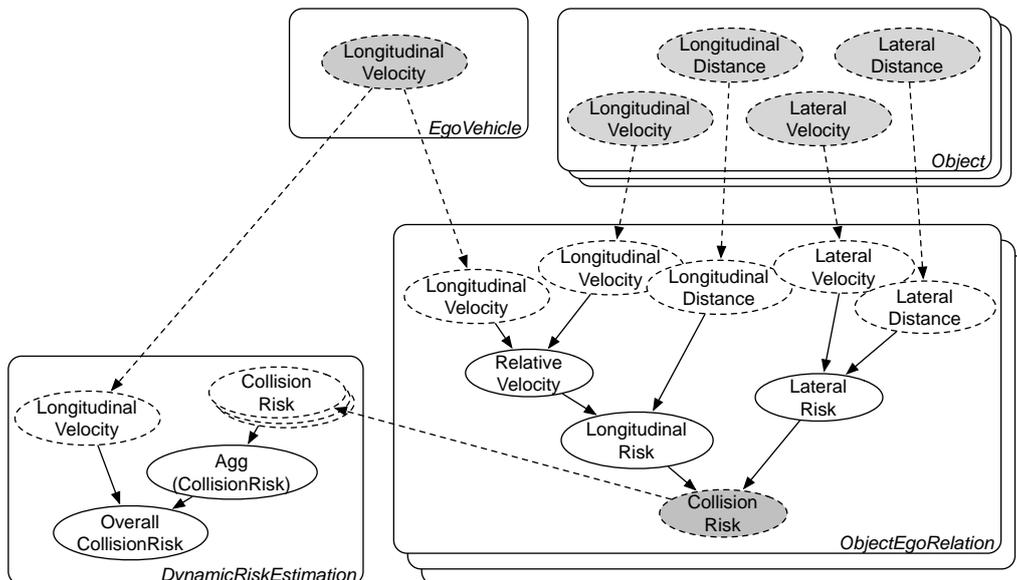


Bild 5.13: Abgeleitetes OOBN für eine Situation mit drei Fahrzeugen im Weltmodell.

Instanziierung des Weltmodells Aus dieser Modellbasis lässt sich zur Laufzeit ein instanziiertes Weltmodell ableiten sobald Informationen der Umfelderkennung vorhanden sind. Dies entspricht einer Konkretisierung der Modellbasis gegeben der Umfeldinformationen. Aus den

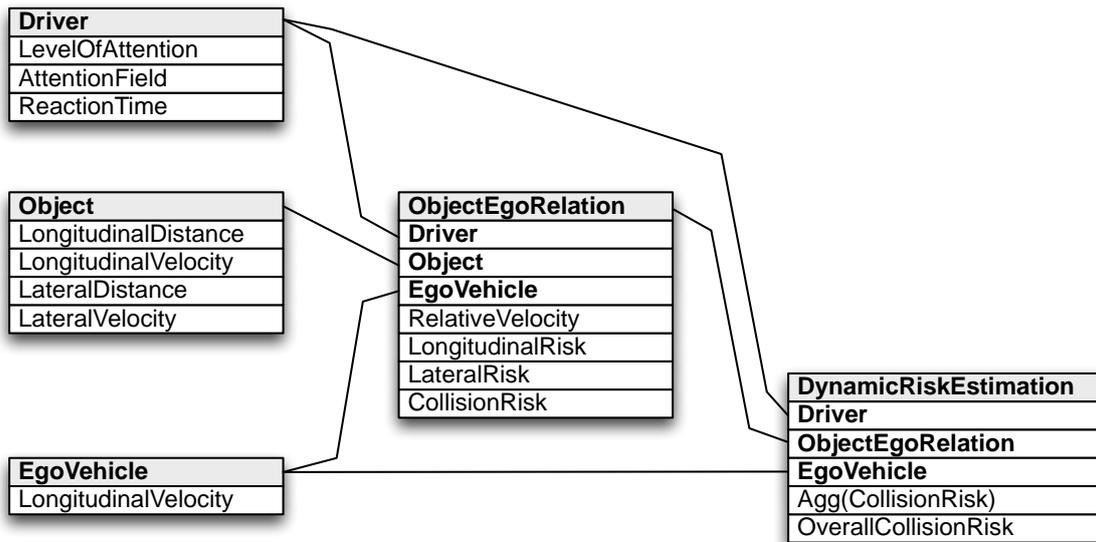


Bild 5.14: Erweiterung der Fahrsituation in relationaler Darstellung zur Betrachtung des Einflusses des Fahrers.

Klassen werden die jeweiligen Instanzen abgeleitet und deren CPDs anhand von Evidenzen der Umfeldmessungen spezifiziert. In Abbildung 5.13 ist das zu dem Weltmodell korrespondierende OOBN dargestellt, welches das probabilistische Modell der Fahrsituation zeigt. Für jede OOBN-Klasse entsprechen grau gestrichelte Knoten den Ausgabe-Knoten und weiß gestrichelte Knoten den Eingabe-Knoten. Diese bilden Referenzen auf Fremdklassen. Weiße Knoten mit durchgezogenen Linien entsprechen den probabilistischen Attributen. In der Abbildung ist ein probabilistisches Attribut beispielsweise *Longitudinal Risk*, welches von den Attributen *Relative Velocity* und *Longitudinal Distance* für jedes Tupel von *EgoVehicle* \times *Object* abhängt.

Zur Bestimmung der Wahrscheinlichkeitsverteilung über die Fahrsituation wird das Kollisionsrisiko gegenüber allen Tupeln durch den Aggregationsknoten *Agg(CollisionRisk)* berechnet. Dessen CPD ist abhängig von der Anzahl aller *ObjectEgoRelation*-Instanzen der Fahrsituation und muss aktualisiert werden, sobald sich die Anzahl der Instanzen verändert. In Abschnitt 6.4.1 wird das spezifizierte Modell anhand unterschiedlicher Szenen evaluiert und auf Eignung zur Bestimmung der Fahrsituation hin untersucht.

Erweiterung des Modells Die in Abbildung 5.11 gezeigte relationale Darstellung der Fahrsituation lässt sich aufgrund der eingesetzten Repräsentation leicht und nachvollziehbar erweitern. Dies soll anhand des folgenden Beispiels konzeptionell dargestellt werden.

Der Fahrer des Ego-Fahrzeugs hat auf die vorherrschende Fahrsituation und insbesondere auf die Kritizität der Situation einen großen Einfluss. Daher kann das Risiko einer Kollision mit einem Objekt auch Fahreradaptiv ausgewertet werden. Die Auslösung einer Warnung kann beispielsweise verzögert geschehen, wenn es sich um einen sehr aufmerksamen oder geübten Fahrer handelt. Dagegen könnte eine Warnung früher ausgelöst werden, falls der Fahrer ein relevantes Objekt nicht beobachtet hat. Zur Erkennung des Fahrstils sowie zur Überwachung des Aufmerksamkeitsfeldes des Fahrers in Relation zum Fahrzeugumfeld existieren geeignete Methoden [Die09; Bär+11; Bär+13].

Abbildung 5.14 zeigt die um die Klasse *Driver* erweiterte relationale Darstellung. Die Klasse besitzt drei Attribute: *LevelOfAttention*, *AttentionField* sowie *ReactionTime*. Die Attribute *AttentionField* und *ReactionTime* werden pro Objekt ausgewertet und bestimmen, wie das objektspezifische Risiko in der Klasse *ObjectEgoRelation* eingeschätzt wird. Das Attribut *LevelOfAttention* beeinflusst schließlich den Zeitpunkt einer Warnung, indem das Gesamtrisiko der Klasse *DynamicRiskEstimation* adaptiv an die Aufmerksamkeit des Fahrers angepasst wird.

Diskussion Anhand dem Anwendungsbeispiel der Methodik wird deutlich, dass der gewählte Ansatz generisch, robust und modular ist. Das vorangegangene Beispiel zeigt (wenn auch nur konzeptionell) auf, dass modellierte Fahrsituationen einfach erweiterbar sind. Die automatische Modellkonstruktion und die zugrundeliegenden OOBNs ermöglichen ein konsistentes Verhalten gegenüber einer variablen Anzahl von Objekten im Fahrzeugumfeld. Der gewählte Ansatz ist in der Lage, die deterministische Vorgehensweise von sicherheitsrelevanten Aspekten abzubilden. Dazu gehören beispielsweise die Kollisionszeit (TTC) oder die Reaktionsreservezeit (TTR). Für die Einschätzung eines Kollisionsrisikos sind dies neben der 3D-Position der Objekte die wichtigsten Parameter.

Des Weiteren ermöglicht der hier präsentierte Ansatz, Unsicherheiten in den Attributwerten zu berücksichtigen. Hierdurch lassen sich Sonderfälle vermeiden, welche in der gängigen Praxis bei der Betrachtung von Unsicherheiten in deterministischen Ansätzen konstruiert werden. Die in dieser Arbeit gewählte OPRML kann neben Attributunsicherheiten auch Struktur- und Existenzunsicherheiten berücksichtigen. Diese wurden im Rahmen der Arbeit nicht näher untersucht.

6 Evaluation und Ergebnisse

Die in den vorausgegangenen Kapiteln vorgestellten Verfahren zur probabilistischen Analyse von Fahrsituationen wurden durchgängig unter realen Bedingungen evaluiert. Ziel der Evaluation ist der Nachweis, dass die Systeme folgende Anforderungen erfüllen:

- Erkennung von Situationen mit ausreichend hoher Performanz,
- Robustheit gegenüber sich verändernden Situationen,
- Erweiterbarkeit bezüglich der verwendeten Modelle, und
- Informationsverarbeitung mit echtzeitfähiger Berechnungsdauer.

6.1 Grundlagen der Evaluation

Die Eignung der Verfahren zur Informationsbereitstellung für Assistenzsysteme ist dann gewährleistet, wenn eine Verarbeitung der Information innerhalb einer angemessenen Zeitspanne durchgeführt werden kann. In Serien-Fahrzeugen werden zur Verarbeitung spezialisierte Steuergeräte (ECU) eingesetzt, wofür bei der Entwicklung eine enge Abstimmung zwischen eingesetzter Hardware und der datenverarbeitenden Software erfolgt. Aufgrund der prototypischen Entwicklung der Verfahren erfolgt die Informationsverarbeitung in dieser Arbeit auf handelsüblichen Rechnersystemen.

Zur Erkennung von Situationen muss das System eine ausreichend hohe Performanz erreichen. Wie in Abschnitt 5.1 dargestellt wurde, basiert die Situationsanalyse (SA) direkt auf den erzeugten Objektlisten der Umfelderkennung. Die Qualität der SA hängt daher von der Performanz der Umfelderkennung ab. In dieser Arbeit wurde auf eine hohe Erkennungsgüte von Objekten im Fahrzeugumfeld Wert gelegt, was durch eine Vielzahl an Tests belegt wurde.

Die im Stand der Technik vorgestellten Verfahren zu SA mit probabilistischen Netzen haben gemeinsam, dass eine Anpassung der SA an die vorherrschende Situation im Sinne eines robusten Verhaltens nicht dynamisch erfolgt. Die Verfahren umgehen diese Herausforderung, indem Annahmen über die maximale Anzahl der vorkommenden Objekte getroffen werden und statische Verarbeitungsmechanismen eingesetzt werden (vergleiche Abschnitt 3.4). In dieser Arbeit wird die Robustheit des Systems gegenüber veränderlichen Situationen darin verstanden, dass eine SA zu jedem Zeitpunkt spezifisch an die tatsächlich in der Umgebung vorkommenden Objekte angepasst ist.

Im Rahmen der Arbeit wurden verschiedene Sensorsysteme in Versuchsfahrzeuge integriert und auf ihre Eignung zur Umfeldwahrnehmung hin untersucht. Die Evaluation der Verfahren zur Umfelderkennung und zur SA erfolgte auf Basis von Aufnahmen der Versuchsfahrzeuge im Labor sowie durch Einsatz der Verfahren bei Testfahrten. Die Evaluation der Performanz der

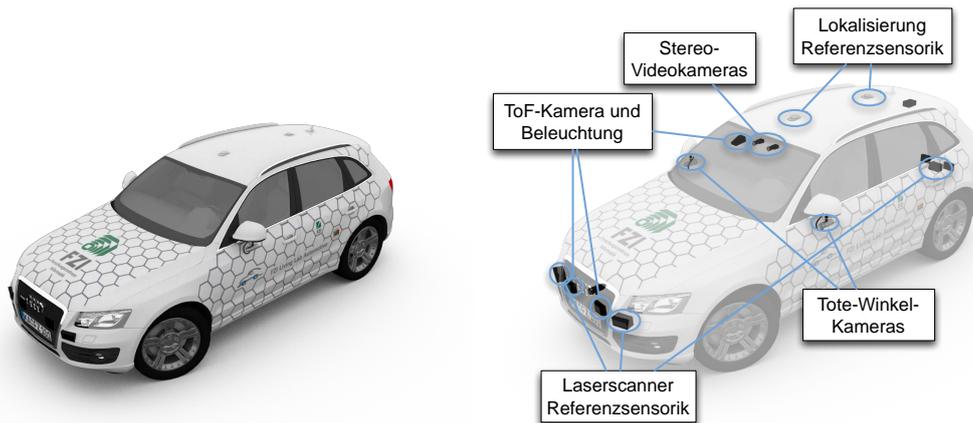


Bild 6.1: Versuchsfahrzeug „CoCar“ und dessen sensorische Abdeckung des Fahrzeugumfelds durch verschiedene, zum Teil redundante Sensoren.

Umgebungserkennung erfolgte auf einem Testgelände, unter realen Bedingungen im Straßenverkehr sowie im Labor und auf Fußgängerwegen. Die Evaluation der SA erfolgte zu großen Teilen unter realen Bedingungen im Straßenverkehr sowie auf Fußgängerwegen. Die aufgezeichneten Daten der Versuchsfahrzeuge umfassen die Sensor-Rohdaten der Lichtlaufzeit (ToF)-, Video- und Referenzsensorik, die fahrzeuginterne Sensorik (beispielsweise des CAN-Busses) sowie die Lokalisierungsdaten der Plattform. Alle aufgezeichneten Daten wurden mit Zeitstempeln versehen, sodass die Daten in Offline-Versuchen zeitsynchron wiederholt abgespielt werden konnten.

6.2 Evaluierungsplattformen

Als Evaluierungsplattform steht eine Fahrsimulationsumgebung zum Test von Fahrfunktionen und zur Überwachung des Fahrers in virtuellen Umgebungen sowie das instrumentierte Testfahrzeug „CoCar“ (*Cognitive Car*) auf Basis eines Audi Q5 zur Verfügung. Das Testfahrzeug ist umfangreich mit Sensorik-Systemen zur Überwachung des Fahrzeugumfelds sowie des Fahrzeuginnenraums ausgestattet (siehe Abbildung 6.1). Weiter ist das Fahrzeug mit Aktuatoren ausgestattet, welche Lenkung, Gas- und Bremspedal des Fahrzeugs per CAN-Bus ansteuern. Somit kann das Fahrzeug voll- oder teilautomatisiert durch die im Kofferraum verbauten Computersysteme angesteuert werden. Die Ausstattung und die Fähigkeiten des Fahrzeugs sind damit auf technischer Ebene mit Fahrzeugen verwandter Forschergruppen wie *Team Annieway*, *Stadtpilot* oder *Stanford Racing Team* sowie der Fahrzeuge von *Google* oder *Daimler* vergleichbar.

Im Versuchsfahrzeug „CoCar“ (siehe Abbildung 6.1) kommt eine Photomischdetektor (PMD) ToF-Kamera mit einer Auflösung von 64×16 Bildpunkten zum Einsatz. Der Sensor ist nach vorne, also entlang der x -Achse des Fahrzeugs, ausgerichtet und liefert Tiefendaten mit einer Bildwiederholrate von 50 Hz bis zu einer maximalen theoretischen Reichweite von 150 m. Das Umfeld vor dem Fahrzeug wird außerdem durch ein Stereo-Videokamerasystem sowie eine weitere Videokamera hinter der Windschutzscheibe und eine Videokamera in der Motorhaube überwacht. Die Daten werden mit einer Bildwiederholrate von 26 Hz und einer



Bild 6.2: PMD- und Videodaten aus Sicht von „CoCar“; oben: Videodaten nach Linsenzerrung mit Überlagerung der PMD-Daten; unten links: Tiefendaten der PMD-Kamera in Falschfarbendarstellung; unten rechts: Amplitudendaten der PMD-Daten.

Auflösung von 1024×480 Bildpunkten geliefert. Für die Überwachung des Bereichs neben dem Fahrzeug sind am linken und rechten Außenspiegel zwei Farb-Videokameras mit einer Auflösung von 640×480 Bildpunkten angebracht. Die Bildwiederholrate beträgt bei diesen Kameras 30 Hz. Eine ToF-Kamera und eine Videokamera hinter der Heckscheibe überwachen schließlich den Fahrzeug-Heckbereich.

Abbildung 6.2 zeigt die Sensordaten für eine Beispielszene aus Sicht von „CoCar“. Für die Darstellung der Videodaten wurde die Linsenverzerrung korrigiert und die Daten in Graustufen konvertiert. Durch Projektion der PMD-Tiefendaten in das Kamerakoordinatensystem der Videokamera wurden die Videodaten mit den Tiefendaten in Falschfarbendarstellung überlagert. In der Falschfarbendarstellung entsprechen warme Farben einer kleinen Entfernung und kalte Farben einer großen Entfernung zur Kamera.

Zusätzlich sind als Referenzsensoren drei ibeo LUX Laserscanner im Fahrzeug integriert. Jeder Sensor tastet die Umgebung mit vier vertikal aufgefächerten Laserzeilen über einen Sichtbereich von 110° ab [ibe13]. Zwei Sensoren sind nach links beziehungsweise rechts vorne und ein Sensor nach hinten ausgerichtet. Die Sensoren übertragen ihre Messdaten an eine ECU, welche eine fusionierte Punktwolke der Umgebung mit einer Datenwiederholrate von 25 Hz liefert. Als zusätzlicher Referenzsensor kann ein Velodyne HDL-32E Laserscanner genutzt werden, welcher auf dem Dach des Fahrzeugs montiert ist und mit 32 vertikal aufgefächerten Laserzeilen über einen Sichtbereich von 360° die Fahrzeugumgebung abtastet [Vel13]. Über den fahrzeuginternen CAN-Bus stehen weitere Informationen wie beispielsweise der momentane Lenkwinkel oder die Fahrzeuggeschwindigkeit zur Verfügung. Eine fest im Fahrzeug verbaute inertielle Messeinheit (IMU) sowie zwei auf dem Dach verbaute GPS-Antennen ermöglichen eine genaue Lokalisierung des Fahrzeugs, welche durch Nutzung von DGPS-Korrekturdaten weiter verbessert wird. Durch die Innenraum-Sensorik ist es außerdem

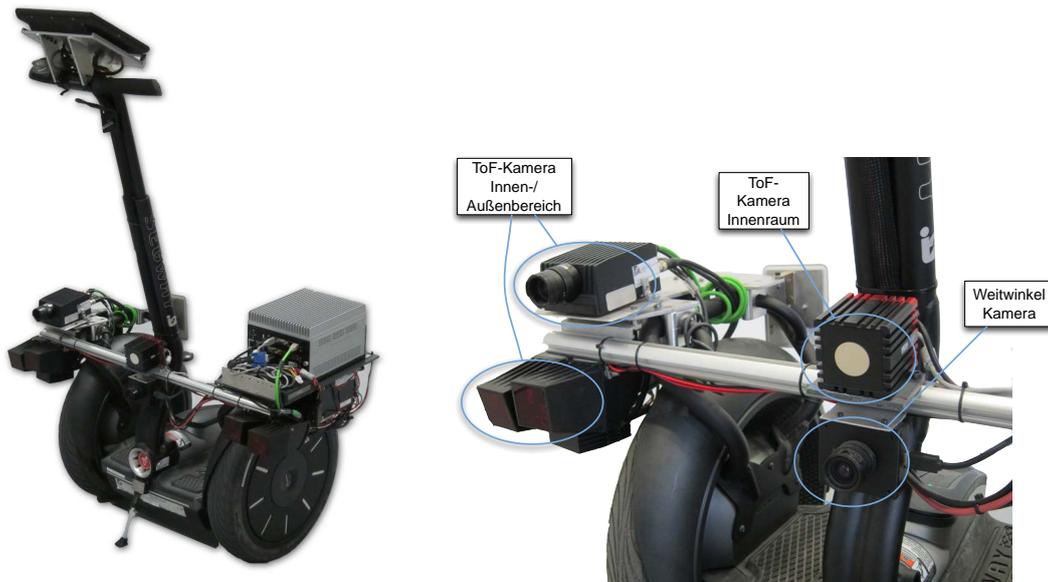


Bild 6.3: Versuchsfahrzeug „SegIT“ und dessen sensorische Abdeckung des Fahrzeugumfelds.

möglich, während der Fahrt Informationen über den Fahrer zu erlangen, wie beispielsweise dessen Aufmerksamkeitsfeld oder dessen momentanes Fahrverhalten [Bär+13].

Eine weitere Plattform zur Evaluation der Verfahren ist der „SegIT“ (*Segway mit IT*), ein mit Sensoren, Computer und Display ausgestatteter Segway PT i2 [Seg13], welcher in Abbildung 6.3 dargestellt ist. Eine Besonderheit dieser Plattform liegt in der Möglichkeit, die entwickelten Verfahren nicht ausschließlich im Straßenverkehr, sondern auch für Assistenzaufgaben im innerstädtischen Bereich wie beispielsweise auf Fußwegen zu evaluieren. Ein Transfer der Verfahren auf weitere Anwendungsgebiete erlaubt eine Aussage über die Spezifität und die Robustheit der getesteten Ansätze. Die Plattform „SegIT“ sowie die autonom fahrenden Varianten „Samson“ und „CityPod“ sind näher in [Str+09; Sch+09; SSZ12] beschrieben.

Auf der Plattform „SegIT“ stehen zwei ToF-Kameras zur Verfügung. Eine Kamera besitzt eine Auflösung von 174×144 Bildpunkten bei einer maximalen Reichweite von 7,5 m und ist für den Einsatz bei Fremdlicht (beispielsweise Sonnenlicht) nicht geeignet. Die zweite ToF-Kamera besitzt eine Auflösung von 64×48 Bildpunkten bei einer Reichweite von 30 m. Diese Kamera besitzt wie die Kamera im Versuchsfahrzeug „CoCar“ eine Unterdrückung von Hintergrundlicht, wodurch der Einsatz auch bei direktem Sonnenlicht ermöglicht wird. Eine Farb-Videokamera liefert Bilddaten mit Sichtbereich entlang der x -Achse (nach vorne) und einer Auflösung von 640×480 Bildpunkten. Zur Verarbeitung der Sensordaten ist ein Computer auf der Plattform vorhanden, die Darstellung der berechneten Informationen und die Steuerung des Computers erfolgen über ein Touchscreen-Display, welches am Lenker der Plattform befestigt ist. Als zusätzliche Informationsquelle steht eine einfache IMU zur Verfügung, welche Daten der beiden Rad-Drehgeschwindigkeiten der Plattform bereitstellt. Im Außenbereich können außerdem Lokalisierungsinformationen mittels eines GPS-Empfängers gemessen werden.

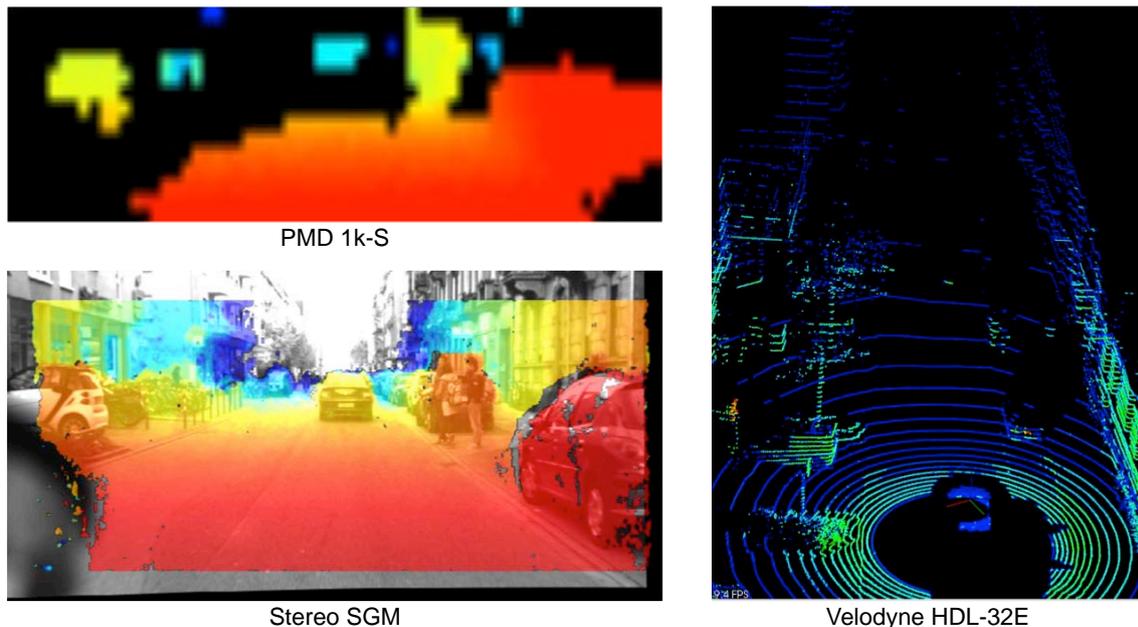


Bild 6.4: Tiefendaten dreier unterschiedlicher Messsysteme: ToF PMD 1k-S Kamera, Stereo SGM-Verfahren, Velodyne HDL-32E Laserscanner.

6.3 Evaluation der Umfelderkennung

Die Bewertung der Performanz der Verfahren zur Umfelderkennung erfolgt anhand aufgezeichneter Sequenzen. Abbildung 6.4 zeigt die Tiefendaten einer Szene, aufgezeichnet mit den in „CoCar“ verbauten Perzeptionssystemen.

Zunächst wird die Erkennungsleistung der ToF-Kamera im Versuchsfahrzeug „CoCar“ bewertet, anschließend folgt eine Evaluation der Verfahren zur Umfelderkennung durch Bewertung der Segmentierungsverfahren der ToF-Kamera. Diese Verfahren ermöglichen unabhängig von spezifischen Objektklassen eine Erkennung von Hindernissen. Anschließend folgt die Evaluation der videobasierten Verfahren für spezifische Objektklassen.

6.3.1 Erkennungsleistung von ToF-Kameras

Soweit nicht anders beschrieben, wird im Folgenden die Performanz der im Versuchsfahrzeug „CoCar“ verbauten ToF-Kamera bewertet. Es handelt sich um eine PMD-Kamera mit einem 1k-S Sensor und einer Auflösung 64×16 Bildpunkten. In Abschnitt 2.1.1 wurde das Messprinzip von ToF-Kameras vorgestellt. Aufgrund der aktiven Beleuchtung der Kameras ist es annähernd unabhängig von der Lichtsituation der Umgebung möglich, Tiefendaten der Szene aufzuzeichnen. Nachteil der aktiven Beleuchtung ist im Gegenzug allerdings, dass Fehler durch mangelnde Ausleuchtung der Szene oder durch Überbelichtung von Objekten die Messqualität beeinflussen, ebenso kann bei starkem Fremdlicht die Messqualität sinken. Die Szene wird durch zwei im Kühlergrill des Fahrzeugs verbaute Infrarotlicht (IR)-Quellen beleuchtet. Die Kamera ist hinter der Windschutzscheibe verbaut.

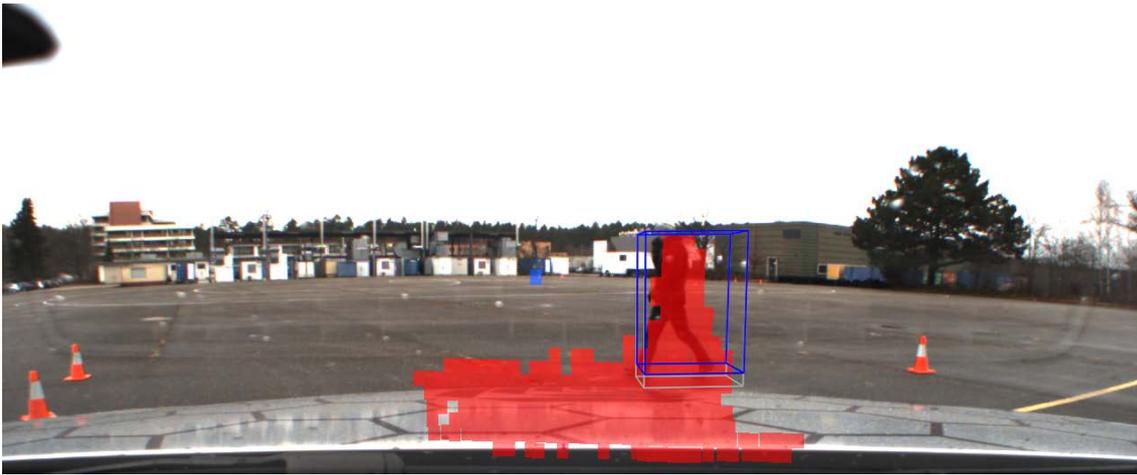


Bild 6.5: Querender Fußgänger in etwa 10 m Abstand zur Fahrzeughinterachse; Videodaten sind mit Tiefendaten der PMD 1k-S Kamera in Falschfarbendarstellung überlagert.

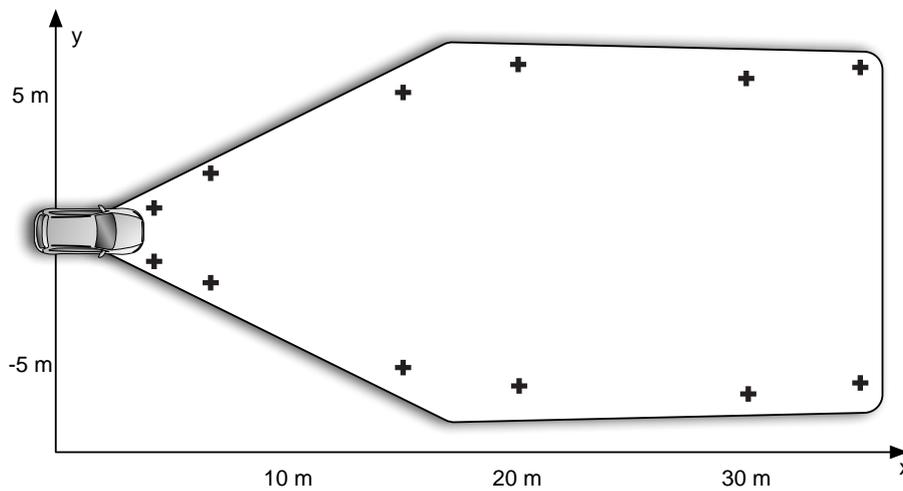


Bild 6.6: Birnenförmiges Ausleuchtungsfeld der PMD-Kamera, gemessen an der Erstdetektion von Fußgängern; die Markierungen stellen die Erstdetektion in den unterschiedlichen Entfernung dar.

Reichweite des Messsystems

Die Reichweite des Messsystems ist abhängig von den Reflektanzeigenschaften der Objekte vor dem Fahrzeug. Üblicherweise erreichbare Messreichweiten sind für Personen etwa 30 m, für vorausfahrende Fahrzeuge etwa 60 m. Transporter, LKWs sowie Objekte mit stark reflektierenden Materialien wie beispielsweise Richtungswegweiser können bis in deutlich höhere Entfernungen detektiert werden.

In Abschnitt 4.2 wurde die Messfläche eines Bildpunktes der PMD-Kamera quantifiziert. Bei der gewählten Optik der Kamera wird eine Person in 25 m Entfernung auf etwa 5 Bildpunkten oder weniger abgebildet. Bei vorausfahrenden Fahrzeugen wirkt sich vor allem die Reflektanzeigenschaft des Kennzeichens für die Detektion positiv aus. Erst im Nahbereich von etwa 20 m ist es möglich, ein Fahrzeug mit seiner vollständigen Kontur in den ToF-Kameradaten auszumachen.

Abbildung 6.6 zeigt das Ausleuchtungsfeld der PMD-Kamera für die Erstdetektion eines

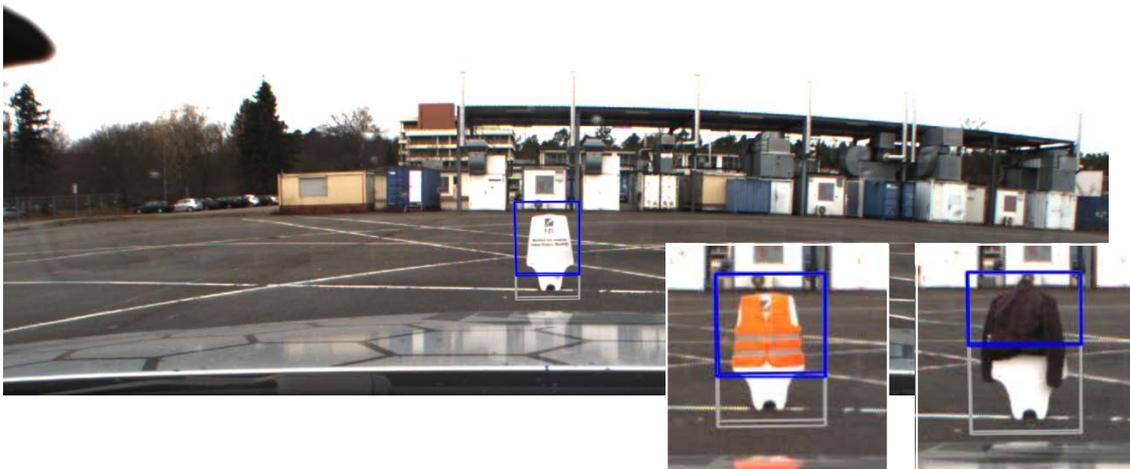


Bild 6.7: Messobjekt „FZI Zwerg“ mit verschiedenen Reflektanzeigenschaften: Original, Reflektorweste, dunkle Jacke.

Fußgängers. Das Ausleuchtungsfeld der aktiven IR-Quellen entspricht erwartungsgemäß einer birnenähnlichen Form. Im Nahbereich der Kamera lässt sich ein Fußgänger damit detektieren, sobald er den Sichtkegel der Kamera betritt. Aufgrund der mittenorientierten Ausleuchtung der IR-Strahler verringert sich die laterale Abdeckung mit zunehmender Entfernung, in einer Entfernung von etwa 20 m aufwärts kann der Fußgänger wegen der geringen Menge an reflektiertem Licht erst in zentraler Position vor dem Fahrzeug gemessen werden.

Der Einfluss der Reflektanz auf die Reichweite der Kamera kann durch Vergleich der Messgüte bei Objekten mit verschiedener Reflektanz verdeutlicht werden. In Abbildung 6.7 ist ein Objekt mit 1 m Höhe und 0,5 m Breite gezeigt. Das Objekt ist zunächst mit einer weißen, im IR-Bereich gut reflektierenden Oberfläche beschichtet. Das zweite Bild zeigt das Objekt mit einer Reflektorweste, welche eine sehr hohe IR-Reflektivität besitzt. Das dritte Bild zeigt das Objekt mit einer dunklen Jacke, welche auch im IR-Bereich ein schlechtes Reflektanzverhalten besitzt.

Die Auswertung in den Abbildungen 6.8 bis 6.10 stellt drei Annäherungsfahren an den „FZI Zwerg“ gegenüber. In der Teilgrafik `exists` ist dargestellt, ob das Objekt durch die ToF-Kamera gemessen werden konnte. Bei Annäherung an das unveränderte Objekt kann eine Messung erstmals bei 45 m erfolgen, eine kontinuierliche Messung ist ab 30 m Entfernung möglich. Ab dieser Entfernung beschreiben auch zwei oder mehr Bildpunkte das Objekt. Diese wurden durch Regionenwachstums-Verfahren zu einem Objekt zusammengefasst (vergleiche Abschnitt 4.2.1). Im zweiten Versuch kann das Objekt mit Reflektorweste bereits bei der initialen Messdistanz von 50 m durchgängig gemessen werden, zwei oder mehr Datenpunkte belegen die Detektion des Objektes. Im dritten Versuch lässt sich das Objekt mit dunkler Jacke ab einer Messdistanz von 45 m erstmals detektieren, eine durchgängige Messung erfolgt allerdings erst ab einer Entfernung von 20 m. Bis zu einer Entfernung von 22 m wird das Objekt fast durchgängig von nur einem Bildpunkt der Kamera gemessen.

Das vorangehende Beispiel zeigt, wie stark das Perzeptionssystem von der Reflektanzeigenschaft der Objekte abhängig ist. Daher wurden im Rahmen der Evaluation weitere Fehlerfälle untersucht, in welchen mangelnde Reflektanz zu fehlerhaften Messungen führen kann.

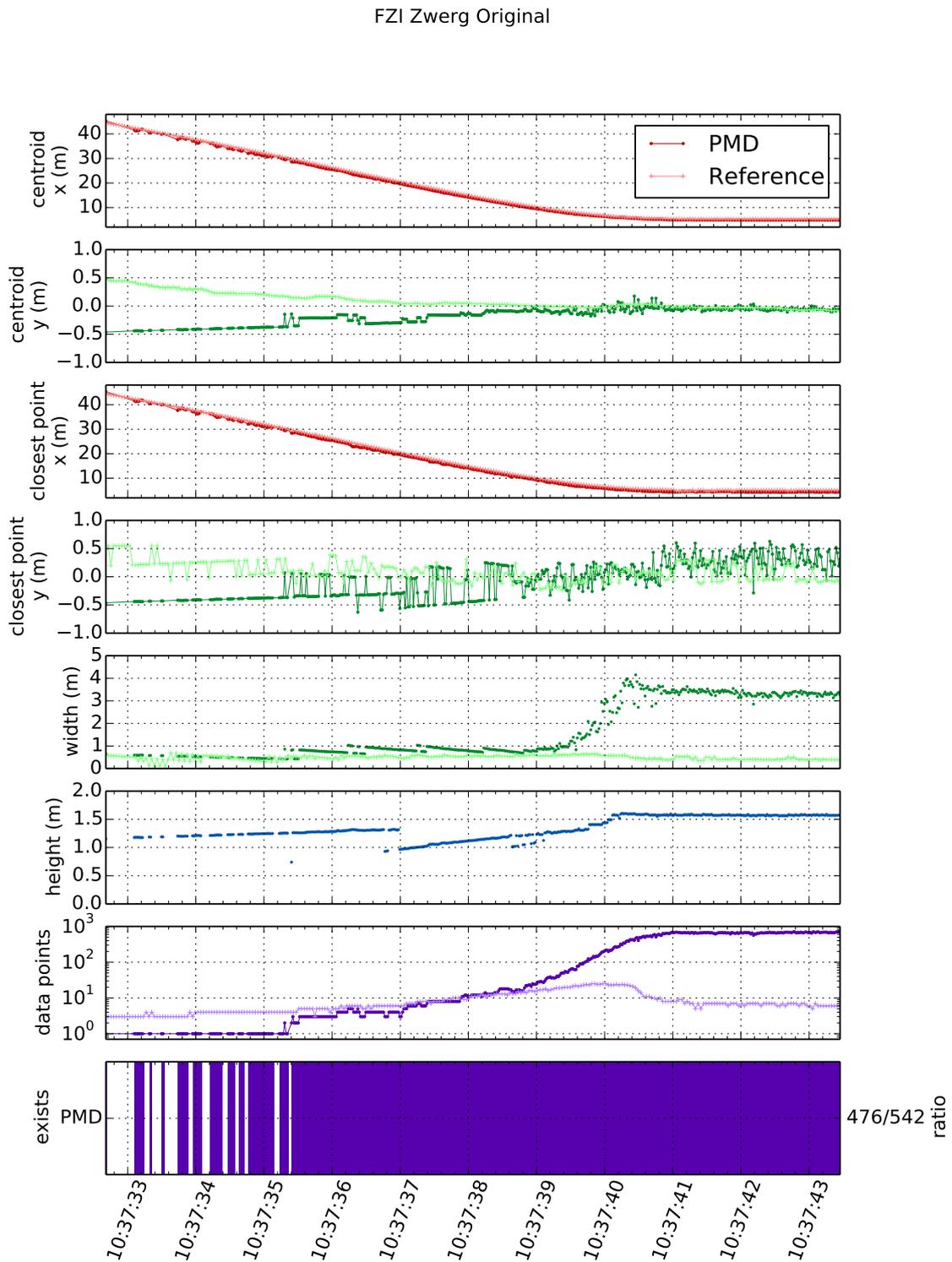


Bild 6.8: Auswertung des Einflusses der Reflektanz bei Annäherungsfahrten an „FZI Zwerg“.

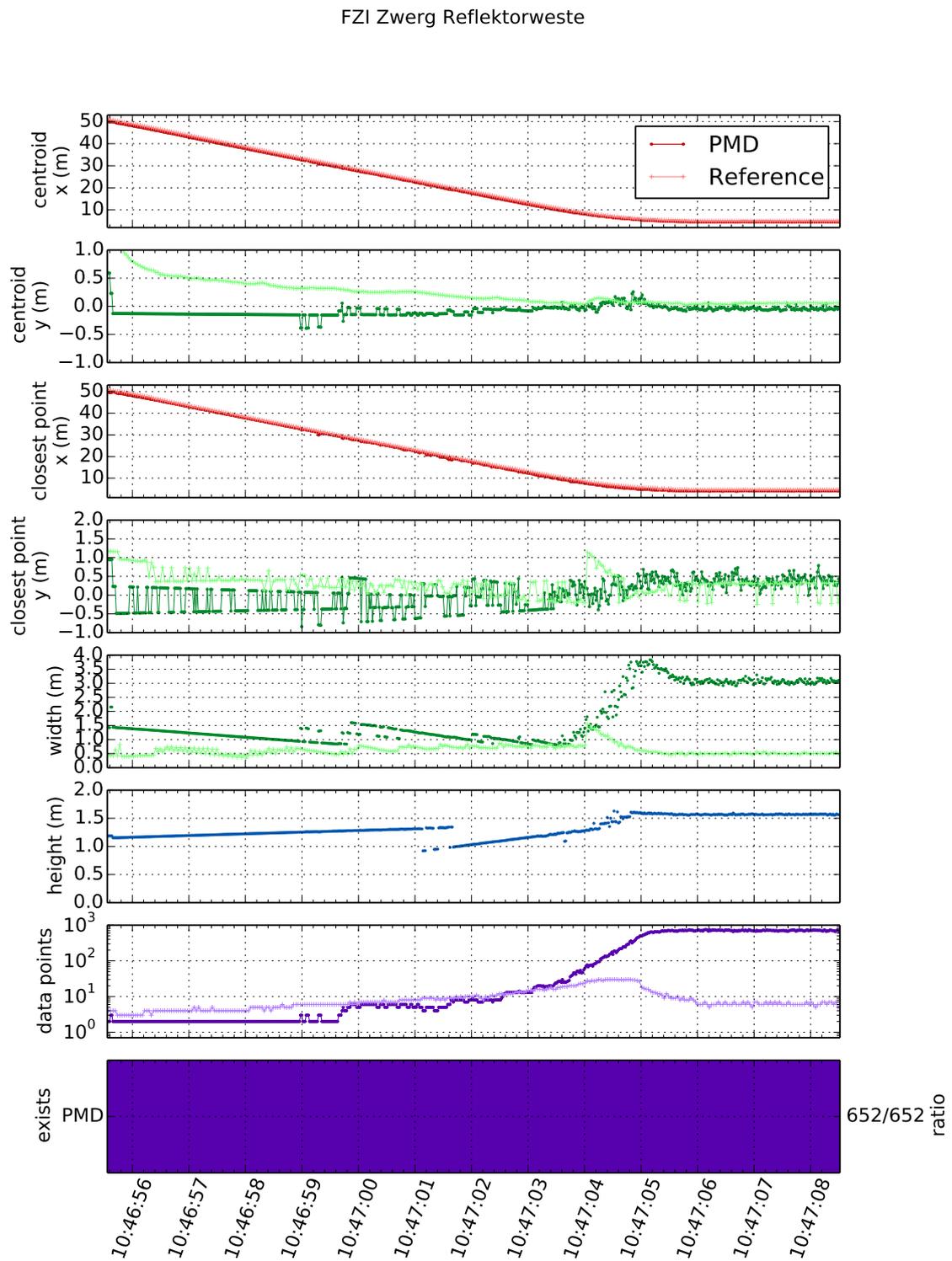


Bild 6.9: Auswertung des Einflusses der Reflektanz bei Annäherungsfahrten an „FZI Zwerg“ mit Reflektorweste.

FZI Zwerg Jacke

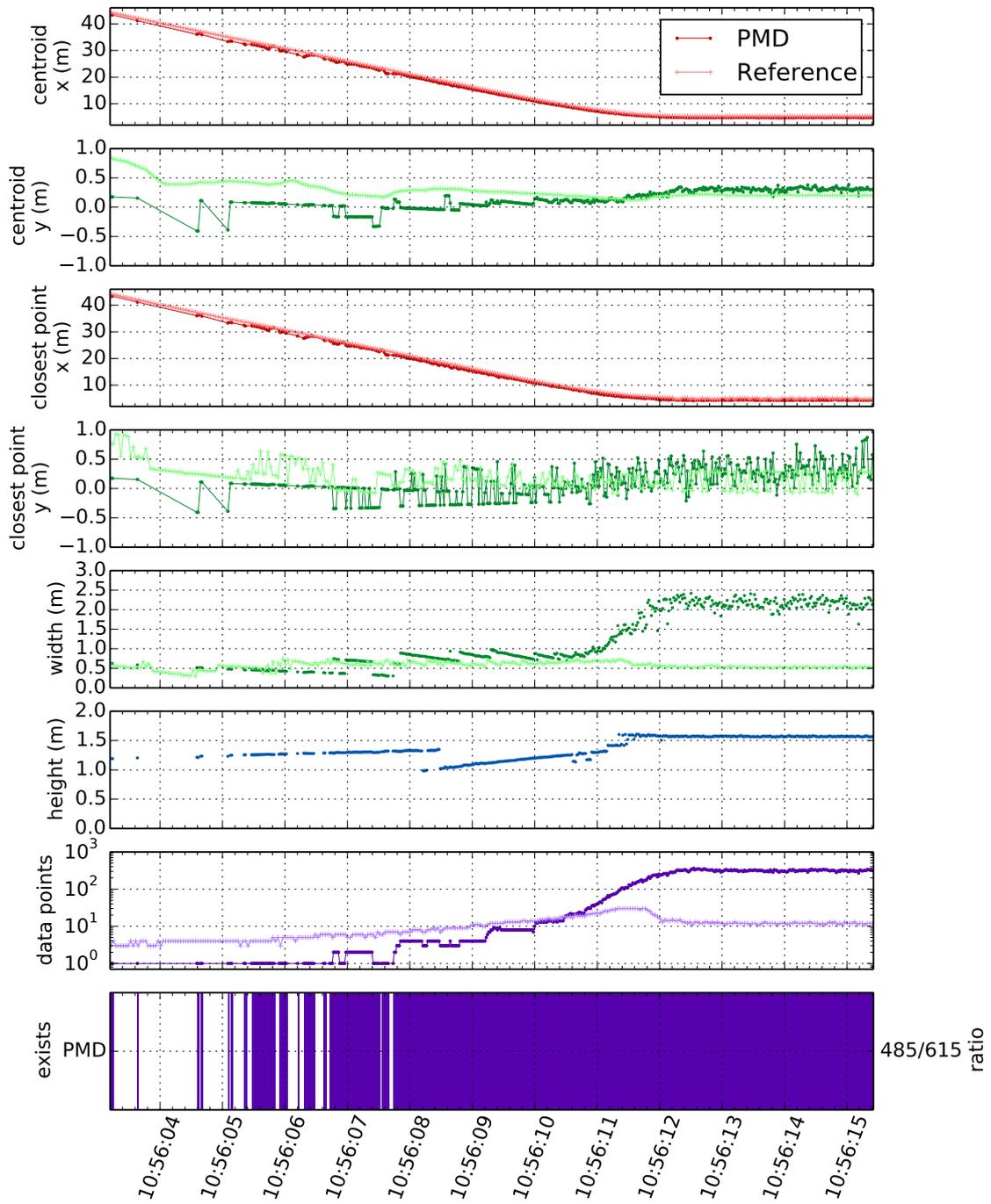


Bild 6.10: Auswertung des Einflusses der Reflektanz bei Annäherungsfahrten an „FZI Zwerg“ mit dunkler Jacke.

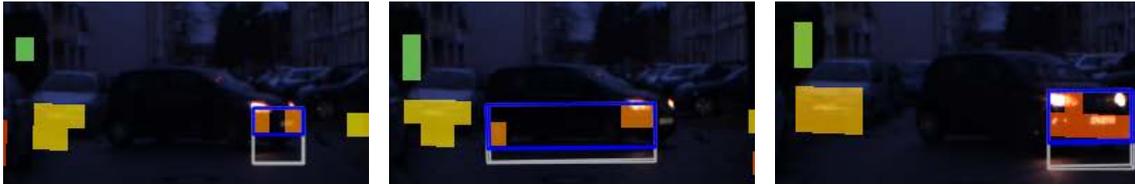


Bild 6.11: Rangierendes Fahrzeug: Nur Radkappen, Kennzeichen und Fahrzeuglichter können durch die PMD-Kamera detektiert werden.

In Abbildung 6.11 ist eine Annäherung an ein quer zum Ego-Fahrzeug positioniertes Fahrzeug gezeigt. Während das Fahrzeug ein Rangiermanöver durchführt, sind wechselseitig lediglich die Radkappen und das rechte Fahrzeuglicht in den PMD-Daten erkennbar. Im dritten Bild der Sequenz hat das Fahrzeug eine Rotation von etwa 60° zum eigenen Fahrzeug. Erstmals kann die Fahrzeugfront vollständig gemessen werden, die seitliche Fahrzeugpartie bleibt weiterhin nicht detektiert. Dies ist besonders dann kritisch, wenn auf Grundlage der Messung eine Ausweichtrajektorie berechnet werden müsste. Eine Kollision mit der Heckpartie des rangierenden Fahrzeugs wäre vermutlich die Folge. Besonders bei Totalreflexion können fehlerhafte Messungen zu falsch geschätzten Objektausdehnungen oder sogar zu ausbleibenden Objektmessungen führen.

Das Auftreten dieser Art von Fehlern lässt sich in probabilistischen Verfahren berücksichtigen. Hier kann eine weiterführende probabilistische Interpretation zu einer Kompensation von Messfehlern beitragen.

Die Grafiken in Abbildung 6.12 zeigen die Auswertung von Messdaten eines querenden Fußgängers in zwei unterschiedlichen Abständen zum Messsystem (siehe auch Abbildung 6.5). Der Fußgänger kann in der Distanz 15 m durchgängig beobachtet werden. Dagegen kann der Fußgänger in einer Distanz von 30 m zum Ego-Fahrzeug lediglich in 345 von 438 Messungen detektiert werden, also in nur 79 % der Messungen. Dabei ist besonders der Randbereich von ausbleibenden Messungen betroffen. In der Auswertung ist auch die mit der Gangart des Fußgängers korreliert variierende Breite des gemessenen Objektes erkennbar. In der Entfernung von 15 m tritt dieser Effekt stark sichtbar auf. Der Fußgänger hat eine Größe von etwa 1,8 m. Dies lässt sich in 15 m Entfernung gut abschätzen. Weniger deutlich kann die Höhe in 30 m Entfernung abgeschätzt werden, hier beträgt die geschätzte Höhe im Mittel ≈ 1 m. Hier kommt zusätzlich zum Tragen, dass der Fußgänger fast durchgängig mit nur 1 bis 2 Datenpunkten gemessen wird, was eine präzise Schätzung der Objektausdehnung ohne Modellwissen unmöglich macht.

Messfehler durch Überbelichtung

Neben der stark abfallenden Intensität bei weit entfernten Objekten können durch die aktive IR-Beleuchtung andere, als kritisch zu bewertende Messfehler auftreten. Bei sehr starker Überbelichtung können zwischen Linse und Chipfläche Reflexionen auftreten, welche ein überbelichtetes Objekt größer als tatsächlich erscheinen lassen. Dies ist besonders dann kritisch, wenn durch diesen Effekt „Phantomobjekte“ vor dem Fahrzeug auftreten und schließlich zu einer Warnung des Fahrers oder gar zu einer automatischen Notbremsung des Fahrzeugs füh-

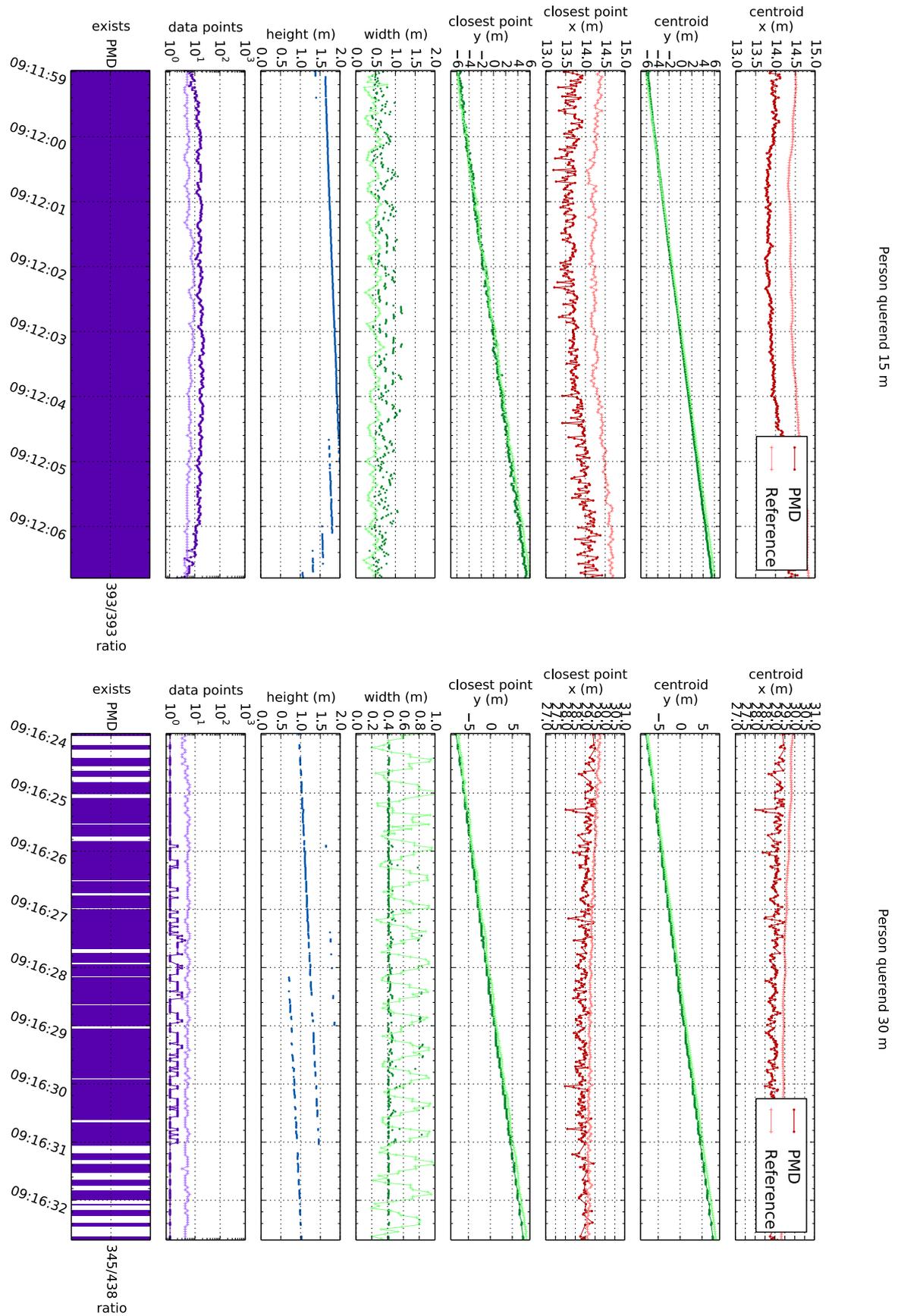


Bild 6.12: Von rechts nach links querender Fußgänger in unterschiedlichen Distanzen vor dem Fahrzeug: 15 m und 30 m.

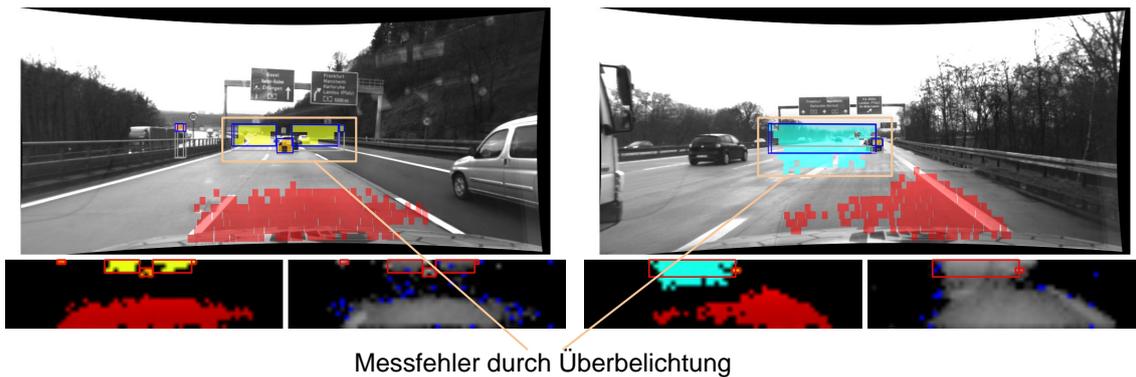


Bild 6.13: Fehlerhafte Messungen durch Überstrahlung bei Richtungswegweisern auf der Autobahn; die Messdaten entsprechen an dieser Stelle im Kamerabild keinem tatsächlichen Objekt.

ren würden. Hier kann eine Schätzung einer Existenzwahrscheinlichkeit wertvolle Hinweise liefern.

Zwei Fälle von Überstrahlung können durch die Evaluation des Systems nachgewiesen werden. Die Überstrahlung bei sehr nahen Objekten sowie die Überstrahlung bei Richtungswegweisern neben der Fahrspur oder über Schnellstraßen. Erster Fall kann an der in Abbildung 6.9 dargestellten Grafik des „FZI Zwerg“ mit Reflektanzweste verdeutlicht werden. Bei etwa 10 m Abstand nimmt die Ausdehnung des Objektes deutlich zu und wird schließlich mit 3 m Breite geschätzt. Bei den Objekten mit geringerer Reflektanz tritt dieser Effekt ebenfalls auf, aber erst in geringerer Distanz. Ähnlich verhält es sich bei Annäherung an vorausfahrende Fahrzeuge, besonders da diese im Nahbereich durch das stark reflektierende Kennzeichen oft zu Überbelichtung führen.

Die neben oder über Schnellstraßen angebrachten Richtungswegweiser sind mit einem retroreflektierenden Material ausgestattet. Dies führt dazu, dass eine große Menge an emittiertem Licht von dieser Fläche reflektiert wird. In Abbildung 6.13 sind zwei Situationen dargestellt, in welchen die Überstrahlung zu Messungen führt, an deren Stelle sich tatsächlich kein Objekt befindet. Diese Messungen entsprechen Falsch-Positiv (FP)-Signalen der Kamera, also Messungen, welche durch nicht (an dieser Stelle) vorhandene Objekte hervorgerufen werden. Abhängig von der Stärke der Reflexion kann es sogar dazu führen, dass tatsächliche Objekte im betroffenen Bereich des Bildes durch die Reflexion überlagert werden, also die gemessene Entfernung nicht mit der Objektentfernung übereinstimmt (siehe Abbildung 6.14).

Messqualität bei Fremdlicht

Das von den Beleuchtungseinheiten emittierte IR-Signal kann abhängig von den Umgebungseigenschaften gegenüber dem Hintergrund-Rauschen so gedämpft sein, dass die Messqualität der Sensorik darunter leidet.

Die in „CoCar“ verbaute PMD-Kamera ist in der Lage, selbst bei direktem Sonnenlicht Messungen über das Umfeld wahrzunehmen. In Abbildung 6.15 ist eine Szene dargestellt, in welcher mehrere Personen sich dem Messsystem nähern. Die Personen können selbst bei größerer Entfernung durch das Messsystem erfasst werden.

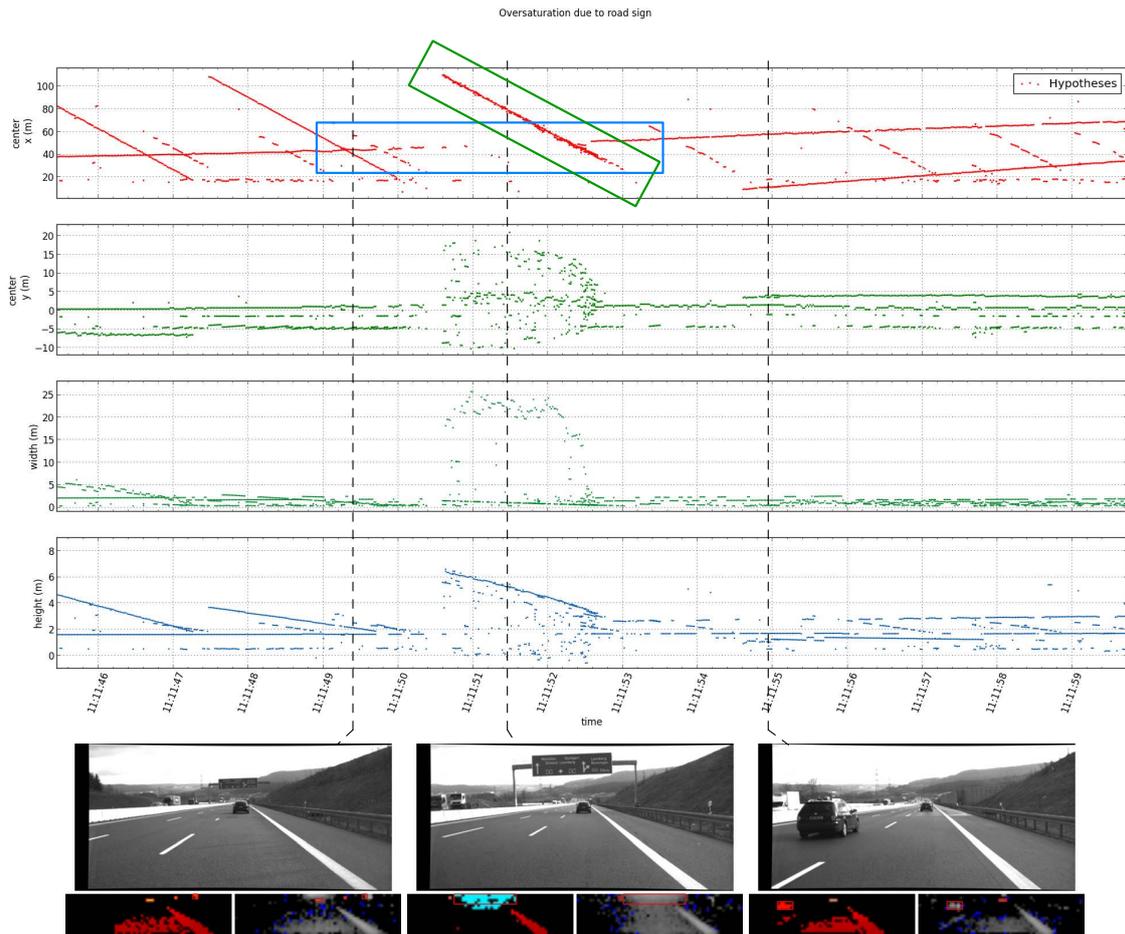


Bild 6.14: Störung der Distanzmessung des vorausfahrenden Fahrzeugs (blaues Rechteck) während Hindurchfahrt unter Richtungswegweiser (grünes Rechteck).



Bild 6.15: Messgüte der PMD-Sensorik bei direktem Sonnenlicht; die Messqualität ist durch das Sonnenlicht nicht beeinflusst.

Neben Sonnenlicht erzeugt auch die IR-Beleuchtung äquivalenter Systeme in anderen Fahrzeugen Fremdlicht für das Messsystem. Abbildung 6.16 zeigt das Messverhalten der Sensorik bei einer Störung durch eine fremde Beleuchtungseinheit. Der Bereich der fremden Beleuchtungseinheit wird bei aktivierter Fremdlichtquelle vom eigenen Messsystem erkannt und die Messdaten der betroffenen Bildpunkte als ungültig markiert. Im benachbarten Messbereich liegende Bildpunkte sind nicht beeinflusst und liefern gültige Messdaten. Daher kann davon ausgegangen werden, dass ein Objekt trotz einer störenden IR-Quelle vom Messsystem detektiert werden kann.

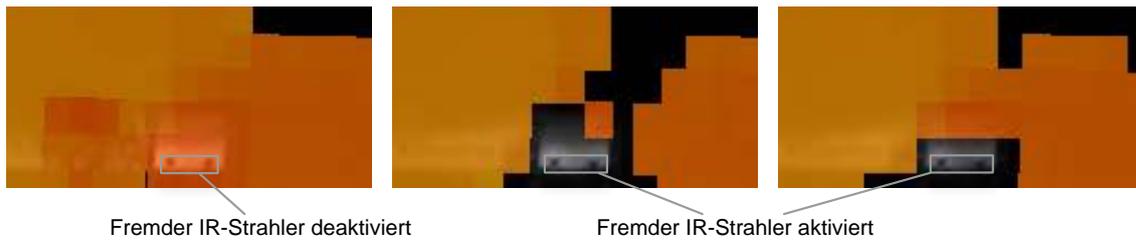


Bild 6.16: Messverhalten der PMD-Sensorik bei Störung durch fremde IR-Beleuchtungseinheit; bei aktiviertem Strahler werden beeinflusste Bildpunkte als ungültig markiert und unterdrückt.

Bewertung der Erkennungsleistung

Die durchgeführte Analyse zur Bewertung der Performanz von ToF-Kameras zeigt, dass die Erkennungsleistung der Sensorik gut geeignet ist für die Detektion von Objekten. Relevante Objekte werden mit wenigen Ausnahmen mit hoher Zuverlässigkeit durch die Messsensorik erkannt. Bei Folgefahrten führen die Kennzeichen von vorausfahrenden Fahrzeugen zu einer hohen Reflektanz, welche zur genauen Bestimmung der Entfernung des Fahrzeugs ausreicht. Allerdings können quer zum Fahrzeug stehende Fahrzeuge aufgrund der Totalreflexion des IR-Signals zum Teil nicht gemessen werden. Durch Überstrahlung verursachte Messfehler beispielsweise an Richtungswegweisern beeinträchtigen die Messqualität erheblich. Daher muss durch die Verfahren zur Objektsegmentierung eine Überprüfung durchgeführt werden, ob Messpunkte mit hoher Reflektanz in großer Entfernung gültig sind.

Die Analyse zeigt, dass Messfehler zu Teilen ausgeglichen werden müssen. Dies kann durch die Berücksichtigung von Attributunsicherheiten sowie Existenzunsicherheiten in den Erkennungs- und Interpretationskomponenten geschehen.

6.3.2 Bewertung der Objektsegmentierung durch Regionenwachstum

Der in Abschnitt 4.2.1 beschriebene Ansatz ist besonders für die Segmentierung gering aufgelöster Tiefendaten entwickelt worden. Der Ansatz beruht darauf, anhand eines Homogenitätskriteriums zu einer Gruppe von Bildpunkten benachbarte Bildpunkte dieser Gruppe zuzuordnen. Das verwendete Ähnlichkeitsmaß entspricht dabei der Differenz zwischen der Tiefe eines Bildpunktes zur mittleren Tiefe der Gruppe. Der Vorteil des Segmentierungsverfahrens liegt darin, dass die genaue Anzahl an Regionen nicht vorgegeben werden muss.

Wie in der Abbildung 6.17 dargestellt ist, bildet die Segmentierung aus den Rohdaten der PMD-Kamera Objekthypothesen. Diese sind durch ihre umschließende Box (BBox) beschrieben. In den farbcodierten Bildern der Tiefendaten sind die zu einem Objekt zusammengehörenden Bildpunkte durch eine rote Rechteck beschrieben. In der Projektion der umschreibenden Box in das Videobild sind die Boxen durch blaue, achsausgerichtete Quader dargestellt. Die Segmentierung erreicht eine gute Trennleistung bei verschiedenen Objekten sowohl in der Stadt als auch auf Autobahnen. In diesen Szenen tritt Über- oder Untersegmentierung von Objekten selten auf. In der Stadt sind besonders nahe beieinander platzierte Objekte wie parkende Fahrzeuge problematisch, häufig reicht der Abstand zwischen den Objekten allerdings für eine gute Segmentierung aus. Die zum Boden gehörenden Bildpunkte werden in Abhängigkeit vom Abstand zur Kamera gefiltert. Nahe Bildpunkte werden nach ihrer Höhe in



Bild 6.17: Beispiele von Segmentierungsergebnissen der Objektsegmentierung durch Regionenwachstum (rote Boxen), projiziert in das Videobild (blaue Boxen).

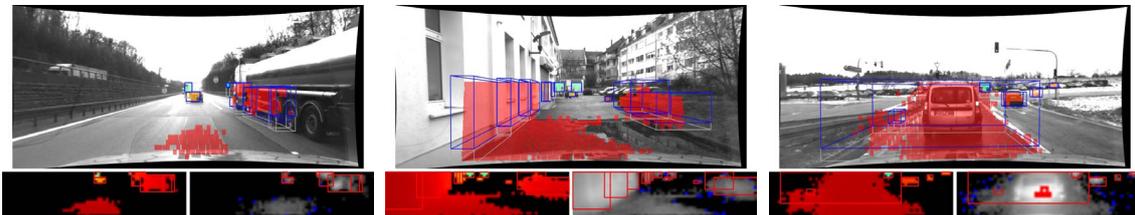


Bild 6.18: Fehlerhafte Objekthypothesen durch Über- und Untersegmentierung.

Fahrzeugkoordinaten gefiltert, während weiter von der Kamera entfernte Bildpunkte nicht gefiltert werden. Dies verhindert, dass Objekte fehlerhaft als Bodenpunkte klassifiziert werden, falls das Fahrzeug eine Nickbewegung vollführt oder die vor dem Fahrzeug liegende Fahrbahn nicht eben ist. Die Segmentierung von Fahrzeugen auf der Autobahn funktioniert problemlos, begünstigt durch die Tatsache, dass das Kennzeichen oder das Heck von Fahrzeugen eine günstige Reflexionsfläche bilden und somit eine gute Schätzung der Fahrzeug-Objekthypothese ermöglichen.

In Abbildung 6.18 sind fehlerhafte Segmentierungen von Objekten dargestellt. Die ersten beiden Bilder zeigen eine Übersegmentierung, da das angewendete Verfahren Objekte mit einer großen Ausdehnung entlang der x -Achse nicht zu einem gemeinsamen Objekt gruppieren kann. Das dritte Bild zeigt eine Untersegmentierung des Objektes, hervorgerufen durch Messfehler der Kamera durch Überbelichtung. Im Nahbereich der Kamera wird eine zu große Hypothese erzeugt. Dies wird durch Reflexion des IR-Signals zwischen Linse und Chip hervorgerufen. Dabei ist zwar die Objekthypothese an sich nicht fehlerhaft, allerdings sind die Attribute der Hypothese nicht akkurat.

Bestimmung der Trennfähigkeit der Segmentierung Eine Untersegmentierung mehrerer Objekte kann auftreten, wenn Objekte sehr eng zueinander platziert sind. Eine Untersegmentierung von Objekten ist dann kritisch, falls ein Objekt, welches in den Fahrkorridor des eigenen Fahrzeugs tritt, durch die Untersegmentierung zunächst einem außerhalb des Korridors platzierten Objekt zugeordnet wird. Eine solche Situation, in welcher sich ein Fußgänger auf der Straße in der Nähe zu einem parkenden Fahrzeug befindet, ist in Abbildung 6.19

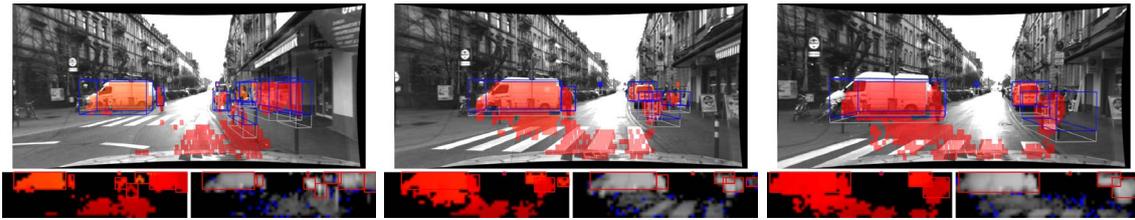


Bild 6.19: Untersegmentierung zweier eng benachbarter Objekte bei Annäherung eines Fußgängers an einen Van.

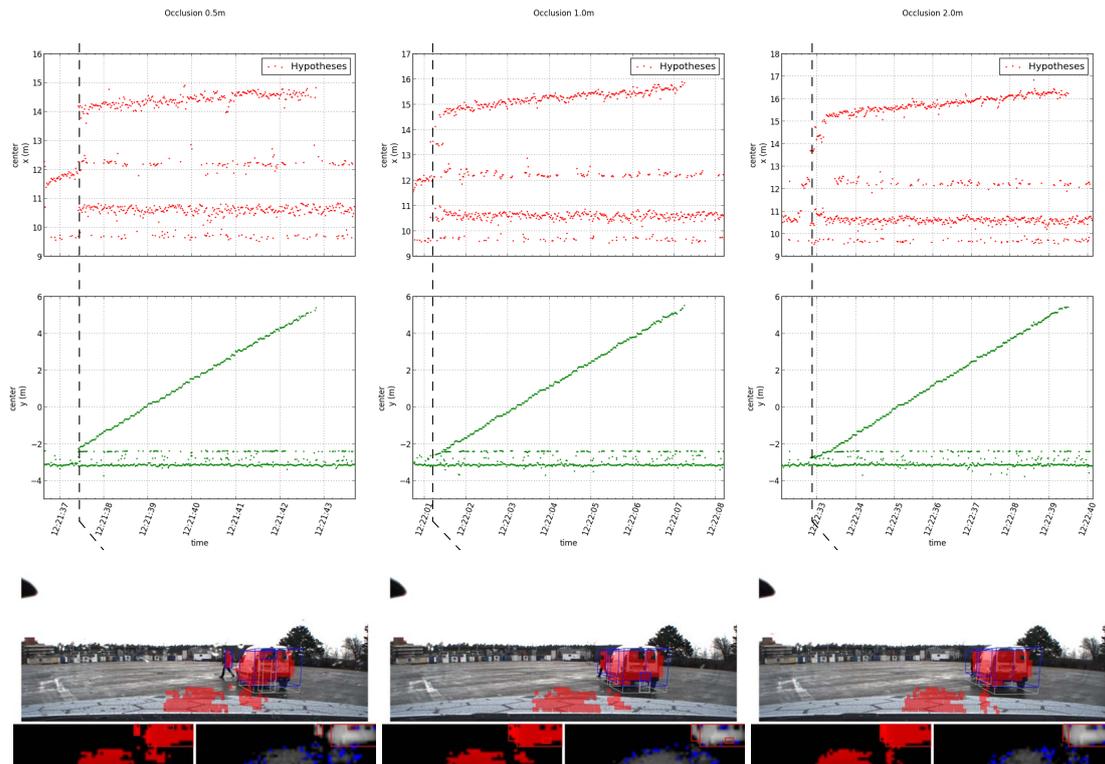


Bild 6.20: Hypothesengenerierung bei Verdeckung nahe beieinander stehender Objekte, 0,5 m (links), 1,0 m (mitte) und 2,0 m x-Abstand (rechts); Bildausschnitt zeigt Zeitpunkt der jeweils ersten gültigen Trennung.

dargestellt. Die Segmentierung der beiden Objekte erfolgt zunächst problemlos. Erst als der Fußgänger in den Nahbereich des parkenden Fahrzeugs tritt, kann eine Segmentierung von beiden Objekten nicht mehr erfolgen, es kommt zu Untersegmentierung.

Das System erreicht bei nahe zueinander platzierten Objekten dennoch eine gute Trennfähigkeit. In Abbildung 6.20 ist dargestellt, wie ein Fußgänger hinter einem parkenden Fahrzeug in den Fahrkorridor tritt. Diese Situation wurde in unterschiedlichen Abständen zwischen Fußgänger und parkendem Fahrzeug evaluiert. In allen Abständen konnte eine frühe Trennung zwischen Fußgänger und Fahrzeug erreicht werden. Die Abbildung zeigt die Trennfähigkeit des Segmentierungsansatzes für die x - respektive y -Achse über die Zeit aufgetragen. Es ist deutlich zu sehen, in welchem Abstand der Fußgänger von dem Fahrzeug unterschieden werden kann. In 0,5 m Entfernung ist dies möglich, sobald der Fußgänger vor das Fahrzeug tritt. Dabei unterscheidet sich der Mittelpunkt des Fußgängers von dem Mittelpunkt des Fahrzeugs entlang der y -Achse um etwa 0,9 m, also in etwa der halben Fahrzeugbreite. Bei den Abständen

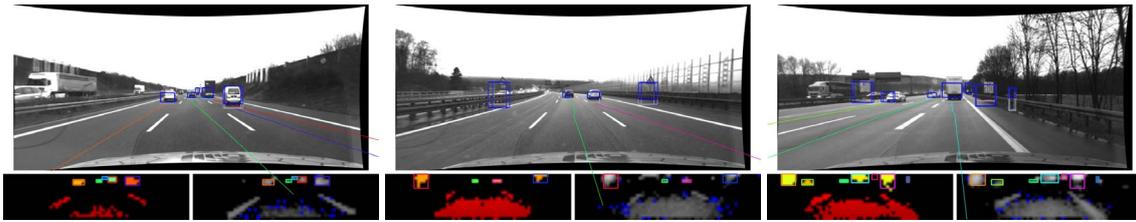


Bild 6.21: Verifikation von Objekten durch temporales Filter auf der Autobahn; Verlauf der Position der Objekte auf Bodenebene projiziert und mit Eigenbewegung kompensiert.



Bild 6.22: Verifikation von Objekten durch temporales Filter in der Stadt; Verlauf der Position der Objekte auf Bodenebene projiziert und mit Eigenbewegung kompensiert.

den 1,0 m und 2,0 m kann der Fußgänger bereits detektiert werden, während er sich noch leicht verdeckt hinter dem Fahrzeug befindet.

Laufzeit des Verfahrens Die Verarbeitungsgeschwindigkeit auf einem 2,7 GHz Core i7 Prozessor bei Nutzung eines Kerns beträgt für die Segmentierung etwa 3 – 5 ms auf der Autobahn und 10 – 15 ms in der Stadt. Der Unterschied ist darauf zurückzuführen, dass in der Stadt im Allgemeinen mehr Bildpunkte der Kamera gültige Messdaten liefern, während auf der Autobahn für die Mehrzahl der Bildpunkte keine gültige Entfernungsmessung erzeugt werden kann.

Verbesserung durch Zustandsschätzer

Die durch Segmentierung gebildeten Objekthypothesen werden durch einen Zustandsschätzer gefiltert und mit weiteren Attributen ergänzt. Hierzu wird das Kalman Filter als Schätzer eingesetzt, dieses ist in Abschnitt 4.5 beschrieben. Auf Basis einer Assoziationsfunktion werden die segmentierten Objekthypothesen den bereits aus zurückliegenden Messungen im System bekannten Objektspuren zugeordnet und anschließend mit einem zugrundeliegenden Bewegungsmodell zeitlich gefiltert. Gleichzeitig wird für jedes Objekt bestimmt, über welchen Zeitraum das Objekt bereits kontinuierlich verfolgt werden konnte. Dies ermöglicht es, zu der Ableitung zusätzlicher Attribute außerdem eine Beobachtungsgüte festzustellen, mit welcher zwischen ungültigen Hypothesen und gültigen Hypothesen unterschieden werden kann. Weiter werden die Attribute wie die Objektausdehnung über die Zeit gefiltert, was eine präzisere Schätzung der Objektgröße durch Wissen aus der Historie erlaubt.

Abbildung 6.21 und 6.22 zeigen verifizierte Objekte, welche während einer Autobahn- und einer Stadtsequenz durch das System verfolgt wurden. Im 2D Videobild ist dazu der Verlauf der Position des Objektes auf die Bodenebene projiziert und mit der Eigenbewegung des Fahrzeugs kompensiert dargestellt. Auf der Autobahn-Sequenz ist deutlich zu erkennen, dass die

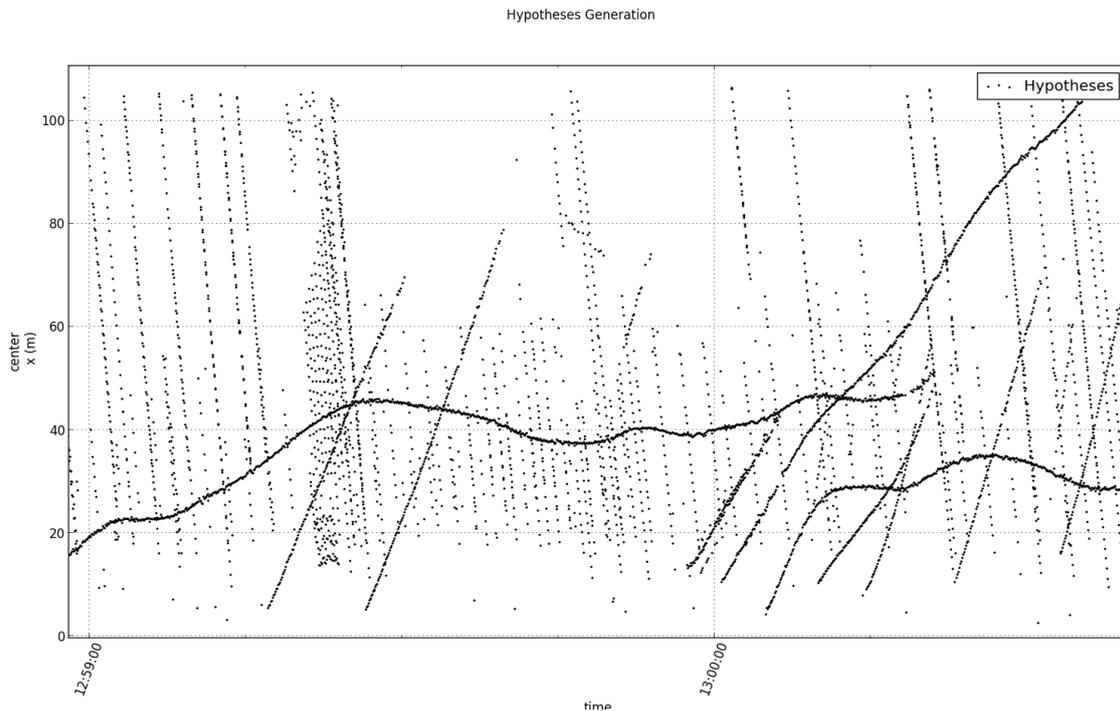


Bild 6.23: Segmentierte Objekthypothesen während einer Folgefahrt über Entfernung (Ordinate) und Zeit (Abszisse) aufgetragen; statische Objekte zeichnen sich als Linien von links oben nach rechts unten ab.

Spur von Fahrzeugen glatt verläuft. Ebenso werden Schilder am Straßenrand auch bei höheren Geschwindigkeiten korrekt zugeordnet und über die Zeit verfolgt. Eine deutlich höhere Dichte von Objekthypothesen in der Stadt-Sequenz verlangt eine besonders gute Zuordnung zwischen Hypothesen und verfolgten Objekten. Hier kommt es vereinzelt zu Objekt-Sprüngen, in welchen eine Zuordnung zwischen Objekten fehlerhaft verläuft.

Abbildung 6.23 zeigt alle segmentierten Objekthypothesen einer Messequenz über die Zeit aufgetragen. Bei der Sequenz handelt es sich um eine Folgefahrt hinter einem PKW auf der Autobahn. Deutlich erkannt werden kann das verfolgte Fahrzeug, weitere überholende Fahrzeuge sowie die statischen Objekte der Infrastruktur. Ebenso zu erkennen sind fehlerhafte Hypothesen, die sich als einzelne Punkte in der Abbildung hervorheben.

In Abbildung 6.24 ist die gleiche Sequenz mit aktiviertem Zustandsschätzer dargestellt. Die Farben der Objektspuren korrespondieren mit der Zuordnung von Objekten. Ein Farbwechsel einer Spur bedeutet einen Verlust der Zuordnung mit anschließender Zuordnung als neues Objekt. Es ist zu erkennen, dass das vorausfahrende Fahrzeug über einen langen Zeitraum (> 1 min) kontinuierlich durch das System verfolgt wird. Weiter ist deutlich erkennbar, dass fehlerhafte Objekthypothesen durch den Zustandsschätzer unterdrückt werden. Gleichzeitig führt diese Unterdrückung allerdings dazu, dass Objekte nach Erstdetektion erst zu einem späteren Zeitpunkt als gültig befunden werden. Leicht zu erkennen ist dies an den statischen Objekten der Infrastruktur, welche bei hoher Entfernung von etwa 105 m erstmalig erkannt werden, nach aktivierter Zustandsschätzung ab etwa 100 m Entfernung für gültig befunden werden. Ein Track-Verlust des zweiten überholenden Fahrzeugs (mittleres Bild in Abbildung 6.24) ist begründet darin, dass das Fahrzeug für einen kurzen Zeitraum nicht mehr

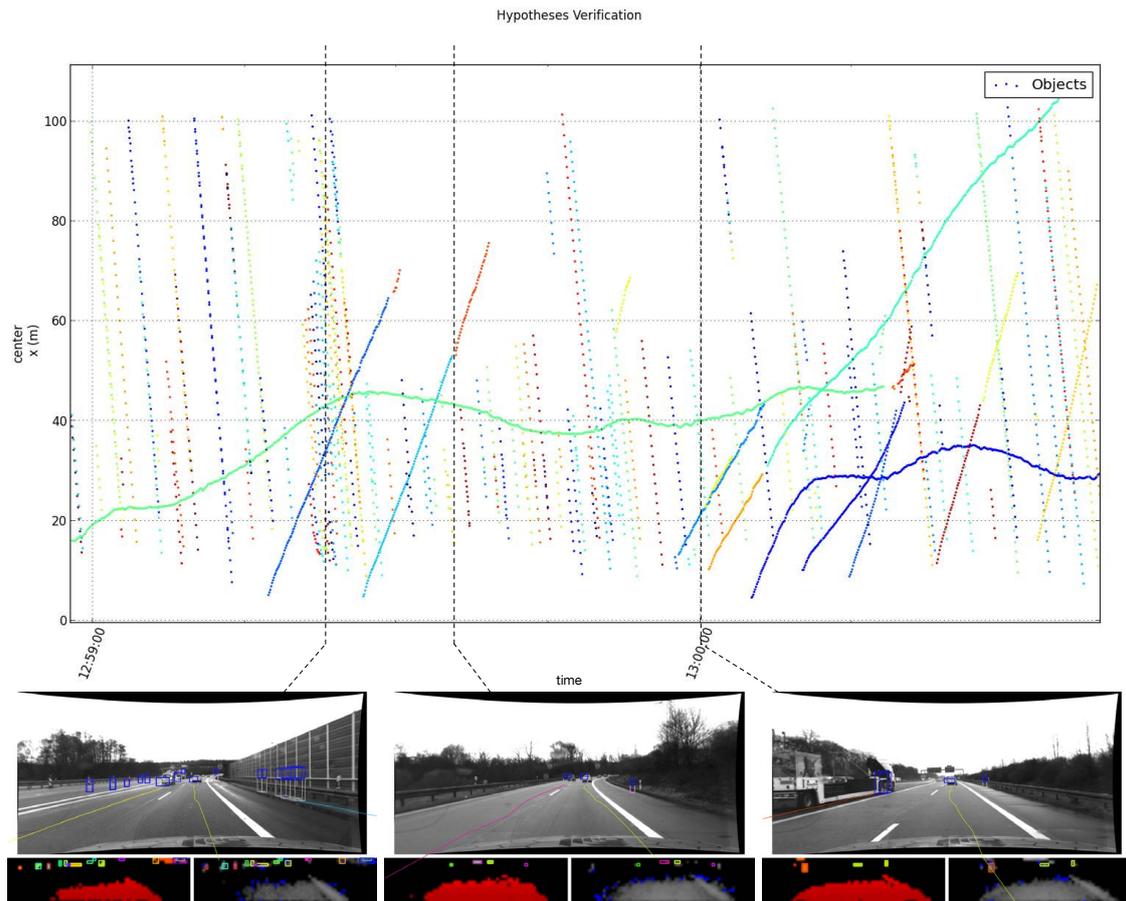


Bild 6.24: Zeitlicher Verlauf der verfolgten, als gültig befundenen Objekthypothesen zu Abbildung 6.23; unterschiedliche Farben markieren verschiedene Tracks.

durch die Sensorik erfasst werden kann. Weitere Track-Verluste treten maßgeblich durch Verdeckungen auf.

Die Verarbeitungsgeschwindigkeit auf einem 2,7 GHz Core i7 Prozessor bei Nutzung eines Kerns beträgt für die Zustandsschätzung etwa 2 – 5 ms bei bis zu 20 verfolgten Objekten.

6.3.3 Bewertung der Histogrammbasierten Objektsegmentierung

Für die Segmentierung höher aufgelösten Tiefendaten eignet sich das Verfahren der Histogrammbasierten Objekterkennung (Abschnitt 4.2.2), welches die Daten in effizienter Weise nach flächigen Strukturen untersucht. Das Verfahren unterdrückt dabei kleine Objekte und stark gekrümmte Objekte.

Ergebnisse des vorgestellten Histogrammbasierten Segmentierungsverfahren sind in Abbildung 6.25 und 6.26 gezeigt. In blau dargestellt sind Objektflächen, welche aus dem U -Raum extrahiert wurden. In Abbildung 6.25 werden flächige Objekte durch das Segmentierungsverfahren deutlich erkennbar extrahiert. Auch die Orientierung von Objekten lässt sich gut erkennen. Besonders bei parallel zur Bewegungsrichtung stehenden Flächen kommt es gegenüber dem Regionenwachstums-Verfahren nicht zu Übersegmentierung. Dies ist beispielsweise bei der Segmentierung von Wänden von Vorteil. Sehr kleine Objekte oder Rauschen im Hintergrund werden durch das Verfahren nicht extrahiert. Die befahrbare Bodenfläche sowie

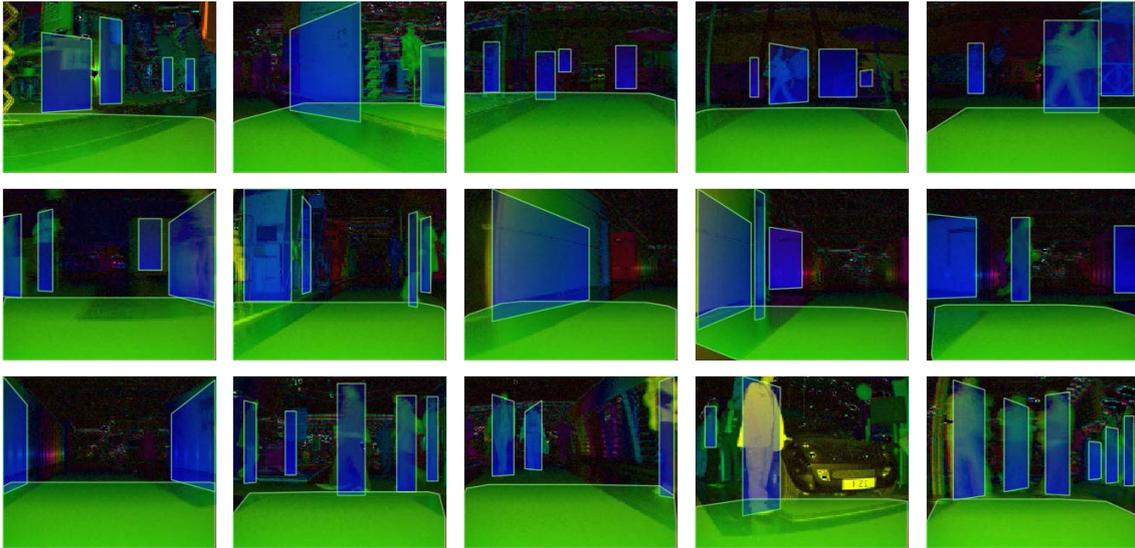


Bild 6.25: Segmentierung von Flächen in einer Sequenz von Tiefendaten mit 25k-Bildpunkten; blaue Flächen stellen Objekte dar, die grüne Fläche die Bodenebene.

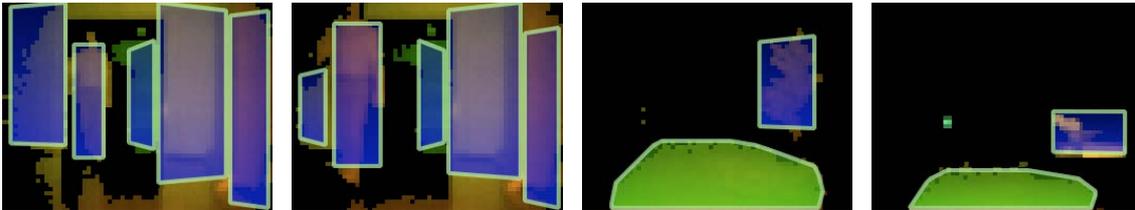


Bild 6.26: Segmentierung von Flächen in einer Sequenz von Tiefendaten mit 3k-Bildpunkten; blaue Flächen stellen Objekte dar, die grüne Fläche die Bodenebene.

der Winkel zwischen Kamera und Bodenfläche wird durch das Verfahren geschätzt. In der Abbildung wurden die Punkte der befahrbaren Fläche bestimmt und anschließend die Menge dieser Punkte durch ein konvexes Polygon umfasst, welches grün dargestellt ist. Die Resultate für gering aufgelöste Tiefendaten sind in Abbildung 6.26 dargestellt. Auch hierfür eignet sich das Segmentierungsverfahren, wobei Effekte durch Übersegmentierung zunehmen. Die Orientierung der Flächen kann gegenüber der höherauflösenden ToF-Kamera nicht akkurat geschätzt werden.

Die Eingangsdaten wurden durch zwei ToF-Kameras mit unterschiedlichen Auflösungen auf der Plattform „SegIT“ aufgezeichnet. Die Eingangsdaten der Abbildung 6.25 besitzen 25k Bildpunkte, die der Abbildung 6.26 3k Bildpunkte. Dabei liegt die Berechnungszeit bei etwa 40 – 60 ms für Eingangsdaten der 25k-Datensätze und bei etwa 20 – 30 ms für 3k-Datensätze.

Das präsentierte Verfahren eignet sich auch zur Segmentierung dichter und semi-dichter Disparitätsdaten von Stereokameras, wie in der Arbeit [Teu+10] beschrieben. Dabei ist besonders die Möglichkeit zur Low-Level Merkmalsfusion mit Radar-Daten hervorzuheben, wie dies in Abbildung 6.27 dargestellt ist. Die aus der Szene gemessenen Radar-Reflexpunkte führen dabei direkt zu einer Verbesserung der Objekt-Segmentierung, anstatt die Attribute der segmentierten Objekte nachträglich zu beeinflussen. Dies wird erreicht, indem jeder Radar-Reflexpunkt in den U -Raum transformiert und die Suche nach Maxima im Hough-Raum durch die Radar-Reflexpunkte gewichtet wird. Dies beeinflusst die Linienextraktion dahingehend,

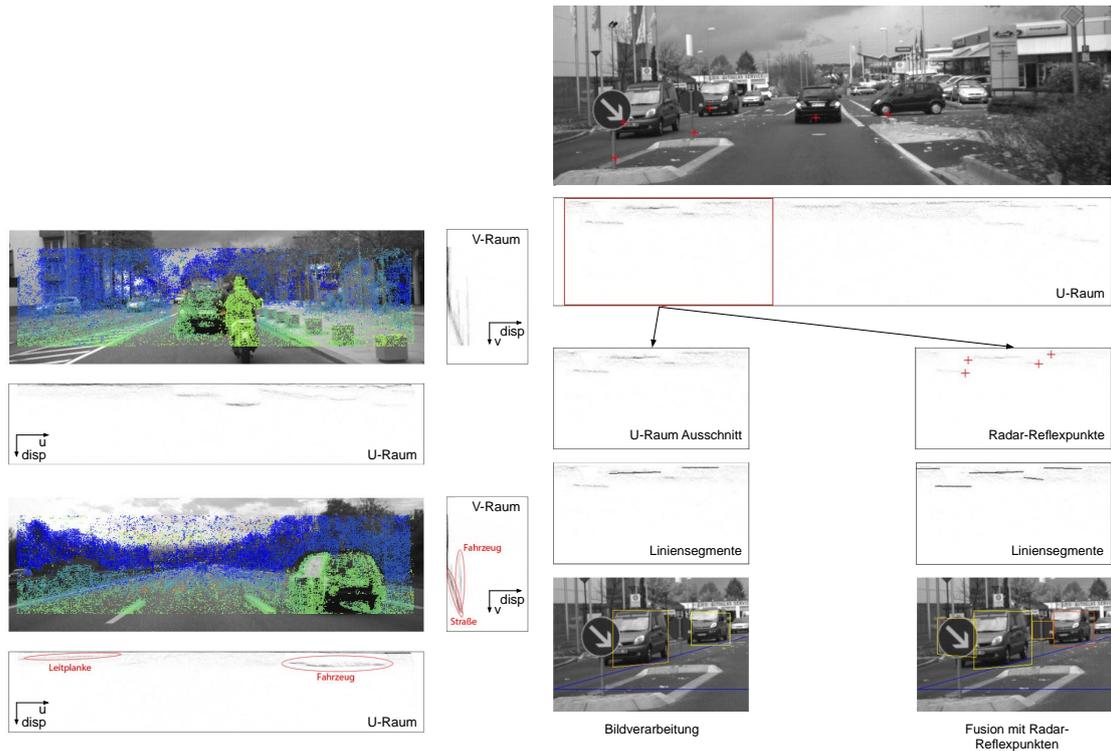


Bild 6.27: Links: Histogrammbasierte Objektsegmentierung auf semi-dichten Disparitätsdaten; rechts: Fusion von Radar-Reflexpunkten zur Verbesserung der Segmentierungsleistung [Teu+10].

dass interessante Objekte mit einer höheren Wahrscheinlichkeit im U -Raum detektiert werden. Gleichzeitig lassen sich die Schwellwerte anheben und so Fremdobjekte stärker unterdrücken.

Abbildung 6.28 stellt die Detektionsgüte des Verfahrens mit und ohne Fusion mit Radar-Daten gegenüber. Die Evaluation erfolgte über 7 Szenen mit unterschiedlicher Komplexität, insgesamt wurden dabei in Einzelbildern 8669 Objekte gezählt.

Besonders in höherer Entfernung kann die Detektionsgüte durch Fusion deutlich gesteigert werden. Diese komplementäre Sensorfusion rechtfertigt den Einsatz verschiedener Datenquellen, was gleichzeitig mit einem höheren Integrationsaufwand im Fahrzeug verbunden ist. Stehen keine Radar-Reflexpunkte oder ähnliche komplementäre Messdaten zur Verfügung, können auf gleichem Weg verifizierte Objekthypothesen der temporalen Zustandsschätzung als Gewichte in den U -Raum eingetragen werden.

6.3.4 Evaluation der videobasierten Fahrzeugerkennung

Die videobasierte Fahrzeugerkennung ermöglicht die Detektion von Fahrzeugen ohne Verwendung von Tiefenbildern. Der Vorteil der kostengünstigeren und weniger komplexen Sensorik bringt den Nachteil einer fehleranfälligeren Datenverarbeitung mit sich. Außerdem muss von deutlich größeren Unsicherheiten ausgegangen werden, insbesondere bei der Schätzung der longitudinalen Attribute, da nur anhand von Modellannahmen auf die Entfernung von Objekten geschlossen werden kann.

Das Verfahren der videobasierten Fahrzeugerkennung bei Tag wurde in [Wen09] evaluiert. Hierbei wurden Aufnahmen von einer Videokamera mit einer Auflösung von 640×480

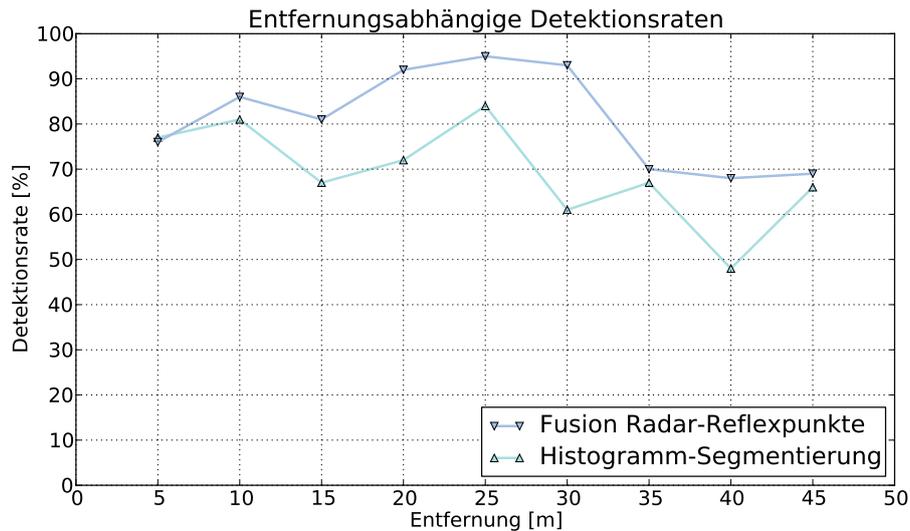


Bild 6.28: Entfernungsabhängige Detektionsrate von Objekten, gemessen mit und ohne Verwendung von Radar-Reflexpunkten.

Tabelle 6.1: Wahrheitsmatrix des SVM-Modell mit 27 Gabor-Filtern der Filterbank *F160*

SVM mit 27 Gabor-Filtern (81-dimensionaler Merkmalsvektor)			
Wahrheitsmatrix	Fahrzeug	Nicht-Fahrzeug	Unentschieden
Fahrzeug	90,14	9,16	0,70
Nicht-Fahrzeug	9,00	90,54	0,46
Unentschieden	0,00	0,00	0,00

Bildpunkten verwendet. Für das Training des Klassifikators wurden Trainingsdaten annotiert. Es stehen insgesamt 4326 Positiv-Beispiele (Fahrzeug-Bilder) und 7759 Negativ-Beispiele zur Verfügung. Alle Eingabebilder werden auf 32×32 Bildpunkte normalisiert. Für die Evaluation wurde eine Gabor-Filterbank mit 160 Filtern (*F160*) bestimmt und eine unterschiedliche Anzahl an Gabor-Filtern mittels des in Abschnitt 4.4.2 beschriebenen Verfahrens ausgewählt.

Die höchste Erkennungsrate wird dabei mit einem SVM-Modell mit 36 Gabor-Filtern (4 Filter pro Subfenster) erreicht. Das Verhältnis zwischen der Anzahl korrekt klassifizierter Objekte und der Anzahl aller klassifizierten Objekte, die sogenannte Treffergenauigkeit, erreicht dabei 91,58 % in der Evaluation der Testdatensequenz.

Zur weiteren Reduktion der Berechnungskomplexität wird schließlich ein SVM-Modell mit 27 Gabor-Filtern (3 Filter pro Subfenster) gewählt, welches eine Treffergenauigkeit von 90,33 % erreicht. Dieses Modell stellt einen guten Kompromiss zwischen Klassifikationsgenauigkeit und Verarbeitungsgeschwindigkeit dar. Die Wahrheitsmatrix des Modells ist in Tabelle 6.1 gegeben. Das Modell wird im weiteren Verlauf zur Evaluation des Verfahrens zur videobasierten Fahrzeugerkennung auf Testsequenzen herangezogen.

Abbildung 6.29 stellt Ergebnisse der Erkennung von Fahrzeugen auf Autobahn-, Stadt- und Bundesstraßen-Sequenzen dar. Die Auswertung der Autobahn-Sequenz zeigt, dass das Verfahren unter Normalbedingungen der Beleuchtungs- und Wettersituation eine hohe Güte bei der Erkennung von vorausfahrenden Fahrzeugen erreicht. Dies funktioniert, solange die Berechnung der Merkmale Schlagschatten und Kanten kaum gestört sind. Fahrzeuge werden bis in



Bild 6.29: Beispiele zur Erkennung von Fahrzeugen in Autobahn-, Stadt- und Bundesstraßen-Sequenzen aus Videodaten.

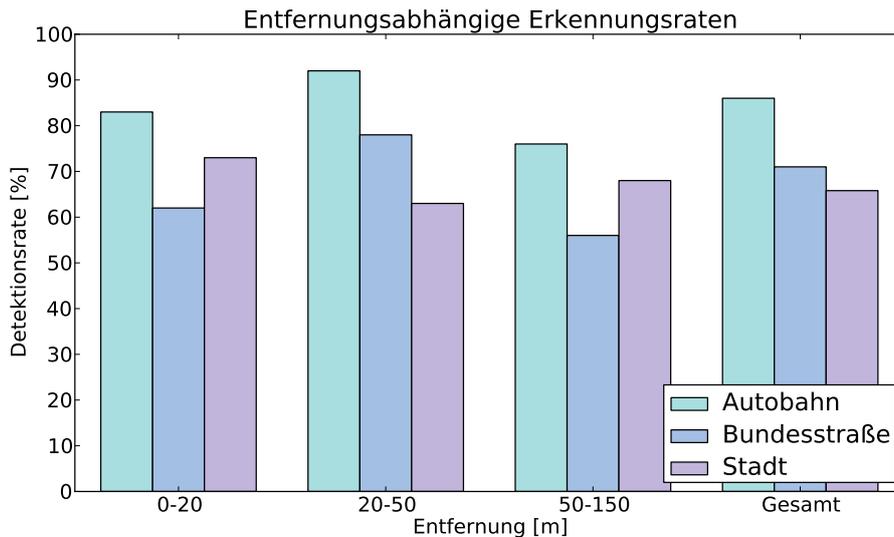


Bild 6.30: Entfernungsabhängige Erkennungsrate von Fahrzeugen auf Autobahn-, Bundesstraße- und Stadt-Sequenzen.

große Entfernungen abhängig von der verwendeten Optik der Videokamera durch das System detektiert. Wie in Tabelle 6.2 aufgeschlüsselt ist, liegt die Richtig Positiv (TP)-Rate in 953 evaluierten Einzelbildern bei 86,08 %, die Falsch Positiv (FP)-Rate bei 8,76 %.

Auch auf Bundesstraßen erfolgt die Erkennung der Merkmale zuverlässig. Die Erkennungs-raten fallen allerdings schlechter aus, da in der Evaluation zusätzlich auch entgegenkommen-de Fahrzeuge berücksichtigt wurden. Die Daten der Stadt-Sequenz stellen eine größere Her-ausforderung an die Merkmalsextraktion dar, da im innerstädtischen Bereich oftmals durch Infrastruktur die Merkmalerkennung beeinträchtigt ist. Besonders Schienen oder der Schlag-schatten von anderen Objekten beeinflussen hierbei die Erkennungsleistung. Die Erkennungs-

Tabelle 6.2: Detektionsgüte der untersuchten Sequenzen Autobahn, Stadt und Bundesstraßen

Autobahn	Anzahl der evaluierten Einzelbilder	953
	Anzahl theoretisch detektierbarer Fahrzeuge	2054
	Anzahl tatsächlich detektierter Fahrzeuge	1768
	Anzahl falsch-positiv detektierter Fahrzeuge	180
	Richtig Positiv (TP)-Rate	86,08 %
Falsch Positiv (FP)-Rate	8,76 %	
Bundesstraße	Anzahl der evaluierten Einzelbilder	215
	Anzahl theoretisch detektierbarer Fahrzeuge	416
	Anzahl tatsächlich detektierter Fahrzeuge	297
	Anzahl falsch-positiv detektierter Fahrzeuge	22
	TP-Rate	71,39 %
FP-Rate	5,23 %	
Stadt	Anzahl der evaluierten Einzelbilder	379
	Anzahl theoretisch detektierbarer Fahrzeuge	1188
	Anzahl tatsächlich detektierter Fahrzeuge	829
	Anzahl falsch-positiv detektierter Fahrzeuge	53
	TP-Rate	65,5 %
FP-Rate	6,39 %	

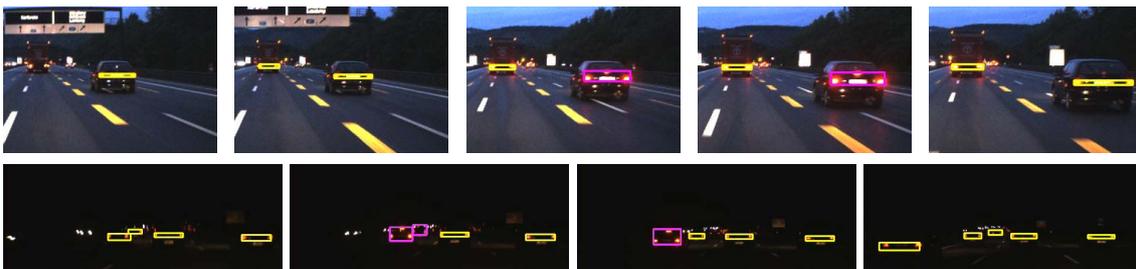


Bild 6.31: Erkennung von Fahrzeughypothesen anhand der Rücklichter (gelbes Rechteck) und des Bremslichts (magenta Rechteck) bei Dämmerung und bei Nacht.

rate nach der Entfernung aufgetragen ist in Abbildung 6.30 gegeben.

Fahrzeugeterkennung bei Nacht

Abschnitt 4.3.2 beschreibt die Erkennung von Fahrzeugen bei Dämmerung und bei Nacht. Wie in Abbildung 6.31 dargestellt ist, erfolgt die Erkennung in verschiedenen Situationen anhand der Fahrzeurrücklichter. Hierzu wurde eine Farbvideokamera mit einer Auflösung von 640×480 Bildpunkten eingesetzt. Die Evaluation der Verfahren ist detailliert in [Car09] beschrieben.

Die Erkennung der Lichtflecken im Videobild mit Hilfe der perspektivischen Filterung führt zu Lichter-Kandidaten. Diese werden durch einen regelbasierten Ansatz zu Lichter-Paaren kombiniert. In der Abbildung sind Fahrzeughypothesen durch eine gelbe BBox beschrieben, wenn es sich um Rücklichter-Kandidaten handelt. Die BBox wird magenta dargestellt, wenn ein drittes Rücklicht detektiert wurde, das Fahrzeug also ein Bremsmanöver durchführt. Die Erkennung von Fahrzeughypothesen erzielt gute Ergebnisse sowohl bei Dämmerung als auch bei Nacht. Die Erkennung erreicht dabei eine TP-Rate von bis zu 93 %. Aufgrund der wenigen zu Verfügung stehenden weiteren Merkmale zur Verifikation von Hypothesen ist allerdings die



Bild 6.32: Fehler bei der Erkennung von Fahrzeughypothesen bei Nacht durch reflektierende Baustellenmarkierungen.

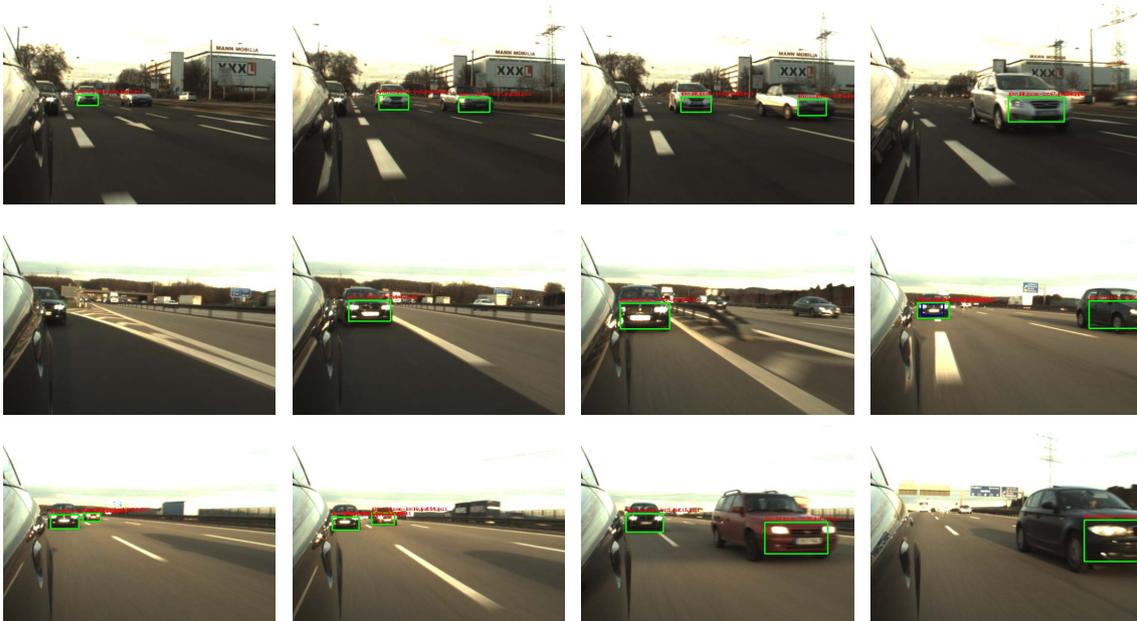


Bild 6.33: Sequenz der Erkennung von Fahrzeugen auf Nebenspuren.

FP-Rate mit 34 % sehr hoch. In einer Videosequenz wie in Abbildung 6.32 dargestellt führen die mit Reflektoren ausgestatteten Baustellenmarkierungen zu falschen Fahrzeughypothesen. Die Zuverlässigkeit mit dem präsentierten Ansatz ließe sich vermutlich durch den Einsatz einer High-Dynamic Range (HDR)-Kamera steigern, was allerdings im Rahmen dieser Arbeit nicht weiter verfolgt wurde.

Evaluation von Fahrzeugen auf Nebenspuren

Die Evaluation des Verfahrens zur videobasierten Erkennung von Fahrzeugen auf Nebenspuren ist in der Arbeit [Zhu10] erfolgt. Für die Überwachung der Nebenspuren wird die Kamera mit einer Auflösung von 640×480 Bildpunkten mit Blickrichtung nach seitlich hinten montiert.

Abbildung 6.33 zeigt Ergebnisse einer Messesequenz. Solange die Fahrzeugfront eines sich nähernden Fahrzeugs vollständig sichtbar ist, erzielt das Verfahren eine hohe Detektionsra-

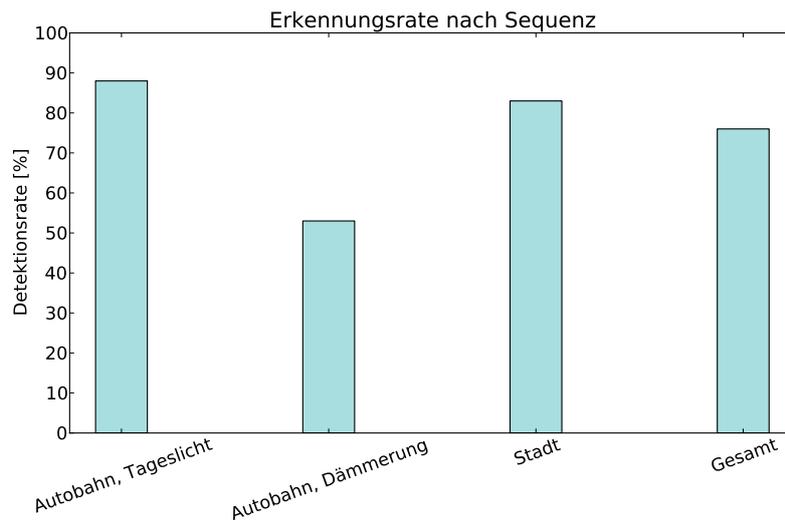


Bild 6.34: Erkennungsrate von Fahrzeugen auf Nebenspuren nach Sequenzen aufgeteilt.



Bild 6.35: Nicht symmetrische Ansicht eines Fahrzeuges aufgrund unterschiedlichem Erscheinungsbild der Fahrzeuglichter.

te. Für sehr nahe Fahrzeuge kann das Verfahren Hypothesen nur anhand der flussbasierten Weiterverfolgung fortführen, da die Erkennung anhand der modellbasierten Merkmale fehlschlägt (vergleiche Abschnitt 4.3.3). Dies funktioniert, wie in der Abbildung zu sehen ist, auch sehr zuverlässig für überholende Fahrzeuge.

Für die Auswertung wurden Messesequenzen einer Stadt- und Autobahnfahrt bei Tageslicht und Dämmerung mit mittlerer Verkehrsdichte aufgezeichnet. Zur quantitativen Bewertung wurden 46 Überholvorgänge in den Videodaten annotiert. Das Ergebnis nach Szenen aufgeteilt ist in Abbildung 6.34 zu sehen.

Das Verfahren erzielt eine gute Qualität bei Tag, allerdings sinkt die Erkennungsgüte deutlich bei Dämmerung. Dies ist erheblich durch die Frontscheinwerfer der Fahrzeuge auf Nebenspuren beeinflusst, da diese aufgrund von Blendeneffekten eine nicht symmetrische Struktur im Kamerabild hervorrufen (siehe Abbildung 6.35). Durch die Verletzung der Symmetrie können diese Fahrzeuge dementsprechend schlecht oder gar nicht durch die Hypothesengenerierung segmentiert werden. Die Erkennung von Fahrzeugen auf Nebenspuren mit Hilfe der Lichtfleck-Suche wurde im Rahmen dieser Arbeit nicht untersucht. Durch die flussbasierte Weiterverfolgung können viele Überholvorgänge erkannt werden. Dies ist bereits ab einem Abstand von etwa 30 – 40 m möglich.

Bewertung der videobasierten Fahrzeugerkennung

Die Evaluation der Verfahren stellt dar, dass in unterschiedlichen Situationen gute Ergebnisse erzielt werden können. Allerdings entspricht die Qualität der Ergebnisse häufig nicht den Anforderungen. Ein Einsatz rein monovideobasierter Ansätze zur Fahrzeugerkennung wird daher nicht weiter verfolgt. Ein zusätzlicher bereits angesprochener Nachteil dieser Ansätze sind die Modellabhängigkeit gegenüber Objektklassen. Die Merkmalsextraktion muss spezifisch angepasst und neue Modelle müssen trainiert werden.

Im Rahmen dieser Arbeit wurde dargestellt, wie eine Fusion von 3D-Tiefendaten und monoskopischer 2D-Merkmale durchgeführt werden kann. Die daraus resultierenden Ansätze verknüpfen die Vorteile der objektunabhängigen Erkennung mit der objektspezifischen Verbesserung der Attribute anhand von Modellwissen. Dabei lässt sich der Rechenaufwand dennoch gering halten, da insbesondere die aufwändigen Verfahren zur ansichtsbasierten Klassifikation nur auf wenigen, spezifisch ausgewählten Regions of Interest (ROI) berechnet werden müssen.

6.4 Evaluation der wissensbasierten Interpretation

Der in dieser Arbeit beschriebene Ansatz zur modellbasierten Interpretation von Fahrsituationen wurde mit Hilfe der Testplattformen „CoCar“ und „SegIT“ evaluiert. Zum Aufbau des Umfeldmodells wurden dabei die Verfahren zur Objektsegmentierung durch Regionenwachstum und Histogramme eingesetzt. Der folgende Abschnitt beschreibt die Evaluation des Ansatzes anhand einer ausgewählten Fahrsituation. Durch die Evaluation wird belegt, dass eine Verknüpfung zwischen Umfeldmodell und formalisiertem Weltmodell zur Berechnung einer Situationsaussage ein aussichtsreiches Konzept darstellt.

6.4.1 Verfahren zur Risikoschätzung

Die Funktion des in Abschnitt 5.3 erstellten Modells wird im Rahmen dieses Kapitels demonstriert. Dazu wurden Versuche mit der „SegIT“ Plattform durchgeführt, welche näher in [Kuh11] beschrieben sind. Eingabe in das System zur Situationsinterpretation sind die Objektliste der Umgebungswahrnehmung, die Parameter der fahrzeuginternen Sensorik sowie Hintergrundinformationen. Die Objektliste enthält die Information über die Anzahl der detektierten Objekte, deren Position im Fahrzeugkoordinatensystem, deren Ausdehnung sowie deren relative Geschwindigkeiten. Die fahrzeuginterne Sensorik liefert die Information über die Eigengeschwindigkeit des Fahrzeugs und die momentane Gierrate des Fahrzeugs. Daraus werden für jedes Objekt der longitudinale und laterale Abstand bestimmt.

Aus diesen Informationen werden zunächst Wahrscheinlichkeitsverteilungen über die Attribute gebildet. Für die Position und Ausdehnung des Objektes wird die dem eigenen Fahrzeug zugewandte Fläche des Objektes betrachtet. Diese wird anhand der äußersten Punkte der Fläche $(x_1, y_1), (x_2, y_2)$ auf die x - y -Ebene projiziert. Der longitudinale Abstand entspricht dem Wert $d_x = \min\{x_1, x_2\}$. Anhand der lateralen Abstände wird geprüft, ob sich das Objekt über dem prognostizierten Fahrweg des Fahrzeugs befindet. Falls dies nicht der Fall ist, wird der

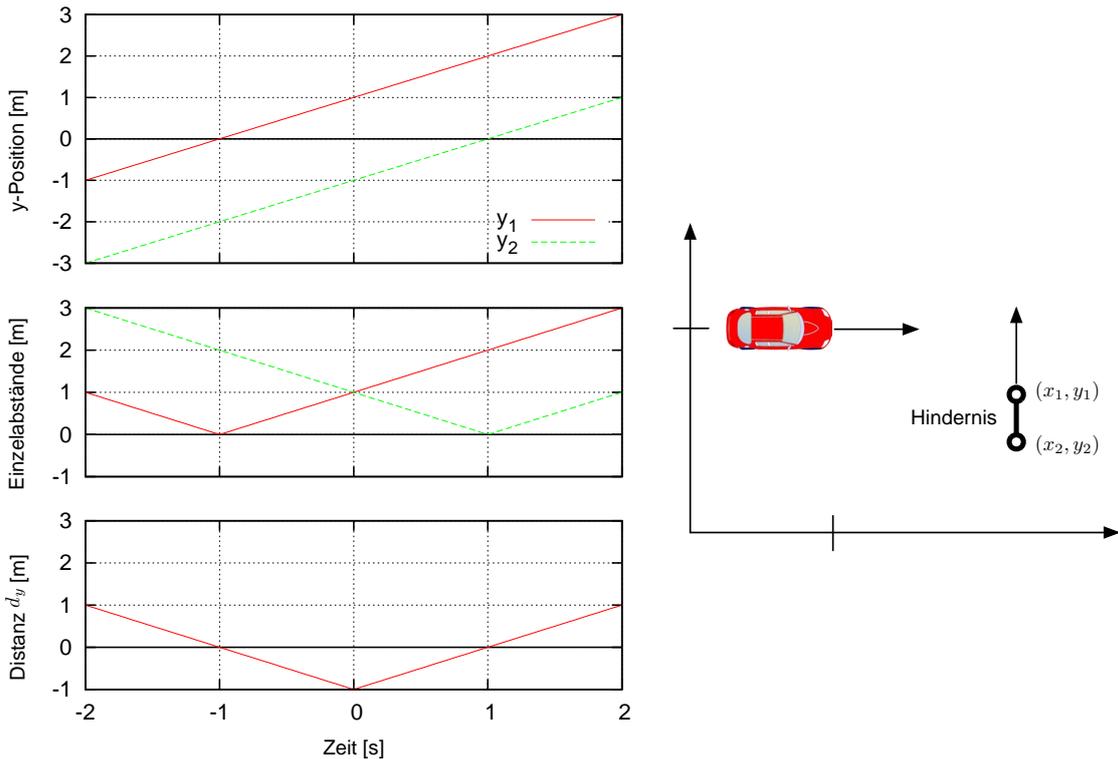


Bild 6.36: Entwicklung des Attributs d_y für ein kreuzendes Objekt; die Distanz d_y ist negativ, während das Objekt den Fahrweg kreuzt.

seitliche Versatz bestimmt. Befindet sich das Objekt über dem prognostizierten Fahrweg, wird der Abstand als negativer Versatz beschrieben:

$$d_y = \begin{cases} \min\{|y_1|, |y_2|\}, & \text{falls } y_1 \cdot y_2 \geq 0 \text{ (Hindernis kreuzt nicht den Fahrweg)} \\ -\min\{|y_1|, |y_2|\}, & \text{falls } y_1 \cdot y_2 < 0 \text{ (Hindernis kreuzt den Fahrweg)} \end{cases} \quad (6.1)$$

In Abbildung 6.36 ist der Zusammenhang zwischen dem Attribut d_y und den Abständen (y_1, y_2) schematisch dargestellt. Zunächst befindet sich das Objekt rechts des Fahrwegs. Das Attribut d_y sinkt mit kleiner werdendem Abstand des Objektes zum Fahrweg. Sobald sich das Objekt auf dem Fahrweg befindet, wird das Attribut negativ. Das Attribut sinkt weiter, solange bis der rechte Begrenzungspunkt näher am Mittelpunkt des Fahrweges ist als der linke. Das Attribut steigt anschließend wieder an und wird positiv, sobald der Fahrweg wieder vollständig frei ist.

Aus diesen Eingangsgrößen werden die Attribute der Instanzen im Weltmodell parametrisiert (vergleiche Abschnitt 5.1.3). Die Attribute der Wissensbasis besitzen diskrete Zustandsvariablen mit einem fest definierten, abzählbaren Wertebereich. Aus den kontinuierlichen Attributen der Umfeldwahrnehmung werden die diskreten Wahrscheinlichkeitsverteilungen geschätzt. Rauscheffekte in der Attributmessung lassen sich durch die Diskretisierung approximieren und Unsicherheiten in der Varianz der Attribute in der diskreten Wahrscheinlichkeitsverteilung abbilden. Die Wahrscheinlichkeitsverteilung $p_i(x)$ jedes diskreten Zustands $i = 1, \dots, N$ zu jedem kontinuierlichen Wert x lässt sich durch lineare Interpolation angeben

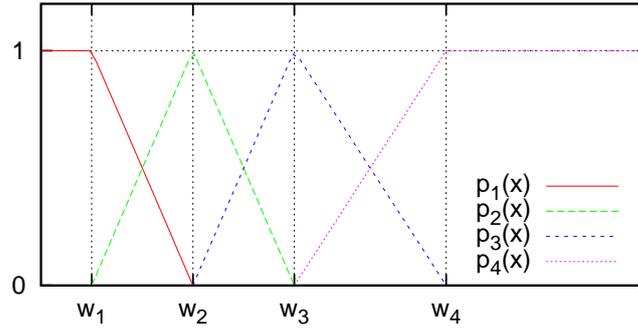


Bild 6.37: Diskretisierung mittels linearer Interpolation; $p_i(x)$ stellt die Wahrscheinlichkeit des Zustands i zum kontinuierlichen Wert x dar; Rauscheffekte in x lassen sich durch die Interpolation approximieren [Kuh11].

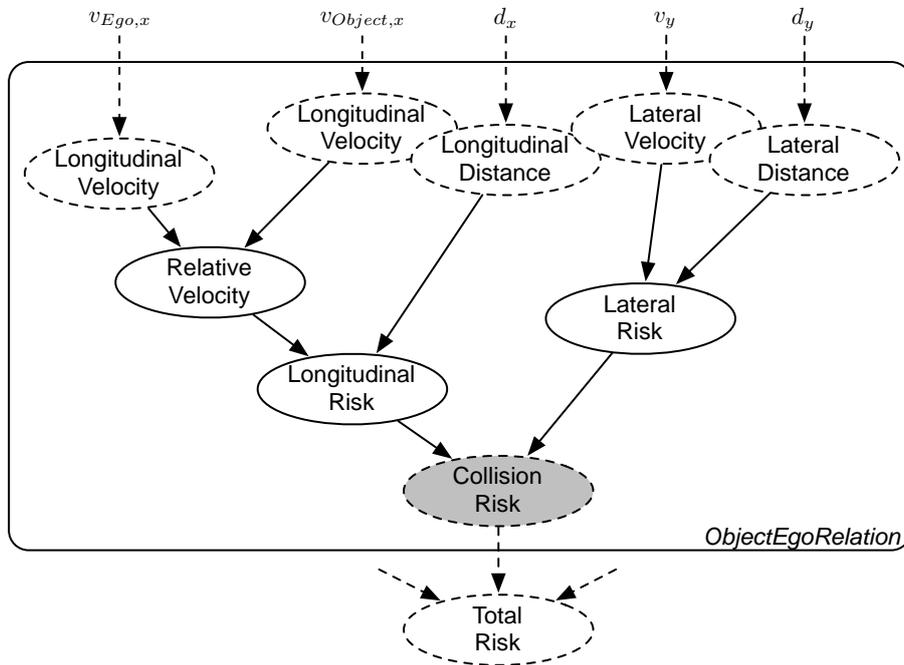


Bild 6.38: Ausschnitt aus Objektorientiertem Bayes'schem Netz (OOBN) zur Bestimmung der Risiko-Relation zwischen einem Objekt und dem Ego-Fahrzeug.

als:

$$p_i(x) = \begin{cases} \frac{x-w_{i-1}}{w_i-w_{i-1}} & \text{für } w_{i-1} \leq x < w_i \\ \frac{w_{i+1}-x}{w_{i+1}-w_i} & \text{für } w_i \leq x < w_{i+1} \\ 1 & \text{für } (i = 1 \text{ und } x < w_1) , \\ 1 & \text{für } (i = N \text{ und } x \geq w_N) \\ 0 & \text{sonst} \end{cases} \quad (6.2)$$

wobei w_i einem diskreten Wert der Wissenbasis entspricht. Abbildung 6.37 zeigt ein Beispiel mit vier Zuständen. Die Wahrscheinlichkeitsverteilung zu dem kontinuierlichen Wert x entspricht dann:

$$P(X)(x) = (p_1(x), p_2(x), p_3(x), p_4(x)). \quad (6.3)$$

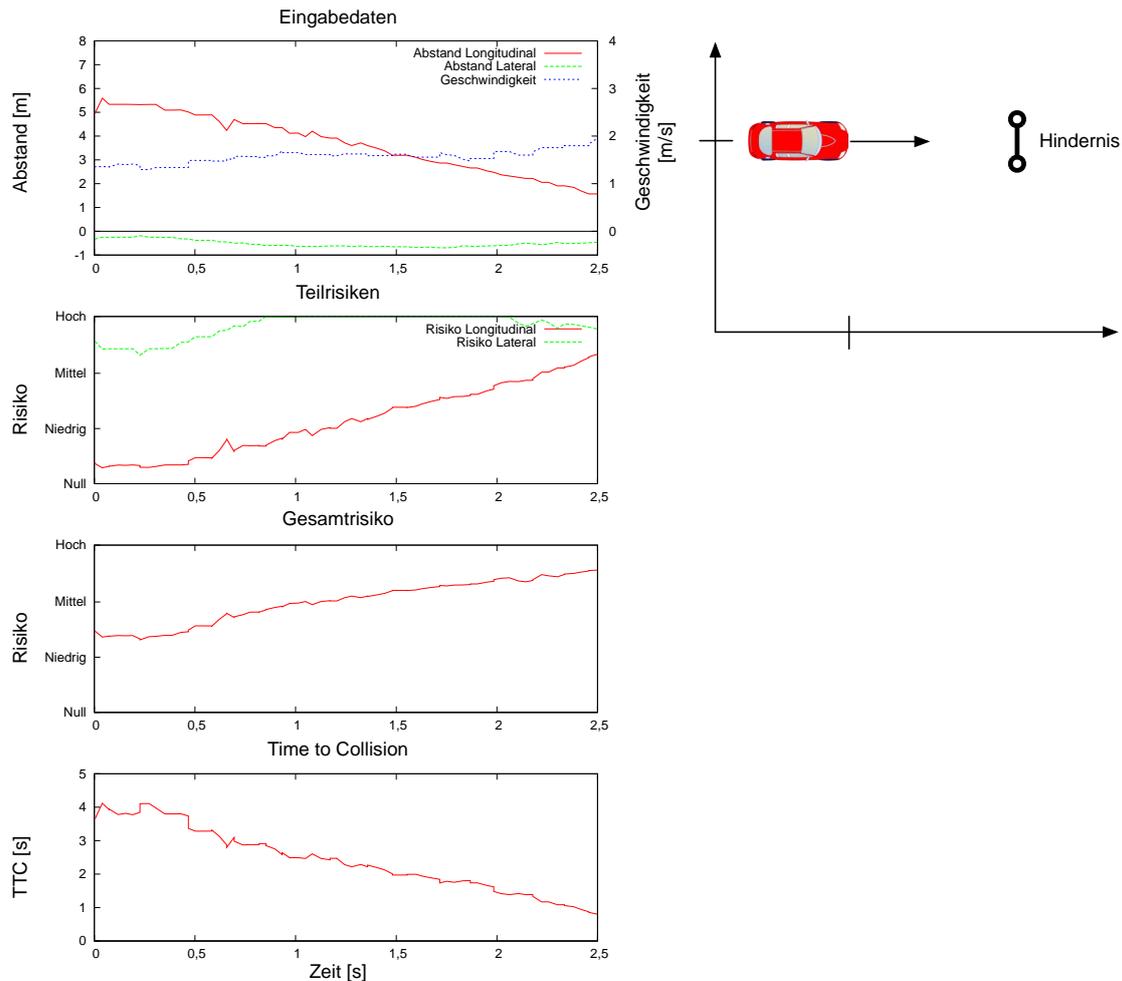


Bild 6.39: Fahr situation und Ergebnisse des Experiments mit einem statischen Objekt auf dem Fahrweg des Ego-Fahrzeugs.

Die diskretisierten Sensordaten können so zur Aktualisierung des Weltmodells eingesetzt werden. Die Modellbasis aus Abschnitt 5.3 enthält die Klassen *EgoVehicle*, *Object*, *ObjectEgoRelation* und *DynamicRiskEstimation*. Für die Klassen *EgoVehicle* und *DynamicRiskEstimation* werden feste Instanzen in das Weltmodell abgelegt, welche nicht von Messungen der Umfelderkennung abhängig sind. Für die Klassen *Object* und *ObjectEgoRelation* wird jeweils eine Instanz pro Messung der Umfelderkennung in das Weltmodell abgelegt und anschließend aktualisiert.

Abbildung 6.38 zeigt die *ObjectEgoRelation*-Klasse in der Darstellung als Objektorientiertes Bayes'sches Netz (OBN). Die Eingangsattribute d_x , $v_{Ego,x}$, $v_{Object,x}$, d_y und v_y werden von den Klassen *EgoVehicle* und *Object* berechnet und bereitgestellt. Die Struktur wird automatisch aus der Modellbasis abgeleitet. Unter gewissen Bedingungen ist es schwierig, die Effekte der Domäne korrekt anhand von Modellwissen abzuleiten. Ansätze zum Parameterlernen zur Schätzung der bedingte Wahrscheinlichkeitsverteilung (CPD) bieten hierfür ein hohes Potential. Die Spezifikation der CPD wurde in dieser Arbeit aus Modellwissen und den kausalen Zusammenhängen abgeleitet. Das Verhalten des Modells wird im Folgenden anhand verschiedener Szenen bewertet.

Zufahrt auf statisches Hindernis auf Fahrweg In Abbildung 6.39 ist eine Fahrsituation gezeigt, bei der sich das Ego-Fahrzeug einem statischen Hindernis auf dem Fahrweg nähert. Das Fahrzeug führt ein Bremsmanöver kurz vor der Kollision mit dem Objekt durch. Das erwartete Ergebnis bei der Bestimmung des Kollisionsrisikos entspricht dabei einer kontinuierlichen Zunahme des Risikos während der Annäherungsfahrt, da das Objekt den Fahrweg des Fahrzeugs über den gesamten Verlauf der Situation blockiert. Die Eingabedaten der „SegIT“ Plattform sind in der obersten Grafik der Abbildung dargestellt. Deutlich zu erkennen ist der kontinuierlich abnehmende longitudinale Abstand sowie der dauerhaft negative laterale Abstand zu dem Objekt. Der negative laterale Abstand verdeutlicht, dass das Objekt sich auf dem Fahrweg des Ego-Fahrzeugs befindet. Die Geschwindigkeit bleibt im Betrachtungszeitraum des Experiments in etwa konstant.

Die zweite Grafik der Abbildung zeigt den durch das System berechneten Verlauf der Teilrisiken. Die Teilrisiken entsprechen den Attributen *LongitudinalRisk* und *LateralRisk* der Instanz der Klasse *ObjectEgoRelation*. Zur besseren Darstellung wurden die vier Zustände der Risikoknoten auf eine kontinuierliche Wahrscheinlichkeitsverteilung übertragen. Wie zu erwarten ist das laterale Teilrisiko konstant hoch, da hierin die laterale Distanz zwischen Objekt und Ego-Fahrzeug ausgewertet wird. Da es sich um ein statisches Objekt handelt, ist die laterale Geschwindigkeit $v_y = 0$. Das longitudinale Risiko korreliert ebenfalls sehr gut mit dem gemessenen longitudinalen Abstand d_x sowie der Geschwindigkeit v_x .

Die dritte Grafik der Abbildung zeigt das ermittelte Risiko im Attribut *CollisionRisk*. Das Risiko entspricht dem gemeinsamen Verlauf der beiden eingehenden Teilrisiken und bildet das steigende Risiko einer Kollision durch die Annäherungsfahrt deutlich erkennbar ab. Die in der unteren Grafik dargestellte Kollisionszeit (TTC) wurde analytisch aus den Eingangswerten bestimmt. Der Vergleich zwischen Gesamtrisiko und der TTC zeigt eine gute Korrelation zwischen beiden Verläufen, die Schätzung der Fahrsituation mittels des gewählten Modells liefert also ein gutes Ergebnis.

Folgefahrt und Ausweichen Das oben beschriebene Experiment wird durch die Hinzunahme einer longitudinalen Geschwindigkeit des Objektes erweitert. Das eigene Fahrzeug folgt dem Objekt bis zu einer gewissen Distanz und beginnt dann mit einem Ausweichmanöver, um eine Kollision mit dem Objekt zu verhindern. Abbildung 6.40 zeigt eine Skizze des Experiments sowie die Ergebnisse als graphische Darstellung.

Anhand den Eingabedaten ist erkennbar, dass sich der Abstand zu dem Objekt kontinuierlich verringert. Die Differenzgeschwindigkeit Δv_x beträgt etwa 2 m/s. Der laterale Abstand befindet sich wie beim vorhergehenden Beispiel zunächst im negativen Bereich, das Objekt wird also im Fahrschlauch des Fahrzeugs erkannt. Ab dem Zeitpunkt $t = 2,2$ s beginnt das eigene Fahrzeug einen Überholvorgang einzuleiten. Der laterale Abstand vergrößert sich und wechselt ins Positive. Zum Zeitpunkt $t = 3,0$ s befindet sich das Objekt noch etwa 1 m vor dem Fahrzeug, verlässt aber durch den lateralen Versatz das Sichtfeld des Sensors.

Sowohl die Teilrisiken als auch das Gesamtrisiko bildet die Situation vollständig ab. Durch die Annäherungsfahrt steigt das longitudinale Risiko kontinuierlich an, so lange sich das Objekt im Sichtbereich des Sensors befindet. Das laterale Risiko ist konstant hoch, bis das Über-

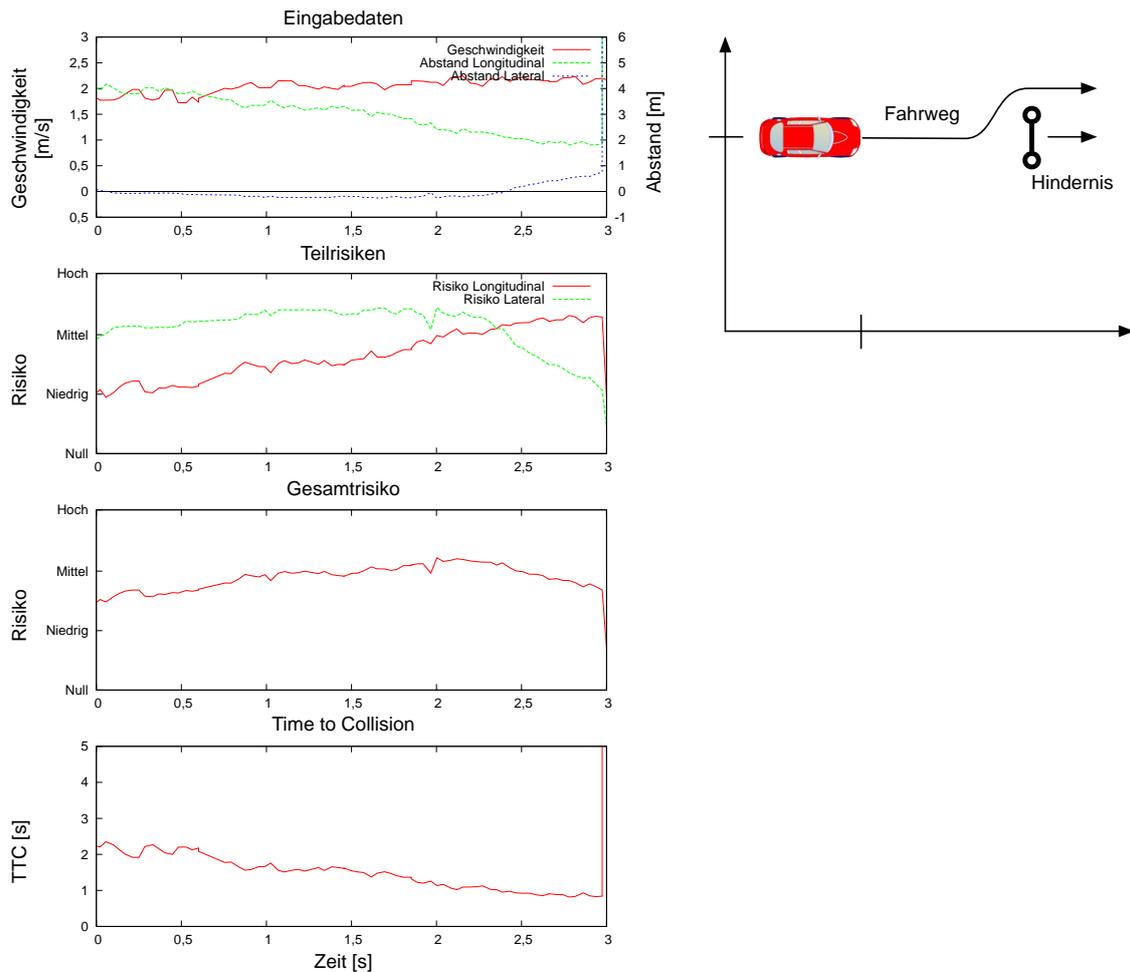


Bild 6.40: Ego-Fahrzeug folgt zunächst einem langsamen Objekt und weicht schließlich aus.



Bild 6.41: Das Experiment „Kreuzendes Hindernis“ aus Sicht eines externen Beobachters.

holmanöver zum Zeitpunkt $t = 2,2$ s beginnt. Hier fällt das laterale Risiko schnell ab, was sich auch auf das Gesamtrisiko auswirkt. Dieses beginnt ebenfalls zu diesem Zeitpunkt leicht zu sinken. Sobald das Objekt den Sichtbereich des Sensors verlässt, sinkt das Gesamtrisiko ab, da kein weiteres Objekt im Sichtbereich des Systems ist. Die vierte Grafik zeigt erneut die erwartete TTC. Da in dessen Berechnung der laterale Versatz nicht berücksichtigt ist, kann der Effekt durch das eingeleitete Überholmanöver hier nicht begünstigend wahrgenommen werden.

Kreuzendes Hindernis Die Bestimmung der Auswirkung des lateralen Abstands wurde anhand des Experiments „Kreuzendes Hindernis“ durchgeführt. Die Fahr-situation ist in Abbildung 6.41 dargestellt. Der laterale Versatz sowie die laterale Geschwindigkeit haben für die

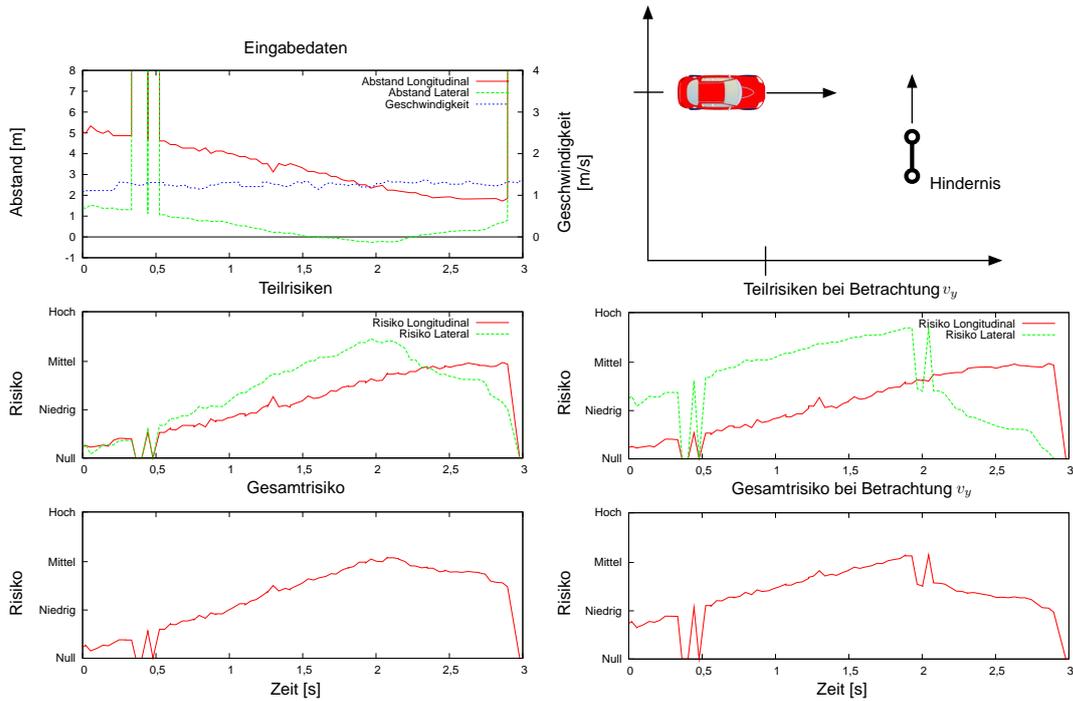


Bild 6.42: Ego-Fahrzeug nähert sich einem kreuzenden Objekt und passiert dieses; links: Auswertung ohne Betrachtung der lateralen Geschwindigkeit, rechts: Ergebnisse bei Betrachtung der lateralen Geschwindigkeit v_y .

Bestimmung des Kollisionsrisikos erhebliche Auswirkung. Das Experiment aus Abbildung 6.42 beschreibt eine Annäherungsfahrt mit etwa 2 m/s an ein Objekt, welches den Fahrweg des Ego-Fahrzeugs während der Annäherung kreuzt. Bei etwa 0,4 s wird das Objekt von der Umfeldsensorik für einen kurzen Zeitpunkt nicht erfasst, bei etwa 2,9 s verlässt das Objekt den Sichtbereich der Sensorik. Da das Objekt lateral kreuzt, verringert sich der longitudinale Abstand ähnlich wie in den oben beschriebenen Experimenten.

In den Graphen auf der linken Seite der Abbildung wird die laterale Objektgeschwindigkeit v_y nicht berücksichtigt. Es ist deutlich zu erkennen, wie das laterale Risiko durch die Annäherung des Objektes zum Fahrweg des Ego-Fahrzeugs schnell ansteigt. Dies bildet sich auch in dem Gesamtrisiko im unteren Graphen ab. Sobald das Objekt den Fahrkorridor des Ego-Fahrzeugs wieder verlässt, sinkt das Risiko leicht ab, da das laterale Risiko zurückgeht. Allerdings wird das Risiko vom System falsch eingeschätzt. Zum Zeitpunkt $t = 1,5$ s und $t = 2,5$ s hat das Objekt jeweils einen ähnlichen lateralen Abstand. Aufgrund des geringeren longitudinalen Abstands wird das Gesamtrisiko allerdings zum Zeitpunkt $t = 2,5$ s deutlich zu groß eingeschätzt.

In den zwei Graphen auf der rechten Seite der Abbildung wird die laterale Objektgeschwindigkeit mit berücksichtigt. Das erzielte Ergebnis entspricht deutlich mehr der Erwartung. Das laterale Risiko steigt durch Berücksichtigung der Objektgeschwindigkeit nun deutlich früher an und fällt nach Überqueren des Fahrwegs schneller ab als zuvor. Das daraus berechnete Gesamtrisiko entspricht dem erwarteten Verlauf umso besser. Weitere Experimente in [Kuh11] zeigen, dass die Hinzunahme der lateralen Objektgeschwindigkeit und eine genaue Schätzung des Fahrkorridors des eigenen Fahrzeugs eine deutlich bessere Schätzung des Kollisionsrisikos

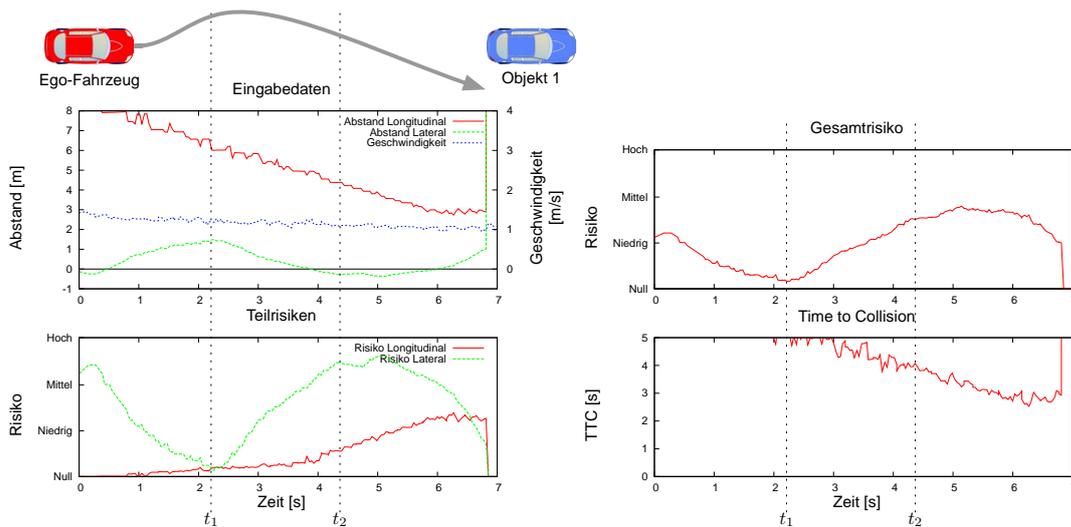


Bild 6.43: Risikoentwicklung bei der Annäherung an ein Objekt; zwischen den Zeitpunkten t_1 und t_2 ist ein deutlicher Anstieg des geschätzten Risikos zu erkennen, da sich das Fahrzeug nach Kursänderung direkt auf das Objekt zubewegt.

ermöglicht.

Verhalten bei mehreren Objekten Abbildung 6.43 zeigt die Entwicklung des geschätzten Risikos für eine geschlängelte Annäherungsfahrt an ein stehendes Objekt. In der Abbildung wird angenommen, dass durch die Umfelderkennung nur das Objekt 1 vor dem Ego-Fahrzeug erkannt wurde. Dargestellt sind im oberen Graphen die Eingangswerte d_x , d_y und v_x . Im zweiten Graph sind die Werte der Attribute *LongitudinalRisk* und *LateralRisk* dargestellt. Besonders gut zu erkennen ist der unterschiedliche Verlauf zwischen dem longitudinalen und lateralen Risiko. Während das longitudinale Risiko im Verlauf der Annäherung kontinuierlich ansteigt, hängt das laterale Risiko maßgeblich von dem Versatz d_y zwischen dem Ego-Fahrzeug und dem Objekt ab. Der dritte Teilgraph stellt den Verlauf der fusionierten Risiko-Schätzung *CollisionRisk* aus longitudinalem und lateralem Risiko dar. Zunächst nähert sich das Ego-Fahrzeug an das Objekt kontinuierlich an, wodurch das longitudinale Risiko kontinuierlich steigt. Da allerdings der laterale Versatz d_y zwischen beiden Fahrzeugen zunimmt, sinkt das Risiko einer Kollision. Ab dem Zeitpunkt t_1 ändert das Ego-Fahrzeug seinen Kurs und fährt nun direkt auf das Objekt zu. Damit steigt auch das laterale Risiko an, was sich in einer deutlichen Steigerung des Kollisionsrisikos ausdrückt. Vor einer tatsächlichen Kollision weicht das Ego-Fahrzeug dem Objekt zum Zeitpunkt t_2 durch laterale Kurskorrektur aus. Sobald der laterale Versatz zwischen Ego-Fahrzeug und Objekt steigt, sinkt das laterale Risiko erneut. Dies wirkt sich auf das Gesamtrisiko entsprechend aus, daher fällt auch dieser Risikowert.

In Abbildung 6.44 ist dasselbe Experiment gezeigt, wobei die Umfelderkennung hier sowohl Objekt 1 als auch Objekt 2 vor dem Ego-Fahrzeug als Hindernisse erkennt. Der Fokus liegt in diesem Experiment also in der Bewertung der Fähigkeit des Ansatzes, eine Situationsaussage auch bei mehreren Objekten zu treffen. Von dem System wird erwartet, dass beide Objekte die momentane Situationsaussage beeinflussen, wobei dem relevanteren Objekt ein höherer Einfluss zukommt. Im Sinne des Kollisionsrisikos bedeutet dies, dass das momentane Risiko vom jeweils situationsbedingt kritischeren Objekt beeinflusst wird.

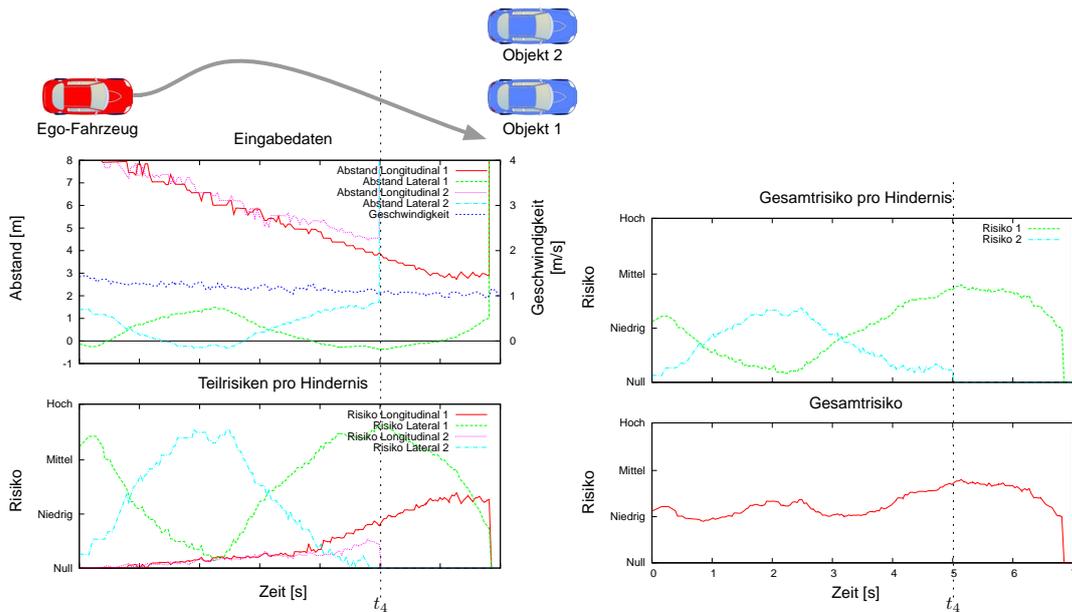


Bild 6.44: Risikoentwicklung bei der Annäherung an zwei Objekte; die Entwicklung der Einzelrisiken pro Hindernis ist deutlich zu erkennen, das Gesamtrisiko ist im Verlauf der Situation deutlich höher.

Die Eingangsdaten sind erneut im oberen Graphen dargestellt. Zusätzlich sind hier die gemessenen Eingangswerte von Objekt 2 eingetragen. Die Schwankungen, verursacht durch den Kurswechsel des Ego-Fahrzeugs, bilden sich sehr deutlich in den Eingangsattributen d_y ab. Der zweite Graph veranschaulicht das Teilrisiko pro Hindernis. Hierbei ist vor allem der Einfluss des lateralen Risikos zu beachten, welches stark mit dem Kurs des Ego-Fahrzeugs korreliert ist. Das longitudinale Risiko verhält sich für beide Objekte annähernd gleich.

Der dritte Graph zeigt das geschätzte Kollisionsrisiko pro Objekt. Zu sehen ist, wie der Einfluss von Objekt 2 durch den Kurswechsel zunächst ansteigt und anschließend wieder fällt, da das Ego-Fahrzeug eine erneute Kurskorrektur vornimmt. Das Gesamtrisiko berechnet sich durch die Aggregationsfunktion. Diese ist in der Modellbasis als Maximum-Funktion spezifiziert (siehe Abschnitt 5.3). Zum Zeitpunkt t_3 wechselt die Berechnung dynamisch von einem Risiko gebildet aus beiden Objekten zu dem Risiko gebildet aus nur einem Objekt. Das Verhalten gilt für beliebig viele Objekte.

Dieses Experiment zeigt, dass der präsentierte Ansatz gut geeignet ist, eine Situationsbestimmung auch gegeben mehrerer Objekte simultan zu berechnen. Das Verfahren stellt sich gut auf eine sich verändernde Anzahl von Hindernissen ein. Dies erfüllt die gestellte Anforderung der Robustheit an Verfahren zur SA.

Verarbeitungsdauer Die Verarbeitungsgeschwindigkeit in den oben beschriebenen Experimenten erfolgt bei bis zu vier Hindernissen in 10 ms oder weniger für den Aufbau und die Auswertung der in Abschnitt 5.3 vorgestellten Modellbasis. Die Verarbeitungsgeschwindigkeit ist höher als die Datenrate der verwendeten Sensorik und eignet sich daher für die Bestimmung des Kollisionsrisikos in direkter Folge an die Verfahren zur Umfeldwahrnehmung.

Das System wurde auf der „SegIT“ Plattform sowie auf „CoCar“ umfangreich evaluiert. Hierzu wurde eine graphische Benutzeroberfläche entwickelt (siehe Abbildung 6.45), welche



Bild 6.45: Graphische Benutzeroberfläche auf „SegIT“-Plattform zur Darstellung der Risikoschätzung; das Kollisionsrisiko ist als Balken am linken Rand dargestellt.



Bild 6.46: Im Projekt EFA2014 wurden mit Hilfe von Vorausschau-Informationen Spurwechsel-Hinweise für energieeffiziente Assistenz berechnet [Sch+12].

die Sensordaten der Plattform sowie das berechnete Risiko *OverallCollisionRisk* darstellt.

6.4.2 Bestimmung einer Blockade-Situation durch Vorausschau-Information

Der präsentierte Ansatz ist geeignet, Situationsaussagen für vorausschauende, energieeffiziente Assistenzsysteme zu erstellen. Dabei liegt ein Schwerpunkt der Betrachtung darin, für einzelne Objekte vor dem Fahrzeug zu bestimmen, ob diese das eigene Fahrzeug bei der Weiterfahrt blockieren (siehe Abbildung 6.46). Ist dies der Fall, wird diese Information dazu verwendet, das eigene Fahrzeug effizient zu verzögern oder falls sinnvoll, dem Fahrer einen Spurwechsel-Hinweis vorzuschlagen.

Wie in [Sch+12] beschrieben, wurde hierfür der probabilistische Ansatz zur Situationsinterpretation aus Abschnitt 5.1 mit der Umfelderkennung aus Abschnitt 4.1 kombiniert. Eingabedaten für das System sind die Tiefendaten der PMD-Kamera, das zeitsynchrone Videobild der 2D-Kamera sowie die Daten des fahrzeuginternen CAN-Busses. Ziel des Experiments ist die Erkennung von Fußgängern vor dem Fahrzeug, welche zu einer Blockade des Fahrzeugs führen. Diese Information kann auch genutzt werden, um nachfolgenden Fahrzeugen zu signalisieren, dass der eigene Fahrweg momentan blockiert ist.

Die einzelnen Verarbeitungsschritte sind in Abbildung 6.47 veranschaulicht. Zunächst erfolgt die Segmentierung von Objekthypothesen durch das Regionenwachstums-Verfahren. Im zweiten Schritt erfolgt die Transformation der Objekthypothesen ins Fahrzeugkoordinatensystem und anschließend die Projektion in das Kamerakoordinatensystem der 2D-Kamera. Der ROI der projizierten Objekthypothese wird mittels Haar-Wavlet-basierter Klassifikatoren die Klasse *Fußgänger* oder *Hintergrund* zugeordnet und der Konfidenzwert der Klassifikation gespeichert. Mit Hilfe der temporalen Verifikation werden die Einzelmessungen der Objektposition, -größe und -klasse über die Zeit fusioniert und das Objektwissen vervollständigt.

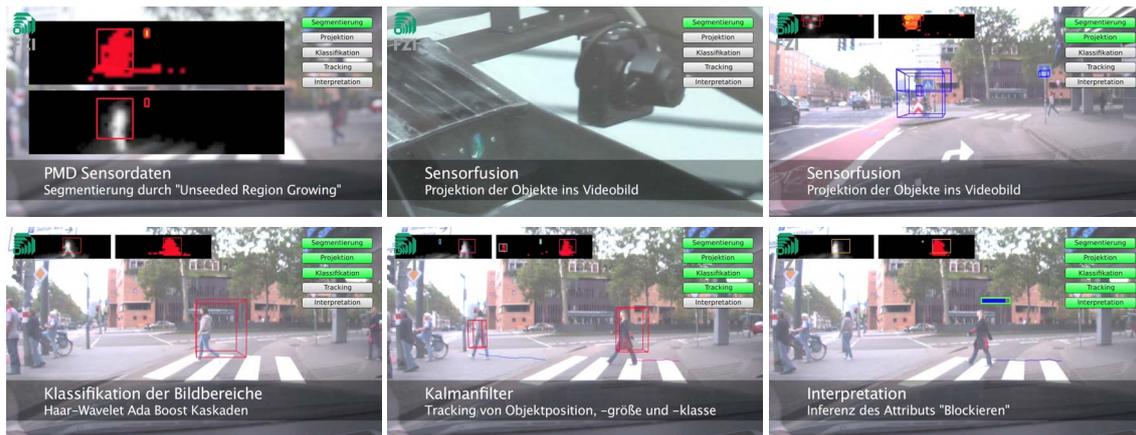


Bild 6.47: Beschreibung der Verarbeitungsschritte des Gesamtsystems; Segmentierung, Projektion, Klassifikation, Tracking und Interpretation bauen aufeinander auf.



Bild 6.48: Ergebnisse der Inferenz des Attributs *IsBlocking* unter Berücksichtigung des Fahrschlauchs des Ego-Fahrzeugs.

Die verifizierten Objekthypothesen, welche der Klasse *Fußgänger* zugeordnet sind, werden als Eingabe an das Interpretationsverfahren geliefert.

Die verwendete Wissensbasis enthält in ähnlicher Art zu der Risiko-Schätzung eine Bewertung jedes Objektes nach dessen longitudinalem und lateralem Abstand. Weiter werden die Attribute *IsBlocking* und *RecuperationPotential* eingeführt. Das Attribut *IsBlocking* bewertet die Position des Objektes anhand dessen lateralen und longitudinalen Abstand in Relation zum Ego-Fahrzeug. Das Attribut *RecuperationPotential* trifft eine Aussage, ob durch die Blockade bei großer Eigengeschwindigkeit Energie mittels Rekuperation gewonnen werden kann. Das Attribut berechnet ein Ergebnis gegeben dem *IsBlocking* Attribut sowie dem Attribut *Relative-Velocity* der longitudinalen Relativgeschwindigkeit. Die resultierende Wahrscheinlichkeitsverteilung des Attributs *RecuperationPotential* wird für jedes Objekt berechnet. In Abbildung 6.48 sind Ergebnisse der Inferenz des Attributs *IsBlocking* für Zeitpunkte einer Testsequenz dargestellt. Die Ergebnisse zeigen das Potential des gewählten Ansatzes. Aufgrund der formalen Repräsentation kann die Modellbasis leicht auf weitere Problemstellungen angepasst werden.

6.4.3 Einsatz zur automatisierten Längsregelung

Auf Basis der in dieser Arbeit beschriebenen Ansätze wurde eine Fahrzeug-Längsführung für automatisiertes Fahren entwickelt [Koh+13]. Diese verbesserte Längsführung berücksichtigt die zulässige Höchstgeschwindigkeit eines Streckenabschnitts, die maximale komfortable Geschwindigkeit während Kurvenfahrten, den Zustand der nächsten Ampel und die Geschwindigkeit für das wichtigste Zielobjekt (MRT). Das MRT bezeichnet das nächste Objekt vor dem

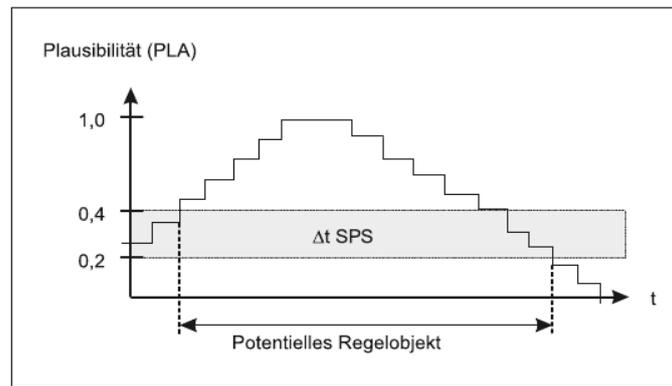


Bild 6.50: Schematischer zeitlicher Verlauf der Zielplausibilität (PLA) über den Verlauf eines potentiellen Objektes; das Attribut ist mit zeitlicher Entwicklung ausgelegt [Len13].



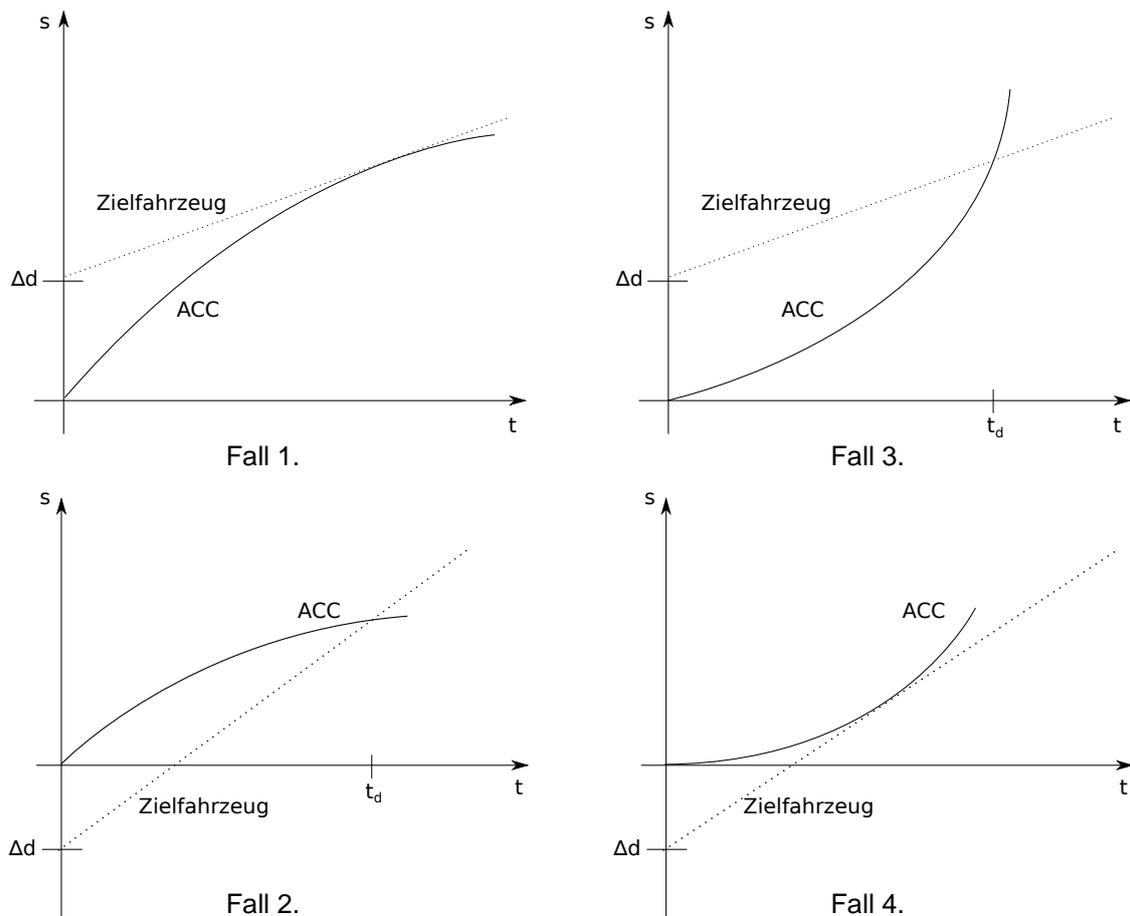
Bild 6.51: Auswahl des MRT anhand des prädizierten Fahrschlauchs; Objekte mit niedriger SPS sind mit einem blauen Rechteck umrandet, Objekte mit gültiger PLA sind mit einem roten Rechteck umrandet.

- negative Geschwindigkeit, da Gegenverkehr für das System als potentielles Zielobjekt nicht berücksichtigt wird.

Die Zielplausibilität (PLA) wird als Attribut mit zeitlicher Entwicklung ausgelegt und liefert für jedes Objekt einen Wert im Intervall $[0, 1]$ anhand der aktuellen SPS sowie der PLA des vorhergehenden Zeitschritts. Objekte mit hoher SPS führen dazu, schneller als potentielles Ziel eingestuft zu werden als Objekte mit mittlerer SPS. Eine niedrige SPS führt schnell dazu, dass Objekte nicht als Ziel infrage kommen. Ein entsprechender OOBN-Ausschnitt ist in Abbildung 6.49 dargestellt.

Abbildung 6.50 stellt den zeitlichen Verlauf der PLA für ein verfolgtes Objekt mit oberer und unterer Grenze dar. Potentielle Ziele sind Objekte, für welche die PLA über dem ersten Schwellwert von 0,4 liegt und danach nicht unter den zweiten Schwellwert von 0,2 gefallen ist. Unter den potentiellen Zielen wird das MRT anhand der kürzesten Entfernung in Längsrichtung ausgewählt. Das Ziel mit der höchsten Relevanz ist in Abbildung 6.51 mit einer roten Box umrandet dargestellt. Nicht für das System relevante Objekte sind mit einer blauen Box gekennzeichnet. Es ist deutlich zu sehen, dass anhand des prädizierten Fahrschlauchs das momentan gültige MRT korrekt gewählt wird.

Auf Basis des ermittelten Zielfahrzeugs wird durch einen Folgeregler schließlich eine Sollbeschleunigung für das eigene Fahrzeug berechnet, um dem Zielfahrzeug mit gleichbleibendem sicheren Abstand zu folgen. Hierzu gehen der Abstand zum Zielfahrzeug sowie die Geschwindigkeiten des Zielfahrzeugs und des eigenen Fahrzeugs in die Berechnung ein. Aus der Relativgeschwindigkeit $\Delta v = v_{\text{ist}} - v_{\text{soll}}$ und dem relativen Abstand $\Delta d = d_{\text{ist}} - d_{\text{soll}}$ wird der

Bild 6.52: Unterschiedliche Fälle für Folgeregler im s - t -Diagramm [Len13]

Folgeregler die Sollbeschleunigung nach vier unterschiedlichen Fällen (siehe Abbildung 6.52) berechnen:

1. $\Delta v > 0, \Delta d > 0$: Annäherung mit höherer Geschwindigkeit \rightarrow Bremsen, so dass $\Delta d = 0$ und $\Delta v = 0$ gleichzeitig erreicht werden;
2. $\Delta v > 0, \Delta d \leq 0$: Sicherheitsabstand unterschritten und Geschwindigkeit zu hoch \rightarrow Es muss bis unter die Zielgeschwindigkeit gebremst werden, um den Abstand zu vergrößern;
3. $\Delta v \leq 0, \Delta d > 0$: Außerhalb des Sicherheitsabstands aber eigene Geschwindigkeit kleiner oder gleich der Zielgeschwindigkeit \rightarrow Um den Abstand zu verkleinern muss vorübergehend über die Zielgeschwindigkeit hinaus beschleunigt werden;
4. $\Delta v \leq 0, \Delta d \leq 0$: Sicherheitsabstand unterschritten, aber eigene Geschwindigkeit kleiner oder gleich der Zielgeschwindigkeit \rightarrow So beschleunigen, dass die Zielgeschwindigkeit und der Abstand zur gleichen Zeit erreicht werden.

Abbildung 6.53 zeigt den Verlauf der Sollbeschleunigung für eine Folgefahrt bei aktivierter Längsführung mit dem Fahrzeug „CoCar“. Während der Folgefahrt befindet sich ein Fahrzeug in etwa 30 m voraus, als es zu einem Wechsel des MRT aufgrund eines einscherenden Fahrzeugs kommt. Da das einscherende Fahrzeug eine Distanz von 20 m hat, wird zunächst

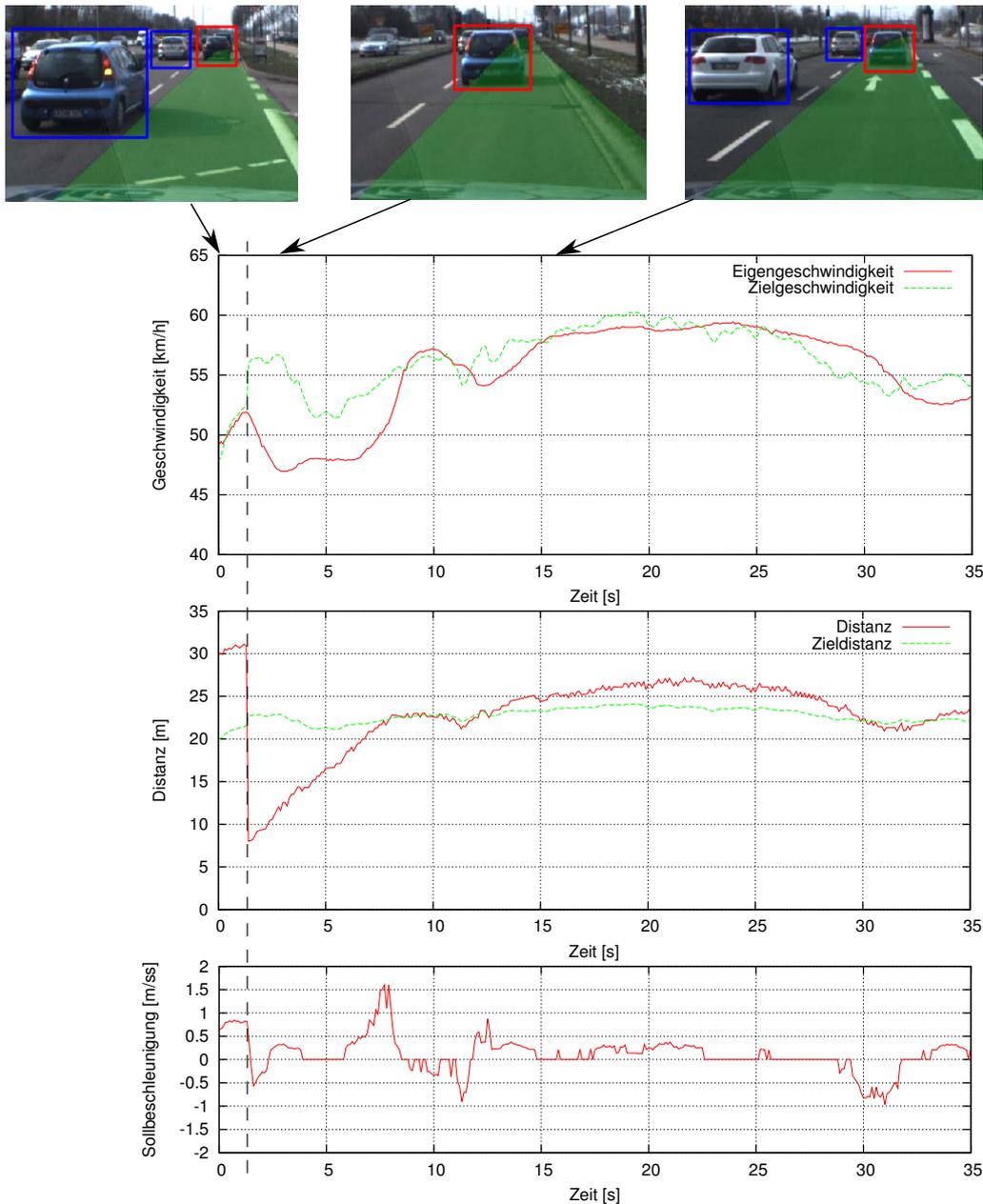


Bild 6.53: Evaluation des Folgereglers mit Fahrzeug „CoCar“ bei einer Folgefahrt und einscherendem Fahrzeug [Len13].

aufgrund der zu geringen Distanz abgebremst. Nach etwa 8 s werden Zieldistanz und Zielgeschwindigkeit gleichzeitig erreicht und „CoCar“ folgt dem neuen Objekt im weiteren Verlauf.

Die dargestellte Anwendung für den Einsatz zur automatisierten Längsführung zeigt, dass die Objektbildung auf PMD-Daten für den Einsatz zur Umfeldwahrnehmung ausreichend robust ist und probabilistische Methoden für die Interpretation relevanter Umfeldinformationen geeignet sind. Die Grenze der Einsatzfähigkeit für dieses Anwendungsbeispiel liegt in der begrenzten Reichweite der Sensorik, wodurch bei Annäherungsfahrten mit Geschwindigkeiten größer als 60 km/h auf ein stehendes Objekt der Sicherheitsfahrer des Fahrzeugs in die Längsführung eingreifen musste, da eine komfortable Verzögerung des Fahrzeugs nicht mehr durch das System erreicht werden konnte.

6.4.4 Bewertung der wissensbasierten Interpretation

Die Evaluation der wissensbasierten Interpretation der Fahrsituation zeigt anhand durchgeführter Experimente, dass der gewählte Ansatz in der Lage ist, eine Einschätzung einer Fahrsituation zu leisten. Anhand des beispielhaft gewählten Risikos einer Kollision wird durch das System eine Schätzung berechnet, welche auf den longitudinalen und lateralen Abständen und Geschwindigkeiten der Objekte in Bezug auf das eigene Fahrzeug basiert.

Die eingesetzte Wissensbasis ist formal beschrieben und erlaubt abhängig von den Objektlisten der Umfelderkennung die Anpassung des Weltmodells. Durch Inferenz ist es möglich, die probabilistischen Attribute der Objekte im Weltmodell auszuwerten und eine Situationsaussage abzuleiten. Dies funktioniert für die getesteten Objektkonstellationen ohne Einschränkungen. Der Aggregationsknoten führt auch bei wechselnder Anzahl von Objekten die geschätzte Aussage zu einer Gesamtaussage zusammen, ein Unterschied zu bisherigen im Stand der Technik beschriebenen Verfahren.

Das Modell zur Bestimmung des Kollisionsrisikos wurde sowohl in fußgängerzentrierten Bereichen als auch im Straßenverkehr evaluiert. Die notwendigen Anpassungen ergeben sich im Wesentlichen in der Überführung der kontinuierlichen Messwerte in diskrete Zustände. Hierfür wurden im Stand der Technik bereits Ansätze präsentiert, wie beispielsweise die Überführung mittels adaptiver Wahrscheinlichkeitsknoten [SW11]. Veränderungen in der Modellbasis wie das Hinzufügen neuer Auswerteknoten oder neuer Klassen sind leicht möglich, wie in den Anwendungsbeispielen gezeigt ist.

7 Zusammenfassung und abschließende Betrachtung

7.1 Zusammenfassung

Ziel dieser Arbeit ist die Erforschung von Verfahren zur Interpretation von Fahrsituationen. Dabei wird besonderes Augenmerk auf eine wiederverwendbare, modellbasierte Repräsentation von Situationen gelegt. Die entwickelte Methodik soll gewährleisten, dass eine durchgängige, robuste Umfeldwahrnehmung und Situationsanalyse (SA) entsprechend gängiger Anforderungen der Fahrerassistenzsystem (FAS)-Domäne möglich ist.

Methodische Schwerpunkte Für die Umfelderkennung wurde ein konzeptueller Ablauf erarbeitet, welcher komplementäre Sensorinformationen verwendet und zu einer robusten Objektbildung führt. Zunächst wird eine Erkennung unbekannter Objekte durch Lichtlaufzeit (ToF)-Kameras durchgeführt. Alternativ werden auf Basis von Videokameras spezifische Objektklassen detektiert. Hierfür werden Methoden vorgestellt, welche aus den Signalinformationen der Kameras Objekthypothesen bilden. Unter den vorgestellten Ansätzen befinden sich auf einzelne Kameratypen spezialisierte sowie generische Verfahren. An die Bildung der Objekthypothesen schließt sich eine Klassifikation von Objekten an. Es werden Ansätze verwendet, die auf Kaskaden- oder Support Vektor Maschinen (SVMs) basieren. Mit Hilfe der vorgestellten Verfahren erfolgt eine Überwachung des Fahrzeugumfelds vor, neben sowie hinter dem Fahrzeug. Durch temporale Zustandsschätzung werden Objekthypothesen über die Zeit verfolgt und zusätzliche Attribute wie beispielsweise die Geschwindigkeit von Objekten ermittelt. Die Integration der Verfahren in ein einheitliches Verarbeitungskonzept ermöglicht es, die entwickelten Ansätze übergreifend und je nach zur Verfügung stehenden Sensorik einzusetzen.

Aus den Objektinformationen der Umfelderkennung wird in dieser Arbeit ein Weltmodell aufgebaut, welches einer formalen Repräsentation der momentanen Umgebung entspricht. Das Weltmodell verwendet Hintergrundwissen und an die Objekterkennung angepasste Modellkonzepte, welche in einer gemeinsamen Wissensbasis gespeichert sind. Die formale Grundlage der Wissensbasis ist die Objektorientierte Probabilistische Relationale Modellierungssprache (OPRML). Es wurde gezeigt, dass diese Sprache zur Beschreibung von Fahrsituationen (also von Objekten, den Veränderungen der Objektzustände sowie der Objektrelationen) gut geeignet ist. Dabei wird insbesondere auf die Berücksichtigung von Unsicherheit Wert gelegt, um Fahrsituationen mit einem Glaubensgrad versehen zu können. Das vorgestellte Konzept aktualisiert Objekte im Weltmodell zur Laufzeit. Das Weltmodell stellt somit die Basis für eine erfolgreiche SA dar.

Die SA erfolgt durch den Einsatz probabilistischer Netze. Diese Netze beschreiben die relevanten Eigenschaften der Fahrsituation in Form probabilistischer Knoten und Abhängigkeiten. Auf Grundlage der graphischen Modelle können mit Inferenzverfahren Schlüsse über die Fahr-

situation gezogen werden. Hierfür wurde in dieser Arbeit ein Verfahren entwickelt, welches aus dem Weltmodell ein probabilistisches Netz erzeugt und zur SA nutzt. Die Erstellung des Netzes basiert dabei auf der formalen Repräsentation der Modellbasis und erfolgt schnell genug, sodass eine Bestimmung der Fahrsituation in Echtzeit möglich ist.

7.2 Bewertung der Prozesskette

Sensordatenverarbeitung Für die Sensordatenverarbeitung wurde eine Verarbeitungskette für komplementäre Sensorsysteme umgesetzt. Die Verfahren zur Segmentierung nutzen Objektwissen oder können ohne Kenntnis über die zu erkennenden Objekte zu einer Objektbildung führen. Die aufgezeigten Verfahren, insbesondere die Verfahren zur Segmentierung von Tiefenbildern anhand eines Homogenitätskriteriums oder anhand von Histogrammen, besitzen dabei eine sehr gute Leistung bei der Erkennung von Objekten. Auf Basis dieser Verfahren wurden in Anwendungsbeispielen weiterführende Experimente durchgeführt. Beispielsweise zur automatisierten Längsführung des Versuchsfahrzeugs „CoCar“ im Stadtverkehr, welche die Erkennungsleistung belegen.

Durch die in die Verarbeitungskette integrierte temporale Zustandsschätzung ist es möglich, für die umgesetzten Segmentierungsverfahren zeitliche Attribute wie die Schätzung der Geschwindigkeit zu generieren. Gleichzeitig trägt der Einsatz des Filters wesentlich zur Verbesserung der Qualität der erzeugten Objektlisten bei, da fehlerhaft erkannte Objekthypothesen durch das System unterdrückt werden. Die Klassifikation von Objekten mittels videobasierter Sensorik konnte in dieser Arbeit nur am Rande betrachtet werden.

Eine Einbindung von Verfahren in die Sensordatenverarbeitung kann nach dem hier vorgestellten Konzept flexibel erfolgen. Die Laufzeit der prototypisch umgesetzten Verfahren orientiert sich an den Anforderungen, Umgebungsdaten von den Kameras in Echtzeit zu verarbeiten.

Modellierung Der zweite wesentliche Beitrag liegt in der Modellierung von Fahrsituationen. In der Arbeit wurde dargestellt, dass sich Fahrsituationen besonders günstig durch Entity-Relationship (ER)-basierte Sprachen beschreiben lassen. Diese können die spezifischen Eigenschaften einer Fahrsituation, nämlich die Kombination des Zustands des Eigenfahrzeugs, der Zustände relevanter Objekte im Fahrzeugumfeld sowie die Veränderung dieser Zustände direkt abbilden. Da die getroffenen Annahmen zur Beschreibung von Fahrsituationen teilweise unter Unsicherheit erfolgen, wurde eine Repräsentation gewählt, welche einen relationalen und einen probabilistischen Teil zur Modellierung vorsieht. In der Arbeit wurde dargestellt, dass die beiden Teile gemeinsam ein probabilistisches Modell über die spezifizierte Situation beschreiben. Die gewählte Struktur erleichtert es durch ihren flexiblen Aufbau, Fahrsituationen durch Expertenwissen kontinuierlich zu erweitern. Der modellbasierte Ansatz erlaubt es, die Zusammenhänge der relevanten Elemente zur Beschreibung von Fahrsituationen übersichtlich auf hoher Ebene darzustellen und erleichtert damit eine Überprüfung der modellierten Situationen.

Interpretation Die Situationsanalyse in dieser Arbeit basiert zu großen Teilen auf der gewählten Repräsentation und ist in der Lage, die dynamischen Veränderungen von Fahrsituationen kontinuierlich zu berücksichtigen. Es wurde aufgezeigt, dass die SA nicht auf Basis starrer Strukturen erfolgen darf, sondern sehr eng an die Umfelderkennung angebunden sein muss. Aspekte von Fahrsituationen können ansonsten nicht korrekt behandelt werden. In der Arbeit wurden dazu Konzepte vorgestellt, welche eine enge Kopplung zwischen Objektlisten und formaler Repräsentation beschreiben und dennoch durch einen modularen Aufbau eine spezifische Entwicklung der einzelnen Komponenten ermöglichen. Hieraus resultiert ein Weltmodell, das als formale Darstellung der aktuellen Fahrsituation verstanden werden kann. Aus dem Weltmodell ist es möglich, automatisch probabilistische Netze zu erzeugen und zur Inferenz über die unsicheren Attribute der Fahrsituation einzusetzen. Die zugrundeliegenden Objektorientierten Bayes'schen Netze (OOBNs) ermöglichen es dabei, die probabilistischen Netze strukturiert aufzubauen und effiziente Inferenz durchzuführen. Damit erreicht die Arbeit das gesetzte Ziel zur probabilistischen Interpretation von Fahrsituationen über einen modellbasierten Ansatz.

Prozesskette Die in dieser Arbeit vorgestellte Prozesskette wurde im Rahmen der Evaluation umgesetzt und an die Verfahren zur Umfelderkennung angebunden. Die Evaluation dieser Methoden zeigt die Ergebnisse der Umsetzung auf verschiedenen Versuchsfahrzeugen wie „CoCar“ und dem „SegIT“ unter realen Bedingungen. Der gewählte Ansatz lässt sich leicht an unterschiedliche Rahmenbedingungen anpassen. Durch die Arbeit konnte der Funktionsnachweis der Situationsanalyse erbracht werden.

Beiträge der Arbeit In der Arbeit wurde ein Ansatz aufgezeigt, welcher in der Lage ist, ausgehend von den Signalinformationen der Kameras und einer formal beschriebenen Wissensbasis eine Situationsbewertung in Echtzeit durchzuführen. In bisherigen Arbeiten wurde eine Situationsbewertung entweder statisch durchgeführt, also ohne eine Anpassung an das Umfeld zu gewährleisten, oder die Situationsbewertung erfolgte analytisch und dadurch schwer an veränderte Randbedingungen erweiterbar. Der hier vorgestellte Ansatz ermöglicht es, sich flexibel an die Fahrsituation zu adaptieren und führt zu einer Analyse modellierter Situationen in Echtzeit.

Diese Arbeit trägt wie folgt zum wissenschaftlichen Stand bei:

1. Sensordatenverarbeitung: Entwicklung einer Verarbeitungskette zur Erkennung relevanter Objekte und Attribute mittels geeigneter komplementärer Sensoren;
2. Situationsmodellierung: Repräsentation von Fahrsituationen in flexibler, erweiterbarer Struktur durch wahrscheinlichkeitsbasierte Logik erster Ordnung;
3. Situationsanalyse: Automatische Konstruktion von graphischen Modellen mit Bezug auf relevante Objekte und Relationen und Situationsinterpretation durch probabilistische Inferenz;
4. Umsetzung der Prozesskette auf Versuchsfahrzeugen zur Bewertung des Funktionsnachweises der Situationsanalyse unter realen Bedingungen.

7.3 Ausblick

Ein System muss in der Lage sein, die Wahrnehmung des Umfelds selbst zu beeinflussen und entsprechend der verfügbaren Ressourcen selbst zu entscheiden, welche Objekte im Fahrzeugumfeld in der momentanen Situation relevant sind und erkannt werden müssen. Dies wird im Allgemeinen als aktives Sehen bezeichnet. Weiter ist es notwendig, die Unsicherheiten in der Wahrnehmung, in der Situationsanalyse und in der Ableitung einer Handlungsempfehlung oder Aktion durchgängig zu berücksichtigen. Insbesondere wenn verschiedene Sensoren zum Einsatz kommen, die möglicherweise widersprüchliche Informationen erzeugen. In dieser Arbeit konnte gezeigt werden, dass probabilistische Abhängigkeiten von Informationen und die kausalen Zusammenhänge technisch abgebildet und genutzt werden können. Allerdings spielt das Lernen der Parameter in Zukunft eine wichtige Rolle. Besonders in Hinblick auf den Einsatz der Systeme in verschiedenen Ländern, welche unterschiedliche Verhaltensweisen von Verkehrsteilnehmern mit sich bringen, muss sich ein System kontinuierlich an die Umgebung adaptieren. Dies führt möglicherweise dazu, dass mit der Zeit Wissen verworfen und neu eingelernt werden muss. Neben dem Lernen der Parameter ist auch das Lernen der Strukturen, also der kausalen Zusammenhänge von großer Bedeutung.

Aus heutigem Stand ist es möglich, eine Situation für einen kurzen Zeitraum präzisieren zu können. Besonders im Hinblick auf automatisiertes Fahren ist es notwendig, dass die technischen Systeme den Zeithorizont, für den zuverlässige Informationen präzisiert werden können, kontinuierlich steigern. Darauf aufbauend können Funktionen umgesetzt werden, welche eine gefährliche Situation frühzeitig erkennen und proaktiv zur Steigerung der Sicherheit im Straßenverkehr beitragen können. Die Vision des unfallfreien Fahrens rückt damit einen Schritt näher.

Damit ein technisches System in der Lage ist, adäquat auf unterschiedliche Situationen zu reagieren, erfordert es einen großen Fortschritt des wissenschaftlichen Standes der Technik im Bereich kognitiver Prozesse. Diese umfassen das Wahrnehmen, das Lernen und das Erinnern. Wenn ein technisches System in der Lage ist, diese hier allgemein beschriebenen Prozesse mit einer hohen Güte zu beherrschen, kann das System als intelligent bezeichnet werden und sich auf unterschiedliche Situationen selbstständig einstellen.

Abbildungsverzeichnis

1.1	In einem kognitiven Regelkreis werden geeignete Aktionen aufgrund wahrge- nommener Informationen berechnet [RN09].	3
1.2	Adaptive FAS setzen sich aus Wahrnehmungs-, Analyse-, Entscheidungs- und Aktionskomponenten zur Informationsverarbeitung zusammen; nach [RHO09].	4
2.1	Taxonomie von Messverfahren zur Erfassung fahrrelevanter Informationen. . . .	10
2.2	Messprinzip bildgebender Systeme; ein fokussiertes Objekt bildet sich durch das optische Zentrum in eine 2D-Bildebene ab.	11
2.3	Durch vier Abtastungen lässt sich die Phasenverschiebung zwischen ausgesen- detem und reflektiertem Signal bestimmen.	12
2.4	Tiefenbild-Kamera in Testfahrzeug „CoCar“: Das System besteht aus zwei Be- leuchtungseinheiten im Kühlergrill und einer Kamera hinter der Windschutz- scheibe.	13
2.5	Zwei Kameras eines Stereokamera-Systems mit paralleler optischer Achse und eine Mono-Videokamera in „CoCar“.	13
2.6	ibeo LUX Sensoren lassen sich durch ihre Bauform kompakt in der Stoßstange von Fahrzeugen integrieren.	14
2.7	Rotierender Velodyne HDL-32E Lidar-Scanner auf dem Fahrzeugdach des Test- fahrzeugs „CoCar“.	15
2.8	Logisches System: Sätze entsprechen formalisierten Aspekten der realen Welt; durch Inferenz lassen sich Folgerungen aus Sätzen ableiten [RN09].	17
2.9	Ein einfaches Bayes'sches Netz (BN) sowie die bedingten Wahrscheinlichkeits- tabellen; Y ist von X bedingt abhängig.	21
2.10	Beispiel der Universitäts-Domäne in relationaler ER-Darstellung; Klassen besit- zen Attribute und sind durch Relationen miteinander verbunden [How10]. . . .	25
2.11	Probabilistisches Modell der Universitäts-Domäne in PRM-Form; Durchgezo- gene Pfeile stellen probabilistische Abhängigkeiten zwischen Attributen dar, gestrichelte Linien Relationen zwischen Klassen; der fette Pfeil entspricht einer Zählfunktion (QNT) zur Aggregation der <i>Impact</i> -Attribute aller <i>Paper</i> -Klassen [How10].	26
2.12	Probabilistisches Modell der Universitäts-Domäne in OPRM-Form; Durchgezo- gene Pfeile stellen probabilistische Abhängigkeiten zwischen Attributen dar, gestrichelte Attribute entsprechen Referenzknoten, grau gefärbte Attribute ent- sprechen Ausgangsknoten; gestrichelte Pfeile bilden Referenzkanten zwischen Ausgangsknoten und erzeugten Referenzknoten [How10].	27

3.1	Komponenten eines Sensorsystems: Sensoren erfassen die Umwelt und bilden daraus Objektwissen; nach Blackman und Popull [BP99].	30
3.2	Definition des fahrzeugfesten Inertialsystems.	31
3.3	Stereodaten werden durch die Stixel-Repräsentation zu Stiften zusammengefasst, welche auf der freien Fahrfläche aufsetzen [PF10].	37
3.4	Komponenten und Ablauf eines <i>Multi-Target-Tracking</i> (MTT) Systems [BP99]. . .	39
3.5	Gating zur Assoziation zwischen Messdatum und Track.	39
3.6	Blockdiagramm des IMM Filters, nach Jordan [Jor06].	43
3.7	Situationsteilbaum zur Identifikation einer Notbremssituation [Sch09].	46
3.8	Vacek [Vac08] setzt fallbasiertes Schließen zur Situationsanalyse ein; der Ablauf des Verfahrens besteht aus vier Schritten: <i>retrieve</i> , <i>reuse</i> , <i>revise</i> und <i>retain</i> . . .	50
3.9	Risiko-Muster entsprechen Objektkonstellationen mit einem Risiko-Wert; In einem Inferenz-Schritt wird die Wissensbasis mit den Mustern verglichen und Risiko-Werte den Objekten zugeordnet.	52
3.10	Wissensgetriebene Interpretation mittels Objekt-Attribut Schablonen [How10].	52
3.11	Veranschaulichung der Anforderung <i>Generisch</i> : Die Domäne sollte bezüglich Klassen und Relationen erweiterbar sein.	55
3.12	Veranschaulichung der Anforderungen <i>Robustheit</i> : Gewährleistung einer einheitlichen Funktionsweise auch bei Veränderungen in der Domäne.	55
3.13	Veranschaulichung der Anforderungen <i>Modularität</i> : Wissensrepräsentation, Modellkonstruktion und Inferenzverfahren stellen separate Komponenten dar. . . .	55
4.1	Konzeptioneller Ablauf der Umfelderkennung mit komplementären Kameras und temporaler Fusion zur Bildung von Objektwissen.	60
4.2	Konzeptioneller Ablauf zur Bildung von Objekthypothesen aus ToF-Tiefendaten durch verschiedene Segmentierungsverfahren.	61
4.3	Ablauf der Objektsegmentierung durch Regionenwachstum; mithilfe eines Homogenitätsmaßes wird für Bildpunkte geprüft, ob diese Teil einer Region sind. .	63
4.4	Messbare Flächen-Konfigurationen sind senkrecht auf der x - y und x - z -Ebene stehende Flächen.	65
4.5	Ablauf der Histogrammbasierten Objektsegmentierung (links) und der modifizierte Liniensuche (rechts).	66
4.6	Tiefenbild in Falschfarbendarstellung mit entsprechenden U/V -Räumen. Die Bodenebene im V -Raum ist rot umrandet; links: Reziproke Erstellung des V -Raums; rechts: Erstellung des V -Raums mit äquidistanten Tiefenwerten.	66
4.7	Unterdrückung des Grauschleiers im U -Raum durch unterschiedliche Akkumulator-Schwellwerte T_h	67
4.8	Liniensextraktion; 1. Bild: Maskierte Linie mit ϵ -Bereich; 2. Bild: Überdeckungshistogramm entlang der Maskierung; 3. Bild: Akkumulatorwerte der Start- und Endpunkte von Linien; 4. Bild: Extrahierte Linie.	68
4.9	Fortlaufend werden extrahierte Linien aus dem Hough-Raum und dem U -Raum entfernt; nach dem dritten Durchlauf wurden alle deutlichen Linien im U -Raum segmentiert.	70

4.10	Line Pattern Hough Transformation (LPHT) zur Bestimmung der vertikalen Ausdehnung von Objekten; oben: Betrachtetes Liniensegment; rechts: Vertikales Überdeckungshistogramm und berechnete Akkumulatorwerte der vertikalen Start-/Endpunktsuche; unten links: resultierende Fläche als weiße Box umrandet.	71
4.11	Ergebnisse der segmentierten Objektflächen und der befahrbaren Bodenfläche; Objektflächen sind mit weißen Boxen umrandet, die befahrbare Fläche ist grau texturiert dargestellt.	71
4.12	Konzeptioneller Ablauf der videobasierten Objekterkennung, das Resultat ist eine Liste segmentierter Objekte.	72
4.13	Konzept der videobasierten Fahrzeugerkennung bei Tag.	73
4.14	Links: Objekte gleicher Breite besitzen mit steigender Entfernung eine geringere Breite in Bildpunkten; rechts: Erwartete minimale und maximale Breite der Unterkante von Fahrzeugen, ermittelt anhand intrinsischer und extrinsischer Kameraparameter.	74
4.15	Verarbeitungsschritte der Fahrzeugerkennung bei Tag: a) Bestimmung der horizontalen Kante (weiß) und vertikalen Symmetrieachse (rot); b) Bestimmung der vertikalen Kanten; c) Korrektur der ROI; d) Erstellte Fahrzeughypothese. . .	76
4.16	Aufgezeichnete Videobilder einer Farbkamera bei Dämmerung und bei Dunkelheit.	77
4.17	Konzeptioneller Ablauf der videobasierten Fahrzeugerkennung bei Nacht.	77
4.18	Links: 3×3 Maske für 2D Gauss-Filterkernel; rechts: Schnitt durch den Filterkernel $g(x) - o_g$ im Intervall $[-l_g, l_g]$	78
4.19	Schätzung der Ähnlichkeit zwischen zwei Lichter-Kandidaten C_1, C_2 zur Bestimmung der besten Fahrzeughypothese.	79
4.20	Sensorische Abdeckung des Fahrzeugumfelds durch Videokameras im Totenwinkel Bereich.	80
4.21	Konzeptioneller Ablauf der videobasierten Erkennung von Fahrzeughypothesen auf Nebenspuren.	81
4.22	Perspektivische Veränderung während der Annäherung eines Fahrzeugs auf der Nachbarspur.	81
4.23	Suche nach horizontalen Fahrzeugkanten anhand von Gradientenmerkmalen. .	82
4.24	Gittermodell mit 8×4 Elementen, Kreise beschreiben im Gitter eingetragene Merkmale.	83
4.25	Flussbasierte Weiterverfolgung durch Gittermodell: Merkmale werden zunächst der Gitterstruktur zugeordnet, jede Gitterzelle wird anhand des mittleren Flusses individuell weitergeführt, was zu einer Verformung des Gittermodells führt. .	84
4.26	Konzeptioneller Ablauf der Klassifikation segmentierter Hypothesen durch ansichtsbasierte Merkmale.	85
4.27	Entzerrung von Texturen durch isometrische Projektion anhand schräger Flächen.	86
4.28	Unterteilung eines Eingabebildes in 9 überlappende Subfenster.	87

4.29	Ablauf der Zustandsschätzung von Objekthypothesen über die Zeit mit einem MTT-System.	89
5.1	Konzeptioneller Aufbau des Weltmodells und des probabilistischen Netzes zur Interpretation der Fahrsituation.	94
5.2	Spezifikation der Frame- und Slot-Konzepte in UML-Notation.	98
5.3	Die eindeutige Benennung von Klassen und Instanzen erlaubt die Unterscheidung und spätere Identifikation der Elemente in der Modellbasis und im Weltmodell; in der unteren Zeile ist die formale Darstellung von zusammengehörender Frame-Slot-Klassen sowie Frame-Slot-Instanzen gezeigt.	99
5.4	Ablauf der wissensbasierten Interpretation mit Trennung der Online- und Offline-Arbeitsschritte.	99
5.5	Schematische Beschreibung des Informationsfluss zum Aufbau des Weltmodells aus Hintergrundwissen und Umfeldinformationen und zur Bestimmung einer Situationsaussage.	100
5.6	Konzeptioneller Zusammenhang von Hintergrundwissen, Umfeldinformationen, informationsverarbeitender Algorithmen, der Modellbasis und der modellbasierten Inferenz zur Situationsinterpretation für FAS.	101
5.7	Aufbau des Weltmodells: Modellkonzepte der Modellbasis dienen als Schablonen für Modellinstanzen im Weltmodell; das Weltmodell ist abhängig von der Objektliste der Umfelderkennung.	103
5.8	Annäherungssituation innerhalb einer Menge von Fahrzeugen.	104
5.9	Relevante Attribute einer Annäherungssituation an ein vorausfahrendes Fahrzeug.	104
5.10	Graphisches Modell zur Bestimmung des Kollisionsrisikos gegenüber einem weiteren Objekt entsprechend zur Abbildung 5.9.	105
5.11	Relationale Darstellung der Klassen und Attribute zur Beschreibung der Fahrsituation; fett geschriebene Attribute stellen Relationen auf Fremdklassen her.	106
5.12	Formale Darstellung der Modellbasis als Klassenstruktur zur Beschreibung der Fahrsituation.	107
5.13	Abgeleitetes OOBN für eine Situation mit drei Fahrzeugen im Weltmodell.	107
5.14	Erweiterung der Fahrsituation in relationaler Darstellung zur Betrachtung des Einflusses des Fahrers.	108
6.1	Versuchsfahrzeug „CoCar“ und dessen sensorische Abdeckung des Fahrzeugumfelds durch verschiedene, zum Teil redundante Sensoren.	112
6.2	PMD- und Videodaten aus Sicht von „CoCar“; oben: Videodaten nach Linsenverzerrung mit Überlagerung der PMD-Daten; unten links: Tiefendaten der PMD-Kamera in Falschfarbendarstellung; unten rechts: Amplitudendaten der PMD-Daten.	113
6.3	Versuchsfahrzeug „SegIT“ und dessen sensorische Abdeckung des Fahrzeugumfelds.	114

6.4	Tiefendaten dreier unterschiedlicher Messsysteme: ToF PMD 1k-S Kamera, Stereo SGM-Verfahren, Velodyne HDL-32E Laserscanner.	115
6.5	Querender Fußgänger in etwa 10 m Abstand zur Fahrzeughinterachse; Videodaten sind mit Tiefendaten der PMD 1k-S Kamera in Falschfarbendarstellung überlagert.	116
6.6	Birnenförmiges Ausleuchtungsfeld der PMD-Kamera, gemessen an der Erstdetektion von Fußgängern; die Markierungen stellen die Erstdetektion in den unterschiedlichen Entfernung dar.	116
6.7	Messobjekt „FZI Zwerg“ mit verschiedenen Reflektanzeigenschaften: Original, Reflektorweste, dunkle Jacke.	117
6.8	Auswertung des Einflusses der Reflektanz bei Annäherungsfahrten an „FZI Zwerg“.	118
6.9	Auswertung des Einflusses der Reflektanz bei Annäherungsfahrten an „FZI Zwerg“ mit Reflektorweste.	119
6.10	Auswertung des Einflusses der Reflektanz bei Annäherungsfahrten an „FZI Zwerg“ mit dunkler Jacke.	120
6.11	Rangierendes Fahrzeug: Nur Radkappen, Kennzeichen und Fahrzeuglichter können durch die Photomischdetektor (PMD)-Kamera detektiert werden.	121
6.12	Von rechts nach links querender Fußgänger in unterschiedlichen Distanzen vor dem Fahrzeug: 15 m und 30 m.	122
6.13	Fehlerhafte Messungen durch Überstrahlung bei Richtungswegweisern auf der Autobahn; die Messdaten entsprechen an dieser Stelle im Kamerabild keinem tatsächlichen Objekt.	123
6.14	Störung der Distanzmessung des vorausfahrenden Fahrzeugs (blaues Rechteck) während Hindurchfahrt unter Richtungswegweiser (grünes Rechteck).	124
6.15	Messgüte der PMD-Sensorik bei direktem Sonnenlicht; die Messqualität ist durch das Sonnenlicht nicht beeinflusst.	124
6.16	Messverhalten der PMD-Sensorik bei Störung durch fremde Infrarotlicht (IR)-Beleuchtungseinheit; bei aktiviertem Strahler werden beeinflusste Bildpunkte als ungültig markiert und unterdrückt.	125
6.17	Beispiele von Segmentierungsergebnissen der Objektsegmentierung durch Regionenwachstum (rote Boxen), projiziert in das Videobild (blaue Boxen).	126
6.18	Fehlerhafte Objekthypothesen durch Über- und Untersegmentierung.	126
6.19	Untersegmentierung zweier eng benachbarter Objekte bei Annäherung eines Fußgängers an einen Van.	127
6.20	Hypothesengenerierung bei Verdeckung nahe beieinander stehender Objekte, 0,5 m (links), 1,0 m (mitte) und 2,0 m x-Abstand (rechts); Bildausschnitt zeigt Zeitpunkt der jeweils ersten gültigen Trennung.	127
6.21	Verifikation von Objekten durch temporales Filter auf der Autobahn; Verlauf der Position der Objekte auf Bodenebene projiziert und mit Eigenbewegung kompensiert.	128
6.22	Verifikation von Objekten durch temporales Filter in der Stadt; Verlauf der Position der Objekte auf Bodenebene projiziert und mit Eigenbewegung kompensiert.	128

6.23	Segmentierte Objekthypothesen während einer Folgefahrt über Entfernung (Ordinate) und Zeit (Abszisse) aufgetragen; statische Objekte zeichnen sich als Linien von links oben nach rechts unten ab.	129
6.24	Zeitlicher Verlauf der verfolgten, als gültig befundenen Objekthypothesen zu Abbildung 6.23; unterschiedliche Farben markieren verschiedene Tracks.	130
6.25	Segmentierung von Flächen in einer Sequenz von Tiefendaten mit $25k$ -Bildpunkten; blaue Flächen stellen Objekte dar, die grüne Fläche die Bodenebene.	131
6.26	Segmentierung von Flächen in einer Sequenz von Tiefendaten mit $3k$ -Bildpunkten; blaue Flächen stellen Objekte dar, die grüne Fläche die Bodenebene.	131
6.27	Links: Histogrammbasierte Objektsegmentierung auf semi-dichten Disparitätsdaten; rechts: Fusion von Radar-Reflexpunkten zur Verbesserung der Segmentierungsleistung [Teu+10].	132
6.28	Entfernungsabhängige Detektionsrate von Objekten, gemessen mit und ohne Verwendung von Radar-Reflexpunkten.	133
6.29	Beispiele zur Erkennung von Fahrzeugen in Autobahn-, Stadt- und Bundesstraßen-Sequenzen aus Videodaten.	134
6.30	Entfernungsabhängige Erkennungsrate von Fahrzeugen auf Autobahn-, Bundesstraße- und Stadt-Sequenzen.	134
6.31	Erkennung von Fahrzeughypothesen anhand der Rücklichter (gelbes Rechteck) und des Bremslichts (magenta Rechteck) bei Dämmerung und bei Nacht.	135
6.32	Fehler bei der Erkennung von Fahrzeughypothesen bei Nacht durch reflektierende Baustellenmarkierungen.	136
6.33	Sequenz der Erkennung von Fahrzeugen auf Nebenspuren.	136
6.34	Erkennungsrate von Fahrzeugen auf Nebenspuren nach Sequenzen aufgeteilt.	137
6.35	Nicht symmetrische Ansicht eines Fahrzeuges aufgrund unterschiedlichem Erscheinungsbild der Fahrzeuglichter.	137
6.36	Entwicklung des Attributs d_y für ein kreuzendes Objekt; die Distanz d_y ist negativ, während das Objekt den Fahrweg kreuzt.	139
6.37	Diskretisierung mittels linearer Interpolation; $p_i(x)$ stellt die Wahrscheinlichkeit des Zustands i zum kontinuierlichen Wert x dar; Rauscheffekte in x lassen sich durch die Interpolation approximieren [Kuh11].	140
6.38	Ausschnitt aus Objektorientiertem Bayes'schen Netz (OOBN) zur Bestimmung der Risiko-Relation zwischen einem Objekt und dem Ego-Fahrzeug.	140
6.39	Fahrsituation und Ergebnisse des Experiments mit einem statischen Objekt auf dem Fahrweg des Ego-Fahrzeugs.	141
6.40	Ego-Fahrzeug folgt zunächst einem langsamen Objekt und weicht schließlich aus.	143
6.41	Das Experiment „Kreuzendes Hindernis“ aus Sicht eines externen Beobachters.	143
6.42	Ego-Fahrzeug nähert sich einem kreuzenden Objekt und passiert dieses; links: Auswertung ohne Betrachtung der lateralen Geschwindigkeit, rechts: Ergebnisse bei Betrachtung der lateralen Geschwindigkeit v_y	144

6.43	Risikoentwicklung bei der Annäherung an ein Objekt; zwischen den Zeitpunkten t_1 und t_2 ist ein deutlicher Anstieg des geschätzten Risikos zu erkennen, da sich das Fahrzeug nach Kursänderung direkt auf das Objekt zubewegt.	145
6.44	Risikoentwicklung bei der Annäherung an zwei Objekte; die Entwicklung der Einzelrisiken pro Hindernis ist deutlich zu erkennen, das Gesamtrisiko ist im Verlauf der Situation deutlich höher.	146
6.45	Graphische Benutzeroberfläche auf „SegIT“-Plattform zur Darstellung der Risikoschätzung; das Kollisionsrisiko ist als Balken am linken Rand dargestellt. . . .	147
6.46	Im Projekt EFA2014 wurden mit Hilfe von Vorausschau-Informationen Spurwechsel-Hinweise für energieeffiziente Assistenz berechnet [Sch+12].	147
6.47	Beschreibung der Verarbeitungsschritte des Gesamtsystems; Segmentierung, Projektion, Klassifikation, Tracking und Interpretation bauen aufeinander auf. .	148
6.48	Ergebnisse der Inferenz des Attributs <i>IsBlocking</i> unter Berücksichtigung des Fahrschlauchs des Ego-Fahrzeugs.	148
6.49	Ausschnitt aus dem Zeitschritt-Modell in OOBN-Darstellung zur Bestimmung der Spursicherheit (SPS) und Zielplausibilität (PLA).	149
6.50	Schematischer zeitlicher Verlauf der Zielplausibilität (PLA) über den Verlauf eines potentiellen Objektes; das Attribut ist mit zeitlicher Entwicklung ausgelegt [Len13].	150
6.51	Auswahl des MRT anhand des prädizierten Fahrschlauchs; Objekte mit niedriger SPS sind mit einem blauen Rechteck umrandet, Objekte mit gültiger PLA sind mit einem roten Rechteck umrandet.	150
6.52	Unterschiedliche Fälle für Folgeregler im $s-t$ -Diagramm [Len13]	151
6.53	Evaluation des Folgereglers mit Fahrzeug „CoCar“ bei einer Folgefahrt und einschierendem Fahrzeug [Len13].	152

Glossar

Assistenzfunktion Software-/Hardwarebaustein, welches ein Algorithmus zur Berechnung einer Assistenzaufgabe ausführt. 45, 46, 72

Detektion Erkennung eines gesuchten Musters in einer Signalinformation. 4, 34, 38, 59, 125

Evidenz Eine als bekannt und damit im Sinne des Glaubengrad als sicher angenommene Information. 20, 21, 46, 103, 107, 108

Fahrsituation Kombination des Zustands des Eigenfahrzeugs, der Zustände relevanter Objekte im Fahrzeugumfeld sowie die Veränderung dieser Zustände. 2–7, 19, 20, 28, 29, 45, 46, 48, 49, 51–57, 59, 93, 97, 100–104, 106–109, 111, 138, 142, 143, 155–157

Fahrzeugkoordinatensystem Fahrzeugfestes Koordinatensystem mit Referenzpunkt am vordersten Punkt des Fahrzeugs in der Mitte der Stoßstange auf der Bodenebene (nach DIN 70000). 16, 30, 31, 62, 74, 83, 138, 147

generisch Eine generische Methode ist für unterschiedliche Situationen verwendbar und insbesondere erweiterbar; gleichwohl sollte eine Wiederverwendbarkeit von bereits modellierten Klassen und Relationen ermöglicht werden. 5, 46–48, 52–54, 56, 100, 102, 109, 155

Histogramm Stochastische Zählung einer Signalinformation anhand einer oder mehrerer Merkmalsklassen des Signals. 37, 64, 66, 87, 88, 138

Kamerakoordinatensystem Kamerafestes Koordinatensystem mit Referenzpunkt im optischen Zentrum der Kamera. 16, 31, 113, 147

Klassifikation Zuordnung eines Wertes aus einer vordefinierten, endlichen Menge von Werten zu einer Signalinformation. 4, 6, 14, 30, 32, 33, 36, 38, 59, 60, 64, 84–88, 91, 155, 156

lateral Messung quer zu einer Längsachse. 4, 10, 48, 72, 104, 105, 117, 138, 142–146, 148, 149, 153

longitudinal Messung entlang einer Längsachse. 10, 37, 104, 105, 132, 138, 142, 144–146, 148, 153

modular Eine modulare Methode erlaubt die Trennung von Einzelkomponenten zur Steigerung der Wiederverwendbarkeit, Erweiterbarkeit und Validierbarkeit. 46, 54, 55, 100, 103, 109, 157

- Objekthypothese** Durch Detektion gebildete Information, welche mit einer gewissen Zuverlässigkeit mit einem Objekt der realen Welt korreliert. 29, 32–36, 56, 59, 64, 72, 88, 93, 125, 126, 128, 129, 132, 147, 148, 155, 156
- Punktwolke** (engl. Point Cloud) Menge von Punkten in einem 3D-Raum, welche eine Szene oder ein Objekt beschreiben. 12, 14, 38
- robust** Eine robuste Methode verhält sich konsistent bezüglich der Variation innerhalb von Situationen, beispielsweise gegenüber einer variablen Anzahl von Objekten im Umfeld des Fahrzeugs sowie gegenüber Attribut-Veränderungen. 46–48, 53, 54, 56, 100, 103, 109, 111, 146
- Segmentierung** Zerlegung und Gruppierung einer Signalinformation in eine Menge separater, bedeutungsvoller Bereiche. 6, 12, 14, 30, 32, 36–38, 59, 60, 62–64, 91, 131, 156
- Tiefenbild** (engl. Range Image) Bildartige Struktur, welche Entfernungswerte anstelle der Grau-/Farbwerte der Szene besitzt. 4, 11–14, 35–38, 56, 62, 63, 132
- Track** Zeitliche Zustandsfolge eines beobachteten Objektes. 34, 39, 40, 129
- Zustandsschätzung** Bestimmung von Merkmalen eines Systems aus einer Datensequenz unter Berücksichtigung von Unsicherheit und Rauschen. 4, 6, 30, 38, 41, 48, 59, 61, 88, 91, 155, 156

Abkürzungsverzeichnis

- ABS** Anti-Blockier-System. 1
- ACC** Abstandsregelnder Tempomat (engl.: Adaptive Cruise Control). 1, 14, 35, 64, 149
- AEB** Autonomes Notbremssystem (engl. Autonomous Emergency Braking). 1, 14
- AL** Aussagenlogik (engl. Propositional Logic). 19
- BBox** Umschließende Box (engl. Bounding Box). 64, 75, 79, 83–85, 125, 135
- BN** Bayes'sches Netz (engl. Bayesian Network). 18, 20–27, 47, 48, 53, 97, 103, 105, 159
- BP** Belief Propagation. 20, 100
- CA** Modell konstanter Beschleunigung (engl. Constant Acceleration Model). 89
- CAN** Controller Area Network. 112, 113, 147
- CBR** Fallbasiertes Schließen (engl. Case-based Reasoning). 49
- CMS** Kollisionsminderungssystem (engl. Collision Mitigation System). 47
- CPD** Bedingte Wahrscheinlichkeitsverteilung (engl. Conditional Probability Distribution). 21, 22, 24, 26, 47, 95–100, 106–108, 141
- CV** Modell konstanter Geschwindigkeit (engl. Constant Velocity Model). 89
- DBN** Dynamisches Bayes'sches Netz. 22, 23
- DGPS** Globales Positionierungssystem mit Differentialsignal (engl. Differential Global Positioning System). 113
- DL** Beschreibungslogik (engl. Description Logic). 49–51, 53
- ECU** Steuergerät (engl. Electronic Control Unit). 1, 111, 113
- ER** Entity-Relationship. 24, 25, 27, 51, 54, 106, 156
- ESP** Elektronisches Stabilitätsprogramm. 1
- FAS** Fahrerassistenzsystem (engl. Advanced Driver Assistance System). 1–4, 9, 10, 14–16, 29, 30, 46, 49, 54–57, 155, 159
- FOL** Logik erster Ordnung (engl. First-Order Logic). 19, 24, 25, 49, 51, 53, 54, 106

- FOPL** Wahrscheinlichkeitsbasierte Logik erster Ordnung (engl. First-Order Probabilistic Logic). 7, 9, 20, 27, 52–54, 56, 157
- FP** Falsch Positiv Klassifikationsrate (engl. False Positive). 123, 134–136
- GNN** Global Nearest Neighbor. 40
- GPS** Globales Positionierungssystem (engl. Global Positioning System). 38, 113, 114
- HDR** High-Dynamic Range. 10, 136
- HOG** Histogram of Oriented Gradients. 87, 88
- IMM** Interacting Multiple Model. 35, 42, 43, 45, 59, 90, 91
- IMU** Inertiale Messeinheit (engl. Inertial Measurement Unit). 113, 114
- IR** Infrarotlicht. 11, 115, 117, 121, 123–126, 163
- JPDA** Joint Probabilistic Data Association. 35, 40
- LKS** Spurhalteassistent (engl. Lane Keeping System). 1, 38, 72
- LPHT** Line Pattern Hough Transformation. 65, 67, 68, 70, 71, 161
- MCMC** Markov chain Monte Carlo. 20, 22
- MHT** Multi-Hypothesen Tracking. 40
- MLN** Markov Logik Netz. 24, 25
- MRT** Wichtigstes Zielobjekt (engl. most relevant target). 148–151
- MTT** Multi-Target Tracking. 39, 42, 59, 88
- OOBN** Objektorientiertes Bayes'sches Netz. 22–25, 27, 48, 97, 103, 107–109, 141, 150, 157, 162
- OPRM** Objektorientiertes Probabilistisches Relationales Modell (engl. Object Oriented Probabilistic Relational Model). 25, 27, 28, 104
- OPRML** Objektorientierte Probabilistische Relationale Modellierungssprache (engl. Object Oriented Probabilistic Relational Modelling Language). 27, 52, 54, 56, 93, 95, 98, 102, 107, 109, 155
- PDA** Probabilistic Data Association. 35, 40
- PL** Wahrscheinlichkeitsbasierte Logik (engl. Probabilistic Logic). 19
- PLA** Zielplausibilität. 150

-
- PMD** Photomischdetektor (engl. Photonic Mixing Device). 36, 61, 62, 64, 112, 113, 115, 116, 121, 123–125, 147, 152, 163
- PRM** Probabilistisches Relationales Modell (engl. Probabilistic Relational Model). 25–27
- ROI** Region of Interest. 32, 60, 62, 64, 74–76, 85, 86, 88, 138, 147
- SA** Situationsanalyse (engl. Situation Assessment). 5, 45–48, 51–57, 61, 92, 93, 101, 102, 111, 112, 146, 155–157
- SHT** Standard Hough Transformation. 67, 68
- SPS** Spursicherheit. 149, 150
- SRG** Seeded Region Growing. 63
- SRL** Statistisches Relationales Lernen (engl. Statistical Relational Learning). 5, 24, 25
- SVM** Support Vektor Maschine. 36, 59, 84, 87, 88, 133, 155
- ToF** Lichtlaufzeit (engl. Time of Flight). 10, 11, 16, 35, 36, 59, 61, 65, 112–117, 125, 131, 155
- TP** Richtig Positiv Klassifikationsrate (engl. True Positive). 134, 135
- TSR** Geschwindigkeitszeichenerkennung (engl. Traffic Sign Recognition). 72
- TTC** Kollisionszeit (engl. Time To Collision). 47, 109, 142, 143
- TTR** Reaktionsreservezeit (engl. Time To React). 47, 109
- URG** Unseeded Region Growing. 63

Quellennachweise

Veröffentlichungen

- [Koh+10] Ralf Kohlhaas, Thomas Schamm, Marcus Strand und J. Marius Zöllner. “Flexible fusion of 2D images and 3D time of flight data using plane segmentation”. In: *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2010 IEEE Conference on*. 2010, S. 242–247.
- [Koh+13] Ralf Kohlhaas, Thomas Schamm, Dominik Lenk und J. Marius Zöllner. “Towards driving autonomously : Autonomous Cruise Control in Urban Environments”. In: *2013 IEEE Intelligent Vehicles Symposium, Workshop on Environment Perception and Navigation for Intelligent Vehicles*. 2013.
- [Nat+08] Koba Natroshvili, Michael Schmid, Martin Stephan, Andreas Stiegler und Thomas Schamm. “Real time pedestrian detection by fusing PMD and CMOS cameras”. In: *Intelligent Vehicles Symposium, 2008 IEEE*. 2008, S. 925–929.
- [Sch+07] Thomas Schamm, Stefan Vacek, Koba Natroshvili, J. Marius Zöllner und Rüdiger Dillmann. “Hinderniserkennung und -verfolgung mit einer PMD-Kamera im Automobil”. In: *Autonome Mobile Systeme*. Springer, 2007, S. 219–225.
- [Sch+08] Thomas Schamm, J. Marius Zöllner, Stefan Vacek, Joachim Schröder und Rüdiger Dillmann. “Obstacle detection with a Photonic Mixing Device-camera in autonomous vehicles”. In: *International Journal of Intelligent Systems Technologies and Applications*. Bd. 5. Nov. 2008, S. 315–324.
- [Sch+09] T. Schamm, M. Strand, T. Gump, R. Kohlhaas, J. M. Zöllner und R. Dillmann. “Vision and ToF-based driving assistance for a personal transporter”. In: *Advanced Robotics, 2009. ICAR 2009. International Conference on*. Juni 2009, S. 1–6.
- [Sch+12] Thomas Schamm, Ralf Kohlhaas, Marcus Strand und J. Marius Zöllner. *Energieeffizientes Fahren 2014 (EFA2014) : Teilvorhaben: Erkennung, Modellierung und Interpretation semistatischer und dynamischer Objekte*. Schlussbericht BMBF-Verbundprojekt EFA2014, Förderkennzeichen 13N10134, 01067014. TIB Hannover, 2012.
- [SCZ10] Thomas Schamm, Christoph von Carlowitz und J. Marius Zöllner. “On-Road Vehicle Detection during Dusk and at Night”. In: *Intelligent Vehicles Symposium (IV), 2010 IEEE*. 2010, S. 418–423.
- [SRZ10] Thomas Schamm, Arne Rönnaun und J. Marius Zöllner. “Fast 3D Time of Flight Data Segmentation using the U-V-Histogram Approach”. In: *Intelligent Robotics and Applications*. Lecture Notes in Computer Science. Heidelberg: Springer, Nov. 2010, S. 618–629.

- [Str+09] Marcus Strand, Thomas Schamm, Ahmed Benazza, Thilo Kerscher, J. Marius Zöllner und Rüdiger Dillmann. “Control of an autonomous personal transporter towards moving targets”. In: *Proceedings of the IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*. Tokyo, Japan, Nov. 2009.
- [SZ11] Thomas Schamm und J. Marius Zöllner. “A Model-based Approach to Probabilistic Situation Assessment for Driver Assistance Systems”. In: *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on*. Okt. 2011.
- [Teu+10] Michael Teutsch, Thomas Heger, Thomas Schamm und J. Marius Zöllner. “3D-Segmentation of Traffic Environments with U/V-Disparity supported by Radar-given Masterpoints”. In: *Intelligent Vehicles Symposium (IV), 2010 IEEE*. Juni 2010, S. 787–792.
- [TS11] Michael Teutsch und Thomas Schamm. “Fast Line and Object Segmentation in Noisy and Cluttered Environments using Relative Connectivity”. In: *IPCV'11 - The 2011 International Conference on Image Processing, Computer Vision, and Pattern Recognition*. Hrsg. von Hamid R. Arabnia, Leonidas Deligiannidis und Gerald Schäfer. Bd. 1. 2011, S. 517–523.
- [Vac+07] Stefan Vacek, Thomas Schamm, Joachim Schröder, J. Marius Zöllner und Rüdiger Dillmann. “Collision avoidance for cognitive automobiles using a 3D PMD camera”. In: *Intelligent Autonomous Vehicles*. Toulouse, France, 2007.

Studentische Arbeiten

- [Car09] Christoph von Carlowitz. “Detektion von Fahrzeugen anhand ihrer Beleuchtung bei Dämmerung und Dunkelheit”. Studienarbeit. Universität Karlsruhe (TH), 2009.
- [Die09] Jürgen Diez. “Kopfpositionserkennung im Fahrzeuginnenraum mit Hilfe von 3D-TOF-Kameras”. Diplomarbeit. Universität Karlsruhe (TH), 2009.
- [Kuh11] Florian Kuhnt. “Probabilistische Kollisionsprädiktion für Segway-Transporter”. Diplomarbeit. Karlsruher Institut für Technologie, 2011.
- [Len13] Dominik Lenk. “Entwicklung einer Fahrzeug-Längsführung für automatisiertes Fahren im urbanen Straßenverkehr”. Masterarbeit. Karlsruher Institut für Technologie, 2013.
- [Wen09] Ruoqing Wen. “Kamerabasierte Erkennung und Klassifizierung von vorausfahrenden Fahrzeugen”. Diplomarbeit. Universität Karlsruhe (TH), 2009.
- [Zhu10] Ning Zhu. “2D-Kamerabasierte Fahrzeugerkennung im Toten-Winkel-Bereich”. Diplomarbeit. Karlsruher Institut für Technologie, 2010.

Literatur

- [AB94] R. Adams und L. Bischof. "Seeded Region Growing". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.6 (1994), S. 641–647.
- [AZ03] N. Apostoloff und A. Zelinsky. "Robust vision based lane tracking using multiple cues and particle filtering". In: *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*. 2003, S. 558–563.
- [Ban04] Olav Bangsø. "Object Oriented Bayesian Networks". Diss. Aalborg University, März 2004.
- [Bär+11] T. Bär, D. Nienhüser, R. Kohlhaas und J. Marius Zöllner. "Probabilistic driving style determination by means of a situation based analysis of the vehicle data". In: *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on*. 2011, S. 1698–1703.
- [Bär+13] Tobias Bär, Denys Linke, Dennis Nienhüser und J. Marius Zöllner. "Seen and Missed Traffic Objects: A Traffic Object Specific Awareness Estimation". In: *2013 IEEE Intelligent Vehicles Symposium, Workshop on Human Interaction with Intelligent Vehicles*. 2013.
- [BCS10] M. Brännström, E. Coelingh und J. Sjöberg. "Model-Based Threat Assessment for Avoiding Arbitrary Vehicle Collisions". In: *Intelligent Transportation Systems, IEEE Transactions on* 11.3 (2010), S. 658–669. ISSN: 1524-9050.
- [BDH09] Yaakov Bar-Shalom, Fred Daum und Jim Huang. "The Probabilistic Data Association Filter". In: *Control Systems, IEEE* 29.6 (Dez. 2009), S. 82–100. ISSN: 1066-033X.
- [BF08] A. Barth und U. Franke. "Where will the oncoming vehicle be the next second?" In: *Intelligent Vehicles Symposium, 2008 IEEE*. Juni 2008, S. 1068–1073.
- [BGD10] S. Brechtel, T. Gindele und R. Dillmann. "Recursive importance sampling for efficient grid-based occupancy filtering in dynamic environments". In: *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. 2010, S. 3932–3938.
- [BP99] S.S. Blackman und R. Popull. *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.
- [Bun12] Bundesanstalt für Straßenwesen. *Rechtsfolgen zunehmender Fahrzeugautomatisierung*. 2012.
- [BZF08] N. Barnes, A. Zelinsky und L.S. Fletcher. "Real-Time Speed Sign Detection Using the Radial Symmetry Detector". In: *Intelligent Transportation Systems, IEEE Transactions on* 9.2 (2008), S. 322–332. ISSN: 1524-9050.
- [Con13] Continental. *Continental Automotive - Lane Departure Warning / Lane Keeping System (LDW / LKS)*. 2013. URL: http://www.conti-online.com/generator/www/de/en/continental/automotive/themes/commercial_vehicles/safety/adas/ldw/ldw_lks_en.html (besucht am 04. 11. 2013).

- [CS00] N. Cristianini und J. Shawe-Taylor. *An Introduction to Support Vector Machines: And Other Kernel-Based Learning Methods*. Cambridge University Press, März 2000. ISBN: 9780521780193.
- [DAR07] DARPA. *DARPA Urban Challenge*. 2007. URL: <http://archive.darpa.mil/grandchallenge/> (besucht am 05. 11. 2013).
- [DBB02] Ismail Dagli, Michael Brost und Gabi Breuel. "Action Recognition And Prediction For Driver Assistance Systems Using Dynamic Belief Networks". In: *In Proceedings of the Conference on Agent Technologies, Infrastructures, Tools, and Applications for E-Services*. Springer-Verlag, 2002, S. 179–194.
- [Deb13] Tom Debus. *Mercedes S 500 Intelligent Drive: Lass die Finger weg, jetzt fahre ich!* Frankfurter Allgemeine Zeitung. Sep. 2013. URL: <http://www.faz.net/-h7b-7haqq> (besucht am 05. 11. 2013).
- [DR07] Pedro Domingos und Matthew Richardson. "Introduction to Statistical Relational Learning". In: Hrsg. von L. Getoor und B. Taskar. MIT Press, 2007. Kap. Markov Logic: A Unifying Framework for Statistical Relational Learning, S. 339–371.
- [DT05] N. Dalal und B. Triggs. "Histograms of oriented gradients for human detection". In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Bd. 1. 2005, S. 886–893.
- [EG08] M. Enzweiler und D. Gavrilu. "Monocular Pedestrian Detection: Survey and Experiments". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. Dez. 2008, S. 2179–2195.
- [EM07] Björn Elias und Petri Mähönen. "Pedestrian Recognition Based on 3D Image Data". In: *Proc. of the International Symposium on Industrial Electronics*. 2007.
- [Ess+09] A. Ess, K. Schindler, B. Leibe und L. van Gool. "Improved Multi-Person Tracking with Active Occlusion Handling". In: *Proceedings of the IEEE International Conference on Robotics and Automation*. Mai 2009.
- [Far+06] Basel Fardi, Jaroslav Dousa, Wanielik Gerd, Björn Elias und Alexander Barke. "Obstacle detection and pedestrian recognition using a 3D PMD camera". In: *Proceedings of Intelligent Vehicles Symposium, 2006 IEEE*. 2006, S. 225–230.
- [FB81] Martin A. Fischler und Robert C. Bolles. "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography". In: *Commun. ACM* 24.6 (Juni 1981), S. 381–395. ISSN: 0001-0782.
- [GEM09] Stefan Gehrig, Felix Eberli und Thomas Meyer. "A Real-Time Low-Power Stereo Vision Engine Using Semi-Global Matching". In: *Computer Vision Systems*. Hrsg. von Mario Fritz, Bernt Schiele und Justus Piater. Bd. 5815. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2009, S. 134–143. ISBN: 978-3-642-04666-7.

- [Göt11] Isabel Götzenberger. “Klassifikation von Bildausschnitten zur Videobasierten Fußgängererkennung für Fahrerassistenzsysteme”. Diplomarbeit. Fachhochschule Kaiserslautern, 2011.
- [Gra+13] R. Graf, H. Deusch, M. Fritzsche und K. Dietmayer. “A learning concept for behavior prediction in traffic situations”. In: *Intelligent Vehicles Symposium (IV), 2013 IEEE*. 2013, S. 672–677.
- [GT07] Lise Getoor und Ben Taskar, Hrsg. *Introduction to Statistical Relational Learning*. The MIT Press, Nov. 2007.
- [Gum+09] T. Gump, D. Nienhuser, R. Liebig und J.M. Zollner. “Recognition and tracking of temporary lanes in motorway construction sites”. In: *Intelligent Vehicles Symposium, 2009 IEEE*. Juni 2009, S. 305–310.
- [HAV11] HAVEit consortium. *HAVEit - Highly Automated Vehicles for Intelligent Transport*. 2011. URL: <http://haveit-eu.org/> (besucht am 04. 11. 2013).
- [HLU05] Zhencheng Hu, F. Lamosa und K. Uchimura. “A complete U-V-disparity study for stereovision based 3D driving environment analysis”. In: *3-D Digital Imaging and Modeling, 2005. 3DIM 2005. Fifth International Conference on* (Juni 2005), S. 204–211. ISSN: 1550-6185.
- [Hof06] Christian Hoffmann. “Fahrzeugdetektion durch Fusion monoskopischer Video-merkmale”. Diss. Universität Karlsruhe (TH), 2006.
- [How10] Catherine Howard. “Knowledge Representation and Reasoning for a Model-Based Approach to Higher Level Information Fusion”. Diss. University of South Australia, 2010.
- [HS09] Catherine Howard und Markus Stumptner. “Automated compilation of Object-Oriented Probabilistic Relational Models”. In: *International Journal of Approximate Reasoning* (Mai 2009).
- [HS88] C. Harris und M. Stephens. “A combined corner and edge detector”. In: *Proceedings of the 4th Alvey Vision Conference*. 1988, S. 147–151.
- [HSK06] J. Hillenbrand, A. M. Spieker und K. Kroschel. “A Multilevel Collision Mitigation Approach—Its Situation Assessment, Decision Making, and Performance Tradeoffs”. In: *Intelligent Transportation Systems, IEEE Transactions on* 7.4 (Dez. 2006), S. 528–540. ISSN: 1524-9050.
- [HZW11] Michael Hülsen, J. Marius Zöllner und Christian Weiss. “Traffic intersection situation description ontology for advanced driver assistance”. In: *Intelligent Vehicles Symposium (IV), 2011 IEEE*. Juni 2011, S. 993–999.
- [ibe13] ibeo Automotive Systems GmbH. *Ibeo Automotive Systems GmbH*. 2013. URL: <http://www.ibeo-as.com/> (besucht am 04. 11. 2013).
- [Jäh05] Bernd Jähne. *Digitale Bildverarbeitung*. ger. 6., überarb. und erw. Aufl. Springer, 2005. ISBN: 3-540-24999-0,978-3-540-24999-3.

- [Jor06] Rüdiger Jordan. “Objekthypothesen für Sicherheitsfunktionen auf Basis eines Radar-Sensors”. Diss. Universität Karlsruhe (TH), 2006.
- [Kas13] Dietmar Kasper. “Erkennung von Fahrmanövern mit objektorientierten Bayes-Netzen in Autobahnszenarien”. Diss. Mathematisch-Naturwissenschaftliche Fakultät der Eberhard Karls Universität Tübingen, 2013.
- [KP97] Daphne Koller und Avi Pfeffer. “Object-Oriented Bayesian Networks”. In: *Proceedings of the Thirteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-97)*. Aug. 1997, S. 302–313.
- [Lat+05] Andreas D. Lattner, Jan D. Gehrke, Ingo J. Timm und Otthein Herzog. “A Knowledge-based Approach to Behavior Decision in Intelligent Vehicles”. In: *Proceedings of Intelligent Vehicles Symposium, 2005 IEEE*. 2005.
- [LCJ08] Kathryn B. Laskey, Paulo C. G. Costa und T. Janssen. “Probabilistic Ontologies for Knowledge Fusion”. In: *In Proceedings of the Eleventh International Conference of the Society of Information Fusion (FUSION 2008)*. Cologne, Germany, Juli 2008.
- [LJT01] Zheng Lin, Jesse Jin und Hugues Talbot. “Unseeded Region Growing for 3D Image Segmentation”. In: *Selected papers from Pan-Sydney Workshop on Visual Information Processing*. Hrsg. von Peter Eades und Jesse Jin. Sydney, Australia: ACS, 2001.
- [Mar10] John Markoff. *Google Cars Drive Themselves, in Traffic*. The New York Times. Okt. 2010. URL: <http://www.nytimes.com/2010/10/10/science/10google.html> (besucht am 04. 11. 2013).
- [Min74] M. Minsky. *A Framework for Representing Knowledge*. MIT-AI Laboratory Memo 306. Juni 1974.
- [Moo10] Joris M. Mooij. “libDAI: A Free and Open Source C++ Library for Discrete Approximate Inference in Graphical Models”. In: *Journal of Machine Learning Research* 11 (Aug. 2010), S. 2169–2173.
- [MR07] Brian Milch und Stuart Russell. “First-order probabilistic languages: Into the unknown”. In: *Proceedings of the 16th international conference on inductive logic programming*. 2007, S. 10–24.
- [MS13] Frank Moosmann und Christoph Stiller. “Joint self-localization and tracking of generic objects in 3D range data”. In: *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. 2013, S. 1146–1152.
- [Mur02] Kevin Patrick Murphy. “Dynamic Bayesian Networks: Representation, Inference and Learning”. Diss. University of California, Berkeley, 2002.
- [NGZ09] D. Nienhüser, T. Gumpp und J.M. Zöllner. “A Situation context aware Dempster-Shafer fusion of digital maps and a road sign recognition system”. In: *Intelligent Vehicles Symposium, 2009 IEEE*. 2009, S. 1401–1406.
- [Pea88] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann series in representation and reasoning. Morgan Kaufmann Publishers, 1988. ISBN: 9781558604797.

- [PF10] D. Pfeiffer und U. Franke. “Efficient representation of traffic scenes by means of dynamic stixels”. In: *Intelligent Vehicles Symposium (IV), 2010 IEEE*. Juni 2010, S. 217–224.
- [Rei79] D. Reid. “An algorithm for tracking multiple targets”. In: *Automatic Control, IEEE Transactions on* 24.6 (Dez. 1979), S. 843–854.
- [RHO09] Elizabeth Rendon-Velez, Imre Horváth und Eliab Z. Opiyo. “Progress with Situation Assessment and Risk Prediction in Advanced Driver Assistance Systems: a Survey”. In: *Proceedings of the 16th ITS World Congress, Sweden* (Sep. 2009).
- [RN09] Stuart J. Russell und Peter Norvig. *Artificial Intelligence: A Modern Approach*. 3. Aufl. Prentice Hall, 2009.
- [SBM06a] Zehang Sun, G. Bebis und R. Miller. “Monocular precrash vehicle detection: features and classifiers”. In: *Image Processing, IEEE Transactions on* 15.7 (Juli 2006), S. 2019–2034. ISSN: 1057-7149.
- [SBM06b] Zehang Sun, G. Bebis und R. Miller. “On-Road Vehicle Detection: A Review”. In: *Transactions on Pattern Analysis and Machine Intelligence* 28.5 (Mai 2006), S. 694–711. ISSN: 0162-8828.
- [Sch09] Jörg Schneider. “Modellierung und Erkennung von Fahrsituationen und Fahrmanövern für sicherheitsrelevante Fahrerassistenzsysteme”. Diss. Fakultät für Elektrotechnik und Informationstechnik, Technische Universität Chemnitz, 2009.
- [Sch+97] Rudolf Schwarte, Zhanping Xu, Horst-Guenther Heinol, Joachim Olk und Bernd Buxbaum. “New optical four-quadrant phase detector integrated into a photogate array for small and precise 3D cameras”. In: *Three-Dimensional Image Capture*. Bd. 3023. 1997, S. 119–128.
- [Seg13] Segway Inc. *Segway PT Models*. 2013. URL: <http://www.segway.com> (besucht am 04. 11. 2013).
- [Sin70] R.A. Singer. “Estimating Optimal Tracking Filter Performance for Manned Maneuvering Targets”. In: *Aerospace and Electronic Systems, IEEE Transactions on* AES-6.4 (Juli 1970), S. 473–483. ISSN: 0018-9251.
- [SJ09] Lin-Lin Shen und Zhen Ji. “Gabor Wavelet Selection and {SVM} Classification for Object Recognition”. In: *Acta Automatica Sinica* 35.4 (2009), S. 350–355. ISSN: 1874-1029.
- [SSZ12] Florian Steinhardt, Marcus Strand und J. Marius Zöllner. “Autonomous Navigation of a Personal Transporter within Moving Human Groups Using Reactive Control”. In: *Intelligent Autonomous Systems 12*. Hrsg. von Sukhan Lee, Hyung-suck Cho, Kwang-Joon Yoon und Jangmyung Lee. Bd. 193. *Advances in Intelligent Systems and Computing*. Springer Berlin Heidelberg, 2012, S. 81–90. ISBN: 978-3-642-33925-7.
- [ST94] J. Shi und C. Tomasi. “Good features to track”. In: *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*. 1994, S. 593–600.

- [SW11] R. Schubert und G. Wanielik. “A Unified Bayesian Approach for Object and Situation Assessment”. In: *Intelligent Transportation Systems Magazine, IEEE* 3.2 (2011), S. 6–19. ISSN: 1939-1390.
- [Vac08] Stefan Vacek. “Videogestützte Umfelderkennung zur Interpretation von Verkehrssituationen für kognitive Automobile”. Diss. Universität Karlsruhe, Deutschland: Fakultät für Informatik, Mai 2008.
- [Vel13] Velodyne. *Velodyne Lidar*. 2013. URL: <http://velodynelidar.com> (besucht am 04. 11. 2013).
- [VJ01] P. Viola und M. Jones. “Rapid object detection using a boosted cascade of simple features”. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2001* 1.C (2001), S. 511–518.
- [WB95] G. Welch und G. Bishop. *An Introduction to the Kalman Filter*. Techn. Ber. TR 95-041. University of North Carolina at Chapel Hill, Department of Computer Science, 1995.
- [Yip94] R.K.K. Yip. “Line Patterns Hough Transform for line segment detection”. In: *TENCON '94. IEEE Region 10's Ninth Annual International Conference. Theme: Frontiers of Computer Technology. Proceedings of 1994*. Bd. 1. 1994, S. 319–323.
- [ZEI11] ZEIT ONLINE. *Der Wagen denkt, der Fahrer lenkt*. 2011. URL: <http://www.zeit.de/auto/2011-08/fahrerassistenzsystem-geschwindigkeit> (besucht am 04. 11. 2013).