



Institut für Informatik der  
Technischen Universität München



# Tracking von Verkehrsteilnehmern im Kontext von Multisensorsystemen

Rudi Lindl

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Arndt Bode

Prüfer der Dissertation:

1. Univ.-Prof. Gudrun J. Klinker, Ph.D.
2. Univ.-Prof. Dr.-Ing. Darius Burschka

Die Dissertation wurde am 05.08.2008 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 20.04.2009 angenommen.



# DANKSAGUNG

---

An erster Stelle möchte ich mich ganz herzlich bei Frau Prof. Gudrun Klinker bedanken, die mich über die letzten drei Jahre in fachlichen, als auch überfachlichen Belangen motiviert, unterstützt und beraten hat. Die unzähligen Gespräche und Diskussionen haben maßgeblich zum Gelingen dieser Dissertation beigetragen und sind in dieser Form und diesem Umfang für eine Industriepromotion sicherlich außergewöhnlich.

Großer Dank gebührt auch Prof. Darius Burschka für seine sofortige Bereitschaft das Zweitgutachten für diese Arbeit zu übernehmen.

Ich danke meinem Kollegen, Kommilitonen und guten Freund Leonhard Walchshäusl für seine Diskussionen und Ratschläge zum Thema Fahrumfelderfassung. In den gemeinsamen Jahren der Dissertation teilten wir deren Höhepunkte und motivierten uns gegenseitig in den schwierigen Zeiten.

Der BMW Forschung und Technik GmbH und im Speziellen Dr. Klaus Gresser bin ich zu Dank verpflichtet, da sie mir diese Arbeit überhaupt erst ermöglicht haben. Meine Betreuer bei BMW, Dr. Marc Walessa und Katrin Vogel, haben durch ihre Diskussionen, Expertise und wertvolles Feedback einen großen Anteil am Gelingen dieser Arbeit. Ich danke allen Kollegen aus dem Sensorik Team für die kollegiale Unterstützung und die tolle Arbeitsatmosphäre. Besonderer Dank gebührt aber auch den Diplomanden und Werkstudenten Florian Homm, Philipp Kemmeter und Georg Kuschk, die durch ihre Arbeiten zu dieser Dissertation beigetragen haben.

Vielen Dank allen, die Kommentare, Anregungen und Korrekturen zu der schriftlichen Ausarbeitung gegeben haben: Gudrun Klinker, Marc Walessa, Katrin Vogel, Adam Burg, Moritz Spohr und Leonhard Walchshäusl.

Ich danke insbesondere meiner Frau Kathrin für ihr Verständnis, ihre Geduld und ihre Liebe. Sie erinnerte mich zusammen mit unserer Tochter Johanna immer wieder daran, dass es neben der Dissertation auch noch ein normales Leben gibt und bot mir Rückhalt und Motivation.

Nicht zuletzt danke ich meinen Eltern. Ohne Ihre Unterstützung wäre weder mein Studium noch diese Arbeit möglich gewesen.

Vielen Dank!



# KURZFASSUNG

---

Fahrerassistenzsysteme, wie das Antiblockiersystem oder die dynamische Fahrzeugstabilisierung unterstützen den Fahrer in schwierigen Situationen und erhöhen somit die allgemeine Verkehrssicherheit. Durch die Forderung nach weiter steigender Sicherheit werden heutzutage zunehmend Systeme entwickelt, die durch Sensoren das Umfeld des Fahrzeuges erfassen, den Fahrer durch Informationen unterstützen oder sogar aktiv in das Fahrgeschehen eingreifen. Hierfür und insbesondere für letzteres ist allerdings ein Höchstmaß an Zuverlässigkeit und Genauigkeit gefordert. Um diesen Ansprüchen zu genügen, wird in dieser Arbeit eine Fusion von mehreren Sensoren betrachtet. Hierzu wird ein System vorgestellt, das durch die Kombination von Radar, Laserscanner und bildgebender Sensorik eine präzise und zuverlässige Erfassung der Positionen, Geschwindigkeiten und Raummaße von Fußgängern und Fahrzeugen ermöglicht. Eine Evaluierung des Mehrwertes dieser multisensoriellen Fahrumfelderfassung im Vergleich zu konventionellen Ein-Sensor-Systemen erfolgt unter verschiedenen Umweltbedingungen und in realen Verkehrsszenarien.



# INHALTSVERZEICHNIS

---

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Zielsetzung . . . . .	2
1.3	Randbedingungen . . . . .	2
1.4	Gliederung der Arbeit . . . . .	2
<b>2</b>	<b>Grundlagen</b>	<b>5</b>
2.1	Fahrerassistenzsysteme . . . . .	5
2.1.1	Überblick . . . . .	6
2.1.2	Aktive Gefahrenbremsung . . . . .	8
2.2	Automotive Sensorik . . . . .	9
2.2.1	Funktionsweise und Verbreitung . . . . .	10
2.2.2	Taxonomie und Vergleich . . . . .	11
2.3	Sensorfusion . . . . .	13
2.3.1	Vorteile . . . . .	14
2.3.2	Fusionsmodelle . . . . .	15
2.3.3	Fusionstaxonomien . . . . .	16
2.3.4	Fusion im Kontext der Fahrumfelderfassung . . . . .	18
2.4	Versuchsträger . . . . .	18
2.4.1	Aufbau und Sensorsystem . . . . .	19
2.4.2	Koordinatensysteme . . . . .	20
2.5	Systemüberblick . . . . .	23
<b>3</b>	<b>Sensor Kalibrierung</b>	<b>25</b>
3.1	Kalibrierung von multispektralen Kameras . . . . .	25
3.1.1	Projektives Kameramodell . . . . .	26
3.1.2	Kalibrierungsverfahren . . . . .	27
3.1.3	Bewertung . . . . .	31
3.2	Referenzdatengewinnung . . . . .	32
3.2.1	Inertialplattform mit Carrier Phase Differential GPS . . . . .	32
3.2.2	Fehlerabschätzung . . . . .	34
3.3	Temporale Kalibrierung . . . . .	38
3.3.1	Entfernungsgebende Sensorik . . . . .	39
3.3.2	Bildgebende Sensorik . . . . .	40
3.4	Zusammenfassung . . . . .	43

<b>4</b>	<b>Beobachtungsmodelle</b>	<b>45</b>
4.1	Objektbeschreibung . . . . .	46
4.1.1	Fahrzeuge . . . . .	46
4.1.2	Fußgänger . . . . .	48
4.2	Laserscanner . . . . .	50
4.2.1	Laserscannerdatensegmentierung . . . . .	50
4.2.2	Merkmalsgewinnung . . . . .	54
4.2.3	Modell für Fußgänger . . . . .	59
4.2.4	Modell für Fahrzeuge . . . . .	60
4.2.5	Anpassungen an den Partikel-Filter . . . . .	62
4.2.6	Bewertung . . . . .	62
4.2.7	Zusammenfassung . . . . .	63
4.3	Radarsensorik . . . . .	64
4.3.1	Bestimmung der Messunsicherheit . . . . .	65
4.3.2	Fahrzeuge . . . . .	65
4.3.3	Fußgänger . . . . .	68
4.3.4	Anpassungen an den Partikel-Filter . . . . .	68
4.3.5	Zusammenfassung . . . . .	68
4.4	Konturmodell Video . . . . .	69
4.4.1	Kopfkontur für Fußgänger . . . . .	70
4.4.2	Fahrzeugmodell . . . . .	72
4.5	Zusammenfassung . . . . .	74
<b>5</b>	<b>Tracking</b>	<b>75</b>
5.1	Grundlagen der Objektverfolgung . . . . .	75
5.1.1	Mathematische Formulierung . . . . .	76
5.1.2	Optimale Algorithmen . . . . .	78
5.2	Klasse der Kalman-Filter . . . . .	78
5.2.1	Standard Kalman-Filter . . . . .	79
5.2.2	Extended-Kalman-Filter . . . . .	81
5.2.3	Unscented-Kalman-Filter . . . . .	82
5.2.4	Fusion . . . . .	83
5.2.5	Kompensation von Latenzzeiten . . . . .	84
5.2.6	Vergleich EKF und UKF . . . . .	85
5.2.7	Zusammenfassung . . . . .	88
5.3	Partikel-Filter . . . . .	89
5.3.1	Kernel-Partikel-Filter . . . . .	90
5.3.2	Annealed-Partikel-Filter . . . . .	91
5.3.3	Fusion . . . . .	92
5.3.4	Bewertung der Partikel-Filter . . . . .	94
5.3.5	Fazit . . . . .	103
5.4	Zusammenfassung . . . . .	103
<b>6</b>	<b>Visualisierung von heterogenen Sensordaten</b>	<b>105</b>
6.1	Grundlagen der Visualisierung . . . . .	106
6.1.1	Daten und Informationsfluss . . . . .	106
6.1.2	Designrichtlinien . . . . .	107



6.2	Visualisierung des Informationsflusses . . . . .	109
6.2.1	Sensorverbauorte und -erfassungsbereiche . . . . .	109
6.2.2	Rohdaten und vorverarbeitete Sensordaten . . . . .	110
6.2.3	Interpretierte Daten . . . . .	114
6.2.4	Zeitlicher Verlauf der Messungen . . . . .	115
6.2.5	Navigation und Interaktion . . . . .	116
6.2.6	Werkzeuge der Implementierung . . . . .	120
6.3	Präsentation und Kommunikation . . . . .	120
6.3.1	Konventionelle Ausgabemedien . . . . .	121
6.3.2	Erweiterte Realität . . . . .	122
6.4	Fahrbahnerkennung mittels Flottenbewegung . . . . .	132
6.5	Zusammenfassung und Diskussion . . . . .	135
<b>7</b>	<b>Systembewertung</b>	<b>139</b>
7.1	Referenzdatengewinnung . . . . .	140
7.1.1	Semiautomatische Referenzdatenerzeugung . . . . .	141
7.1.2	Algorithmusgestützte Referenzdatenerzeugung . . . . .	145
7.1.3	Referenzsensor . . . . .	146
7.1.4	Hybride Referenzdatenerzeugung . . . . .	148
7.1.5	Zusammenfassung und Bewertung . . . . .	148
7.2	Evaluierung . . . . .	149
7.2.1	Metriken . . . . .	149
7.2.2	Bewertung des Trackingsystems . . . . .	151
7.2.3	Bewertung des Gesamtsystems . . . . .	158
7.2.4	Charakteristische Fehlerfälle . . . . .	165
7.2.5	Laufzeitanalyse . . . . .	167
7.3	Zusammenfassung . . . . .	169
<b>8</b>	<b>Fazit</b>	<b>171</b>
8.1	Zusammenfassung . . . . .	171
8.2	Wissenschaftlicher Beitrag . . . . .	174
8.3	Ausblick . . . . .	174
	<b>Literaturverzeichnis</b>	<b>177</b>



# EINLEITUNG

---

## 1.1 Motivation

Seit Ende der 70er-Jahre wurden, angefangen vom Bremsen-Antiblockiersystem (ABS) bis zur dynamischen Fahrzeugstabilisierung (DSC), überwiegend Fahrerassistenzsysteme eingeführt, die eine Steigerung der allgemeinen Sicherheit in schwierigen Fahrsituationen adressieren oder den Komfort erhöhen. Bedingt durch den technischen Fortschritt in der Sensortechnologie und der Forderung nach mehr Sicherheit im Straßenverkehr werden heutzutage zunehmend Systeme entwickelt, die das Umfeld des Fahrzeuges interpretieren, den Fahrer informieren und eventuell sogar aktiv in das Fahrgeschehen eingreifen.

Als erstes kommerzielles Beispiel für Systeme, die das Verkehrsgeschehen aktiv erfassen, wurde 1995 *Adaptive Cruise Control (ACC)* von Mitsubishi in Japan eingeführt [MS05]. Dieses Assistenzsystem hält eine vom Fahrer vorgegebene Geschwindigkeit und passt sich an die Geschwindigkeit eines vorausfahrenden Fahrzeuges an. Weitere aktuell erhältliche Systeme warnen vor Kollisionen und unterstützen den Fahrer in einem Bremsmanöver, indem sie gegebenenfalls dessen unvollständige oder zu zaghafte Ausführung ergänzen. Gemeinsame Grundlage dieser Systeme ist eine maschinelle Erfassung der aktuellen Verkehrssituation durch zusätzliche Sensorik. Zukünftige Assistenzfunktionen sind auf ein möglichst umfassendes Abbild des Fahrzeugumfeldes angewiesen. So benötigen Kreuzungs- oder Spurwechselassistenten zum Beispiel einen großen Erfassungsbereich, um Verkehrsteilnehmer in Seitenstraßen oder der Nachbarspur zu erkennen. Systeme zum Schutz von Fußgängern müssen hingegen zwischen unterschiedlichen Objektklassen unterscheiden und entsprechend reagieren.

Aus diesen beiden Beispielen wird deutlich, dass die Anforderungen, die an ein Fahrumfelderfassungssystem gestellt werden, nicht nur in eine Dimension wachsen. So müssen für zukünftige Fahrerassistenzsysteme eine größere Anzahl an unterschiedlichen Objektklassen (Fahrzeuge, Fußgänger oder Motorradfahrer), in komplexeren Szenarien (Autobahn, Landstraße, Kreuzung oder Stadt) und unter schwierigeren Umweltbedingungen (Tag, Nacht, Regen, Schnee oder Nebel) erkannt und verfolgt werden. Oftmals ist hierbei zusätzlich ein Höchstmaß an Zuverlässigkeit und Genauigkeit gefordert, da ein fehlerhaftes Verhalten sowohl die Akzeptanz beim Fahrer als auch die allgemeine Verkehrssicherheit negativ beeinflussen würden.

Aktuelle Sensorsysteme, wie sie zum Beispiel dem ACC zugrunde liegen, können diesen immensen Ansprüchen nur bedingt genügen. So ist das ACC, welches im Allgemeinen auf einem Radarsensor basiert, für den Einsatz auf einer Autobahn oder auf großen innerstädtischen Ausfallstraßen optimiert und erkennt nur Fahrzeuge, die sich in der selben Spur wie das Eigenfahrzeug befinden. Fußgänger, Stadt- oder Kreuzungsszenarien werden hingegen nicht adressiert.

Eine Kombination von mehreren Sensoren verspricht an dieser Stelle einen beträchtlichen Vorteil gegenüber Systemen, die auf einem einzelnen Sensor beruhen. So könnte durch die Sensorfusion zum Einen ein größerer Erfassungsbereich modelliert werden und zum Anderen die Präzision und Zuverlässigkeit der Fahrumfelderfassung in komplexen Szenarien und unter widrigen Umweltbedingungen gesteigert werden. Allerdings ist die Wahl der zu fusionierenden Sensorik und die Entwicklung geeigneter Algorithmen - insbesondere unter automotiven Aspekten - noch immer Bestandteil aktueller Forschung.

### 1.2 Zielsetzung

Ziel dieser Dissertation ist es, durch die Kombination (Fusion) von Radar, Laserscanner und bildgebender Sensorik eine präzise und zuverlässige Erfassung der Positionen, Geschwindigkeiten und Raummaße von Fußgängern und Fahrzeugen zu realisieren und den Mehrwert dieser multisensoriellen Fahrumfelderfassung im Vergleich zu konventionellen Ein-Sensor-Systemen zu untersuchen. Die Zielsetzung ist dabei stets unter automotiven Aspekten wie Echtzeitfähigkeit und unter realistischen Umweltbedingungen zu betrachten. Zusätzlich soll evaluiert werden, welche Verfahren der Objektverfolgung unter diesen Randbedingungen zu bevorzugen sind.

### 1.3 Randbedingungen

Diese Dissertation war Teil eines Projektes bei der BMW Forschung und Technik GmbH, das im Rahmen des europäischen Forschungsprojektes COMPOSE<sup>1</sup> zukünftige Fahrerassistenzsysteme mit hohen Ansprüchen an die unterliegende Fahrumfelderfassung demonstrieren sollte. Erklärtes Ziel des Projektes war es, durch Kombination unterschiedlicher Sensorik eine präzise und zuverlässige Fahrumfelderfassung von Fußgängern und Fahrzeugen zu erreichen und den Mehrwert dieser Sensorfusion im Vergleich zu konventionellen Ein-Sensor-Systemen zu bewerten. Aufbauend auf dieser Sensorik wurde als Applikation eine autonome Notbremsung (siehe Abschnitt 2.1.2) realisiert. Der gesamte Projektumfang wurde im Wesentlichen von zwei Dissertationen abgedeckt. Die Dissertation von Leonhard Walchshäusl [Wal08] beschäftigt sich hauptsächlich mit der Klassifikation von Verkehrsteilnehmern, wohingegen diese Arbeit vor allem das Tracking von Verkehrsteilnehmern thematisiert. Die Implementierung des Gesamtsystems, die zugrunde liegende Softwarearchitektur, die Visualisierung von heterogenen Sensordaten sowie die Evaluierung in Bezug auf die Gesamtsystemleistung sind im Rahmen des Projektes in Kooperation mit Leonhard Walchshäusl entstanden. Aufgrund dieser engen Zusammenarbeit wurden die Kapitel 2, 6 sowie 7 in Abschnitten gemeinsam verfasst.

### 1.4 Gliederung der Arbeit

In Kapitel 2 werden grundlegende Definitionen und Begriffe der Fahrumfelderfassung und der Sensorfusion eingeführt. Es wird ein Überblick über aktuelle Fahrerassistenzsysteme gegeben und die Funktionsweise automotiver Sensorsysteme näher betrachtet. Abschließend wird die Sensorkonfiguration des in dieser Arbeit verwendeten Versuchsträgers beschrieben.

---

<sup>1</sup> COMPOSE (collision mitigation and protection of road users) ist ein Teilprojekt des EU-Förderprojektes PREVENT zur Erhöhung der Verkehrssicherheit.

Um Bildausschnitte der Kameras mit den Daten des Laserscanners oder der Radare zu verknüpfen, ist eine exakte Kalibrierung der einzelnen Sensoren erforderlich. Hierzu werden in Kapitel 3 Möglichkeiten dargestellt und diskutiert, wie zum Einen ein multispektrales Kamerasystem, bestehend aus Ferninfrarot- und Videokameras kalibriert und zum Anderen die Latenzzeiten von bild- und entfernungsgebender Sensorik gemessen werden können.

In Kapitel 4 wird beschrieben, wie Fußgänger und Fahrzeuge durch Laserscanner, Radar und Video Sensoren beobachtet werden können. Hierbei werden für den Laserscanner Erweiterungen der Vorverarbeitungsschritte vorgestellt.

Anschließend werden in Kapitel 5 verschiedene Kalman- und Partikelfiltervarianten in Hinblick auf automotive Aspekte betrachtet und Erweiterungen beschrieben, wie die Messungen mehrerer Sensoren fusioniert werden können.

Aufgrund der großen Datenmenge, die in einer multisensoriellen Fahrumfelderfassung verarbeitet werden müssen, kann die Entwicklung und Diagnose des Gesamtsystems schnell unübersichtlich und ineffektiv werden. In Kapitel 6 werden deshalb für diese Daten Visualisierungsmöglichkeiten vorgestellt, die den Systementwickler, unter anderem durch den Einsatz von erweiterter Realität, im Umsetzungsprozess unterstützen.

In Kapitel 7 werden Möglichkeiten vorgestellt und diskutiert, wie multisensorielle Referenzdaten effizient generiert werden können. Diese Daten dienen als Grundlage einer anschließenden Bewertung der Tracking- sowie Gesamtsystemleistung, die unter verschiedenen realistischen Umweltbedingungen den Mehrwert einer multisensoriellen Fahrumfelderfassung untersucht.

Abschließend werden in Kapitel 8.1 die wichtigsten Ergebnisse zusammengefasst und ein Ausblick hinsichtlich möglicher Erweiterungen und Verbesserungen gegeben.



# GRUNDLAGEN

---

Fahrumfelderfassungssysteme stellen zum Einen eine Grundlage für viele moderne Fahrerassistenzsysteme (FAS) dar. Zum Anderen sind sie auf automotive Sensorik angewiesen. In diesem Kapitel wird zunächst ein Überblick über aktuelle Fahrerassistenzsysteme gegeben und die Funktionsweise automotiver Sensorsysteme näher betrachtet. Anschließend werden wichtige theoretische und praktische Aspekte einer multisensoriellen Fahrumfelderfassung erläutert. Dazu zählen sowohl Sensorfusionskonzepte als auch formale Grundlagen. Abschließend wird die Sensorkonfiguration des in dieser Arbeit verwendeten Versuchsträgers beschrieben.

## 2.1 Fahrerassistenzsysteme

Die Vorläufer der Fahrerassistenzsysteme, wie zum Beispiel der elektrische Anlasser, das Automatikgetriebe oder die Servolenkung, hatten zum Ziel, die Führungsaufgabe des Fahrers durch den Entfall von Maschinistentätigkeiten zu erleichtern [Fre04]. Seit Ende der 70er-Jahre wurden, angefangen vom Bremsen-Antiblockiersystem (ABS) bis zur dynamischen Fahrzeugstabilisierung (DSC), überwiegend Systeme eingeführt, die eine Steigerung der allgemeinen Sicherheit in schwierigen Fahrsituationen adressieren. Diese Fahrerassistenzsysteme beruhen im Allgemeinen auf einer Überwachung von fahrzeuginternen Zustandsgrößen wie Lenkwinkel, Querbeschleunigung oder Raddrehzahlen. Bedingt durch den technischen Fortschritt in der Sensortechnologie werden heutzutage zunehmend Systeme entwickelt, die die Umgebung des Fahrzeuges bzw. das aktuelle Verkehrsgeschehen mit berücksichtigen. Hierdurch können zum Einen bereits bestehende Systeme erweitert werden und zum Anderen „vorausschauende“ Anwendungen entwickelt werden, die vor gefährlichen Situationen warnen oder gegebenenfalls Gegenmaßnahmen ergreifen. Als erstes Beispiel solcher Systeme wurde 1995 „Adaptive Cruise Control“ (ACC) von Mitsubishi in Japan eingeführt [MS05]. Hierbei wird zusätzlich zur eigenen Geschwindigkeit, der Abstand zu einem vorausfahrenden Fahrzeug geregelt. Neben diesen aktiven Sicherheitssystemen wird fortlaufend auch die passive Sicherheit im Fahrzeug verbessert. Dazu zählen Maßnahmen des Fahrzeugaufbaus, wie zum Beispiel Sicherheitsgurte, Knautschzonen, Überrollbügel, Kindersitze und Airbags.

Im Folgenden wird zunächst ein Überblick aktueller Fahrerassistenzsysteme gegeben, und es werden mögliche Einteilungen vorgestellt. Als ein Beispiel für ein zukünftiges FAS wird abschließend die aktive Gefahrenbremsung für Fußgänger und Fahrzeuge herausgegriffen und ausführlicher betrachtet. Dieses fungiert als Zielapplikation, um die Leistungsfähigkeit der im Rahmen dieser Arbeit entwickelten Fahrumfelderfassung an einem realen Fahrerassistenzsystem erlebbar zu machen.

### 2.1.1 Überblick

Obwohl die im vorherigen Abschnitt aufgeführten Beispiele nur einen kleinen Ausschnitt aktueller Fahrerassistenzsysteme darstellen, sind sie dennoch ein Beleg dafür, wie vielfältig diese Systeme sein können. Eine Möglichkeit, Fahrerassistenzsysteme zu ordnen ist die Einteilung hinsichtlich der Art und Weise ihrer Unterstützung [Fre04]:

**Systeme, die warnen oder informieren:** Durch visuelle, haptische oder akustische Signale wird der Fahrer gewarnt oder informiert. So wird er zum Beispiel beim unbeabsichtigten Verlassen der eigenen Spur durch LDW bzw. LCW (lane departure warning bzw. lane change warning) alarmiert. Darüber hinaus werden dem Fahrer durch Navigationssysteme (NAVI) oder Kommunikation zwischen mehreren Fahrzeugen (Car2Car) Informationen über die aktuelle Verkehrslage weitergegeben. Bei Nachtfahrten unterstützen ihn Nachtsichtsysteme wie beispielsweise Wärmebildkameras, mit deren Hilfe wärmeemittierende Objekte wie Menschen oder Tiere visuell hervorgehoben werden können, schon bevor sie vom Scheinwerfer des Fahrzeuges erfasst werden (NiVi = night vision).

**Systeme mit langsamer Reaktion, die führend unterstützen:** Bei dieser Art von FAS wird unter wohldefinierten Randbedingungen moderat in die Fahrzeugführung eingegriffen. So regelt der Einparkassistent (PA = park assist) beispielsweise die Querführung und lenkt das Fahrzeug automatisch in eine Parklücke. Im höheren Geschwindigkeitsbereich übernehmen FAS wie das ACC teilautonom die Längsführung des Fahrzeuges um den Abstand zum vorausfahrenden Fahrzeug zu regulieren. Ein weiteres System, das sowohl in die Längs- als auch in die Querführung des Fahrzeuges eingreift, ist der Spurwechselassistent (LCA = lane change assistance), der den Fahrer bei einem Spurwechselmanöver unterstützt. Für derartige Systeme gilt im Allgemeinen, dass sie in kritischen Situationen die Kontrolle wieder an den Fahrer übergeben.

**Systeme mit schneller Reaktion, die stabilisieren bzw. in das Fahrverhalten eingreifen:** In kritischen Fahrsituationen müssen Systeme wie das DSC oder ABS sehr schnell analysieren und gegebenenfalls geeignete Gegenmaßnahmen treffen, indem sie das Fahrzeug aktiv stabilisieren. Bremsassistenten (iBA = intelligent braking assistance) erkennen zum Beispiel anhand der Bewegungsgeschwindigkeit und -beschleunigung des Bremspedals die Intention des Fahrers für eine Vollbremsung und ergänzen gegebenenfalls deren unvollständige oder zu zaghafte Ausführung. Einen Schritt weiter gehen Systeme wie die aktive Gefahrenbremsung (AGB) oder die Kollisionsvermeidung (CA = collision avoidance), die auch ohne eine Initiierung des Fahrers intervenieren, um Unfälle zu vermeiden oder zumindest deren Schwere zu verringern.

Eine weitere Einteilung von Fahrerassistenzsystemen kann anhand der jeweiligen Funktionsausprägung in Komfort- und Sicherheitssysteme vorgenommen werden [Fre04, Naa04]. Allerdings ist dabei eine eindeutige Zuordnung für manche Systeme wie beispielsweise das ACC nicht möglich, da sie sowohl den Komfort, als auch die Sicherheit erhöhen. In Abbildung 2.1 ist für die oben beschriebenen Fahrerassistenzsysteme eine Einteilung sowohl nach Komfort- und Sicherheitssystem als auch nach ihrer Unterstützungscharakteristik dargestellt.

Für viele Fahrerassistenzsysteme ist der Zeitpunkt des Eingriffes charakteristisch. Deshalb kann hiermit eine weitere Einteilung von Fahrerassistenzsystemen erfolgen, wie sie in Abbildung 2.2 dargestellt ist [Fre04]: Die Mehrzahl an Systemen, wie z.B. die Navigation oder ACC, unterstützen den Fahrer zu einem Zeitpunkt, in dem sich das Fahrzeug in einer völlig unkritischen Fahrsituation befindet. Systeme wie beispielsweise ABS bzw. DSC intervenieren zu einem späteren Zeitpunkt, wenn sich das



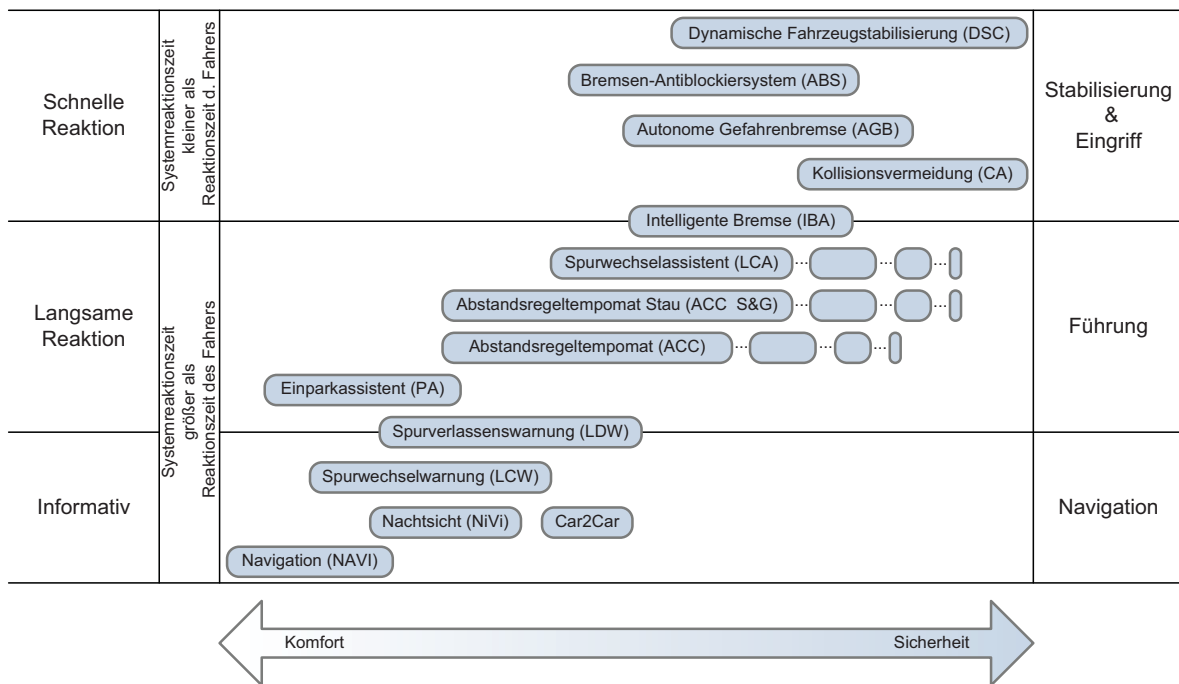


Abbildung 2.1: Einteilung von ausgewählten Fahrerassistenzsystemen nach Komfort und Sicherheit bzw. nach der Art ihrer Unterstützung [Fre04, Naa04]. Zukünftige Ausbaustufen einzelner Systeme sind durch leere Ovale gekennzeichnet. Das FAS „autonomes Fahren“ wurde aus Übersichtsgründen nicht eingezeichnet, da es den gesamten Diagrammbereich abdecken würde.

Fahrzeug bereits in einem instabilen Zustand (Blockieren der Räder bzw. Schleudern) befindet, und versuchen diesen zu stabilisieren. Ist ein Unfall hingegen nicht mehr zu vermeiden, kann die Energie des Zusammenstoßes durch Pre-Crash-Systeme wie die aktive Gefahrenbremse minimiert werden. Zum Unfallzeitpunkt werden die Insassen durch Airbags geschützt. Post-Crash-Systeme können nach einem Unfall in Anspruch genommen werden (z.B. Notruf).

Der perfekte Beifahrer ist eines der großen Ziele aktueller Fahrerassistenzsysteme. Nur wenn Notwendigkeit besteht, assistiert er dem Fahrer bezüglich des aktuellen Fahrgeschehens. Ähnlich anspruchsvoll ist die Vision vom autonomen Fahren, bei der die Fahraufgabe komplett dem Fahrzeug überlassen wird, und der Fahrer somit selbst zum „Beifahrer“ wird. Zwar könnte hiermit menschliches Versagen als einer der wesentlichen Unfallfaktoren eliminiert werden – allerdings müsste dieser Vorteil durch ungeklärte Haftungsfragen und hohe Ansprüche in Bezug auf Sensorverlässlichkeit und Systemleistung teuer erkauft werden. Zudem müsste das nichtdeterministische Verhalten anderer Verkehrsteilnehmer berücksichtigt werden. Unter wohldefinierten Randbedingungen und dem exzessiven Einsatz von Sensorik bzw. Rechenleistung ist autonomes Fahren im wissenschaftlichen Umfeld bereits Realität, wie die DARPA Urban Challenge [DAR07] eindrucksvoll belegt.

Bei aktiven Systemen, die autonom in das Fahrgeschehen eingreifen und somit die Verantwortung vom Fahrer auf das Fahrzeug übertragen, spielen rechtliche Aspekte eine große Rolle. Die zentrale zulasungsrechtliche Hürde liegt hier vor allem im Wiener Übereinkommen über den Straßenverkehr von 1968: „Jeder Fahrzeugführer muss unter allen Umständen sein Fahrzeug beherrschen“. Eine ausgie-

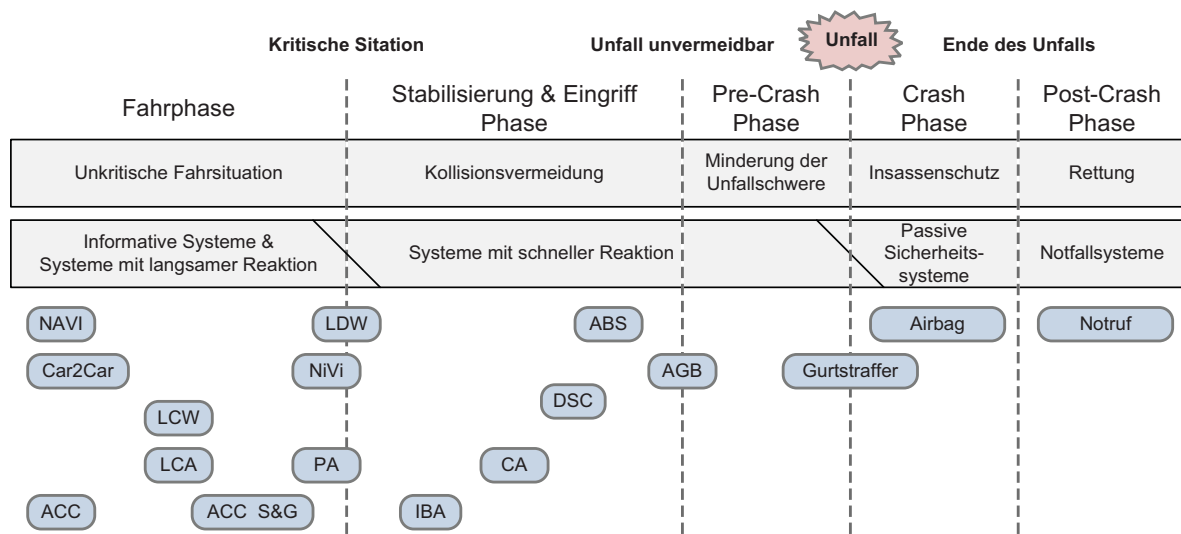


Abbildung 2.2: Einteilung von ausgewählten Fahrerassistenzsystemen nach dem Zeitpunkt ihres Eingriffs [Fre04].

bige Diskussion dieser Fragestellung ist in [vdHvW01, SG06] zu finden. Nicht zuletzt aus psychologischen Gründen, wie zum Beispiel dem Verlust der Fahrfreude und technischen Berührungsängsten, wird in absehbarer Zukunft vollautonomes Fahren im realen Straßenverkehr eine Vision bleiben. An teilautonomen Systemen, wie zum Beispiel der „aktiven Gefahrenbremsung“, wird jedoch schon heute intensiv geforscht.

### 2.1.2 Aktive Gefahrenbremsung

Die deutsche Verkehrsstatistik von 2006 [Sta06] belegt, dass 60% aller tödlichen Unfälle auf einen Zusammenstoß mit einem fremden Fahrzeug zurückzuführen sind. An weiteren 14% der Unfälle sind Fußgänger beteiligt. In beiden Szenarien ist menschliches Versagen mit Abstand die häufigste Unfallursache. Fahrerassistenzsysteme wie der Bremsassistent oder die aktive Gefahrenbremsung besitzen großes Potential, die Zahl der Unfälle zu reduzieren bzw. die Unfallfolgen zu mindern, da sie schneller und effizienter als der Fahrer reagieren können. Um trotz dieses autonomen Eingriffes, die Souveränität des Fahrers zu bewahren, dürfen derartige Systeme erst dann agieren, wenn für den Fahrer keine Möglichkeit mehr existiert, den Unfall durch Ausweichen oder Bremsen selbst zu vermeiden. Obwohl sich unter diesen Randbedingungen nicht mehr alle Unfälle verhindern lassen, ist es zumindest möglich, die Schwere der Verletzungen zu reduzieren, da ein Großteil der kinetischen Energie durch das Bremsmanöver absorbiert wird. An die AGB werden sehr hohe Anforderungen in Bezug auf die Häufigkeit von Fehlauflösungen gestellt, da schon eine einzige ungerechtfertigte autonome Vollbremsung sowohl die Akzeptanz für das System, als auch die allgemeine Verkehrssicherheit erheblich reduzieren würde.

Die Qualität der aktiven Gefahrenbremsung ist hauptsächlich von zwei Faktoren abhängig: Sie wird zum Einen von der Zuverlässigkeit und Genauigkeit der zugrundeliegenden Fahrumfelderfassung beeinflusst. Geisterziele, Fehlklassifikationen oder falsche Zustandsschätzungen können zu einer Fehleinschätzung und somit zu einer ungerechtfertigten Initiierung des Bremsmanövers führen. Zum An-

deren wird die Qualität des Gesamtsystems, selbst unter der Annahme einer perfekten Fahrumfelderfassung, durch die Interpretation der aktuellen Situation und der daraus abgeleiteten Entscheidung bestimmt. So ist zum Beispiel das Verhalten von Fußgängern nur sehr schwer zu präzisieren.

Mertz [MDG<sup>+</sup>05] versucht diese Unsicherheit zu vermindern, indem für sich bewegende Objekte mehrere zufällige Trajektorien erzeugt werden, die in Abhängigkeit von Mittelwert und Varianz der zugehörigen Zustandsgrößen verteilt sind. Aus dem Verhältnis von den Trajektorien eines Objektes, die zu einer Kollision führen, zur Gesamtheit der Trajektorien eines Objektes, kann eine Kollisionswahrscheinlichkeit berechnet werden. Bei diesem Ansatz wird somit versucht, durch eine Monte-Carlo-Simulation die reale Kollisionswahrscheinlichkeit zu approximieren. Allerdings ist die Qualität und der Rechenaufwand direkt proportional zur Größe der Stichprobe.

In Abhängigkeit von der Geschwindigkeit des Eigenfahrzeugs bzw. des Fußgängers und der maximalen Beschleunigungen definiert Fürstenberg [FS05] einen sogenannten „Bereich ohne Entrinnen“<sup>1</sup>. Befindet sich ein Fußgänger innerhalb dieses Bereiches, ist ein Unfall unvermeidbar. Kämpchen [Käm07] berechnet alle physikalisch möglichen Trajektorien für das Eigen- und ein Fremdfahrzeug. In der Berechnung werden keine einschränkenden Annahmen über die Geschwindigkeit und Beschleunigung der beteiligten Objekte getroffen. Zusätzlich wird die Orientierung, Länge und Breite der Fahrzeuge berücksichtigt, was eine Anwendung dieses Verfahrens auch in Kreuzungsszenarien ermöglicht. Erst wenn alle Pfade zu einer Kollision führen, wird eine autonome Gefahrenbremsung initiiert.

In den meisten Verfahren werden Verkehrsteilnehmer jedoch isoliert von ihrer Umwelt betrachtet. Broadhurst et al. [BBK04] haben eine Simulationsumgebung entwickelt, die das komplette Verkehrsgeschehen in komplexen Szenarien berücksichtigt und kritische Situationen erkennt. Neben Simulationen und regelbasierten Ansätzen, existieren in der Literatur Verfahren [JJ02, KJG04, LWP07], die eine probabilistische Situationsinterpretation verwenden. Unsicherheiten des Trackings und der Odometrie des Eigenfahrzeugs sowie deren zeitlicher Verlauf werden in einer gemeinsamen Wahrscheinlichkeitsdichtefunktion abgebildet, aus der die Wahrscheinlichkeit für einen Zusammenstoß abgeleitet werden kann.

Die aktive Gefahrenbremsung wird auch weiterhin Bestandteil von Forschungsaktivitäten bleiben, da sowohl die aktuellen Algorithmen der Situationsinterpretation, als auch die verfügbaren Sensoren nicht den hohen Qualitätsansprüchen genügen. Nicht nur für die AGB, sondern auch für viele andere Fahrerassistenzsysteme, ist die zugrundeliegende automotive Sensorik ein entscheidender Faktor.

## 2.2 Automotive Sensorik

Sensoren<sup>2</sup> oder Messfühler sind technische Bauteile, die spezifische Eigenschaften ihrer Umgebung, aufgrund physikalischer oder chemischer Effekte (z.B. Wärmestrahlung, Temperatur, Feuchtigkeit, Druck, Schall, Helligkeit oder Beschleunigung) erfassen und in weiterverarbeitbare Größen (z.B. elektrische Spannung) umwandeln können. In modernen Fahrzeugen sind eine Vielzahl von Sensoren in Gebrauch. Diese müssen hohen Anforderungen genügen. Sie sollten unempfindlich gegenüber elektromagnetischen Interferenzen sein, in einem Temperaturbereich von  $-40\text{ °C}$  bis zu  $+125\text{ °C}$  funktionieren, geringe Fehlerraten aufweisen, vibrations- und schmutzunempfindlich sein und vieles mehr. Fleming [FTSW01] kategorisiert automotive Sensoren anhand ihres Einsatzgebietes im Fahrzeug und

<sup>1</sup> engl. „Region of no escape“

<sup>2</sup> Sensor kommt von lateinisch sentire („fühlen“, „empfinden“).

ihrer Kontrollfunktion. Er unterscheidet dabei die Bereiche Antriebsstrang, Fahrwerk und Fahrzeugaufbau. Sensoren des Antriebsstrangs sind für den Energieverbrauch, die Fahrbarkeit und die Fahrzeugleistung zuständig. Als Beispiele seien Druck- und Drehzahlsensoren für die Motorsteuerung genannt. Sensoren des Fahrwerks adressieren Fahreigenschaften und Sicherheitsbelange. Typische Vertreter sind Raddrehzahl- und Giergeschwindigkeitssensoren für die Fahrzeugstabilisierung und das Antiblockiersystem. Sensoren des Fahrzeugaufbaus decken zum Einen Komfortfunktionen ab, wie beispielsweise Klimaregelung (Temperaturfühler, Feuchtigkeits- bzw. Regensensor) und automatisches Nachtfahrlicht (Helligkeitssensor). Zum Anderen bilden sie die Basis für Sicherheitsfunktionen, wie z.B. Airbag (Piezosensor), Einparkhilfe (Ultraschall) und Spurwechselwarnung (Radar, Laserscanner). Wird in dieser Arbeit der Begriff (*automotiver*) *Sensor* verwendet, bezieht sich die Bezeichnung auf die Kategorie der sicherheitsrelevanten Sensoren des Fahrzeugaufbaus und zwar speziell auf automotiv Sensorik, die in der Lage ist, Informationen über das Fahrumfeld zu erlangen.

Im Folgenden wird zunächst die grobe Funktionsweise von der, für diese Arbeit relevanten, automotiven Sensorik erläutert. Der nächste Abschnitt kategorisiert die Sensoren anhand ihrer Erfassungseigenschaften und stellt einen Vergleich hinsichtlich der Sensorperformanz auf.

### 2.2.1 Funktionsweise und Verbreitung

Dieser Abschnitt erläutert die Funktionsweise der für diese Arbeit relevanten automotiven Sensoren. Details zum Funktionsprinzip finden sich in den angefügten Referenzen.

**Videokamera** In den späten 60er Jahren wurden CCDs (charge coupled device) entwickelt, die heute günstig gefertigt werden können und als Videokamerasensor weit verbreitet sind. Grundsätzlich funktioniert eine einzelne Detektorzelle dieser elektronischen Bausteine ähnlich einer Photodiode. In jeder Zelle werden Photonen gesammelt, in elektrische Ladung umgewandelt und ausgelesen. Kamerasensoren auf CMOS (complementary metal oxide semiconductor) Basis ermöglichen ein schnelleres Auslesen der Bildinformationen was höhere Bildwiederholraten ermöglicht. Videokameras werden aktuell für Spur- und Verkehrszeichenerkennung im Automobil eingesetzt [Bor06].

**Infrarotkamera** Infrarotkamera arbeiten im nicht sichtbaren Lichtspektrum. Der Infrarotbereich liegt zwischen 780 nm und 14  $\mu\text{m}$ , wobei aufgrund atmosphärischer Absorption nicht der gesamte Bereich für einen sinnvollen Sensoreinsatz zur Verfügung steht [RC02]. Zwei unterschiedliche Technologien sind derzeit als Nachtsichtsystem im automotiven Einsatz:

FIR (far infrared) Sensoren arbeiten im Spektralbereich von 8  $\mu\text{m}$  bis 14  $\mu\text{m}$  und erkennen mittels ungekühlter Mikrobolometer<sup>3</sup>-Matrizen Wärmestrahlung, die von jedem Körper mit einer Temperatur von mehr als 0 K ausgestrahlt wird. Aktuell werden Auflösungen von bis zu 320  $\times$  240 Bildpunkten erreicht. Leistungsfähigere Systeme benötigen eine Kühlung und bleiben wegen ihrer hohen Kosten militärischen Anwendungen vorbehalten. Das erste automotiv FIR-System wurde 1999 von Cadillac und Raytheon eingeführt [KKNK03].

NIR (near infrared) Systeme arbeiten im nahen Infrarotspektrum von 780 nm bis 1  $\mu\text{m}$ . Sie müssen dabei die Szene aktiv im Infrarotbereich beleuchten, da die emittierte Energie der relevanten Objekte in diesem Spektrum nicht signifikant ist. Als Sensor fungieren herkömmliche CCDs auf Siliziumbasis. Vorteile von NIR-Systemen liegen in der für einen Menschen natürlicheren

---

<sup>3</sup> Mikrobolometer sind thermische Detektoren die Infrarotstrahlung durch wärmebedingte Änderung ihres elektrischen Widerstandes messen können.

Darstellung von Objekten. Im Gegensatz zu FIR-Systemen werden auch Straßenmarkierungen und Verkehrsschilder abgebildet. Nachteile sind die geringere Reichweite und das gesondert erforderliche Beleuchtungsmodul.

**RADAR** Das Akronym RADAR steht für „Radio Detection and Ranging“ und bezeichnet ein Verfahren zur Ortung von Objekten mittels elektromagnetischer Wellen im Gigahertzbereich (Funk). Dazu wird über eine Sendeeinheit ein Signal emittiert, welches besonders stark von metallischen Objekten zurückgestrahlt wird. Diese Reflektionen werden von einer Empfangseinheit detektiert. Über verschiedene Kriterien wie Laufzeit und Frequenzverschiebung des Signals können Informationen wie Entfernung, Winkel und radiale Geschwindigkeit über das erfasste Objekt gewonnen werden. Seit 1998 haben sich frequenzmodulierte Dauerstrichradare in Oberklassenfahrzeugen für adaptive Abstandshaltung etabliert. Nahbereichsradare mit 24 GHz haben in der Regel einen größeren Öffnungswinkel und werden für Spurwechselassistenten und Kollisionenwarnungen eingesetzt [GJE<sup>+</sup>04]. Aktuelle automotiv Radarentwicklungen zielen vermehrt in Richtung bildgebender Verfahren.

**LIDAR** Das LIDAR (Light Detection and Ranging) Verfahren ist mit der Radarsensorik verwandt. Statt elektromagnetischer Wellen im Gigahertzbereich werden Lichtimpulse emittiert und empfangen, die im ultravioletten, sichtbaren oder infraroten Spektralbereich liegen. Aufgrund der hohen Wellenfrequenz des Primärsignals sind die Auswirkungen des Dopplereffekts kaum messbar. Geschwindigkeiten können somit nicht direkt, sondern nur indirekt über zeitlich versetzte Messungen bestimmt werden. Jedoch ist eine wesentlich genauere Winkelauflösung als mit Radar-Systemen möglich, da Lichtsignale stärker fokussiert werden können als Mikrowellen. Im Gegensatz zur Radarsensorik fanden LIDAR-Systeme bisher nur vereinzelt Einzug ins Automobil. Gründe dafür sind die niedrige Schlechtwetterperformanz und die begrenzten Verbauorte im Fahrzeug. Details zur LIDAR-Technologie im Kontext von Fahrerassistenzsystemen findet sich in [RG05].

### 2.2.2 Taxonomie und Vergleich

Eine mögliche Untergliederung von automotiven Sensorsystemen kann hinsichtlich ihrer Funktionsweise erfolgen. Aktive Sensoren emittieren ein elektromagnetisches Signal, dessen Reflektionen an Objekten von einer Empfangseinheit aufgenommen wird. Über Signallaufzeiten und -veränderungen können Rückschlüsse auf Aspekte der Realität geschlossen werden. Typische Vertreter sind Ultraschall, Radarsensoren und Laserscanner. Passive Sensoren verfügen lediglich über eine Empfangseinheit. Mittels dieser wird die Eigenstrahlung von Objekten wie beispielsweise sichtbares Licht oder Wärme detektiert. Passive Sensoren sind beispielsweise Video- und Wärmebildkameras.

Eine weitere Unterteilungsmöglichkeit der Perzeptionssensorik kann hinsichtlich des Messresultats erfolgen. Rein bildgebende Sensoren erzeugen ein Abbild der Realität in dem strukturelle Objektdetails sichtbar werden. Rein entfernungsgebende Sensoren geben Aufschluss über statische und dynamische Zustandsdaten der erfassten Objekte, wie beispielsweise Position, Orientierung, Geschwindigkeit und Entfernung. Aktuelle Sensoren sind selten reine entfernungs- bzw. bildgebende Sensoren. Es liegen vielmehr beide Messresultate in unterschiedlicher Qualität vor und die Kategorieübergänge sind fließend (siehe Abbildung 2.3).

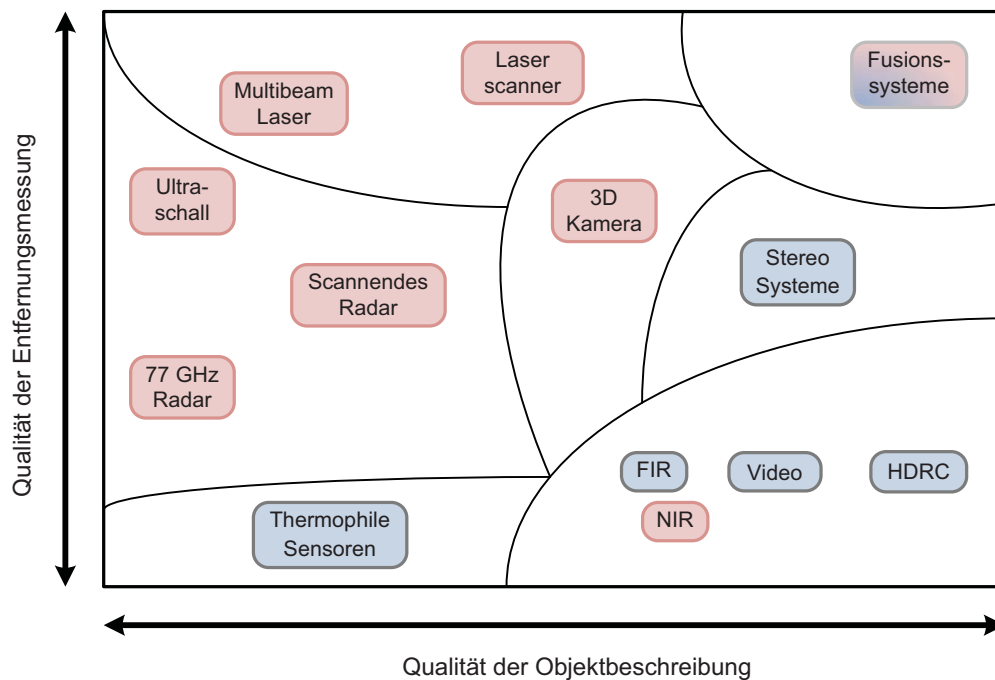


Abbildung 2.3: Kategorisierung automotiver Umfelderfassungssensorik hinsichtlich Bild- und Entfernungswiedergabe. Aktive Sensoren sind rot hinterlegt und passive Sensoren blau.

Videokameras oder HDRC (high dynamic range cameras) bilden beispielsweise vorwiegend die strukturellen Aspekte der Realität ab. Dennoch lassen sich Distanzen und somit Zustandsdaten mit begrenzter Genauigkeit schätzen, sofern Modellannahmen einfließen. Laserscanner, als vorwiegend entfernungsgebende Sensoren, liefern ebenfalls strukturelle Aspekte der Objekte, wie beispielsweise Oberflächenbeschaffenheiten. Ein idealer Perzeptionssensor würde sowohl die Struktur als auch die statischen und dynamischen Zustandsdaten von Umfeldobjekten in sehr hoher Qualität wiedergeben. Zurzeit vermögen lediglich Fusionsysteme, die über geeignete Algorithmik unterschiedliche Sensordaten kombinieren, dem idealen Sensor einen Schritt weit näher zu kommen.

Die Qualität, mit welcher die Realität abgebildet wird, entscheidet nicht allein über die Performanz eines Perzeptionssensors. Ein wesentliches Kriterium perzeptionstauglicher Sensorik ist eine hohe zeitliche Erfassungsrate. Speziell für Sensoren, die in sicherheitsrelevanten Fahrerassistenzsystemen zum Einsatz kommen, sind Messfrequenzen von 10 – 100 Hz zwingend notwendig. Weitere wichtige Gesichtspunkte im automotiven Kontext sind Sensorabmessungen, Kosten, Wettertauglichkeit und Reichweite (siehe Tabelle 2.1 für einen Sensorvergleich). Ein automotives Radar beispielsweise liefert im Vergleich zu einem Laserscanner eine begrenzte laterale Messgenauigkeit, verhält sich aber zugleich aufgrund des Messprinzips sehr robust bei Schlechtwetterbedingungen.



Tabelle 2.1: Vergleich automotiver Sensorsysteme zur Fahrumfelderfassung hinsichtlich ihrer Performanz ( $\oplus\oplus$  sehr gut,  $\oplus$  gut,  $\circ$  befriedigend,  $\ominus$  ausreichend,  $\ominus\ominus$  mangelhaft). Die betrachteten Stereosysteme können aus Videokameras/IR aufgebaut sein. Teile des Diagramms stammen aus [WGS<sup>+</sup>04]. Die restlichen Bewertungen basieren auf der Beurteilung des Autors.

	Video	Nahes IR	Fern IR	Radar	Scan Radar	3D Kamera	Stereo-systeme	Laser-scanner	Ultra-schall
max. Reichweite	100 m	80 m	>3 m	150 m	100 m	20 m	80 m	150 m	4 m
Longitudinale Genauigkeit	$\ominus$	$\ominus$	$\ominus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus\oplus$	$\oplus$
Laterale Genauigkeit	$\oplus\oplus$	$\oplus\oplus$	$\oplus\oplus$	$\ominus$	$\circ$	$\oplus$	$\oplus\oplus$	$\oplus$	$\circ$
Objektbeschreibung	$\oplus\oplus$	$\oplus\oplus$	$\oplus\oplus$	$\circ$	$\circ$	$\circ$	$\oplus\oplus$	$\oplus$	$\circ$
Schlechtwetterverhalten	$\ominus$	$\ominus$	$\ominus$	$\oplus\oplus$	$\oplus\oplus$	$\ominus$	$\ominus$	$\ominus$	$\oplus$
Verhalten bei Dunkelheit	$\ominus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\ominus/\oplus$	$\oplus$	$\oplus$
Verbaubarkeit	$\oplus$	$\circ$	$\oplus$	$\oplus\oplus$	$\oplus$	$\circ$	$\ominus$	$\circ$	$\oplus\oplus$
Kosten	$\oplus\oplus$	$\circ$	$\ominus$	$\circ$	$\circ$	$\circ$	$\circ/\oplus$ $\ominus$	$\ominus$	$\oplus\oplus$
Fahrerassistenzsysteme	LDW	NiVi	NiVi	ACC, ACC S&G	ACC, ACC S&G	iBA	ACC, iBA	ACC, iBA	PA

## 2.3 Sensorfusion

Augenblicklich verfügbare automotiv Sensoren, wie sie im vorherigen Abschnitt 2.2 beschrieben wurden, können einzeln betrachtet den hohen Anforderungen, die neuartige Fahrerassistenzsysteme, wie beispielsweise die AGB, an die zugrundeliegende Sensorik in Bezug auf Zuverlässigkeit, Verfügbarkeit und Genauigkeit stellen, nur schwer genügen. Deshalb wird durch Theorien und Methoden der Sensorfusion versucht, Informationen aus mehreren Sensoren derart gewinnbringend zu kombinieren (fusionieren), dass das gemeinsam erzielte Ergebnis in gewisser Weise besser ist, als das Ergebnis der einzelnen Sensoren. Obwohl im Allgemeinen davon ausgegangen werden kann, dass zusätzliche Informationen einen positiven Einfluss auf das Ergebnis haben, sollte stets bedacht werden, dass sie unter Umständen keinen signifikanten oder sogar einen negativen Beitrag zum Gesamtergebnis leisten [Das00].

Im Folgenden wird zunächst auf Vorteile eingegangen, die von einem Multisensorsystem erwartet werden können. Die nachfolgenden Abschnitte beschreiben unterschiedliche Architekturen und Einteilungsschemata. Abschließend wird die Projektlandschaft zu automotiven Fusionssystemen beleuchtet, sowie ausgewählte Beispiele gegeben, in denen Sensorfusion in einem FAS zur Anwendung gebracht wurde.

### 2.3.1 Vorteile

In Bezug auf folgende Punkte können von einem Fusionssystem im Vergleich zu einem Einzelsensorsystem Vorteile erwartet werden:

**Zuverlässigkeit** Multisensorsysteme zeichnen sich durch eine inhärente Redundanz aus, die die Zuverlässigkeit des Gesamtsystems erhöht. Existieren zum Beispiel redundante Informationen von mehreren Sensoren, können hardwarebedingte Sensorausfälle bis zu einem gewissen Umfang komplett kompensiert oder zumindest ein eingeschränkter Betrieb des Systems gewährleistet werden. Darüber hinaus kann die Abhängigkeit des Gesamtsystems von externen Umwelteinflüssen wie Wetter oder Beleuchtung minimiert werden, wenn einzelne Sensoren nicht nur redundante Informationen liefern, sondern auch, bedingt durch verschiedene physikalische Messprinzipien oder Parametrisierungen, unterschiedlich auf diese externen Störeinflüsse reagieren.

**Genauigkeit und Konfidenz** Durch redundante Messungen mehrerer Sensoren, können in der Kombination höhere Genauigkeiten für einzelne Messgrößen erzielt werden. Bild- und entfernungsgebende Sensorik ergänzen sich zum Beispiel sehr gut, da sie wechselseitig, in den Messgrößen Entfernung und laterale Ablage, geringe Varianzen aufweisen. Einzelne Sensoren können sich aber auch gegenseitig bestätigen, was zu einer höheren Konfidenz führt. Auf diese Weise lassen sich Geisterziele unterdrücken und Fehlklassifikationen verringern.

**Erfassungsbereich und räumliche Auflösung** Die räumliche Abdeckung eines Multisensorsystems ist durch die Vereinigung der Erfassungsbereiche der beteiligten Sensoren bestimmt. Dadurch lassen sich, weitgehend unabhängig von den Einschränkungen der jeweiligen Sensorik, komplexe Erfassungsbereiche modellieren. Durch eine Überlagerung der Erfassungsbereiche von Sensoren mit redundantem Messprinzip, kann innerhalb dieses gemeinsamen Bereiches eine erhöhte räumliche Auflösung erreicht werden.

**Zeitliche Auflösung** Neben der räumlichen Auflösung profitiert die zeitliche Auflösung des Gesamtsystems von zusätzlicher Sensorik. Hierbei werden mehrere Sensoren asynchron und zeitlich verschränkt zueinander betrieben. Im Idealfall, wenn alle Sensoren identische Taktraten aufweisen, steigt die zeitliche Auflösung linear mit der Anzahl der Sensoren. Allerdings ist zu beachten, dass Latenzzeiten, die durch das Messprinzip der jeweiligen Sensorik bedingt sind, hiervon nicht beeinflusst werden können. Im Allgemeinen wird jedoch die Systemantwortzeit eines Multisensorsystems verringert, da mehr Daten in der gleichen Zeit zur Verfügung stehen.

**Messprinzip** Mittels geeigneter Algorithmen, lassen sich mit mehreren Sensoren völlig neue oder erweiterte Messprinzipien realisieren. Mit zwei Kameras und Methoden der Stereobildverarbeitung können zum Beispiel Entfernungen direkt gemessen werden. Dies ist mit einer Kamera allein nicht unmittelbar möglich.

Alle oben genannten Vorteile lassen sich nur selten in einem System gleichzeitig erreichen. Vielmehr werden im Allgemeinen, je nach gewünschter Anwendung und zur Verfügung stehender Sensorik, gezielt bestimmte Aspekte herausgegriffen, die durch den Einsatz mehrerer Sensoren optimiert werden.



### 2.3.2 Fusionsmodelle

Maschinelle Sensordatenfusion wurde erstmals in den 80er Jahren im militärischen Bereich angewendet [Var97]. Aus dieser Zeit stammen die ersten Fusionsmodelle, wie das JDL-Modell<sup>4</sup> oder der Boyd-Steuerungszyklus [Boy87]. Unter einem Modell ist in diesem Zusammenhang eine Beschreibung von Prozessen zu verstehen, die für ein funktionelles Fusionssystem realisiert werden müssen. Aspekte der Softwareumsetzung oder Implementierungsdetails sind dabei nicht von Interesse.

Im erweiterten JDL-Modell [SBW99], das aktuell einen de facto Standard für Fusionssysteme im militärischen Bereich darstellt, wird die Informationsverarbeitung in fünf Ebenen unterteilt, wobei eine zunehmende Informationsverdichtung stattfindet [RPL07].

- **Vorverarbeitung** (Ebene 0): Die Anfangsphase beinhaltet eine zeitliche und räumliche Registrierung der Daten sowie Vorverarbeitungsmaßnahmen auf Signal- oder Pixelebene.
- **Objekterkennung und Merkmalsextraktion** (Ebene 1) bezieht sich auf die Schätzung und Vorhersage von kontinuierlichen (Position und Kinematik) oder diskreten (Objektverhalten und -identität) Objektmerkmalen.
- **Situationsanalyse** (Ebene 2): Alle detektierten Objekte werden in einen größeren Kontext gebracht, indem Objektbeziehungen analysiert werden. Dadurch wird eine Situationsinterpretation bezüglich des aktuellen Missionsstatus möglich.
- **Bedrohungsanalyse** (Ebene 3): Situationsabhängig werden, im Sinne einer Risikominimierung, unterschiedliche Handlungsoptionen evaluiert.
- **Prozessoptimierung** (Ebene 4): Diese Phase ist ein Metaprozess, der phasenübergreifend agiert und Laufzeitanforderungen überwacht sowie den Ressourcenbedarf (Sensoren und Rechenkomponenten) optimiert.

Das JDL-Modell beschreibt keinen Ablauf in einem Datenfusionssystem, sondern gibt lediglich die verschiedenen Hierarchieebenen wieder, in denen Informationen verarbeitet werden. Neben diesem Defizit wirken speziell die Ebenen 2 und 3 oftmals deplaziert im nicht militärischen Gebrauch. Ein etabliertes Fusionsmodell ist das Omnibus<sup>5</sup> Modell [BO00], das Aspekte aus älteren Modellen aufgreift und zu einem neuen Modell kombiniert (siehe auch Abbildung 2.4). Die Prozesse „beobachten“, „ausrichten“, „entscheiden“ und „handeln“ werden als zyklisches Ablaufdiagramm – entsprechend dem Boyd-Steuerungszyklus [Boy87] – dargestellt und anhand des Wasserfallmodells [Bed92] detailliert. Die Prozessübergänge markieren verschiedene Fusionsstufen (Rohdatenfusion, Merkmalsfusion, Soft-decision Fusion, Hard-decision Fusion), die aus dem Dasarathy Modell [Das97] abgeleitet sind. Darüber hinaus wird eine Militärterminologie vermieden.

Die vorgestellten Fusionsmodelle ermöglichen die Unterteilung eines komplexen Gesamtsystems in kleinere Teilsysteme. Analog dazu wird die Problemstellung in möglicherweise leichter zu bewältigende Teilprobleme zerlegt. Keines der Fusionsmodelle dient jedoch als optimale Vorlage für beliebige Aufgabenstellungen bezüglich der Sensorfusion. Vielmehr fungieren sie als Richtlinien, die stets an die speziellen Bedürfnisse angepasst werden müssen.

<sup>4</sup> Das Modell wurde von der *Joint Directors of Laboratories (JDL) Data Fusion Working Group* entwickelt, die 1986 vom amerikanischen Militär zur Vereinheitlichung der Datenfusionsterminologie ins Leben gerufen wurde.

<sup>5</sup> omnibus = umfassend (AE)

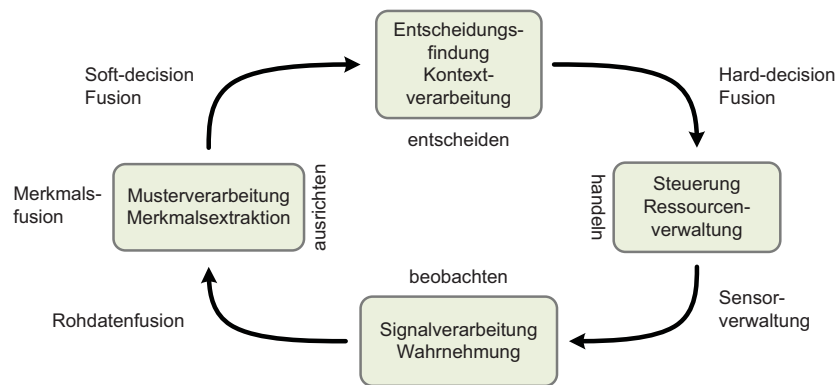


Abbildung 2.4: Omnibus Prozessmodell für Fusionssysteme. Zwischen den zyklisch ablaufenden Prozessen Signalverarbeitung, Musterverarbeitung, Entscheidungsfindung und Steuerung werden die entsprechenden Fusionsarten markiert [BO00].

### 2.3.3 Fusionstaxonomien

Eine von Fusionsmodellen unabhängige Einteilung kann über den Zeitpunkt und die Art und Weise der Fusion gegeben werden. So kann eine Sensorfusion auf verschiedenen Abstraktionsebenen der Datenverarbeitungskette erfolgen. Unterschieden wird dabei zwischen einer Fusion auf Signal-, Merkmals- oder Objektebene [RPL07]. In diesem Zusammenhang wird in der Literatur häufig von früher oder später Fusion gesprochen, wobei sich „früh“ bzw. „spät“ auf den Fusionszeitpunkt in der Datenverarbeitungskette und somit auf die jeweilige Abstraktionsebene bezieht.

**Fusion auf Signalebene (Rohdatenfusion)** Die Fusion erfolgt zu einem frühen Zeitpunkt in der Datenverarbeitungskette, in dem Sensorrohdaten vorliegen. Als Rohdaten werden meist Daten bezeichnet, die unmittelbar von einem Sensor stammen und noch nicht weiterverarbeitet wurden. Im Sensorsystem selbst kann jedoch schon eine Vorverarbeitung stattgefunden haben (z.B. Rauschunterdrückung bei einem Kamerasystem).

**Fusion auf Merkmalsebene** Bei einer Merkmalsfusion werden unabhängig für jeden Sensor Attribute, wie beispielsweise Kanten im Bild, aus den Messdaten extrahiert. Anschließend werden die Merkmalsdaten der unterschiedlichen Sensoren kombiniert, um eine kollektive Objektbeschreibung abzuleiten.

**Fusion auf Objektebene** Bei einer Fusion auf Objektebene wird die gesamte Datenverarbeitungskette für jeden Sensor eigenständig durchlaufen. Erst am Ende des Objektflusses werden die Objektdaten zu einer gemeinsamen Aussage kombiniert. Weitere oftmals synonym verwendete Bezeichnungen für diese Art der Fusion sind späte oder entscheidungsbasierte Fusion.

Bei der Frage nach dem optimalen Fusionszeitpunkt besteht ein Zielkonflikt hinsichtlich Effektivität und Effizienz. Auf späteren Abstraktionsebenen wird durch die Datenverarbeitung zwar die Datenmenge, aber zugleich auch der Informationsgehalt vermindert. Dadurch ist eine Fusion auf Objektebene in der Regel laufzeiteffizienter als eine Fusion auf Signalebene, aber zugleich weniger effektiv. Aus informationstheoretischer Sicht sollte frühestmöglich fusioniert werden, wobei eine pauschale Aussage nicht getroffen werden kann, da der Fusionsnutzen stark von den beteiligten Sensoren und den eingesetzten Algorithmen abhängt.

Ein höherer Informationsgehalt ist durch die Kombination von redundanten oder komplementären Informationsquellen erzielbar. Sensorredundanz kann Unsicherheiten vermindern, Genauigkeiten erhöhen und die Ausfallsicherheit anheben. Sensorheterogenität erweitert den Merkmalsraum, indem durch verschiedenartige Messverfahren unterschiedliche Aspekte aus der Umgebung extrahiert werden [Var97]. Dieser beiden Prinzipien bedient sich Durrant Whyte [DW88] zur Aufgliederung der Fusionsarten hinsichtlich der Art und Weise, wie die Sensoren untereinander interagieren. Er unterscheidet komplementäre, konkurrierende und kooperative Fusionsansätze.

**Komplementäre Fusion** Arbeiten zwei oder mehr Sensoren unabhängig voneinander und beobachten unterschiedliche räumliche, zeitliche oder physikalische Aspekte der Realität, kann eine komplementäre Fusion die einzelnen Sensordaten kombinieren. Von algorithmischer Seite ist diese Fusion einfach zu realisieren, da die Daten lediglich zeitsynchron vereinigt werden müssen. Ein Beispiel für eine komplementäre Fusion sind zwei Videokameras, die voneinander disjunkte Bereiche eines Raumes beobachten. Durch eine Vereinigung der beiden komplementären Sichtbereiche vergrößert sich der erfasste Gesamtbereich entsprechend.

**Konkurrierende Fusion** Sensoren, die die gleichen Aspekte und Merkmale der Realität beobachten, können als konkurrierend aufgefasst werden. Durch diese redundante Perzeption können die beteiligten Sensoren kongruieren, aber auch im Widerspruch zueinander stehen. Geeignete Algorithmen werten die Einzelschätzungen aus und bilden daraus eine gemeinschaftliche Entscheidung. Beispielsysteme für eine konkurrierende Fusion finden sich in Flugzeugen (Fly-by-wire Systeme) bzw. auch in Fahrzeugen (Pedalwertgeber in der Motorsteuerung). Um die Fehlertoleranz zu erhöhen werden essentielle Sensoren und Kommunikationspfade mehrfach redundant ausgelegt. Durch einen verschränkten Datenabgleich können sowohl Sensorausfälle kompensiert als auch Übertragungsfehler erkannt und korrigiert werden.

**Kooperative Fusion** Bei einer kooperativen Fusion werden die beteiligten Sensoren zu einem Kollektiv zusammengeschlossen. Geeignete Algorithmen extrahieren neue Informationen über die gemeinsame Beobachtung, die einem Einzelsensor im Kollektiv verborgen wären. Stereoskope Kamerasysteme beispielsweise erlauben über Epipolargeometrie und Triangulation Entfernungsbestimmungen, die über ein monokulares System nicht direkt möglich sind. Robuste Algorithmen für eine kooperative Fusion zu finden ist schwierig, da kleine Einzelsensorunzuverlässigkeiten das Gesamtergebnis negativ beeinflussen können. Dadurch wird, im Gegensatz zur redundanten Perzeption einer konkurrierenden Fusion, die Robustheit und Genauigkeit des Multisensorsystems vermindert.

Bei einem realen Fusionssystem liegen in der Regel keine dieser Aufgliederungen in Reinform vor. Es existieren vielfältige Mischformen. In einem Fusionssystem zur Objekterkennung mit Stereobildverarbeitung beispielsweise, kann in überlappenden Bildbereichen kooperative (Entfernungsbestimmung über Triangulation) und in disjunkten Bildbereichen komplementäre Fusion (Erhöhung des Abdeckungsbereichs durch Vereinigung) stattfinden. Dies erfolgt auf Rohdaten- und Merkmalebene. Ist an dem System zusätzlich entfernungsgebende Sensorik beteiligt, könnte überdies konkurrierend und entscheidungsbasiert fusioniert (redundante Fusion auf Objektebene) werden.

### 2.3.4 Fusion im Kontext der Fahrumfelderfassung

Von 1987 bis 1995 wurden in dem europäischen Forschungsprojekt PROMETHEUS<sup>6</sup> die Grundlagen für aktuelle Systeme zur Fahrumfelderfassung im Fahrzeug gelegt. Basierend auf der Fusion von mehreren Kameras ist, anlässlich der Abschlusspräsentation von PROMETHEUS, 1994 ein Fahrzeug autonom mehrere hundert Kilometer im realen Straßenverkehr auf einer Autobahn in der Nähe von Paris gefahren [Dic02]. Seitdem beschäftigen sich nationale und internationale Nachfolgeprojekte in enger Zusammenarbeit zwischen Zulieferern, Automobilherstellern und Forschungsinstituten mit innovativer Sensorik, einer Erweiterung der Fahrumfelderfassung auf komplexe Szenarien und der Entwicklung neuartiger Fahrerassistenzsysteme. Hierbei wurden insbesondere von den Projekten PREVENT [Eur07a], INVENT [Bun07], RADARNET [Eur07b] und CARSENSE [INR02] Aspekte der Sensorfusion näher betrachtet. Zum Einen wurden unterschiedlichste Fusionsarchitekturen wie Frühe-, Späte oder Feedback-Fusion entwickelt und zum Anderen verschiedene Sensorkonfigurationen bestehend aus Radar, Laserscanner und bildgebender Sensorik untersucht.

Als Ergebnis der genannten Forschungsaktivitäten befinden sich inzwischen erste Produkte auf dem Markt, die Methoden der Sensorfusion für Fahrerassistenzsysteme verwenden. Daimler beispielsweise ergänzt in der S-Klasse seit 2005 den Fernbereichsradar, der bisher für ACC verwendet wurde, um Radarsensorik für den Nahbereich. Durch die zusätzlich gewonnene Genauigkeit und den vergrößerten Erfassungsbereich wurden Einparkassistent, ACC S&G, Kollisionswarnung sowie ein Bremsassistent realisiert. Seit Herbst 2006 ist eine autonome Teilbremsung verfügbar, die ausgelöst wird, wenn der Fahrer die optische und akustische Kollisionswarnung ignoriert. In den aktuellen Lexus Modellen wird hingegen eine Kombination aus NIR-Stereobildverarbeitung und Radarsensorik für ACC S&G sowie für eine Kollisionswarnung verwendet. Durch die Fusion lässt sich ein größeres Spektrum an detektierbaren Objekten erreichen.

## 2.4 Versuchsträger

Im Rahmen des europäischen Forschungsprojektes COMPOSE<sup>7</sup> zur Verkehrssicherheit, demonstrierte BMW, in Zusammenarbeit mit weiteren Projektpartnern, zukünftige Fahrerassistenzsysteme mit hohen Ansprüchen an die unterliegende Fahrumfelderfassung. Somit war neben der Entwicklung und dem Vergleich von Fusionskonzepten, die Systemintegration ins Fahrzeug ein Bestandteil dieser Arbeit. Dabei sind sowohl Echtzeitanforderungen als auch Randbedingungen wie vordefinierte Koordinatensysteme und limitierte Ressourcen zu berücksichtigen.

Als Versuchsfahrzeug wurde eine BMW Limousine verwendet, um genügend Verbauplatz für die Sensorik und die Rechensysteme bereitzustellen. Alle für die Fahrumfelderfassung relevanten Sensoren wurden in die Frontpartie des Fahrzeugs integriert und deren Detektionsbereiche auf den Bereich vor dem Fahrzeug ausgerichtet. Das Rechensystem ist im Kofferraum verbaut und basiert auf handelsüblichen und fahrzeugtauglichen Komponenten. Durch die unterschiedlichen Öffnungswinkel und Verbauplätze der Sensoren ergeben sich redundante und exklusive Abdeckungsbereiche. Eine Datenfusion ist nur in den redundanten Bereichen möglich. Durch die Ausrichtung der Sensoren werden vornehmlich Applikationen wie beispielsweise ACC und AGB adressiert.

<sup>6</sup> Akronym für „Program for European Traffic with Highest Efficiency and Unprecedented Safety“.

<sup>7</sup> COMPOSE (collision mitigation and protection of road users) ist ein Teilprojekt des EU-Förderprojektes PREVENT.

### 2.4.1 Aufbau und Sensorsystem

Die im Versuchsfahrzeug verbauten Sensoren lassen sich in entfernungs- und bildgebende Sensoren gliedern. Unter die entfernungsgebenden Sensoren fallen ein Laserscanner und je zwei Nah- und Fernbereichsradare. Das bildgebende Sensorsystem setzt sich aus einer Videokamera mit Teleobjektiv, einer Videokamera mit Standardobjektiv und einer Wärmebildkamera zusammen. Die Verbauorte der Sensoren befinden sich hinter der Windschutzscheibe und im Bereich der Schürze (siehe Abbildung 2.5 für einen Verbauüberblick und Tabelle 2.2 für Sensordetails).

Während einer Versuchsfahrt kann zwischen zwei Systembetriebsmodi gewählt werden: Im Aufzeichnungsmodus werden alle Sensordaten, die über digitale (CAN und Ethernet) und analoge (Composite Video) Kommunikationspfade an einen zentralen Rechner übermittelt werden, auf einen Festspeicher geschrieben. Eine zeitliche Sensorsynchronisation erfolgt softwareseitig über die Ankunftszeiten der Daten. Der zweite Betriebsmodus ist der Perzeptionsmodus, in dem alle Sensordaten dem Fahrumfelderfassungssystem zugeführt werden. Dieses ermittelt eine relevante Objektliste bestehend aus detektierten Fahrzeugen und Fußgängern und übersendet sie an die Applikationsschicht.

Abbildung 2.6 zeigt Momentaufnahmen der Sensoren von einer Versuchsfahrt. Um den maximalen Datendurchsatz des Festspeichers nicht zu überschreiten, werden die Bilddaten vor der Speicherung auf eine Viertel-NTSC Größe skaliert und von RGB-Farb- nach Grauwertbilder konvertiert. Durch diese Maßnahmen wird die Datenrate auf ca. 7 Megabyte pro Sekunde begrenzt.

Tabelle 2.2: Details zur Sensorik des Versuchsfahrzeugs (horizontaler und vertikaler Öffnungswinkel, Reichweite für Fahrzeug/Fußgängererkennung, Winkelauflösung, Abtastfrequenz und Kurzbeschreibung).

	hor. $\sphericalangle$	vert. $\sphericalangle$	Reichweite	Auflösung	Frequenz	Beschreibung
FIR Kamera	$\pm 17.6^\circ$	$\pm 13.4^\circ$	80/60 m	320×240 (NTSC/4)	29.97 Hz	Mikrobolomertechnik, Shutterabgleich
Video (Tele)	$\pm 8.5^\circ$	$\pm 6.5^\circ$	100/80 m	320×240 (NTSC/4)	29.97 Hz	CMOS Farbbildkamera mit Teleobjektiv
Video (Normal)	$\pm 21.2^\circ$	$\pm 16.2^\circ$	60/30 m	320×240 (NTSC/4)	29.97 Hz	CMOS Farbbildkamera
Laserscanner	$\pm 61^\circ$	$\pm 3.6^\circ$	120/50 m	0.25° (lateral)	10 Hz	Vier vertikale Strahlebenen mit einem Versatz von 1.8°.
Fernbereichsradar	$\pm 3.5^\circ$	$\pm 2.5^\circ$	150 m	1° (lateral)	14 Hz	FMCW <sup>8</sup> -Verfahren (77 GHz)
Nahbereichsradar	$\pm 40^\circ$	$\pm 8.5^\circ$	25 m	ca. 2° (lateral)	14 Hz	Kurzpulsradar (24 GHz)

<sup>8</sup> Frequency Modulated Continuous Wave

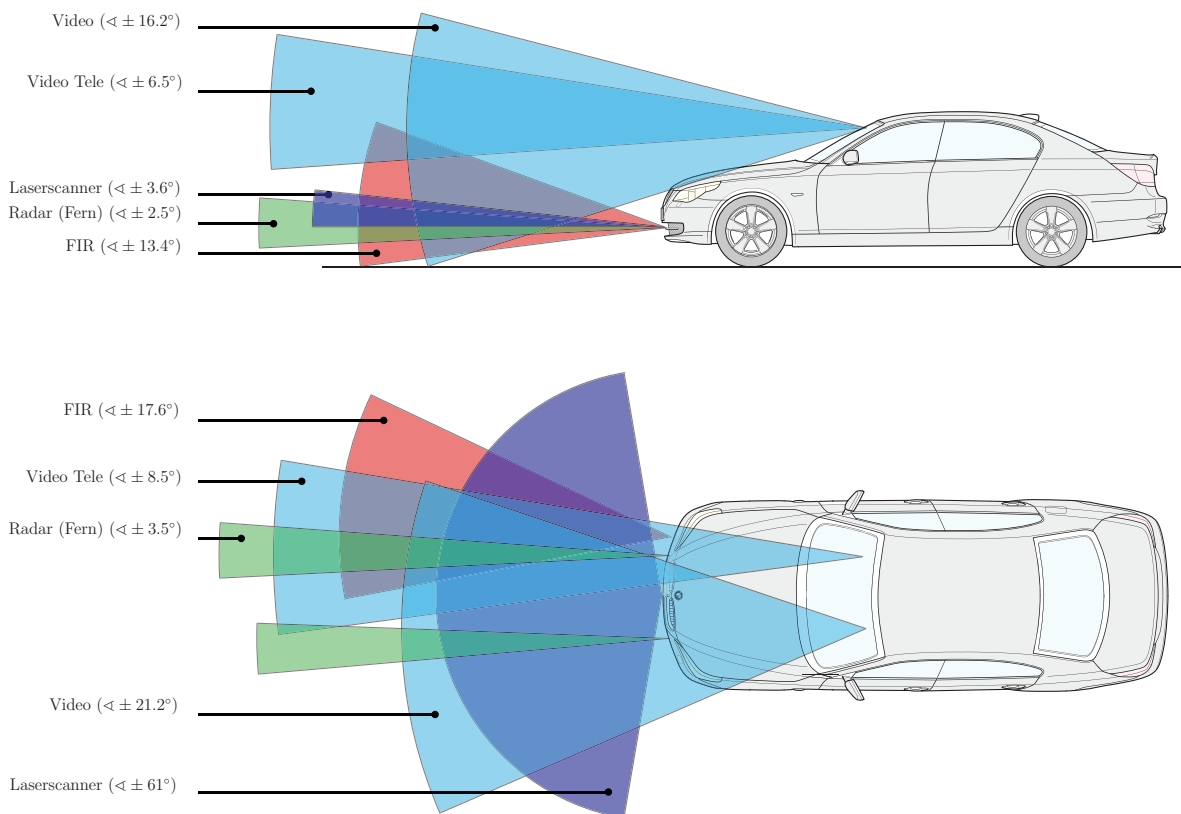


Abbildung 2.5: Sensorverbaupositionen und -sichtbereiche am Versuchsfahrzeug. Die vertikalen bzw. horizontalen Öffnungswinkel werden durch die obere bzw. untere Abbildung illustriert.

### 2.4.2 Koordinatensysteme

Um eine Verknüpfung zwischen Daten zu ermöglichen, ist es besonders in Multisensorsystemen wichtig, die Messwerte in einer einheitlichen mathematischen Form zu repräsentieren. Die Messwerte entfernungsgebender Sensoren können als Messpunkte im dreidimensionalen Raum aufgefasst werden. Ein einzelner Messwert  $\mathbf{x}^S$  wird durch einen vierdimensionalen Vektor  $(x, y, z, 1)^T$  in homogenen Koordinaten dargestellt. Die Extradimension ist nützlich weil dadurch affine Transformationen wie Rotationen, Skalierungen und auch Translationen durch Multiplikation mit  $4 \times 4$  Matrizen repräsentiert werden können [HZ03]. Mittels dieser Transformationsmatrizen können Messwerte in andere Koordinatensysteme überführt werden. Ein Koordinatensystem fungiert als Basis im Raum – relativ zu dieser können homogene Messwerte dargestellt werden.

Abbildung 2.7 zeigt die relevanten rechtshändigen Koordinatensysteme einer multisensoriellen Fahrfelderfassung. Unmittelbar nach einer Messung liegen die Daten zunächst im jeweiligen Sensorkoordinatensystem  $S_i$  vor. Die  $x$ -Achse weist entlang der zentralen Sichtachse des Sensors und der Ursprung liegt im Projektionszentrum. Sind mehrere Sensoren beteiligt, ist es wünschenswert, die verschiedenen Sensordaten in einem gemeinsamen Koordinatensystem darzustellen, um eine Datenfusion zu ermöglichen. Ein übergeordnetes Fahrzeugkoordinatensystem  $F$  kann dieses leisten. Die





Abbildung 2.6: Momentaufnahme der Sensordaten des Versuchsfahrzeugs während einer Aufnahme-fahrt. Abbildung (a) und (b) zeigen die Bilder der beiden Videokameras mit Normal- bzw. mit Teleobjektiv. In Abbildung (c) wird das Bild der Wärmebildkamera dargestellt. Abbildung (d) zeigt die Daten des verbauten 4-Ebenen Laserscanners in einer virtuellen Umgebungsdarstellung.

x-Achse des, durch die DIN Norm 70000 [Nor94] beschriebenen fahrzeugfesten Koordinatensystems weist in Fahrzeuglängsrichtung, die y-Achse in Fahrzeuguerrichtung. Die z-Achse verläuft vertikal und ist orthogonal zur x- und y-Achse. Der Ursprung des Koordinatensystems ist die Projektion des vordersten Stoßfängerpunktes auf die Fahrbahnebene. Diese Wahl des fahrzeugfesten Koordinatensystemursprungs ermöglicht eine einfachere Vermessung und Kalibrierung der Sensorkoordinatensysteme. Die Koordinatensysteme  $H_k$  bzw.  $H_z$  sind lokale Koordinatensysteme von kuboidischen bzw. zylindrischen Hypothesen, die datengetrieben initialisiert werden. Sowohl bei  $H_k$  als auch bei  $H_z$  liegt der Koordinatensystemursprung im, auf die Fahrbahnebene projizierten, Schwerpunkt der Hypothesen. Die x-Achse weist bei  $H_k$  entlang der Längsrichtung in Bewegungsrichtung und bei  $H_z$  entlang der Bewegungsrichtung. Die Beziehungen zwischen den verschiedenen Koordinatensystemen können als „Spatial Relationship Graph“ (SRG) [RH95, NWB04, PHBK06] beschrieben und visualisiert werden. Die Knoten repräsentieren dabei Koordinatensysteme und die Kanten starre oder dynamische

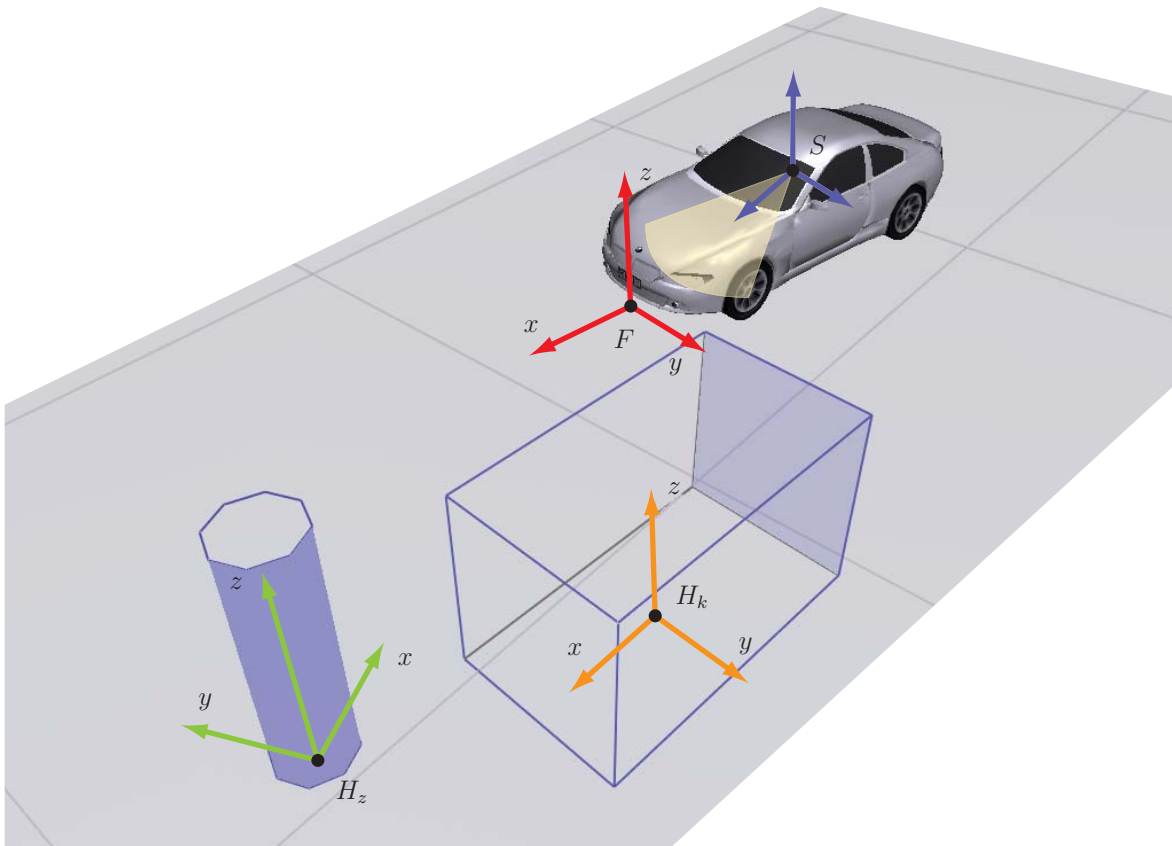


Abbildung 2.7: Die Abbildung zeigt den Ursprung und die Ausrichtung der verwendeten rechtshändigen Fahrzeug- ( $F$ ), Sensor- ( $S$ ), und Objektkoordinatensysteme ( $H_k$  bzw.  $H_z$ ).

räumliche Beziehungen zwischen den Systemen, die mittels Transformationsmatrizen  $\mathbf{T}$  darstellbar sind. Eine gute Einführung findet sich in [ABW01]. Ein typischer SRG einer Fahrumfelderfassung ist in Abbildung 2.8 dargestellt. Das Weltkoordinatensystem  $W$  ist das globale Koordinatensystem und beherbergt das eigene Fahrzeug  $F$  sowie das detektierte Fahrumfeld  $B$ . Relativ zum Fahrzeugkoordinatensystem sind die automotiven Sensoren  $S_1$  bis  $S_n$  angebracht. Die räumlichen Verknüpfungen  $\mathbf{T}_{S_i}^F$  zwischen Eigenfahrzeug und verbauten Sensoren sind starr und müssen über einen Kalibriervorgang bestimmt werden. Die Beziehung  $\mathbf{T}_F^W$  zwischen Welt- und Fahrzeugkoordinatensystem ist dynamisch und ändert sich über die Zeit infolge der Eigenfahrzeugbewegung, die durch Verfahren der Odometrie ermittelt wird.

Soll nun beispielsweise ein Messwert  $\mathbf{x}^S$  im Sensorkoordinatensystem  $S$  in einen Messwert  $\mathbf{x}^F$  in Fahrzeugkoordinaten  $F$  überführt werden, muss er lediglich mit der Transformationsmatrix  $\mathbf{T}_F^S$  multipliziert werden. Über die Vorschrift  $(\mathbf{T}_B^A)^{-1} = \mathbf{T}_A^B$  können die Kanten des Szenengraphen in beiden Richtungen durchlaufen werden. Eine Umrechnung von einem Messwert  $\mathbf{x}^{S_1}$ , der im Sensorkoordinatensystem  $S_1$  vorliegt, in ein anderes Sensorkoordinatensystem  $S_2$  berechnet sich somit über  $\mathbf{x}^{S_2} = \mathbf{T}_F^{S_2} \mathbf{T}_{S_1}^F \mathbf{x}^{S_1} = (\mathbf{T}_{S_2}^F)^{-1} \mathbf{T}_{S_1}^F \mathbf{x}^{S_1}$ . Alternativ zu Transformationsmatrizen kann auch eine VQS-Darstellung mit Quaternionen [Vic01] verwendet werden, die speziell Rotationen kompakter kodieren kann.



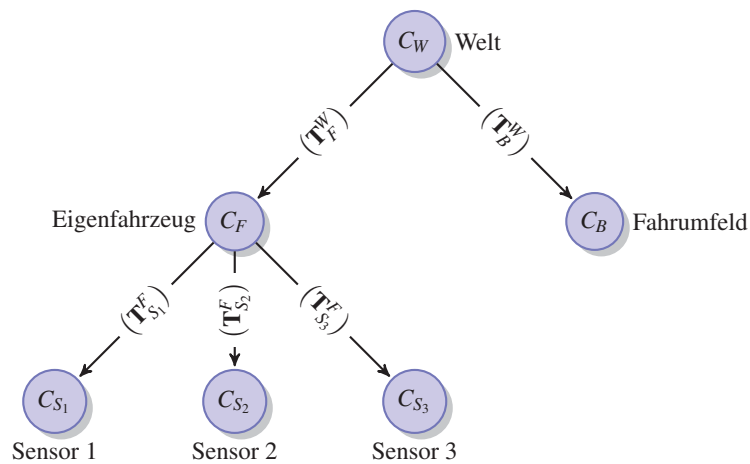


Abbildung 2.8: Dieser „Spatial Relationship Graph“ veranschaulicht die räumlichen Beziehungen von Sensoren, Eigenfahrzeug und Fahrumfeld.

## 2.5 Systemüberblick

Dieser Abschnitt gibt einen kurzen Überblick über die einzelnen Komponenten des Gesamtsystem und deren Zusammenspiel. Das Systemdesign und der zyklische Ablauf des Systems ist weitgehend an das Omnibus Fusionsmodell (siehe Abbildung 2.4) angelehnt. Das adaptierte Diagramm ist in Abbildung 2.9 dargestellt. Im Folgenden wird der Datenfluss beschrieben und die Begriffe Klassifikation bzw. Tracking kurz erklärt.

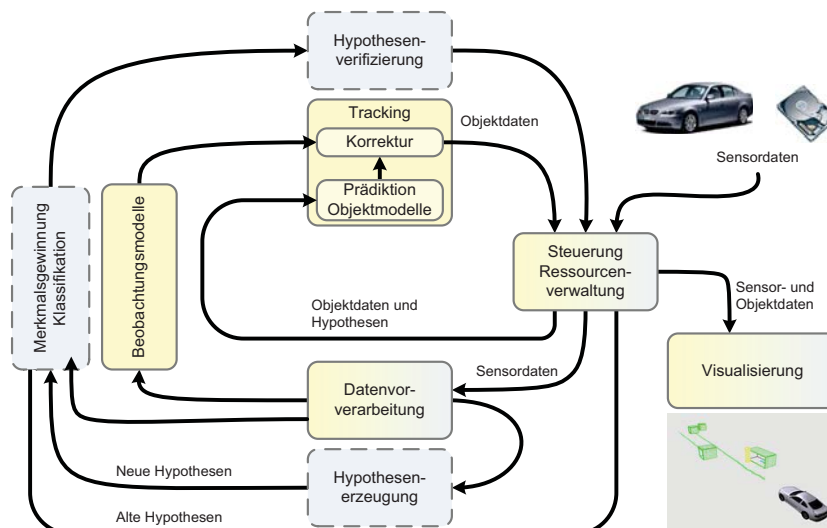


Abbildung 2.9: Systemüberblick der in dieser Arbeit adressierten Fahrumfelderfassung. In der Abbildung sind das Klassifikations- (gestrichelt) und Trackingmodul (dunkel hinterlegt) als die wesentlichen Komponenten des Gesamtsystems besonders hervorgehoben.

**Datenfluss:** Der Datenfluss ist in die zwei Komponenten Klassifikation und Tracking unterteilt. Dabei werden Sensordaten, die entweder unmittelbar von der Sensorik stammen oder in aufzeichneter Form vorliegen, in Objektdaten umgewandelt und in die jeweiligen Objektzustände eingebracht. Werden neue Sensordaten über die Eingabesteuerung aufgenommen, werden diese zunächst vorverarbeitet und aggregiert. Anschließend werden basierend auf diesen vorverarbeiteten Daten neue Objekthypothesen initialisiert. Über einen Klassifikationsprozess wird anhand von Merkmalen, die durch die Lage und Geometrie der Hypothesen gewonnen werden, eine Objekteinteilung vorgenommen. Zeitgleich werden bereits bekannte Hypothesen und klassifizierte Objekte mit den vorverarbeiteten Sensordaten in Verbindung gebracht. Hierzu ist in den Beobachtungsmodellen beschrieben, wie die jeweiligen Objekte in den verwendeten Sensoren abgebildet werden und mit welchen Verfahren bestimmte Zustandsgrößen gemessen werden können. Anschließend werden die Zustandsdaten der Hypothesen und der klassifizierten Objekte mit diesen Messwerten aktualisiert. Eine Ressourcenverwaltung entfernt Objekte, die nicht länger durch Messwerte gestützt werden.

**Klassifikation:** Ein Klassifikator ist eine beliebige Funktion  $D: \mathbb{R}^n \rightarrow \Omega$ , die einem Merkmal  $\mathbf{x} \in \mathbb{R}^n$  eine Klassenbezeichnung  $\omega \in \Omega = \{\omega_1, \dots, \omega_c\}$  zuweist. Eine Klasse  $\omega$  beinhaltet ähnliche Objekte und grenzt sich zu einer anderen Klasse  $\omega'$  dadurch ab, dass deren Objekte unähnlich zu den eigenen Objekten sind. Objekte können anhand von charakteristischen Merkmalen beschrieben werden. In der statistischen Mustererkennung werden ausschließlich numerische Merkmale verwendet. Numerische Merkmale können zu einem  $n$ -dimensionalen Merkmalsvektor  $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{R}^n$  zusammengefasst werden.

Fahrzeuge und Fußgänger sind ein Beispiel für zwei Klassen. Charakteristische Merkmale um Fußgänger von Fahrzeugen unterscheiden zu können, sind beispielsweise Gewicht und Abmessungen. Weniger geeignete Merkmale wären die Farbe und das Baujahr bzw. der Geburtstag.

**Tracking:** Der Begriff Tracking<sup>9</sup> beschreibt den Prozess, mit dem Objekte über mehrere Zeitschritte hinweg verfolgt werden. Dabei wird versucht aus Messungen bzw. Beobachtungen des Objektes Zustandsgrößen wie zum Beispiel Position, Geschwindigkeit, Breite oder Orientierung abzuleiten bzw. zu schätzen. Ein weiterer wichtiger Aspekt des Trackings ist die Einbeziehung des zeitlichen Verlaufes bzw. der letzten Beobachtungen des zu verfolgenden Objektes. Dementsprechend wird ein Objekt im Kontext seiner Historie betrachtet. Auf diese Weise können zum Einen Störeinflüsse wie zum Beispiel Messfehler kompensiert werden und zum Anderen Zustandsgrößen wie die Beschleunigung abgeleitet werden, die sich einer direkten Beobachtung entziehen. In der Literatur wird der Prozess des Trackings auch als Filterung bezeichnet, da die Auswirkungen von Messfehlern abgeschwächt werden.

Das Tracking wird oft in die drei Komponenten Objektmodell (a priori Wissen), Beobachtungsmodell (likelihood) und den Prozess der Filterung aufgeteilt. Hierbei umfasst das a priori Wissen charakteristische Merkmale, Zustandsgrößen oder Eigenschaften, welche das zu verfolgende Objekt möglichst eindeutig festlegen. Das Beobachtungsmodell beschreibt, wie die Sensorinformationen für die jeweiligen Objekte zu behandeln und auszuwerten sind. Letztendlich vereint bzw. gewichtet die Filterung die ersten beiden Komponenten und ist für die eigentliche Zustandsschätzung des zu verfolgenden Objektes verantwortlich.

In den Kapiteln 4 und 5 wird im Detail auf die drei Komponenten des Trackings eingegangen. Für eine ausführliche Betrachtung der Klassifikationskomponente sei auf die Arbeit von Leonhard Walchshäusl [Wal08] verwiesen.

---

<sup>9</sup> engl. für das Verfolgen bzw. das Nachführen

# SENSOR KALIBRIERUNG

---

Es ist ein wesentlicher Aspekt der Sensorfusion, Messwerte von Sensoren, die im Allgemeinen an unterschiedlichen Orten verbaut sind, gewinnbringend in Beziehung zueinander zu setzen. Eine Voraussetzung hierfür ist die Bestimmung der relativen Position und Orientierung der einzelnen Sensoren zueinander, was auch als *extrinsische Kalibrierung* bezeichnet wird. In Verbindung mit einer *intrinsischen* Kalibrierung, die interne Sensorparameter wie beispielsweise die Abbildungseigenschaften einer Kamera bestimmt, können Messwerte von einem Sensor in das Koordinatensystem eines anderen Sensors umgerechnet werden. Hiermit ist es zum Beispiel möglich, Bildausschnitte der Kameras mit Daten des Laserscanners oder der Radarsensorik zu verknüpfen bzw. zu attributieren und somit gemeinsam zu interpretieren.

Eine weitere Voraussetzung für eine korrekte Fusion ist die Berücksichtigung von Latenzzeiten, die zum Beispiel durch einen asynchronen Systemaufbau, sensorinterne Vorverarbeitung oder durch die Kommunikation zwischen den Komponenten entstehen können. Im Allgemeinen ist die genaue Zusammensetzung der Sensorlatenzen nicht bekannt, da zum Einen häufig nicht alle relevanten Informationen von den Sensorherstellern veröffentlicht werden und die Latenzzeiten zum Anderen sehr systemspezifisch sind. Werden diese Zeiten allerdings vernachlässigt kann es insbesondere bei hohen Relativgeschwindigkeiten zu Widersprüchen zwischen den Messwerten der einzelnen Sensoren kommen. Deshalb sollten die Latenzzeiten der Sensoren in einer *temporalen Kalibrierung* experimentell bestimmt und in der Fusion berücksichtigt werden.

Im Folgenden wird zunächst auf die Besonderheiten der Kalibrierung eines multispektralen Kamerasystems<sup>1</sup> eingegangen. Anschließend wird beschrieben, wie Referenzdaten erzeugt werden können, die im weiteren Verlauf sowohl die Grundlage für die extrinsische als auch die temporale Kalibrierung des Multisensorsystems bilden. In einer Fusion von asynchronen Sensoren dürfen Latenzzeiten nicht vernachlässigt werden. Deshalb werden Verfahren diskutiert, um experimentell für entfernungs- und bildgebende Sensoren die Latenzzeiten zu bestimmen.

## 3.1 Kalibrierung von multispektralen Kameras

Das Ziel einer Kamerakalibrierung ist die Berechnung der Abbildungseigenschaften sowie die Bestimmung der relativen Position und Orientierung der einzelnen Kameras zueinander. Im Folgenden wird zunächst das dem Kalibrierungsprozess zugrundeliegende mathematische Modell einer Kamera

---

<sup>1</sup> Ein multispektrales Kamerasystem besteht aus mindestens zwei Kameras, die in unterschiedlichen Spektren operieren, wie zum Beispiel sichtbares Licht und Wärmestrahlung.

eingeführt. Anschließend werden zwei Verfahren diskutiert, die den besonderen Anforderungen von multispektralen Sensorsystemen entsprechen.

### 3.1.1 Projektives Kameramodell

In diesem Abschnitt wird das häufig verwendete Modell einer Lochkamera [HZ03] vorgestellt. Dieses Modell ist zum Einen die Grundlage für die intrinsische Kalibrierung der verwendeten Kameras und stellt zum Anderen das mathematische Handwerkszeug zur Verfügung, um Messwerte anderer Sensoren mit den Bilddaten der Kamera zu fusionieren. Zunächst wird von einem einfachen Modell ausgegangen, das dann schrittweise verfeinert wird, um den Besonderheiten aktueller Kameras gerecht zu werden.

Für eine projektive Lochkamera ist die Abbildung von einem 3D Punkt  $\mathbf{x} = (x, y, z)^T$  aus dem Kamerakoordinatensystem auf einen 2D Punkt  $\tilde{\mathbf{x}} = (u, v)^T$  in der Bildebene wie folgt durch die Brennweite  $f$  bestimmt:

$$(x, y, z)^T \mapsto \left(f \frac{x}{z}, f \frac{y}{z}\right)^T \quad (3.1)$$

Das einfache Modell einer Lochkamera muss für reale digitale Kameras um zwei Erweiterungen ergänzt werden. Zum Einen sind bei digitalen Kameras im Allgemeinen die Höhe und Breite eines Pixels auf der Bildebene nicht identisch. Deshalb wird die Brennweite  $f$  durch eine für  $u$  und  $v$  spezifische Brennweite  $f_u$  bzw.  $f_v$  ersetzt. Zum Anderen wird ein Offset  $p_u$  bzw.  $p_v$  eingeführt, der den Abstand angibt, um den das Koordinatensystem eines Bildsensors vom Ursprung der Bildebene der einfachen Lochkamera abweicht. Rein theoretisch entspricht dieser Offset bei den meisten Kameras der Hälfte von Bildbreite und Bildhöhe in Pixeln. Durch Fertigungstoleranzen von Sensoren und Objektiven, ist dies in der Praxis jedoch selten der Fall. Zusammen ergibt sich folgende erweiterte Abbildungsvorschrift:

$$(x, y, z)^T \mapsto \left(f_u \frac{x}{z} + p_u, f_v \frac{y}{z} + p_v\right)^T \quad (3.2)$$

In homogenen Koordinaten lässt sich die Abbildung eines 3D Punktes  $\mathbf{x} = (x, y, z, 1)^T$  aus dem Kamerakoordinatensystem auf einen 2D Punkt  $\tilde{\mathbf{x}} = (u, v, 1)^T$  in der Bildebene wie folgt beschreiben, wobei  $\mathbf{K} \in \mathbb{R}^{3 \times 4}$  die Kamerakalibrierungsmatrix bezeichnet:

$$\tilde{\mathbf{x}} = \mathbf{K} \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \quad (3.3)$$

$$\mathbf{K} = \begin{bmatrix} f_u & 0 & p_u & 0 \\ 0 & f_v & p_v & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (3.4)$$

Die Bestimmung der Matrix  $\mathbf{K}$  wird oft auch als *intrinsische* Kalibrierung bezeichnet. In dem bisherigen Kameramodell wurden Verzerrungen, die auf das Objektiv der Kamera zurückzuführen sind, nicht betrachtet. Für die meisten Kameras sind hierbei vor allem radiale Verzerrungen erster und zweiter Ordnung dominant [Zha99]. Komplexere Modelle wie die tangential- oder Prisma-Verzerrung

[WCH92] werden an dieser Stelle nicht weiter betrachtet, da sie im Rahmen der in dieser Arbeit geforderten Genauigkeit vernachlässigt werden können. Im Folgenden repräsentieren  $\mathbf{x} = (x, y, 1, 1)^T$  bzw.  $\mathbf{x}_d = (x_d, y_d, 1, 1)^T$  die verzerrungsfreien bzw. verzerrten normalisierten<sup>2</sup> Bildkoordinaten. Zusammen mit den Koeffizienten  $k_1$  und  $k_2$  ist die radiale Verzerrung dementsprechend wie folgt definiert.

$$\mathbf{x}_d = \left( 1 + k_1 (x^2 + y^2) + k_2 (x^2 + y^2)^2 \right) \mathbf{x} \quad (3.5)$$

Der 2D Punkt  $\tilde{\mathbf{x}}$  auf der Bildebene ergibt sich analog zu 3.3:

$$\tilde{\mathbf{x}} = \mathbf{K} \mathbf{x}_d \quad (3.6)$$

Das bisher beschriebene Modell einer Kamera ist noch nicht vollständig, da in der Abbildung die Lage und Orientierung der Kamera in Bezug auf ein Weltkoordinatensystem bzw. in Bezug auf andere Sensoren nicht berücksichtigt wird. Diese fehlenden Größen werden als *extrinsische* Parameter bezeichnet und setzen sich aus einer Rotation  $\mathbf{R} \in \mathbb{R}^{3 \times 3}$  sowie einer Translation um den Vektor  $\mathbf{t} \in \mathbb{R}^3$  zusammen. In der Darstellung als Szenegraph [PHBK06] entspricht dies genau der Kante von der Kamera zur Wurzel des Graphen. Die Transformation von Welt- nach Kamerakoordinaten lässt sich in homogenen Koordinaten mit einer Matrix darstellen. Zusammen mit der Kamerakalibrierungsmatrix  $\mathbf{K} \in \mathbb{R}^{3 \times 4}$  ergibt sich schließlich für einen 3D Punkt  $\mathbf{x}$  aus dem Weltkoordinatensystem, folgende Abbildung  $\tilde{\mathbf{x}}$  auf der Bildebene.

$$\tilde{\mathbf{x}} = \begin{array}{c} \text{intrinsische Parameter} \\ \left[ \begin{array}{ccc|cc} f_u & p_u & 0 & & \\ & f_v & p_v & 0 & \\ & & 1 & 0 & \end{array} \right] \end{array} \begin{array}{c} \text{extrinsische Parameter} \\ \left[ \begin{array}{cc|c} \mathbf{R} & \mathbf{t} & \\ \hline 0 & 1 & \end{array} \right] \end{array} \mathbf{x} = \mathbf{P} \mathbf{x} \quad (3.7)$$

$$(3.8)$$

Im weiteren Verlauf dieser Arbeit wird die vollständige Projektionsmatrix  $\mathbf{P} \in \mathbb{R}^{3 \times 4}$  verwendet, die sich aus einer Komposition von intrinsischen und extrinsischen Parametern zusammensetzt.

### 3.1.2 Kalibrierungsverfahren

Nachdem im vorherigen Abschnitt das Kameramodell vorgestellt wurde, das dem Kalibrierungsprozess zugrunde liegt, werden im Folgenden die Verfahren beschrieben, um die 12 Freiheitsgrade dieses Modells bzw. der Projektionsmatrix  $\mathbf{P}$  zu bestimmen. Darüber hinaus werden zwei Kalibrierkörper<sup>3</sup> eingeführt, die den Bedürfnissen eines multispektralen Kamerasystems gerecht werden.

Einige intrinsischen Werte könnten durchaus aus den Herstellerangaben des verwendeten Kameratyps entnommen werden. In der Praxis sind diese Werte jedoch starken Schwankungen unterworfen, so dass

<sup>2</sup>  $(x, y, z, 1)^T \mapsto (x/z, y/z, 1, 1)^T$

<sup>3</sup> Ein Kalibrierkörper ist ein Gegenstand, der durch vordefinierte Eigenschaften wie Form oder Textur charakteristische Messwerte im Sensor erzeugt. Durch die Beziehung aus Messwert und dem Zusatzwissen der vordefinierten Eigenschaften des Körpers können Rückschlüsse auf interne Parameter des jeweiligen Sensors gezogen werden. Dieser Vorgang wird als Kalibrierung bezeichnet.

bedingt durch Fertigungstoleranzen zwei baugleiche Kameras unterschiedliche Werte aufweisen können. Ähnliches gilt für die extrinsischen Parameter. Hier kann zwar in erster Näherung der Verbauort der Kameras aus CAD-Modellen übernommen oder mit einfachen Hilfsmitteln vermessen werden. Für eine präzise Bestimmung der extrinsischen Parameter insbesondere der Orientierung ist dieses Verfahren allerdings nicht geeignet. Hierfür ist ein gesondertes Kalibrierungsverfahren notwendig.

#### Schachbrett aus Peltierelementen

In dieser Arbeit wurde die „Camera Calibration Toolbox“ [Bou07] verwendet, die eine komfortable Implementierung der Verfahren von Heikkila [JH97] und Zang [Zha99] zur Kalibrierung eines Multikamerasystems bietet. Der Kalibrierkörper ist ein planares Schachbrett, welches in  $m \times n$  Felder mit gut sichtbarem Kontrast aufgeteilt ist. Bei zwei Kameras aus unterschiedlichen Spektralbereichen kommt erschwerend hinzu, dass der Kalibrierkörper in beiden Spektren sichtbar sein sollte. Für die in dieser Arbeit verwendete Sensorik bedeutet dies, dass er im Videobild einen Kontrastunterschied und im Ferninfrarotbild einen Temperaturunterschied aufweisen muss.

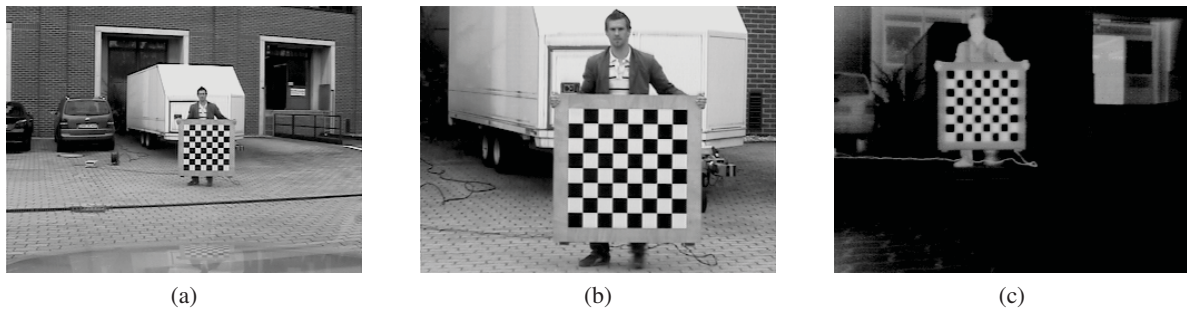


Abbildung 3.1: Kalibrierung einer multispektralen Kamera mit einem Schachbrett aus Peltierelementen. Die Kanten des Schachbretts sind im Videobild (a) bzw. (b) gut zu erkennen. Durch Peltierelemente, die abwechselnd die Kacheln des Schachbrettes beheizen bzw. kühlen, ist auch im Bild der Ferninfrarotkamera (c) ein guter Kontrast gewährleistet.

Der Aufbau eines solchen Kalibrierkörpers aus Peltierelementen ist in [Nev05] ausführlich beschrieben und wird als gegeben vorausgesetzt. An dieser Stelle werden nur die wichtigsten Eigenschaften beschrieben: Das Schachbrett besteht aus 64 Aluminiumquadern, die abwechselnd schwarz und weiß lackiert sind (siehe Abb. 3.1a). Im Videobild ergibt sich somit ein guter Kontrast. Um auch im Ferninfrarotbild einen guten Kontrast zu erreichen, wurden an den Rückseiten der schwarzen Quader heizende Peltierelemente angebracht und analog dazu an weißen Flächen kühlende Elemente befestigt. Abbildung 3.1 zeigt eine Szene einer Kalibrierung mit einem Schachbrett aus Peltierelementen. Im Ferninfrarot- wie auch im Videobild sind die wichtigen Kanten zwischen den einzelnen Feldern deutlich sichtbar.

Mit der „Calibration Toolbox“ werden zuerst für jede Kamera einzeln die intrinsischen Parameter bestimmt. Hierfür haben sich 20 bis 30 Posen des Schachbretts in unterschiedlichen Entfernungen als ausreichend erwiesen. In einem zweiten Schritt werden die extrinsischen Parameter ermittelt. Die verwendete Toolbox ist allerdings auf zwei Kameras beschränkt. Um dennoch mehrere Kameras extrinsisch kalibrieren zu können, werden für alle Kameras Paare mit einem festen Partner, der später als Referenzkoordinatensystem  $C_1$  dient, gebildet. Die extrinsische Kalibrierung berechnet anschließend



für jedes Paar die Rotationsmatrix  $\mathbf{R}$  und den Translationsvektor  $\mathbf{t}$  in Bezug auf das Referenzkoordinatensystem  $C_1$ . Alle Kameras sind nun in Bezug auf  $C_1$  registriert. Die Schätzung der intrinsischen Parameter wird im Allgemeinen durch die zusätzlichen Informationen aus der extrinsischen Kalibrierung verbessert. Die so erhaltenen Projektionsmatrizen werden für eine FIR- und zwei Videokameras in den folgenden Kapiteln als  $\mathbf{P}_{fir}$ ,  $\mathbf{P}_{video1}$  und  $\mathbf{P}_{video2}$  bezeichnet.

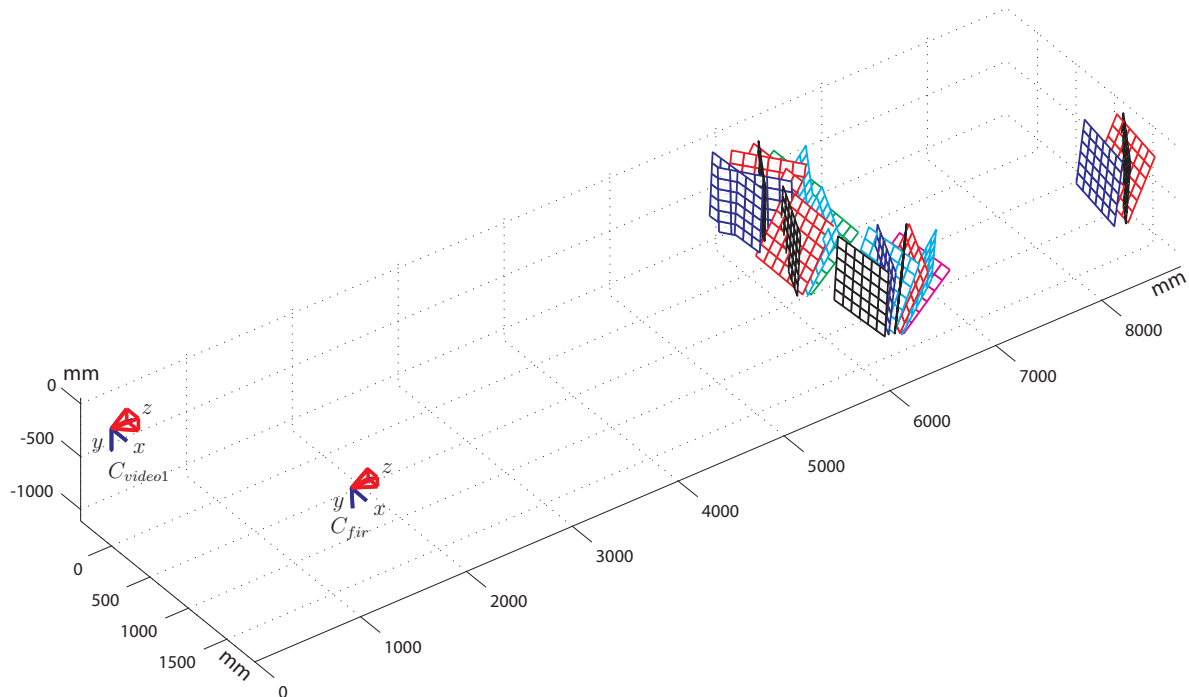


Abbildung 3.2: Extrinsische Kalibrierung der Versuchskameras  $C_{video1}$  (Videokamera) und  $C_{fir}$  (Wärmebildkamera) mittels eines Schachbrettes aus Peltierelementen. Die Abbildung gibt zum Einen die räumliche Anordnung (extrinsische Kalibrierung) der beiden Kameras zueinander wieder und zum Anderen die unterschiedlichen Positionen des Schachbrettes während der Durchführung der Kalibrierung.

### Punktlichtquelle

Verfahren zur Punktkalibrierung benötigen keine geometrischen Kalibrierkörper (z.B. Schachbrettkörper wie in Abschnitt 3.1.2), sondern lediglich eine Vorrichtung, durch die einzelne korrespondierende Punkte in den beteiligten Kameras erzeugt werden können. Diese Verfahren kommen vor Allem dann zum Einsatz, wenn ein geometrischer Kalibrierkörper nicht zeitgleich sichtbar in allen Kamerabildern positioniert werden kann, oder viele Kameras zueinander kalibriert werden sollen. In einer Publikation von Svoboda [SMP05] wird ein automatisches Punktkalibrierungsverfahren vorgestellt, womit die intrinsischen und extrinsischen Parameter einer unbegrenzten Menge (mindestens drei) von visuellen Kameras ermittelt werden können. Über einen modifizierten Laserpointer werden durch den Benutzer korrespondierende Punkte in allen Kamerabildbereichen erzeugt und mittels Bildverarbeitungsalgorithmen automatisch extrahiert. Alle Punkte werden anhand von Epipolarbedingungen validiert und

einander zugeordnet. Dabei ist es nicht notwendig, dass die Punkte stets in allen Kameras sichtbar sind. Vielmehr sollte das Sichtvolumen der Kameras ausgefüllt und eine ausreichende Überlappung vorhanden sein.

In den folgenden Abschnitten wird der Versuch beschrieben, die Vorgehensweise von Svoboda auf multispektrale bildgebende Sensoren zu übertragen. Dafür musste eine neue Kalibriervorrichtung erstellt werden, um Korrespondenzen ebenso in Wärmebildkameras zu generieren. Der Kalibrierungsvorgang selbst wurde nach der Nutzeranleitung von Svoboda [Svo03] und mit Hilfe des zugehörigen, leicht modifizierten und öffentlich zugänglichen Quellcodes durchgeführt.

**Kalibriervorrichtung** Bei der Kalibrierung von multispektralen Kameras muss darauf geachtet werden, dass korrespondierende Punkte in allen Kamerabildern zur Verfügung stehen. Sind im bildgebenden Sensorsystem Wärmebildkameras integriert, ist ein Laserpointer ungeeignet zur Korrespondenzerzeugung, da von diesem nicht ausreichend Wärmeenergie abgestrahlt wird. Deshalb wurde eine neuartige prototypische Kalibriervorrichtung erstellt, die in Abbildung 3.3a zu sehen ist. Durch einen von innen beleuchteten Kunststoffdiffusor wird sowohl Licht- als auch ausreichend Wärmestrahlung emittiert. Der Diffusor wurde kugelförmig gewählt, um in allen Kamerabildern eine elliptische Abbildung zu gewährleisten. Dadurch kann mittels einer subpixelgenauen Ellipsenextraktion im Bildbereich das Zentrum des Diffusors als korrespondierender Punkt verwendet werden. In den Abbildungen 3.3b und 3.3c ist zu sehen, wie sich die Kalibriervorrichtung in den multispektralen Kamerabildern darstellt.

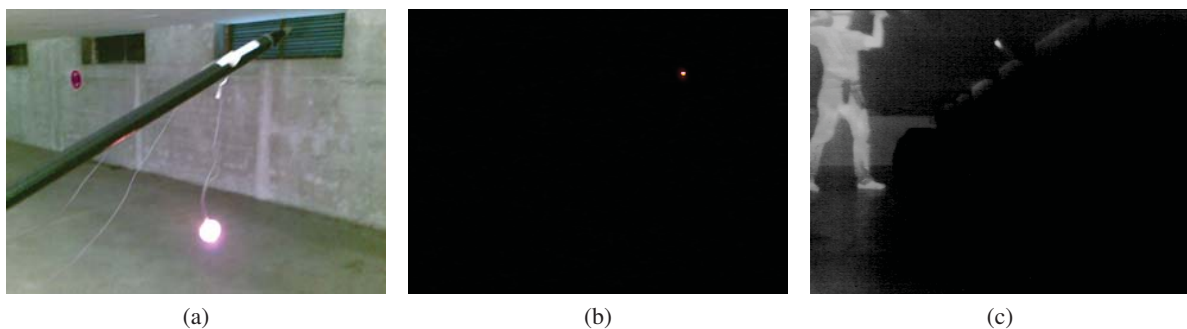


Abbildung 3.3: (a) Prototypische Kalibriervorrichtung für die Punktkalibrierung von multispektralen Kamerasystemen. An einem Stab ist eine aktive Leuchtquelle befestigt, die von einem kugelförmigen Kunststoffdiffusor umgeben ist. Es wird sowohl Licht- als auch Wärmestrahlung emittiert. (b) Abbildung der Kalibriervorrichtung in der Videokamera und (c) Wärmebildkamera. Im Wärmebild ist deutlich der Benutzer zu erkennen.

**Durchführung und Ergebnis** Mit der vorgestellten Kalibriervorrichtung können korrespondierende Punkte in multispektralen Kamerabildern generiert werden. Obwohl dieser Arbeitsschritt manuell durchgeführt werden muss und mindestens 100 Aufnahmen in verschiedenen Entfernungen benötigt werden, ist lediglich ein geringer Zeitaufwand erforderlich. In einer dunklen Umgebung oder bei entsprechenden Kameraeinstellungen (verminderte Verschlusszeiten) können diese Punkte für visuelle Kameras automatisch segmentiert werden. Im Wärmebild ist die Segmentierung aufwändiger, da der Benutzer, der die Kalibrierung vornimmt, zusätzliche Intensitätswerte erzeugt (siehe Abbildung 3.3c).



In Abbildung 3.4 ist das Ergebnis einer durchgeführten Punktkalibrierung zu sehen. Die Datenpunkte spiegeln die rückprojizierten Positionen des Kalibriergeräts wider. Darüber hinaus ist die Position und Blickrichtung der Kameras dargestellt.

Die erzielten Genauigkeiten sind allerdings unzureichend. Dies liegt nicht an der Kalibriervorrichtung oder an den Segmentierungsschritten, sondern vielmehr an der ungünstigen Verbaukonstellation der Kameras. So verlaufen die optischen Achsen der drei Versuchsträgerkameras nahezu parallel zueinander (siehe auch Abbildung 2.5). Diese Konstellation kann zu numerischen Instabilitäten führen und sollte für eine Punktkalibrierung vermieden werden [KTÅ00]. Für andere multispektrale Kameraanordnungen (z.B. im Fahrzeuginnenraum) ist die Kalibriervorrichtung möglicherweise gut geeignet. Für die vorliegende Verbaukonstellation ist das Verfahren jedoch nicht zweckmäßig.

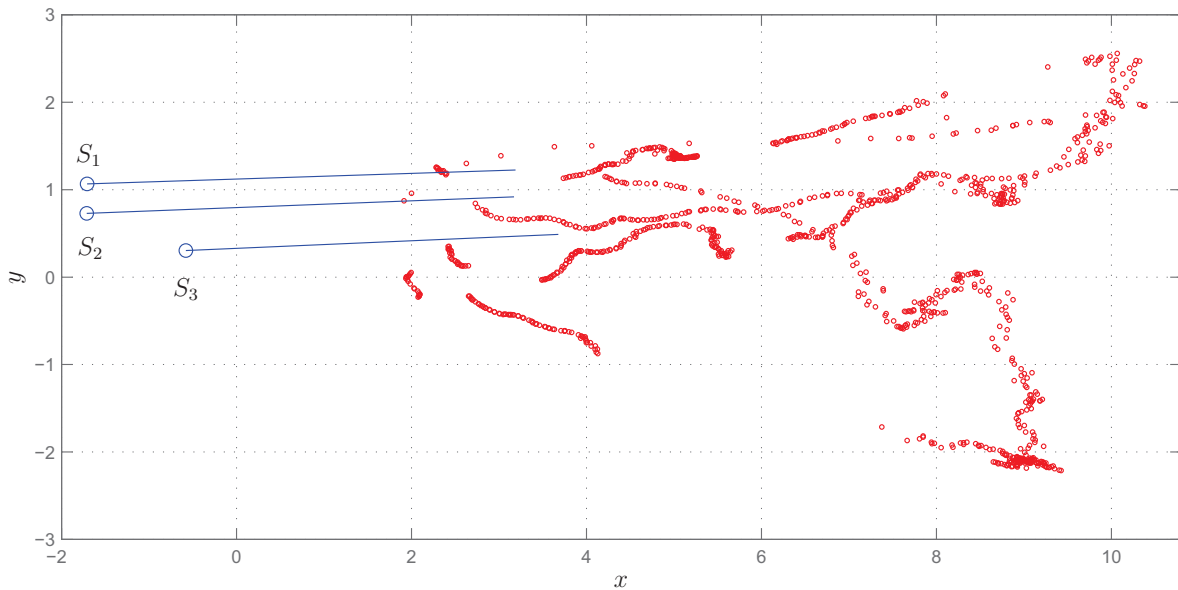


Abbildung 3.4: Beispielkalibrierung der drei Versuchskameras  $S_1$  (Videokamera),  $S_2$  (Videokamera mit Teleobjektiv) und  $S_3$  (Wärmebildkamera) mittels der Punktkalibrierung. Die Datenpunkte spiegeln die rückprojizierten Positionen des Kalibriergeräts wider.

### 3.1.3 Bewertung

In einem multispektralen Kamerasystem werden besondere Anforderungen an die Kalibrierungsverfahren gestellt. So müssen zum Beispiel spezielle auf die jeweiligen Spektralbereiche abgestimmte Kalibrierkörper verwendet werden, um die intrinsischen Parameter der einzelnen Kameras bestimmen zu können. Für eine extrinsische Kalibrierung ist es zusätzlich erforderlich, dass der Kalibrierkörper in beiden Spektralbereichen gleichzeitig verwendet werden kann. Für ein System bestehend aus Wärmebild- und Videokamera wurden zwei Verfahren vorgestellt, die diesen Ansprüchen genügen. Ein Schachbrett aus Peltierelementen überträgt das Grundprinzip der weitverbreiteten Kalibrierung mit einem Schachbrett in den ferninfraroten Spektralbereich. Bestehende Verfahren zur intrinsischen und extrinsischen Kalibrierung können somit direkt angewendet werden. Die erzielten Genauigkeiten sind für die Ansprüche einer Fahrumfelderfassung ausreichend. Eine präzise Fertigung eines derartigen Kalibrierkörpers ist jedoch sowohl zeit- als auch kostenintensiv. Darüber hinaus müssen die

Daten manuell segmentiert werden und es ist ein mehrstufiger Prozess zur extrinsischen Kalibrierung erforderlich. Das zweite Verfahren mit einer punktförmigen Lichtquelle bietet in diesen Punkten entscheidende Vorteile. In einem vollautomatischen Prozess können beliebig viele Kameras kalibriert werden. Einzige Anforderung an den Kalibrierkörper ist, dass sowohl Licht- als auch Wärmestrahlung emittiert werden, die ein punktförmiges Abbild im Sensor erzeugen. Dies wurde durch einen kugelförmigen Kunststoffdiffusor und einer handelsüblichen Lichtquelle erreicht. Durch diesen einfachen Aufbau ist eine mobile und einfache multispektrale Kamerakalibrierung möglich.

Allerdings können im Gegensatz zu dem Schachbrettkalibrierkörper keine zusätzlichen geometrischen Eigenschaften (Zellenbreite und planarer Körper) herangezogen werden. Bei ungünstigen Verbaukonstellationen der Kameras, kann dies zu Mehrdeutigkeiten und somit zu numerischen Instabilitäten führen [KTÅ00]. Nahezu parallele Kamerasichtachsen, wie sie im Rahmen dieser Arbeit zum Einsatz kommen, führen zu einer diesen ungünstigen Kameraanordnungen. Somit konnte im Vergleich zum Schachbrettkalibrierkörper nur eine erheblich geringere Genauigkeit der Kalibrierung erreicht werden.

## 3.2 Referenzdatengewinnung

Die Position und Orientierung von Referenzgegenständen in Bezug auf das eigene Fahrzeug wird sehr häufig benötigt. So lassen sich zum Beispiel aus diesen Informationen Latenzzeiten oder die extrinsischen Parameter der verwendeten Sensorik bestimmen. Im Folgenden wird ein Verfahren vorgestellt, mit dem sehr präzise Referenzdaten erzeugt werden können.

### 3.2.1 Inertialplattform mit Carrier Phase Differential GPS

Grundlage des Verfahrens sind zusätzliche zeitlich synchronisierte Sensoren, die sowohl im Eigenfahrzeug, als auch im Referenzobjekt verbaut sind. Im Eigenfahrzeug befindet sich eine Inertialplattform<sup>4</sup> (RT3003 [Oxf08]), die mittels Carrier Phase Differential GPS<sup>5</sup> (CPD-GPS) die Ausrichtung nach Norden  $\phi_{RT}^{GPS}$ , sowie die Position  $\mathbf{p}_{RT}^{GPS} \in \mathbb{R}^3$  bezüglich eines globalen Weltkoordinatensystems bestimmt. Durch die Kombination aus Inertialplattform und CPD-GPS können zum Einen Drifts in den Gyroskopen und zum Anderen Positionsschwankungen in der GPS basierten Ortbestimmung kompensiert werden. Damit stehen sowohl in statischen als auch sehr dynamischen Situationen sehr präzise Messungen zur Verfügung. Für eine weiterführende Betrachtung zur Fusion von Gyroskopen sei auf [PHK08] verwiesen.

Im Allgemeinen entspricht der Verbauort der Inertialplattform bzw. deren Koordinatensystem  $C_{RT}$  nicht dem Fahrzeugkoordinatensystem  $C_{ego}$ . In den nachfolgenden Berechnungen wird dies durch die zusätzliche Transformation  $\mathbf{T}_{RT}^{ego}$  ausgeglichen, die in einem eigenständigen Kalibrierungsprozess bestimmt wird. Eine schematische Darstellung des Versuchsaufbaus bzw. der beteiligten Koordinatensysteme ist in Abbildung 3.5 zu sehen. Eine zweite Inertialplattform mit CPD-GPS Empfänger, die am Zielobjekt angebracht ist, liefert Daten über die Position  $\mathbf{p}_{ref}^{GPS} \in \mathbb{R}^3$  und die Ausrichtung  $\phi_{ref}^{GPS}$  des Referenzobjektes. Mit dem zusätzlichen Sensor können weitere Attribute wie die Geschwindigkeit  $v_{ref}$ ,

<sup>4</sup> Eine Inertialplattform besteht aus mehreren Gyroskopen und Beschleunigungssensoren. Durch eine Kombination mit einem CPD-GPS Empfänger können die Position, Geschwindigkeit, Beschleunigung sowie die Orientierung der Plattform sehr präzise gemessen werden.

<sup>5</sup> Satellitengestütztes System zur Bestimmung der Position auf der Erdoberfläche.

die Beschleunigung  $a_{ref}$  und die Gierrate  $\dot{\phi}_{ref}$  gemessen werden. Durch die zeitliche Synchronisation zwischen dem Referenzobjekt und dem Eigenfahrzeug ist es letztendlich möglich aufgezeichnete Sensordaten mit den Messwerten des Referenzsensors zu attributieren. Für eine Betrachtung der Qualität der zeitlichen Synchronisation sei auf Abschnitt 3.2.2 verwiesen.

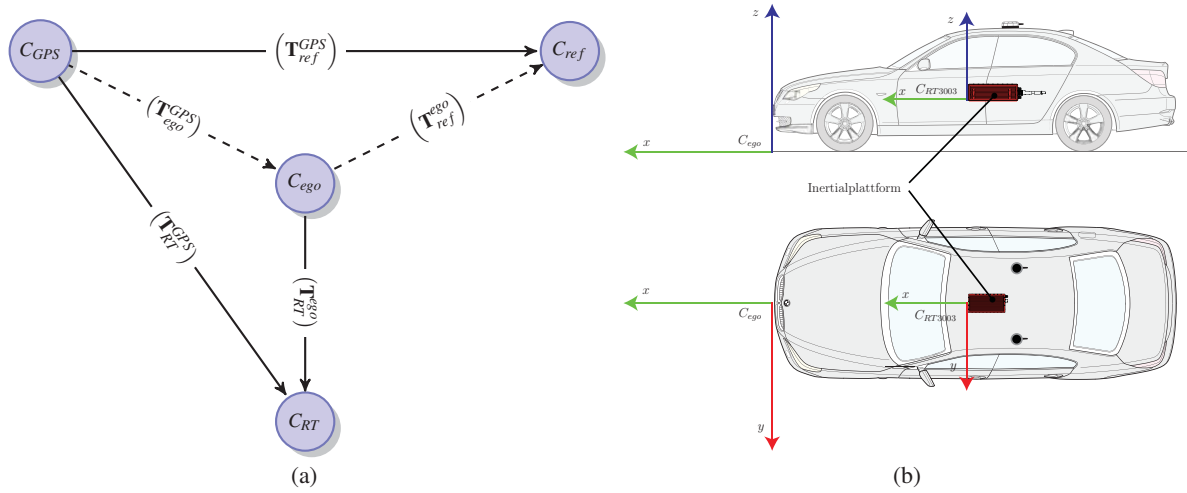


Abbildung 3.5: Referenzdatengewinnung mit einem dedizierten Sensor (Inertialplattform mit CPD-GPS Empfänger). In Abbildung (a) ist der Spatial-Relationship-Graph für den Versuchsaufbau dargestellt. Die beiden Kanten  $(\mathbf{T}_{ego}^{GPS})$  bzw.  $(\mathbf{T}_{ref}^{GPS})$  entsprechen den Messungen der Referenzsensorik für die Position und Orientierung des Eigenfahrzeuges bzw. des Zielobjektes. Über die Kante  $(\mathbf{T}_{ref}^{ego})$  lässt sich die Position des Zielobjektes bezüglich des Eigenfahrzeuges bestimmen. In Abbildung (b) ist die Verbauposition der dedizierten Sensorik im Eigenfahrzeug abgebildet.

Aus den Messwerten der dedizierten Sensorik können Zustandsgrößen wie die Orientierung, die Geschwindigkeit, die Beschleunigung und die Gierrate direkt abgeleitet werden. Die Position des Referenzobjektes  $\mathbf{p}_{ref}^{ego}$  relativ zu dem Koordinatensystem des Eigenfahrzeug  $C_{ego}$  lässt sich aus den Messwerten  $\phi_{RT}^{GPS}$ ,  $\mathbf{p}_{RT}^{GPS}$ ,  $\mathbf{p}_{ref}^{GPS}$  und der starren Transformation  $\mathbf{T}_{RT}^{ego}$  wie folgt berechnen:

$$\mathbf{R} = \begin{pmatrix} \cos(\phi_{RT}^{GPS}) & -\sin(\phi_{RT}^{GPS}) & 0 \\ \sin(\phi_{RT}^{GPS}) & \cos(\phi_{RT}^{GPS}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.9)$$

$$\mathbf{T}_{RT}^{GPS} = \begin{bmatrix} \mathbf{R} & \mathbf{p}_{RT}^{GPS} \\ 0 & 1 \end{bmatrix} \quad (3.10)$$

$$\mathbf{p}_{ref}^{ego} = \mathbf{T}_{RT}^{ego} (\mathbf{T}_{RT}^{GPS})^{-1} \mathbf{p}_{ref}^{GPS} \quad (3.11)$$

### 3.2.2 Fehlerabschätzung

Grundlage für die Abschätzung der Güte der ermittelten Referenzdaten sind die Genauigkeiten der verwendeten Inertialplattformen. In Tabelle 3.1 sind alle wichtigen Messgrößen mit dem zugehörigen RMSE aufgelistet.

	Position	Geschwindigkeit	Beschleunigung	Orientierung	Gierrate
	$\mathbf{p}$ (RMS)	$v$ (RMS)	$a$ (RMS)	$\phi$ (RMS)	$\dot{\phi}$ (RMS)
Referenzsensor	0.02 m	$0.02 \frac{m}{s}$	$0.01 \frac{m}{s^2}$	$0.1^\circ$	$0.1 \frac{^\circ}{s}$

Tabelle 3.1: Genauigkeiten der verwendeten Inertialplattform *RT3003* für das Referenzobjekt sowie das Eigenfahrzeug.

Der zu erwartende Fehler für die Orientierung, Geschwindigkeit und die Beschleunigung kann direkt aus den Herstellerangaben für die Referenz übernommen werden (siehe Tabelle 3.1). Für eine Abschätzung der Ortsungenauigkeit muss allerdings berücksichtigt werden, dass die Fehler untereinander korreliert sind. Nach 3.11 ist die relative Position der Referenz abhängig von der GPS-Position und der Orientierung des Eigenfahrzeugs sowie der GPS-Position der Referenz. Deshalb ist der zu erwartende Fehler abhängig von der Positionsungenauigkeit beider Inertialplattformen, sowie der Qualität der Orientierungsbestimmung des Eigenfahrzeuges. Dieser Fehler kann allgemein mit dem Modell der Fehlerfortpflanzung bestimmt werden.

#### Fehlerfortpflanzung

Die im vorherigen Abschnitt 3.2.1 beschriebene Methode zur Gewinnung von Referenzdaten basiert ausschließlich auf einfachen affinen Koordinatentransformationen. Durch die Anwendung eines Fehlerfortpflanzungsmodells soll im Folgenden der Einfluss dieser Transformationen auf die zu erwartenden Genauigkeiten der Referenzdaten bestimmt werden. Die Fehlerfortpflanzung beschreibt den Fehler, der durch eine Funktion  $f$  entsteht, die auf Fehlerbehafteten Variablen beruht. Im Folgenden wird von Gaussverteilten Fehlern mit der zugehörigen Kovarianzmatrix  $\Sigma$  ausgegangen. Für eine ausführliche Betrachtung von Fehlermodellen sei auf [Bau07] verwiesen.

Für lineare Funktionen, die sich auch in einer Matrix  $\mathbf{A}$  darstellen lassen, kann der resultierende Fehler wie folgt beschrieben werden:

$$\Sigma_f = \mathbf{A} \Sigma \mathbf{A}^T \quad (3.12)$$

Bei nicht linearen Funktionen weicht der resultierende Fehler häufig signifikant von einer Gaussverteilung ab. Trotzdem kann für kleine Fehler näherungsweise von einer Gaussverteilung ausgegangen werden. Im Folgenden sei  $v$  ein Zufallsvektor aus  $\mathbb{R}^M$  mit dem Mittelwert  $\bar{v}$  und der zugehörigen Kovarianzmatrix  $\Sigma$ . Wenn zusätzlich für die nicht lineare Funktion  $f: \mathbb{R}^M \mapsto \mathbb{R}^N$  angenommen werden kann, dass sie in der Nachbarschaft von  $\bar{v}$  differenzierbar und annähernd linear ist, dann ist der resultierende Fehler  $\Sigma_f$  in erster Näherung bestimmt durch:

$$\Sigma_f = \mathbf{J}_f \Sigma \mathbf{J}_f^T \quad (3.13)$$

Hierbei wird die Jacobi-Matrix  $\mathbf{J}_f$  der Funktion  $f$  an der Stelle  $\bar{v}$  ausgewertet.

Eine weitere Möglichkeit, den Fehler für beliebig komplexe Funktionen bzw. Verteilungen zu berechnen, ist die Monte Carlo Simulation [RC04]. Entsprechend der zugrundeliegenden Fehlerverteilung werden Stichproben  $x_i$  gezogen und der zugehörige Funktionswert  $f(x_i)$  gebildet. Aus dem Ergebnis der Abbildung kann eine neue Verteilungsfunktion bestimmt werden. Die Qualität dieser Approximation ist allerdings abhängig von der Anzahl der verwendeten Stichproben.

Die im vorherigen Abschnitt beschriebenen Modelle zur Fehlerfortpflanzung geben dem Entwickler Hilfsmittel an die Hand, um für beliebige Funktionen den zu erwartenden Fehler zu bestimmen. Die in Abschnitt 3.2.1 beschriebene Methode zur Gewinnung von Referenzdaten basiert ausschließlich auf einfachen affinen Koordinatentransformationen. Deshalb wird im Folgenden für diese spezielle Art von Funktionen beschrieben, wie die vorgestellten Hilfsmittel verwendet werden können, um die Auswirkungen der Transformationen auf die jeweiligen Fehler zu bestimmen. An dieser Stelle werden nur die wichtigsten Aspekte des Verfahrens vorgestellt. Eine ausführliche Beschreibung findet sich in [Bau07].

Eine affine Koordinatensystemtransformation  $\mathbf{T} = [\mathbf{R}|t]$  ist im Allgemeinen aus einer Rotation  $\mathbf{R}$  sowie einer Translation  $t$  zusammengesetzt. Im Folgenden wird davon ausgegangen, dass der zugehörige Rotations- bzw. Translationsfehler  $(\Delta r, \Delta t)$  normalverteilt ist und durch die Kovarianzmatrix  $\Sigma \in \mathbb{R}^{6 \times 6}$  mit einem Mittelwert von 0 bestimmt ist. In Abbildung 3.6 sind alle für die Referenzdatengewinnung benötigten Transformationen dargestellt. Die Position und Orientierung des Eigenfahrzeuges ist durch eine fehlerbehaftete Transformation  $\mathbf{T}_{ego}^{GPS}$  aus dem GPS Koordinatensystem  $C_{GPS}$  bestimmt. Analog dazu ist die Position der Referenz gegeben durch die Transformation  $\mathbf{T}_{ref}^{GPS}$ . Zusammen lässt sich die Position der Referenz bezüglich des Eigenfahrzeuges durch folgende Transformation berechnen:

$$\mathbf{T}_{ref}^{ego} = (\mathbf{T}_{ego}^{GPS})^{-1} \circ \mathbf{T}_{ref}^{GPS} \quad (3.14)$$

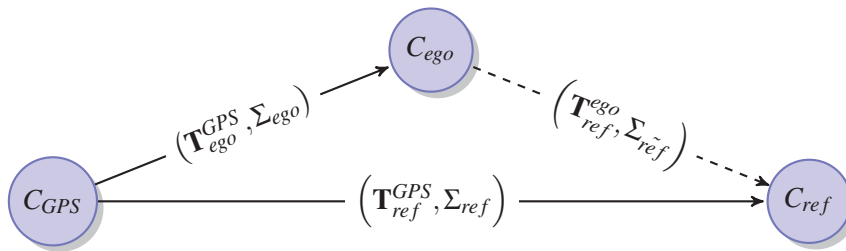


Abbildung 3.6: Spatial-Relationship-Graph für die Berechnung von Referenzdaten mit zwei CPD-GPS Empfängern. Die Position des Referenzobjekts in Bezug auf das eigene Koordinatensystem ist durch die Transformationen  $\mathbf{T}_{ref}^{GPS}$  sowie  $\mathbf{T}_{ego}^{GPS}$  eindeutig bestimmt.

Der resultierende Fehler  $\Sigma_{ref}$  dieser Transformation ist durch die Invertierung ( $\mathbf{T}^{-1}$ ) und Konkatination ( $\mathbf{T}_1 \circ \mathbf{T}_2$ ) von einzelnen Transformationen bestimmt. Im Allgemeinen können aus diesen beiden Grundoperationen beliebig komplexe affine Transformationen zusammengesetzt werden. Deshalb ist es ausreichend, die Fehlerfortpflanzung für diese Basistransformationen zu bestimmen.

**Invertierung von Transformationen** Gegeben sei eine Transformation  $\mathbf{T} = [\mathbf{R}|t]$  mit der zugehörigen Kovarianzmatrix  $\Sigma$ . Der resultierende Fehler  $\Sigma_I$  der in Abbildung 3.7 dargestellten Invertierung  $\mathbf{T}^{-1}$  lässt sich wie folgt berechnen:

$$\Sigma_I = \mathbf{J}_{[-\mathbf{R}^T|t]} \Sigma \mathbf{J}_{[-\mathbf{R}^T|t]}^T \quad (3.15)$$

Die Jacobimatrix  $\mathbf{J}_{[\mathbf{R}|t]}$  ist hierbei wie folgt definiert:

$$\mathbf{J}_{[\mathbf{R}|t]} = \begin{bmatrix} \mathbf{I} & \text{Skew}(t) \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \quad (3.16)$$

$$\text{Skew}(t) = \text{Skew}((x_t, y_t, z_t)^T) = \begin{pmatrix} 0 & z_t & -y_t \\ -z_t & 0 & x_t \\ y_t & -x_t & 0 \end{pmatrix} \quad (3.17)$$

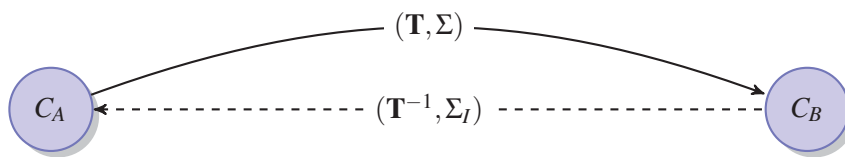


Abbildung 3.7: Die Invertierung der Transformation  $\mathbf{T}$  mit dem zugehörigen Fehler  $\Sigma$  ergibt die neue Transformation  $\mathbf{T}^{-1}$  mit dem neuen Fehler  $\Sigma_I$ .

**Konkatenation von Transformationen** Gegeben sind zwei Transformationen  $\mathbf{T}_1 = [\mathbf{R}_1|t_1]$  bzw.  $\mathbf{T}_2 = [\mathbf{R}_2|t_2]$  mit den zugehörigen Kovarianzmatrizen  $\Sigma_1$  und  $\Sigma_2$ . Der resultierende Fehler  $\Sigma_C$  der Konkatenation ( $\mathbf{T}_1 \circ \mathbf{T}_2$ ) lässt sich wie folgt berechnen:

$$\Sigma_C = \mathbf{R}_1 \left( \mathbf{J}_{[\mathbf{R}_2|t_2]} \mathbf{R}_1^T \Sigma_1 \mathbf{R}_1 \mathbf{J}_{[\mathbf{R}_2|t_2]}^T + \Sigma_2 \right) \mathbf{R}_1^T \quad (3.18)$$

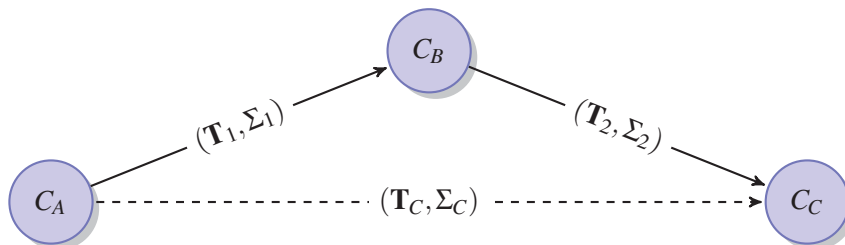


Abbildung 3.8: Die Konkatenation der Transformationen  $\mathbf{T}_1$  und  $\mathbf{T}_2$  mit den zugehörigen Fehlern  $\Sigma_1$  und  $\Sigma_2$  ergibt die neue Transformation  $\mathbf{T}_C$  mit  $\Sigma_C$ .

### Abschätzung der Ortsungenaugkeit

Im vorherigen Abschnitt wurden die mathematischen Grundlagen gelegt, um eine Fehlerabschätzung der ermittelten Referenzdaten durchführen zu können. Je nach Versuchsaufbau können sämtliche Latenzzeiten der Sensoren sowie der Aufzeichnung vernachlässigt werden. So ist für statische Versuche zum Beispiel die Schätzung der Position und der Orientierung des Eigenfahrzeugs sowie des Referenzobjektes die einzige Fehlerquelle. Mit den Gleichungen 3.15 und 3.18 kann die resultierende Ungenauigkeit der Referenzdaten bestimmt werden.

Für unterschiedliche Positionen  $\mathbf{x}_{ref}$  des Referenzobjektes ist in der Tabelle 3.2 der zu erwartende Fehler für die Entfernung  $x$  und die laterale Ablage  $y$  dargestellt. Der laterale Fehler nimmt fast linear mit der Entfernung  $r$  zu. Durch die starke Abhängigkeit von der Orientierungsschätzung ist dieses Verhalten sehr ähnlich zu der Umrechnung von Polar- in Kartesische Koordinaten mit dem einzigen Unterschied, dass hier zusätzlich die Fehler in der Wahl des Koordinatenursprungs berücksichtigt werden.

Position Referenzobjekt	$\sigma_x$	$\sigma_y$
(0m, 25m)	0.03 m	0.05 m
(0m, 50m)	0.03 m	0.09 m
(0m, 100m)	0.03 m	0.17 m
(30m, 100m)	0.06 m	0.17 m

Tabelle 3.2: Genauigkeiten der Referenzdatengewinnung in unterschiedlichen Entfernungen zum Eigenfahrzeug. Die angegebenen Fehler wurden durch das Modell einer Fehlerfortpflanzung bestimmt und beziehen sich auf das kartesische Koordinatensystem des Eigenfahrzeuges.

Für nicht stationäre Versuchsaufbauten ist die zeitliche Synchronisation der beiden GPS Empfänger entscheidend. Die in dieser Arbeit verwendeten CPD-GPS Empfänger liefern alle 10 ms Messwerte über Position und Orientierung, die mit der globalen GPS Zeit synchronisiert sind. Mit einem einfachen Versuchsaufbau konnte gezeigt werden, dass beide Empfänger sehr genau ( $\leq 3$  ms) synchronisiert sind: Hierzu fahren zwei Fahrzeuge mit identischer und konstanter Geschwindigkeit  $v \geq 50 \frac{\text{m}}{\text{s}}$  auf einer Teststrecke. Mit dem Laserscanner wird die Entfernung zu dem vorausfahrenden Fahrzeug gemessen und mit der Referenz verglichen. Da die Relativgeschwindigkeit beider Fahrzeuge annähernd Null ist, können sämtliche Latenzzeiten des Laserscanners vernachlässigt werden. Für die Daten des Laserscanners werden der Mittelwert und die Standardabweichung des Fehlers bestimmt und mit den Werten einer stationären Aufnahme verglichen. Ein Unterschied in der Synchronisation der beiden GPS-Empfänger von 3 ms würde einem Offset von 0.15 m in der relativen Entfernung zu dem Referenzfahrzeug entsprechen. Ein derartiger Effekt konnte in den durchgeführten Messungen nicht bestätigt werden. Somit beträgt die maximale Latenz zwischen den beiden CPF-GPS Empfängern weniger als 3 ms und kann für die im Rahmen dieser Arbeit untersuchten Geschwindigkeitsbereiche vernachlässigt werden.



### 3.3 Temporale Kalibrierung

Eine der Voraussetzungen für eine korrekte Fusion ist die Berücksichtigung von Latenzzeiten, die zum Beispiel durch einen asynchronen Systemaufbau, sensorinterne Vorverarbeitung oder durch die Kommunikation zwischen den Komponenten entstehen können (siehe auch Abbildung 3.9). Im Allgemeinen ist die genaue Zusammensetzung der Sensorlatenzzeiten nicht bestimmbar, da sie zum Einen sehr systemspezifisch sind und die Sensorhersteller zum Anderen häufig nicht die relevanten Informationen veröffentlichen. Werden diese Zeiten allerdings vernachlässigt, kann es insbesondere bei hohen Relativgeschwindigkeiten zu Widersprüchen zwischen den Messwerten der einzelnen Sensoren kommen. Im Folgenden wird beschrieben, wie zumindest der durchschnittliche Gesamtbeitrag aus Sensor- und Kommunikationslatenzzeit in einer „temporalen Kalibrierung“ experimentell bestimmt werden kann.

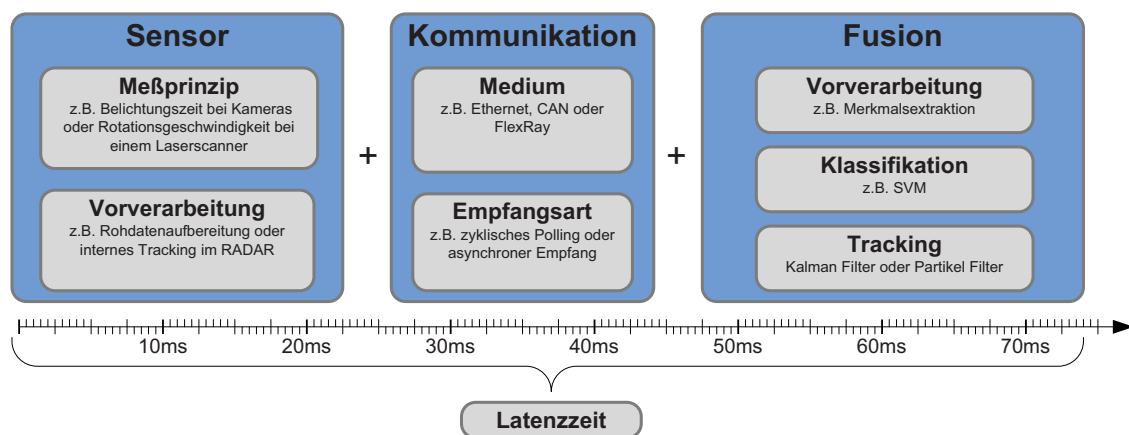


Abbildung 3.9: Mögliche Ursachen, die in einer Fahrumfelderfassung zu Latenzzeiten führen können. Die dargestellte Zeitskala dient nur zur Veranschaulichung und hat keinen Bezug zu der verwendeten Sensorik.

Grundlage für die Bestimmung der Gesamtlatenzzeit ist das in Abschnitt 3.2 beschriebene System zur Referenzdatengewinnung. Die DGPS-gestützte Inertialplattformen liefern alle 10 ms die aktuelle Position und Orientierung sowohl des eigenen Fahrzeuges als auch von beliebigen anderen Referenzobjekten. Die Ausgaben beider Plattformen sind präzise auf die GPS-Zeit synchronisiert und besitzen die im vorherigen Abschnitt beschriebenen Genauigkeiten von unter 2 cm. Dadurch, dass beide Plattformen synchronisiert sind, kann die durch die Sensorik ermittelte Position direkt mit der durch die Referenz bestimmten Position verglichen werden. Bei hohen Differenzgeschwindigkeiten ist ein signifikanter Unterschied zwischen gemessener und realer Position zu erwarten, der direkt proportional zur gesamten Latenzzeit ist. In Abbildung 3.10 ist dieser Versatz gut zu erkennen. Latenzzeiten können mit diesem Versuchsaufbau, bedingt durch die Abtastrate der Plattformen von 100 Hz, nur mit einer Genauigkeit von 10 ms ermittelt werden. Die Differenzgeschwindigkeit sollte so gewählt werden, dass der Einfluss der Latenzzeit im Vergleich zu dem Messrauschen signifikant ist.

Im Folgenden wird beschrieben wie sich experimentell für entfernungs- und bildgebende Sensorik die Latenzzeiten bestimmen lassen.



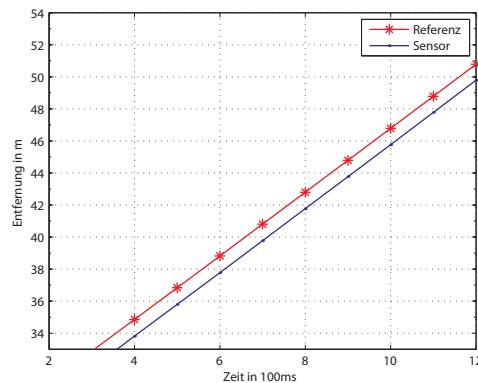


Abbildung 3.10: Einfluss der Latenzzeit auf die Entfernungsbestimmung. Das Referenzobjekt entfernt sich mit einer Differenzgeschwindigkeit von  $20 \text{ m/s}$  von dem eigenen Fahrzeug. Bedingt durch die Latenzzeit ist ein signifikanter Unterschied in der Entfernungsbestimmung des Laserscanners und der Referenz zu erkennen.

### 3.3.1 Entfernungsgebende Sensorik

Bei entfernungsgebender Sensorik lassen sich Entfernungen im Allgemeinen deutlich genauer messen als die laterale Ablage. Aus diesem Grund sollte für die Ermittlung der Latenzzeiten ein Szenario gefunden werden, in dem die Latenzzeit die Entfernungs-messung beeinflusst. Hierfür sind folgende zwei Versuchsaufbauten geeignet: In dem ersten Fall bewegt sich das Eigenfahrzeug mit konstanter Geschwindigkeit  $v$  auf ein stationäres Referenzobjekt zu. In dem zweiten Szenario werden die Rollen von Referenzobjekt und Eigenfahrzeug vertauscht. Im Ergebnis sind beide Varianten gleichwertig.

Im Folgenden sei  $f$  die Frequenz der zu untersuchenden entfernungsgebenden Sensorik. Dementsprechend wird alle  $\frac{1}{f}$  s eine neue Messung der Entfernung  $d_i^s$  des Referenzobjektes durchgeführt. Die Reihe  $F = (d_1^s, d_2^s, \dots, d_n^s)$  gibt den zeitlichen Verlauf dieser Messungen wieder. Analog dazu bezeichnet die Reihe  $G_r = (d_1^r, d_2^r, \dots, d_n^r)$  den realen Verlauf der Entfernung, der durch das Referenzsystem ermittelt wurde (siehe Abschnitt 3.2). Gesucht ist die zeitliche Verschiebung  $t_{delay}$ , die den Fehler zwischen den beiden Reihen  $F$  und  $G$  minimiert. Der Fehler  $E$  in Abhängigkeit der zeitlichen Verschiebung  $t$  ist dabei wie folgt definiert:

$$E(t) = \frac{1}{n} \sum_{i=1}^n (F(i) - G(i+t))^2 \quad t \in \mathbf{Z} \quad (3.19)$$

Unter der Annahme einer konstanten Differenzgeschwindigkeit  $v$  ist ein quadratischer Verlauf der diskreten Funktion  $E(t)$  zu erwarten. Die Funktion  $p \mapsto at^2 + bt + c$  ist demzufolge ausreichend, um den Verlauf von  $E$  anzunähern. Letztendlich entspricht das Minimum von  $p$  der gesuchten zeitlichen Verschiebung  $t_{delay}$ .

In Abbildung 3.11 ist das Ergebnis dieser Berechnung für den Laserscanner und die Radarsensoren aufgetragen. Es ist deutlich zu erkennen, dass die Radarsensorik eine größere Latenzzeit aufweist als der Laserscanner. Dies ist durch die komplexere interne Datenvorverarbeitung zu erklären.

Bedingt durch das Messprinzip des Laserscanners ist, in Abhängigkeit von dem Azimutwinkel des Referenzobjektes, eine unterschiedliche Latenzzeit zu erwarten. Diese Hypothese konnte in den durch-

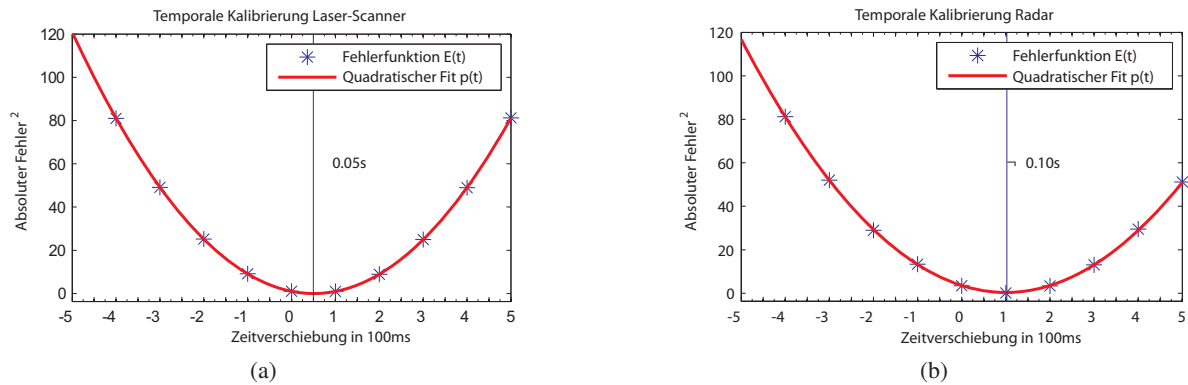


Abbildung 3.11: Zeitliche Kalibrierung entfernungsgebender Sensorik. Durch die diskrete Fehlerfunktion  $E(t)$  wird eine Parabel gelegt, deren Scheitelpunkt die optimale zeitliche Verschiebung  $t_{delay}$  bestimmt. In (a) ist die Fehlerfunktion für den Laserscanner aufgetragen, deren Minimum bei 50 ms liegt. Aus Abbildung (b) kann abgelesen werden, dass die Latenzzeit der Radarsensorik 100 ms beträgt.

geführten Messungen bestätigt werden. Die im weiteren Verlauf dieser Arbeit angegebene und verwendete Latenzzeit des Laserscanners ist auf einen Azimutwinkel von  $\phi = 0^\circ$  bezogen. In der Tabelle 3.3 sind für die verwendeten entfernungsgebenden Sensoren die entsprechenden Latenzzeiten dargestellt.

#### 3.3.2 Bildgebende Sensorik

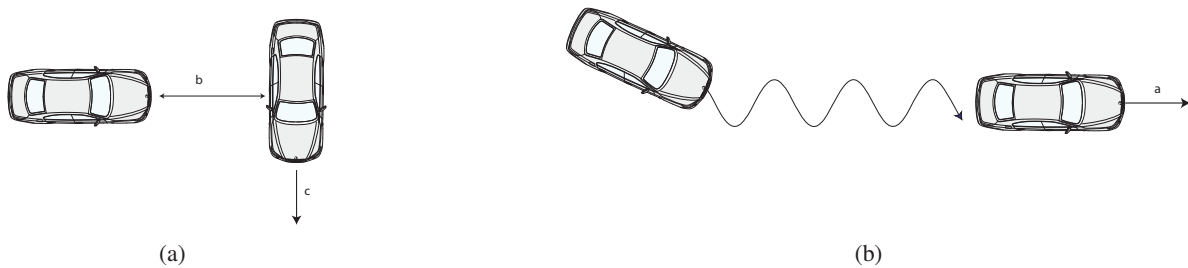


Abbildung 3.12: Zwei Versuchsaufbauten zur Bestimmung der Latenzzeit von bildgebender Sensorik. In Abbildung (a) bewegt sich das Referenzfahrzeug aus Sicht des Eigenfahrzeuges mit konstanter Geschwindigkeit von links nach rechts durch die Szene. Im zweiten Szenario, dargestellt in Abbildung (b), bewegt sich das Eigenfahrzeug auf einer sinusförmigen Trajektorie. Das Referenzfahrzeug fährt hingegen mit konstanter Geschwindigkeit geradlinig vor dem Eigenfahrzeug.

Im Gegensatz zur entfernungsgebenden Sensorik wie Radar oder Laserscanner besitzen Kameras nur bedingt die Fähigkeit Entfernungen zu messen. Um die Latenzzeiten von bildgebender Sensorik zu bestimmen kann der im vorherigen Abschnitt beschriebene Versuchsaufbau nicht verwendet werden,

da sich in diesem Szenario die Latenzzeit ausschließlich in einer Diskrepanz zwischen realer und gemessener Entfernung niederschlägt.

Im Folgenden werden deshalb zwei Versuchsaufbauten beschrieben, bei denen die laterale Ablage bzw. der Winkel zu dem Zielobjekt direkt von der Latenzzeit beeinflusst wird. Beide Szenarien sind in Abbildung 3.12 schematisch dargestellt.

**Stationäres Eigenfahrzeug** In diesem Versuchsaufbau bewegt sich das Referenzfahrzeug aus Sicht des stationären Eigenfahrzeuges mit konstanter Geschwindigkeit und näherungsweise konstantem Abstand von links nach rechts durch die Szene. Als Messung dient der Mittelpunkt  $c_{cam}$  des Referenzfahrzeuges im Bild (siehe Abbildung 3.13a). Die Diskrepanz zwischen gemessener und realer Position  $c_{ref}$  lässt sich analog zu dem im Abschnitt 3.3.1 beschriebenen Verfahren verwenden, um die Latenzzeit zu bestimmen. Bei diesem Verfahren sollte allerdings sichergestellt werden, dass der Einfluss der Latenzzeiten auf die Messung signifikant größer ist, als das zugehörige Messrauschen. Aus diesem Grund sollte für die Versuchsdurchführung ein möglichst geringer Abstand und eine hohe Differenzgeschwindigkeit gewählt werden. Allerdings ist die Zeitspanne, in der der Effekt der Latenzzeit beobachtet werden kann, hierzu gegenläufig. Als guter Kompromiss hat sich bei einem Öffnungswinkel der Kamera von  $40^\circ$  ein Abstand von 25 m und eine Differenzgeschwindigkeit von  $v_{diff} = 18 \frac{m}{s}$  erwiesen. In Abbildung 3.13b ist die Messung und das Ergebnis der Auswertung dargestellt.

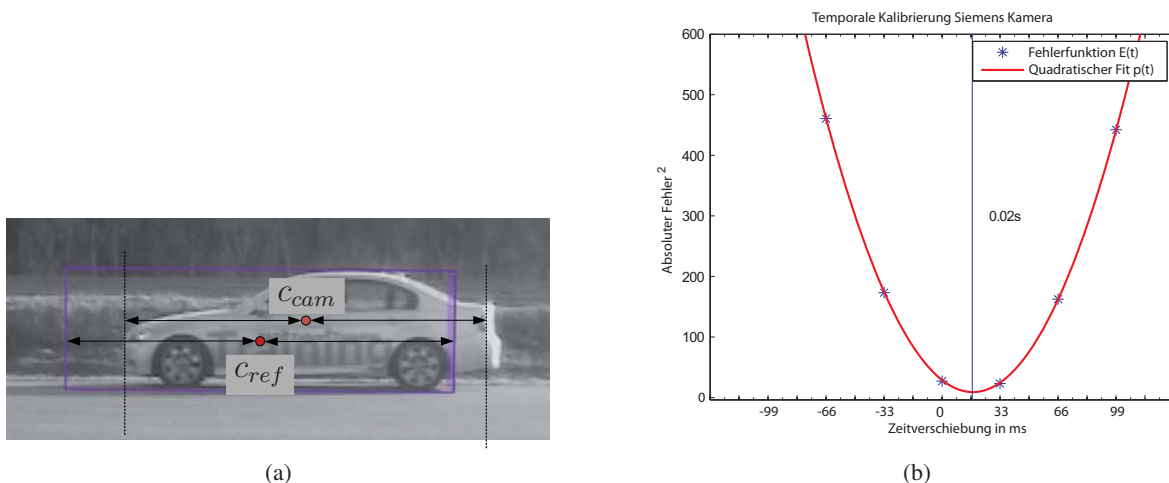


Abbildung 3.13: Bestimmung der Latenzzeit einer Videokamera für den stationären Versuchsaufbau. In Abbildung (a) ist der Einfluss der Latenzzeit auf die Videokamera zu sehen. Die reale Position des Referenzfahrzeuges wurde in das Bild projiziert (blauer Kasten mit zugehörigen Mittelpunkt  $c_{ref}$ ). In der Abbildung ist eine Diskrepanz zwischen der erwarteten und der gemessenen Position  $c_{cam}$  des Fahrzeuges zu erkennen. In Abbildung (b) ist die zugehörige Fehlerfunktion (siehe Abschnitt 3.3.1) für mehrere Messungen aufgetragen. Die Latenzzeit beträgt 20 ms.

**Sinusförmige Bewegung** In diesem Szenario bewegt sich das Eigenfahrzeug auf einer sinusförmigen Trajektorie mit konstanter Geschwindigkeit. Das Referenzfahrzeug hingegen fährt auf einer geradlinigen Spur mit konstantem Abstand vor dem Eigenfahrzeug. Für die weiteren Berechnungen wird mit zwei unterschiedlichen Methoden die Mitte der Rückansicht des vorausfah-

renden Fahrzeuges im Bild bestimmt. Zum Einen wird die Position manuell oder mit Verfahren der Mustererkennung im Bild gemessen. Für mehrere Bilder ergibt sich somit die Reihe  $F = (d_1^s, d_2^s, \dots, d_n^s)$ , die den zeitlichen Verlauf dieser Messungen wiedergibt. Zum Anderen wird die reale Position des Fahrzeughecks, die durch das Referenzsystem ermittelt wurde, in das Bild projiziert. Analog zu  $F$  bezeichnet die Reihe  $G_r = (d_1^r, d_2^r, \dots, d_n^r)$  den realen Verlauf dieser Projektion. In Abbildung 3.14b ist der zeitliche Verlauf von  $F$  und  $G$  dargestellt. Es ist gut zu erkennen, wie die sinusförmige Bewegung des Eigenfahrzeuges einen sinusförmigen Verlauf der Position im Bild erzeugt. Unter der Annahme einer korrekt kalibrierten Kamera sind die Graphen von  $F$  und  $G$  bis auf eine zeitliche Verschiebung identisch. Mit Hilfe der diskreten Kreuzkorrelation von  $F$  und  $G$  kann diese zeitliche Verschiebung berechnet werden. Die diskrete Kreuzkorrelation ist definiert als:

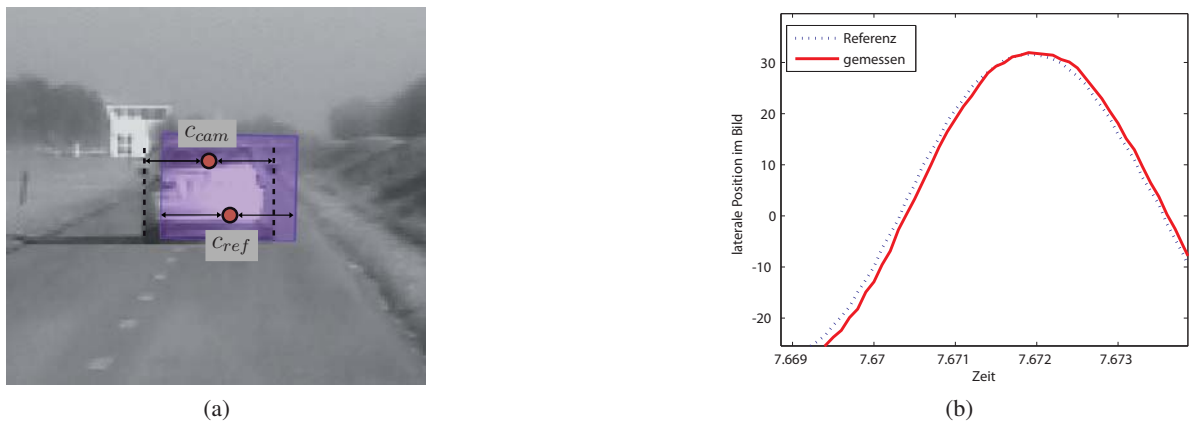


Abbildung 3.14: Bestimmung der Latenzzeit einer Videokamera anhand einer sinusförmigen Bewegung.

In Abbildung (a) ist der Einfluss der Latenzzeit auf die Videokamera zu sehen. Die reale Position des Referenzfahrzeuges wurde in das Bild projiziert (blauer Kasten mit zugehörigem Mittelpunkt  $c_{ref}$ ). In der Abbildung ist eine Diskrepanz zwischen der erwarteten und der gemessenen Position  $c_{cam}$  des Fahrzeuges zu erkennen. In Abbildung (b) ist ein Ausschnitt des zeitlichen Verlaufs von  $c_{ref}$  und  $c_{cam}$  dargestellt. Durch eine Kreuzkorrelation der beiden Kurven wurde eine Latenzzeit von 20 ms ermittelt.

$$(F \star G)_\tau \stackrel{\text{def}}{=} \sum_j F_j^* G_{\tau+j} \quad (3.20)$$

wobei  $F_i^*$  die komplexe Konjugation von  $F_i$  bezeichnet. Die Latenzzeit  $t_{delay}$  zwischen  $G$  und  $F$  berechnet sich aus der Kreuzkorrelation wie folgt:

$$t_{delay} = \underset{\tau}{\operatorname{argmax}} (F \star G)_\tau \quad (3.21)$$

Alle Kameras sind hardwareseitig synchronisiert. Deshalb wurde die Bestimmung der Latenzzeit nur für eine Kamera durchgeführt und kann auf die anderen übertragen werden. Für das gesamte Perzeptionssystem sind in Tabelle 3.3 alle experimentell bestimmten Latenzzeiten aufgeführt.

Sensor	durchschnittliche Latenzzeit in ms
Laserscanner	50 ms
Radar	100 ms
Kameras	20 ms

Tabelle 3.3: Durchschnittliche Latenzzeiten der entfernungs- und bildgebenden Sensorik.

### 3.4 Zusammenfassung

In diesem Kapitel wurden Aspekte der Kalibrierung betrachtet, die im Zusammenhang mit einem Multisensor System von besonderem Interesse sind. Für eine korrekte Fusion müssen die relative Position und Orientierung der einzelnen Sensoren zueinander sowie deren interne Abbildungseigenschaften bekannt sein. So ist es zum Beispiel nur durch eine genaue intrinsische und extrinsische Kalibrierung der Sensoren möglich, Bildausschnitte der Kameras mit Daten des Laserscanners oder der Radarsensorik zu verknüpfen und somit gemeinsam zu interpretieren.

In einem multispektralen Kamerasystem werden besondere Anforderungen an die Kalibrierungsverfahren gestellt. Es wurden deshalb zwei Möglichkeiten dargestellt und diskutiert, wie ein multispektrales Kamerasystem, bestehend aus Ferninfrarot- und Videokameras, kalibriert werden kann. Die Variante mit einem Kalibrierkörper aus Peltierelementen liefert deutlich genauere Ergebnisse im Vergleich zu einer Kalibrierung mit einer Punkt-Lichtquelle. Allerdings ist bei ersterer die Durchführung komplexer und nur bedingt für den mobilen Einsatz geeignet.

In einer Fahrumfelderfassung sind hochgenaue Referenzdaten in vielen Stadien der Entwicklung hilfreich. So bilden sie zum Beispiel die Grundlage für eine abschließende Bewertung des Gesamtsystems (siehe Kapitel 7). Sie können aber auch dazu verwendet werden, um die Latenzzeiten der einzelnen Sensoren zu bestimmen. Aufgrund der vielfältigen Anwendungsmöglichkeiten wurde ein Verfahren beschrieben, wie auf Basis von CPD-GPS Empfängern sehr präzise Referenzdaten gewonnen werden können. In einer Fehlerabschätzung wurde dargelegt, welche Genauigkeiten bei dieser Art der Referenzdatengewinnung zu erwarten sind.

Insbesondere in zeitkritischen Fahrerassistenzsystemen, wie beispielsweise der aktiven Gefahrenbremsung, müssen Latenzzeiten berücksichtigt werden. Diese können durch einen asynchronen Systemaufbau, sensorinterne Vorverarbeitung oder durch die Kommunikation zwischen den einzelnen Komponenten entstehen. Die genaue Zusammensetzung lässt sich in vielen Fällen nicht bestimmen. Es wurden verschiedene Möglichkeiten vorgestellt, wie sich zumindest der Gesamtbeitrag für bild- und entfernungsgebende Sensorik messen lässt. Mit diesen Verfahren wurden die Latenzzeiten für alle verwendeten Sensoren empirisch bestimmt.



# BEOBSACHTUNGSMODELLE

In dieser Arbeit wird das Tracking von Verkehrsteilnehmern in die drei Komponenten Beobachtungsmodell (likelihood), Objektmodell (a priori Wissen) und den Prozess der Filterung aufgeteilt (siehe auch Gesamtsystemübersicht in Abschnitt 2.5). Hierbei umfasst das a priori Wissen charakteristische Merkmale, Zustandsgrößen oder Eigenschaften, welche das zu verfolgende Objekt möglichst eindeutig festlegen. Das Beobachtungsmodell beschreibt, wie die jeweiligen Objekte in den verwendeten Sensoren abgebildet werden und mit welchen Verfahren bestimmte Zustandsgrößen gemessen werden können. Letztendlich vereint bzw. gewichtet ein Filterverfahren die Ergebnisse der ersten beiden Komponenten und ist für die eigentliche Zustandsschätzung des zu verfolgenden Objektes verantwortlich.

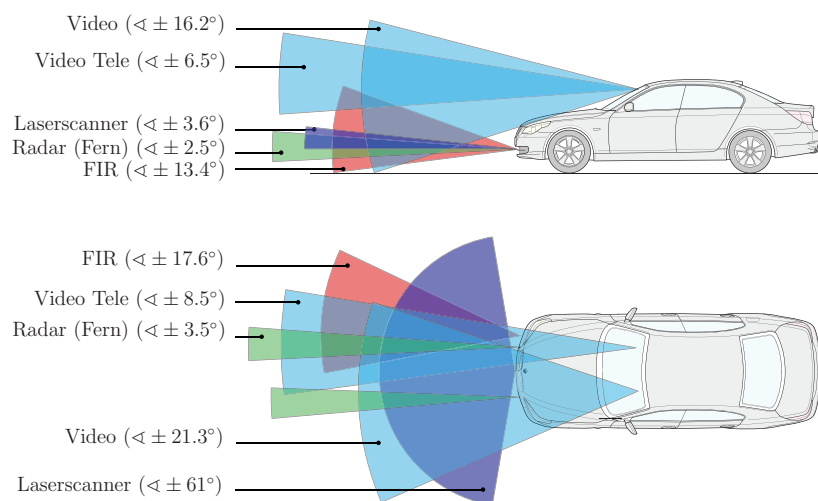


Abbildung 4.1: Sensorverbaupositionen und -sichtbereiche am Versuchsfahrzeug. Die vertikalen bzw. horizontalen Öffnungswinkel werden durch die obere bzw. untere Abbildung illustriert.

In diesem Kapitel werden zunächst die einzelnen Objektklassen mitsamt ihrer Objekt- und Bewegungsmodelle vorgestellt. Nachfolgend wird für die in Abbildung 4.1 dargestellten Sensoren beschrieben, wie Fußgänger und Fahrzeuge beobachtet werden können. Zunächst werden die entfernungsgebenden Sensoren Laserscanner und Radar behandelt. Für den Laserscanner wird hierbei detailliert auf

die Segmentierung und den Prozess der Merkmalsgewinnung eingegangen, und es werden Erweiterungen vorgestellt. Aus der spezifischen Verteilung der Messwerte für ein Fahrzeug wird für die Radarsensorik ein Modell eingeführt, das die charakteristischen Rückstrahleigenschaften von Fahrzeugen explizit berücksichtigt. Abschließend werden für die bildgebenden Sensoren (Ferninfrarot- und Videokamera) Konturmodelle beschrieben, die dem charakteristischen Erscheinungsbild von Fußgängern und Fahrzeugen nachempfunden sind.

Im Allgemeinen existieren zwischen den drei Komponenten des Trackings (a priori Wissen, likelihood und Filterverfahren) Abhängigkeiten. So müssen zum Beispiel die Beobachtungsmodelle an die Besonderheiten der verwendeten Filterverfahren angepasst werden. Aus diesem Grund werden in Vorgriff auf das Tracking (siehe Abschnitt 5) bereits in diesem Kapitel Anpassungen der vorgestellten Modelle an die in dieser Arbeit betrachteten Filterverfahren behandelt.

## 4.1 Objektbeschreibung

Das Ziel der in dieser Arbeit vorgestellten Fahrumfelderfassung ist es, sowohl Fahrzeuge als auch Fußgänger zu erkennen und zu verfolgen. Im weiteren Verlauf dieser Arbeit wird angenommen, dass für jedes zu verfolgende Objekt bereits bekannt ist, zu welcher Klasse es gehört. Für eine detaillierte Betrachtung dieses Klassifikationsproblems im Kontext von heterogener Sensorik sei auf die Arbeit von Leonhard Walchshäusl [Wal08] verwiesen.

In diesem Abschnitt wird erläutert, wie die einzelnen Objektklassen modelliert und durch charakteristische Zustandsgrößen beschrieben werden können. Darüber hinaus werden die zugehörigen Bewegungsmodelle eingeführt, die das zeitliche Verhalten der einzelnen Objekte definieren.

### 4.1.1 Fahrzeuge

Für Fahrzeuge wird die Annahme getroffen, dass sie sich durch die Angabe einer Position  $\mathbf{p} = (x, y)^T$ , eines Orientierungswinkels  $\varphi$  sowie der Fahrzeuglänge und -breite geometrisch beschreiben lassen. Der Ursprung des zugehörigen Fahrzeugkoordinatensystems  $\mathbf{C}_{veh}$  liegt in der Mitte der Hinterachse auf der Fahrbene ( $z = 0$  Ebene). In Abbildung 4.2 sind alle geometrischen Zustandsgrößen eines Fahrzeuges dargestellt.

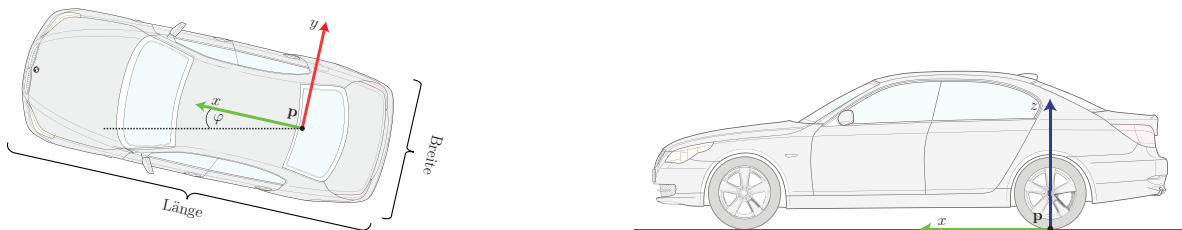


Abbildung 4.2: Verwendetes Fahrzeugmodell. Ein Fahrzeug ist durch seine Position  $\mathbf{p} = (x, y)^T$ , seine Orientierung  $\varphi$ , seine Länge  $l$  und Breite  $w$  definiert. Der Ursprung des lokalen Fahrzeugkoordinatensystems liegt in der Mitte der Hinterachse auf der Fahrbene ( $z = 0$  Ebene).



Neben den geometrischen Größen, wie Position oder Orientierung kann ein Fahrzeug zusätzlich durch dynamische Größen wie die Geschwindigkeit  $v$ , die Beschleunigung  $a$  und die Gierrate  $\dot{\varphi}$  definiert werden. Da Fahrzeuge im Allgemeinen entlang ihrer Orientierung fahren, wird im weiteren Verlauf dieser Arbeit ein Fahrzeugmodell verwendet, das die Position, Geschwindigkeit und Beschleunigung an die Orientierung koppelt. Für weitere mögliche Fahrzeugmodelle sei auf [Käm07, BSL93] verwiesen.

Für Fahrzeuge wird folgender Zustandsvektor  $\mathbf{x}_t^{veh}$  verwendet:

$$\mathbf{x}_t^{veh} = (x, y, v, a, \varphi, \dot{\varphi}, w, l)^T \quad (4.1)$$

Hierbei kennzeichnen die einzelnen Komponenten folgende Zustandsgrößen:

- $x$  = x-Position des Objektes auf der Fahrbene in  $[m]$
- $y$  = y-Position auf der Fahrbene in  $[m]$
- $v$  = Betrag der Geschwindigkeit entlang der Orientierung  $\varphi$  in  $[m/s]$
- $a$  = Betrag der Beschleunigung entlang der Orientierung  $\varphi$  in  $[m/s^2]$
- $\varphi$  = Orientierung des Objektes auf der Fahrbene in  $[rad]$
- $\dot{\varphi}$  = Gierrate des Objektes in  $[rad/s]$
- $w$  = Objektbreite in  $[m]$
- $l$  = Objektlänge in  $[m]$

Für die dynamische Modellierung von Fahrzeugen wird angenommen, dass die Beschleunigung  $a$  und die Gierrate  $\dot{\varphi}$  für einen Zeitschritt  $\Delta t$  konstant sind. Daraus ergibt sich folgende Differentialgleichung:

$$\dot{\mathbf{x}}^{veh}(t) = \begin{pmatrix} v \cos(\varphi) \\ v \sin(\varphi) \\ a \\ 0 \\ \dot{\varphi} \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (4.2)$$

Durch die Integration der Gleichung 4.2 ergibt sich das folgende nichtlineare dynamische Modell für Fahrzeuge:

$$\hat{\mathbf{x}}_t^{veh} = f(\mathbf{x}_{t-1}^{veh}) = \mathbf{x}_{t-1}^{veh} + \int_0^{\Delta t} \dot{\mathbf{x}}^{veh}(\tau) d\tau \quad (4.3)$$

$$= \mathbf{x}_{t-1}^{veh} + \begin{pmatrix} \frac{v+a\Delta t}{\phi} \sin(\varphi + \phi\Delta t) + \frac{a}{\phi^2} \cos(\varphi + \phi\Delta t) - \frac{v}{\phi} \sin(\varphi) - \frac{a}{\phi^2} \cos(\varphi) \\ -\frac{v+a\Delta t}{\phi} \cos(\varphi + \phi\Delta t) + \frac{a}{\phi^2} \sin(\varphi + \phi\Delta t) + \frac{v}{\phi} \cos(\varphi) - \frac{a}{\phi^2} \sin(\varphi) \\ a\Delta t \\ 0 \\ \phi\Delta t \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (4.4)$$

Bei sehr kleinen Gierraten ( $\phi \approx 0$ ) kann die Gleichung 4.4 nicht verwendet werden, da an dieser Stelle numerische Instabilitäten bzw. Singularitäten auftreten. Für diesen Fall wird deshalb folgende Näherung verwendet.

$$\hat{\mathbf{x}}_t^{veh} = \mathbf{x}_{t-1}^{veh} + \begin{pmatrix} \cos(\varphi)v\Delta t + \frac{1}{2} \cos(\varphi)a\Delta t^2 \\ \sin(\varphi)v\Delta t + \frac{1}{2} \sin(\varphi)a\Delta t^2 \\ a\Delta t \\ 0 \\ \phi\Delta t \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (4.5)$$

#### 4.1.2 Fußgänger

Für Fußgänger wird die Annahme getroffen, dass sie sich durch die Angabe einer Position  $\mathbf{p} = (x, y)^T$  sowie einer Höhe und Breite geometrisch durch einen Zylinder beschreiben lassen. Der Ursprung des zugehörigen Fußgängerkoordinatensystems  $\mathbf{C}_{ped}$  ist durch die orthogonale Projektion des Kopfmittelpunktes auf die Fahrbene ( $z = 0$  Ebene) bestimmt. Die explizite Angabe einer Orientierung ist für die in dieser Arbeit verwendete Sensorkonfiguration nicht sinnvoll, da sie nicht beobachtet werden kann. In Abbildung 4.3 sind alle geometrischen Zustandsgrößen eines Fußgängers dargestellt.

Analog zu den Fahrzeugmodellen werden auch Fußgänger durch zusätzliche dynamische Parameter beschrieben. Allerdings kann sich ein Fußgänger beliebig in alle Richtungen bewegen. Deshalb ist hier eine Einschränkung bzw. Koppelung der Bewegungsrichtung nicht sinnvoll. Darüber hinaus wird

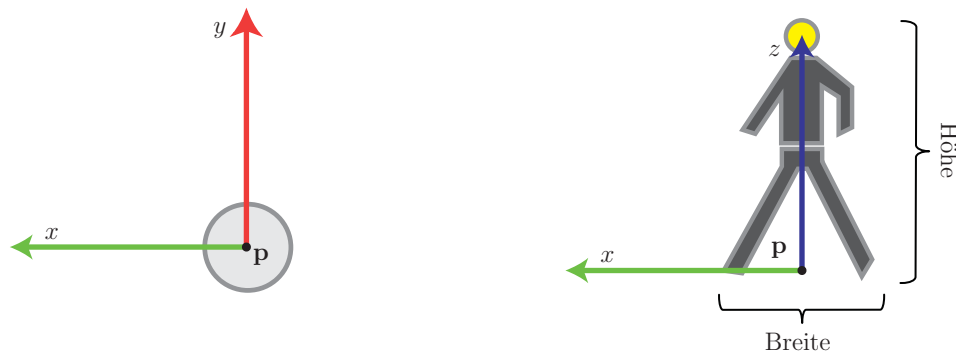


Abbildung 4.3: Verwendetes Fußgängermodell. Ein Fußgänger ist durch seine Position  $\mathbf{p} = (x, y)^T$  sowie durch seine Breite  $w$  und Höhe  $h$  definiert. Der Ursprung des lokalen Fußgängerkoordinatensystems liegt unter dem Mittelpunkt des Kopfes auf der Fahrbene ( $z = 0$  Ebene).

auf eine Modellierung von Beschleunigungen verzichtet, da diese selten und mit sehr begrenzter Amplitude und Dauer auftreten. Einzig der Geschwindigkeitsvektor  $\mathbf{v} = (v_x, v_y)^T$  wird verwendet, um die Dynamik eines Fußgängers zu beschreiben. Für eine detaillierte Betrachtung unterschiedlicher Modelle der Fußgängerdynamik sei auf [Mol95] verwiesen. Für Fußgänger ergibt sich folgender, im Vergleich zu Fahrzeugen einfacher Zustandsvektor  $\mathbf{x}_t^{ped}$ :

$$\mathbf{x}_t^{ped} = (x, y, v_x, v_y, w, h)^T \quad (4.6)$$

Die einzelnen Komponenten kennzeichnen folgende Zustandsgrößen:

- $x$  = x-Position des Objektes auf der Fahrbene in [m]
- $y$  = y-Position auf der Fahrbene in [m]
- $v_x$  = Geschwindigkeitskomponente entlang der x-Achse in [m/s]
- $v_y$  = Geschwindigkeitskomponente entlang der y-Achse in [m/s]
- $w$  = Objektbreite in [m]
- $h$  = Objekthöhe in [m]

Für die dynamische Modellierung von Fußgängern wird angenommen, dass die Geschwindigkeit  $v$  für einen Zeitschritt  $\Delta t$  konstant ist. Daraus ergibt sich folgendes lineares dynamisches Modell:

$$\hat{\mathbf{x}}_t^{ped} = f(\mathbf{x}_{t-1}^{ped}) = \mathbf{x}_{t-1}^{ped} + \begin{pmatrix} v_x \Delta t \\ v_y \Delta t \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (4.7)$$

## 4.2 Laserscanner

Für eine Fahrumfelderfassung ist ein Laserscanner im Augenblick einer der vielseitigsten und genauesten Sensoren [FD04]. Mit seiner Hilfe können zum Beispiel Fahrbahnmarkierungen erkannt und Informationen über die Straßenart gewonnen werden [SDS01]. Darüber hinaus wird dieser Sensor verwendet, um Verkehrsteilnehmer zu klassifizieren und zu verfolgen (vergleiche [Wal08, KBD05, MBN04]).

Im Folgenden wird das Messmodell für diesen Sensor behandelt. Zunächst werden Erweiterungen in der Vorverarbeitung und der Prozess der Merkmalsgewinnung vorgestellt. Darauf folgt eine Beschreibung der unterschiedlichen Beobachtungsmodelle für Fußgänger und Fahrzeuge. Abschließend werden die eingeführten Verbesserungen bewertet.

### 4.2.1 Laserscannerdatensegmentierung

Das Ziel einer Rohdatensegmentierung ist es, zum Einen zusammenhängende Bereiche zu bestimmen, die einem Objekt aus der Realität entsprechen und zum Anderen für die nachfolgenden Verarbeitungsschritte das Datenaufkommen zu reduzieren. Viele Segmentierungsalgorithmen für räumliche Rohpunkte stammen aus der Robotik. Dort bestimmen autonom agierende Roboter ihre Position, indem sie räumliche Umgebungsdaten mit elektronischen Karten abgleichen. In diesem Zusammenhang wird häufig zwischen Verfahren unterschieden, die entweder auf Kanten oder auf Regionen arbeiten [HGS07].

Laserscannerrohpunkte, die zu einem Objekt gehören, liegen im Allgemeinen in enger Nachbarschaft zueinander. Mit dieser Annahme lässt sich die Frage der Objektzugehörigkeit auf ein einfaches Entfernungsmaß reduzieren. Dementsprechend werden Messwerte,  $\mathbf{x}_i$  die einen geringen euklidischen Abstand zu benachbarten Messpunkten  $\mathbf{x}_j$  aufweisen, in einer Zusammenhangskomponente  $\mathcal{R}_i$  gruppiert. Hierfür existieren sehr einfache und effiziente Algorithmen, wie das „Region-Growing“, das seinen Ursprung in der Bildverarbeitung hat [ST88].

#### Region-Growing Algorithmus

Formal lässt sich das „Region-Growing“ auch als Graphenproblem darstellen: Sei  $G = (V, E)$  ein ungerichteter Graph mit den Knoten  $V = \{\text{Menge aller Messpunkte}\}$  und den Kanten  $E = \{(v_1, v_2) \mid v_1, v_2 \in V \wedge \|v_1 v_2\| \leq \tau\}$ . Die Zusammenhangskomponenten von  $G$  entsprechen dann den gesuchten Regionen  $\mathcal{R}_i$ , die zum Beispiel mit einer Tiefensuche ermittelt werden können. Hierbei ist allerdings auf eine effiziente Implementierung zu achten, in der die Menge der Kanten nicht a priori berechnet wird, sondern dynamisch entsteht.

Die Qualität der Segmentierung ist im Wesentlichen von der Wahl des Entfernungsschwellwertes  $\tau$  abhängig. Wird dieser zu gering gewählt, so zerfallen Objekte in mehrere Zusammenhangskomponenten. Ist dieser hingegen zu groß, so verschmelzen benachbarte Objekte. Hierbei ist zu beachten, dass bedingt durch die limitierte Winkelauflösung  $\beta$  des Laserscanners entfernte Objekte schlechter getrennt werden können. Aus diesem Grund sollte  $\tau$  abhängig von der Entfernung  $r$  gewählt werden.

$$\tau = \max(\tilde{\tau}, r \tan(\beta)) \quad (4.8)$$

Ein minimaler Abstand  $\tilde{\tau}$  verhindert, dass sehr nahe Objekte zerfallen. Auf diese Weise ist sichergestellt, dass alle Regionen einen entfernungsunabhängigen Mindestabstand zueinander besitzen.

Nachdem alle Messwerte in zusammenhängende Bereiche zusammengefasst worden sind, werden für jede Region Merkmale wie der Schwerpunkt  $\mathbf{x}_c$  oder die Ausdehnung des kleinsten umschließenden Rechtecks bestimmt. Alle Rohdatenpunkte sowie die daraus abgeleiteten Merkmale können als Grundlage für weitere Verarbeitungsschritte wie Klassifikation oder Tracking dienen.

**Aggregation von Regionen** Bei modernen Laserscannern, die die Umgebung in mehreren Ebenen abtasten, kann das „Region-Growing“ entweder für alle Ebenen gemeinsam oder für jede Ebene einzeln durchgeführt werden. Ein Vorteil der ersten Variante ist, dass keine Mehrdeutigkeiten bzw. Überschneidungen entstehen, da der Mindestabstand  $\tau$  zwischen allen Regionen gewahrt bleibt. Auf die Berechnung einzelner Merkmale kann die gemeinsame Betrachtung der Ebenen einen positiven oder negativen Einfluss haben. Sind zum Beispiel Messfehler (Ausreißer) über alle Ebenen gleichverteilt, so ist die Berechnung des Schwerpunktes robuster. Sind Messfehler punktuell auf einzelne Ebenen beschränkt, wie es häufig bei LKWs oder PKWs zu beobachten ist (z.B. Messung der Fahrzeugbereifung und Teile des Unterbodens), so werden Merkmale wie der Schwerpunkt verfälscht. Wird das „Region-Growing“ für jede Ebene unabhängig durchgeführt, kann der Einfluss einzelner verrauschter Ebenen auf die Qualität der Merkmale minimiert werden. Allerdings entstehen Mehrdeutigkeiten und eine größere Anfälligkeit in Bezug auf globales Rauschen.

Um die Vorteile aus beiden Vorgehensweisen zu kombinieren wird ein zweistufiger Ansatz verwendet. In der ersten Stufe werden für jede Ebene Regionen gebildet. Anschließend werden diejenigen Regionen vereint (aggregiert), die sehr ähnliche Merkmale besitzen. Auf diese Weise können Mehrdeutigkeiten stark reduziert werden. Das Ähnlichkeitsmaß sollte so gewählt werden, dass Ebenen mit punktuellen Messfehlern nicht kombiniert werden. Hierfür kann zum Beispiel der euklidische Abstand zwischen den Schwerpunkten der einzelnen Regionen verwendet werden. Dieses Merkmal ist allerdings nicht invariant gegenüber Rotationen, was einen Ausgleich derartiger Fehler verhindert. Im Folgenden wird deshalb ein Verfahren vorgestellt, welches das Konzept der Regionen erweitert und die Berechnung weiterer Merkmale ermöglicht.

### Inkrementelle Linienextraktion

Einer der schnellsten und genauesten Algorithmen [NMTS05] zur Linienextraktion aus Rohdaten ist der inkrementelle Line-Track Algorithmus [SKK<sup>+</sup>97]. Dieser stammt ursprünglich aus der Bildverarbeitung [FP03] und wurde auf die Aggregation von Laserscannerrohdaten übertragen. Analog zum „Region-Growing“-Algorithmus werden dabei Rohpunkte zusammengefasst, die einen geringen euklidischen Abstand zueinander aufweisen. Diese Regionen werden jetzt als Segmente  $\mathcal{S}_i = [\mathbf{k}_1, \dots, \mathbf{k}_n]$  bezeichnet, da zusätzlich Ausgleichsgeraden in die Punktwolken eingepasst werden, die durch Kontrollpunkte  $\mathbf{k}_j$  repräsentiert werden. Dadurch können weitere Merkmale, wie Orientierung und Ausdehnung, abgeleitet werden. Die Eingabe für den Algorithmus ist eine Menge an winkelsortierten Rohdatenpunkten  $\mathbf{x}_i$ , die ebenenweise abgearbeitet werden. Dadurch kann die dritte Vektorkomponente vernachlässigt werden und es gilt  $\mathbf{x}_i \in \mathbb{R}^2$ .

Der detaillierte Ablauf der inkrementellen Linienextraktion (ILE) ist wie folgt: Zunächst wird durch zwei Ausgangspunkte  $\mathbf{x}_1$  und  $\mathbf{x}_2$  eine Gerade mit der Gleichung  $y = a + bx$  gelegt. Anschließend werden so lange weitere Punkte  $\mathbf{x}_i$  zu den zwei Ausgangspunkten hinzugenommen und eine aktualisierte

Ausgleichsgerade durch die anwachsende Punktmenge  $\mathcal{P}$  gelegt, bis ein Fehlerschwellwert  $\sigma_e$  überschritten wird. Als Fehler wird in der Regel der durchschnittliche Abstand  $D = \frac{1}{n} \sum_{i=1}^n d_i$  der Rohpunkte  $\mathbf{x} \in \mathcal{P}$  zur Geraden gewählt. Der orthogonale Abstand  $d_i$  eines Punktes  $(x_i \ y_i)^T$  zur Geraden berechnet sich über  $d_i = \frac{|y_i - (a + bx_i)|}{\sqrt{1+b^2}}$ . Die Summe der Abstände  $R = \sum_{i=1}^n d_i$  wird zudem verwendet, um die Parameter  $a$  und  $b$  der Ausgleichsgeradengleichung zu ermitteln [Wei03]. Dazu wird  $R^2$  minimiert, indem über die partiellen Ableitungen die Extremwerte der Funktion bestimmt werden.

$$\begin{aligned} \frac{\partial R^2}{\partial a} &= \frac{2}{1+b^2} \sum_{i=1}^n [y_i - (a + bx_i)](-1) = 0 \\ \frac{\partial R^2}{\partial b} &= \frac{2}{1+b^2} \sum_{i=1}^n [y_i - (a + bx_i)](-x_i) + \sum_{i=1}^n \frac{[y_i - (a + bx_i)](-1)(2b)}{(1+b^2)^2} = 0 \end{aligned}$$

Nach Einsetzen und Umformen ergeben sich die Parameter der Geradengleichung zu:

$$a = \bar{y} - b\bar{x} \quad (4.9)$$

$$b = -B \pm \sqrt{B^2 + 1} \quad (4.10)$$

$$B = \frac{1}{2} \frac{(\sum_{i=1}^n y_i^2 - n\bar{y}^2) - (\sum_{i=1}^n x_i^2 - n\bar{x}^2)}{n\bar{x}\bar{y} - \sum_{i=1}^n x_i y_i} \quad (4.11)$$

Wenn der Fehlerschwellwert  $\sigma_e$  überschritten wird, dann wird eine neue Ausgleichsgerade innerhalb des aktuellen Segments  $\mathcal{S}$  angelegt, die direkt am Endpunkt der alten Ausgleichsgerade anknüpft. Dazu wird  $\mathcal{P}$  neu initialisiert. Ein Segment besteht somit aus einem kreisfreien geschlossenen Streckenzug, der durch eine Menge von Kontrollpunkten  $\mathcal{S} = [\mathbf{k}_1, \dots, \mathbf{k}_n]$ , die die Schnittstellen der Einzelstrecken markieren, repräsentiert werden kann. Ein Kontrollpunkt wird berechnet, indem der Rohpunkt, bei dem der Fehler über dem Schwellwert liegt, auf die aktuelle Gerade projiziert wird.

Das alte Segment  $\mathcal{S}_j$  wird abgeschlossen und ein neues Segment  $\mathcal{S}_{j+1}$  erzeugt, wenn der euklidische Abstand  $\sigma_d = \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2$  zwischen dem aktuellen Rohdatenpunkt  $\mathbf{x}_k$  und dem neu hinzugenommenen Rohdatenpunkt  $\mathbf{x}_{k+1}$  einen Schwellwert  $\tau_{\text{dist}}$  überschreitet. Der Algorithmus terminiert, wenn alle Rohpunkte aufgebraucht sind. Damit der Algorithmus für den automotiven Einsatz verwendet werden kann, wurden zahlreiche Anpassungen vorgenommen.

Ein ausschlaggebender Parameter für die Performanz ist der maximal zulässige Abstand  $\tau_{\text{dist}}$  zwischen zwei Rohpunkten. Wird dieser Schwellwert zu klein gewählt, zerfallen die Segmente unter Umständen zu stark und die Ein-Segment-Ein-Objekt Beziehung wird häufig verletzt. Eine zu großes  $\tau_{\text{dist}}$  hat zur Folge, dass Segmente leichter zusammengeführt werden und dadurch Objekte, die in der realen Welt getrennt sind, als ein einzelnes Segment repräsentiert werden. Es hat sich herausgestellt, dass  $\tau_{\text{dist}}$  zunächst klein gewählt werden sollte und anschließend in einem zweiten Schritt die Segmente mit den zusätzlichen Richtungsinformationen der Zusammenhangskomponenten verlängert, konkateniert und aggregiert werden sollten.

**Segmentverlängerung** Manche Rohpunkte können beim ersten Segmentierungsschritt keinem Segment zugeordnet werden, da der Abstand, den sie zu den benachbarten Rohpunkten aufweisen, größer als der Schwellwert  $\tau_{\text{dist}}$  ist. Oftmals sind solche vereinzelt Rohpunkte Bestandteil eines angrenzenden Segments. Damit möglichst die Gesamtausdehnung aller Objekte erfasst wird, werden die

Segmente nun an den beiden Endpunkten verlängert. Dazu wird für alle einzelnen Rohpunkte  $\mathbf{x}$  untersucht, ob sie einen Abstand  $\tau_{\text{verl}}$  zur verlängerten Endstrecke der angrenzenden Segmente unterschreiten. Erfüllt ein Rohpunkt diese Bedingung, wird der Rohpunkt als neuer Kontrollpunkt an das Segment angehängt  $\mathcal{S} = [\mathbf{k}_1, \dots, \mathbf{k}_n] \rightarrow \mathcal{S}' = [\mathbf{k}_1, \dots, \mathbf{k}_n, \mathbf{x}]$ , wodurch sich das Segment entsprechend vergrößert.

**Segmentkonkatenation** Bei diesem Schritt werden kollineare Segmente zusammengefügt. Dadurch können kurze Unterbrechungen überbrückt werden, die durch Verdeckungen oder Messfehler verursacht werden. Zunächst werden dazu die Segmente nach dem Schwerpunkt  $\mathbf{m} = \frac{1}{n} \sum_1^n \mathbf{k}_i$  in Polarkoordinaten winkelsortiert. Anschließend werden zwei benachbarte Segmente  $\mathcal{S}_j$  und  $\mathcal{S}_{j+1}$  auf Kollinearität überprüft. Dazu muss der Winkel zwischen den beiden Endvektoren  $\mathbf{v}_j = \mathbf{k}_n^{(j)} - \mathbf{k}_{n-1}^{(j)}$  bzw.  $\mathbf{v}_{j+1} = \mathbf{k}_2^{(j+1)} - \mathbf{k}_1^{(j+1)}$  der Segmente einen Schwellwert  $\tau_\alpha$  unterschreiten:

$$\arccos \frac{\mathbf{v}_j \cdot \mathbf{v}_{j+1}}{\|\mathbf{v}_j\| \|\mathbf{v}_{j+1}\|} \leq \tau_\alpha \quad (4.12)$$

Gleichzeitig darf der Abstand  $\|\mathbf{k}_n^{(j)} - \mathbf{k}_1^{(j+1)}\|$  zwischen den beiden Segmentendpunkten einen Schwellwert  $\tau'_{\text{dist}}$  nicht überschreiten. Es gilt jedoch  $\tau'_{\text{dist}} > \tau_{\text{dist}}$ , da durch die Kollinearitätsbedingung eine zusätzliche Einschränkung gegeben ist. Sind die Bedingungen erfüllt, wird ein neues konkateniertes Segment  $\mathcal{S}' = [\mathbf{k}_1^{(j)}, \dots, \mathbf{k}_n^{(j)}, \mathbf{k}_1^{(j+1)}, \dots, \mathbf{k}_n^{(j+1)}]$  erstellt. Die Ausgangssegmente  $\mathcal{S}_j$  und  $\mathcal{S}_{j+1}$  werden entfernt.

**Segmentaggrierung** Moderne Laserscanner tasten die Umgebung ebenenweise ab (siehe auch Abschnitt 2.4). Werden die Rohpunkte in den Ebenen gemeinsam zu Segmenten zusammengefasst, führt dies oftmals zu fehlerhaften Segmentverläufen, da eine einzelne verrauschte Ebene das gesamte Segment beeinflusst. Es hat sich deshalb herausgestellt, dass es von Vorteil ist, jede Ebene getrennt zu segmentieren und erst im Anschluß überlappende Segmente zu aggregieren. Zunächst werden dafür alle Segmente  $\mathcal{S}_j$  anhand des jeweiligen Segment Schwerpunktes  $\mathbf{m}$  winkelsortiert. Segmente  $\mathcal{S}_j$ , deren Schwerpunkte einen geringen euklidischen Abstand zueinander aufweisen, werden zu Segmentmengen  $\mathcal{Z}_i = \{\mathcal{S}_1, \dots, \mathcal{S}_n\}$  zusammengefasst. Segmente  $\mathcal{S}$ , die nicht aggregiert werden können, werden in eine einelementige Segmentmenge  $\mathcal{Z}_i = \{\mathcal{S}\}$  eingefügt. Auch nach der Aggregation können alle Segmente noch unabhängig voneinander betrachtet werden.

## Zusammenfassung

Die Erweiterungen des vorgestellten Segmentierungsalgorithmus haben viele Vorteile im Vergleich zum einfachen „Region-Growing“-Vorgehen. In Abbildung 4.4 ist das Segmentierungsergebnis einer typischen Straßenszene mit sowohl dem klassischen „Region-Growing“- (rechteckige Bereiche) als auch der erweiterten Segmentierung (Linienzüge) zu sehen. Die Entfernungsschwellwerte  $\tau$  wurden bei beiden Algorithmusvarianten auf einen identischen Wert von 0.7 m festgelegt.

Die Ergebnisse sind wie folgt: Fahrzeuge werden durch die Segmentverlängerung vollständiger segmentiert. Dieser Vorteil wird insbesondere bei Strukturen erzielt, die nahezu in Messrichtung verlaufen und dadurch hohe Messwertabstände verursachen (vergleiche Lieferwagen bei Position (1)). Durch die

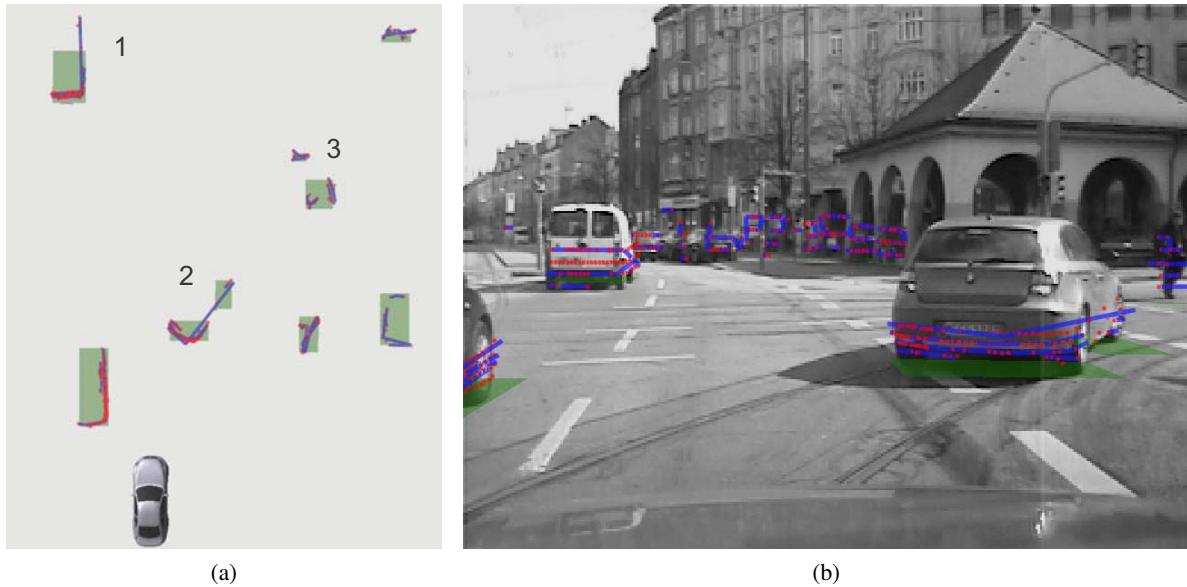


Abbildung 4.4: Exemplarischer Vergleich zwischen dem „Region-Growing Algorithmus“ und der erweiterten inkrementellen Linienextraktion für Laserscannerrohdaten (rote Punkte). Die Ergebnisse des „Region-Growing“-Algorithmus sind durch grüne rechteckige Bereiche und die Resultatsegmente der inkrementellen Linienextraktion durch blauen Linienzüge in einer Vogelperspektiven- (a) und einer Videoansicht (b) der Szene markiert.

Segmentkonkatenation werden Unterbrechungen von kollinearen Segmenten korrigiert. Diese Unterbrechungen bzw. Rohdatenverluste werden neben Teilverdeckungen durch z.B. Bäume oder Ampeln, ebenso durch flache Auftreffwinkel auf stark reflektierende Materialien verursacht (vergleiche abbiegendes Fahrzeug bei Markierung (2)). Dennoch können Objekte, die eine geringe räumliche Nähe zueinander aufweisen, wie beispielsweise Fußgänger, voneinander getrennt werden, solange keine Kollinearitäten zwischen den Segmenten auftreten (vergleiche die drei Fußgänger bei Markierung (3)).

#### 4.2.2 Merkmalsgewinnung

Für die in dem vorherigen Abschnitt eingeführten Segmente wurden keine Annahmen über das Erscheinungsbild der zu verfolgenden Objekte getroffen. Für diese Art der Segmentierung ist das wesentliche Kriterium der minimale Abstand zwischen zwei benachbarten Objekten. In dem Schritt der Merkmalsgewinnung wird a priori Wissen verwendet, um basierend auf den Segmenten Messwerte zu generieren, die die charakteristischen Eigenschaften der zu verfolgenden Objekte besser beschreiben. Grundlage dieser Messwerte ist das jeweils typische Erscheinungsbild von Fußgängern und Fahrzeugen in dem Laserscanner.

Im Folgenden wird die Vereinfachung angenommen, dass Fahrzeuge durch Quader und Fußgänger durch Zylinder angenähert werden können. Dadurch ergeben sich bei Fahrzeugen je nach sichtbaren Seitenflächen zwei unterschiedliche Erscheinungsbilder, die als L-Shape und I-Shape bezeichnet werden: Ein L-Shape entsteht, wenn zwei Seitenflächen des Fahrzeuges sichtbar sind und wird durch zwei



sich im rechten Winkel schneidende Gerade beschrieben. Ein I-Shape entspricht nur einer sichtbaren Fläche und ist dementsprechend durch eine Gerade bestimmt. Für Fußgänger wird angenommen, dass sie sich unabhängig vom Blickwinkel durch ein I-Shape beschreiben lassen. In Abbildung 4.5 sind für Fahrzeuge und Fußgänger Beispiele dieses Einteilungsschemas gegeben. Diese Art der Einteilung und Merkmalsgewinnung basiert im Wesentlichen auf den Arbeiten von Fürstenberg [FD05] und Kämpchen [Käm07].

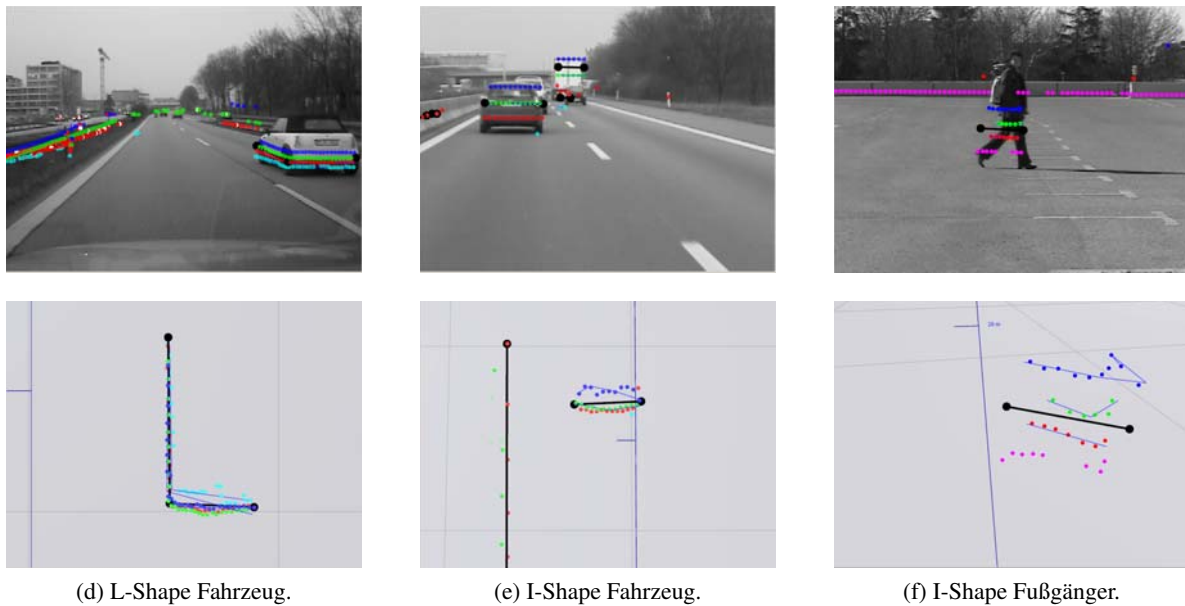


Abbildung 4.5: Typisches Erscheinungsbild von Fahrzeugen und Fußgängern im Laserscanner. Sind zwei Flächen eines Fahrzeuges zu sehen, entsteht ein L-Shape, wie es in Abbildung (a) dargestellt ist. In den Abbildungen (b) und (c) ist ein I-Shape für ein Fahrzeug oder einen Fußgänger abgebildet. Ein I-Shape entsteht, wenn nur eine Fläche vom Sensor gesehen werden kann. Die schwarzen Linien entsprechen einem L-Shape oder I-Shape. Die farbigen Punkte symbolisieren die Rohpunkte des Laserscanners.

Der Prozess der Merkmalsgewinnung ist zweigeteilt: In einem ersten Schritt muss bestimmt werden, ob ein Segment einem I oder L-Shape entspricht. Für weitere Details zu diesem Klassifikationsproblem sei an dieser Stelle auf die Arbeit von Leonhard Walchshäusl [Wal08] verwiesen. In dem zweiten Schritt werden entsprechend der Klasse aus den Segmenten Merkmale bestimmt, die im weiteren Verlauf als Messwerte dienen.

### I-Shape

Wurde ein I-Shape klassifiziert, so wird durch alle Rohpunkte  $x_i$  des Segments eine Ausgleichsgerade  $g$  gelegt.

$$g : y = mx + b \quad (4.13)$$

Diese Gerade  $g$  bestimmt die Orientierung  $\varphi$  des I-Shapes. Um den Anfangspunkt  $\mathbf{a}$  und den Endpunkt  $\mathbf{b}$  des I-Shapes zu bestimmen, werden alle Rohpunkte auf  $g$  projiziert. Das Minimum bzw. Maximum

dieser Projektion entspricht dem gesuchten Anfangs- bzw. Endpunkt des Geradenstückes. Durch diese beiden Punkte kann die Länge  $l$  des I-Shapes gemessen werden. Die Qualität der Längenmessung ist von zwei Faktoren abhängig: Zum Einen von der Genauigkeit der Orientierung  $\phi$  und zum Anderen von der Position der beiden Punkte, die auf die Gerade projiziert werden. Als Positionsmesswert  $\mathbf{c}$  wird die Mitte des Geradenstückes herangezogen. Hierfür könnte alternativ auch der Schwerpunkt aller Rohpunkte verwendet werden. In der Entfernungsmessung ist kein Vorteil zu erwarten, da durch die Ausgleichsgerade der Einfluss von Messfehlern, vergleichbar zum Schwerpunkt, kompensiert wird. Für die Messung der lateralen Ablage ist der Schwerpunkt robuster gegenüber Ausreißern, die sich am Rand des I-Shapes befinden. Dies ist häufig bei Regen und starker Gischt zu beobachten. Der Mittelpunkt des Geradenstückes ist hingegen besser geeignet, den Einfluss von fehlenden Messwerten auf einzelnen Ebenen des Laserscanners zu minimieren, was bei Fahrzeugen im Bereich des Unterbodens vorkommen kann. Für normale Szenarien konnte in der Praxis kein signifikanter Unterschied festgestellt werden.

Die Ausgleichsgerade  $g$  ist die Grundlage für alle Messwerte. Bei der Wahl des Algorithmus zur Berechnung einer Regressionsgeraden sollte darauf geachtet werden, dass das Ergebnis robust gegenüber Messfehlern ist. Insbesondere die Messung der Orientierung und der Länge ist hiervon stark betroffen. Die einfache lineare Regression, wie sie auch in Abschnitt 4.2.1 beschrieben ist, ist allerdings sehr anfällig gegenüber Ausreißern (siehe Abbildung 4.6a). Weitaus besser geeignet ist eine Implementierung nach dem RANSAC-Verfahren [FB81], bei dem nur Punkte zur Berechnung zugelassen werden, die einen maximalen Abstand  $\theta$  von der Ausgleichsgeraden nicht überschreiten. Da es sich bei RANSAC um einen randomisierten Algorithmus handelt, wird nicht immer die optimale Lösung gefunden. Vielmehr kann nur eine Wahrscheinlichkeit  $P_w$  angegeben werden, dass eine suboptimale Ausgleichsgerade gefunden wird. Diese Wahrscheinlichkeit ist einerseits von der Ausreißerquote  $P_a$  und andererseits von der Anzahl der Iterationen abhängig. Bei einer Quote von  $P_a = 40\%$  und einer gewünschten Wahrscheinlichkeit  $P_w = 1\%$  müssen mindestens 11 Iterationen durchgeführt werden. Für eine fehlerbehaftete Punktemenge ist in Abbildung 4.6 der Unterschied der beiden Verfahren sehr gut zu erkennen. Für eine quantitative Evaluierung dieser Erweiterung sei auf Abschnitt 4.2.6 verwiesen.

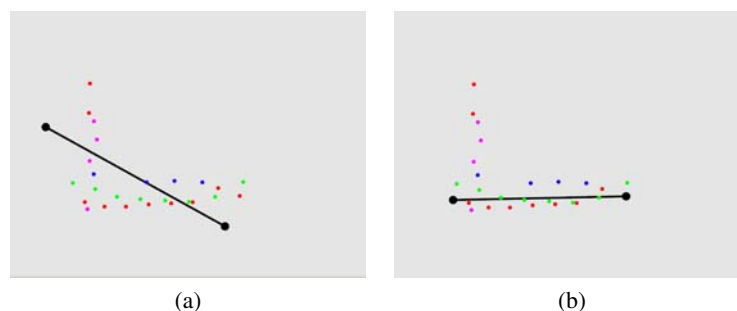


Abbildung 4.6: Unterschiedliche Verfahren zur Berechnung eines I-Shapes (schwarze Linie) bei einer fehlerbehafteten Punktemenge. In Abbildung (a) wurde eine einfache lineare Regression verwendet, um die Ausgleichsgerade zu bestimmen. Der Einfluss von Messfehlern ist gut zu erkennen. Wird hingegen ein Implementierung nach dem RANSAC-Verfahren [FB81] verwendet, ist das I-Shape, wie in Abbildung (b) dargestellt, deutlich robuster gegenüber Ausreißern in der Punktemenge.

In Abbildung 4.7 sind alle Messwerte schematisch dargestellt. Die Messung  $\mathbf{z}_I = (\mathbf{c}, \varphi, l)$  für ein I-Shape setzt sich aus der Position  $\mathbf{c}$ , der Orientierung  $\varphi$  und der Länge  $l$  zusammen. Die einzelnen Komponenten der Messung lassen sich aus der Ausgleichsgeraden  $g$  und den Rohpunkten  $\mathbf{x}_i$  wie folgt berechnen:

$$g: y = mx + b \quad (4.14)$$

$$\mathbf{c} = \frac{\mathbf{a} + \mathbf{b}}{2} \quad (4.15)$$

$$\varphi = \arctan(m) \quad (4.16)$$

$$l = \|\mathbf{a} - \mathbf{b}\| \quad (4.17)$$

$$\mathbf{a} = \underset{\mathbf{x}_i}{\operatorname{argmin}} \langle \mathbf{x}_i, (1, m)^T \rangle \quad (4.18)$$

$$\mathbf{b} = \underset{\mathbf{x}_i}{\operatorname{argmax}} \langle \mathbf{x}_i, (1, m)^T \rangle \quad (4.19)$$

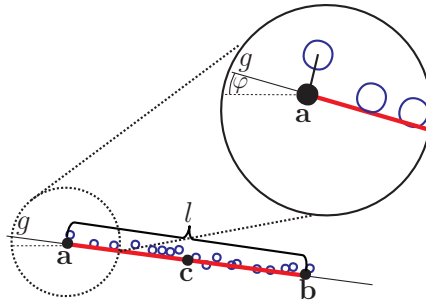


Abbildung 4.7: Schematische Darstellung eines I-Shapes (rote Linie). Durch alle Rohpunkte (blaue Kreise) wird eine Ausgleichsgerade  $g$  gebildet. Die Projektion der Rohpunkte auf diese Gerade bestimmt den Anfangs- und Endpunkt  $\mathbf{a}$  bzw.  $\mathbf{b}$  des I-Shapes und somit dessen Länge  $l$ . Die Mitte  $\mathbf{c}$  des Geradenstückes definiert die Position und durch die Steigung von  $g$  ist die Orientierung  $\varphi$  des I-Shapes gegeben.

Für eine übersichtlichere Darstellung und eine einfachere Berechnung wurde die Messung in Kartesischen Koordinaten dargestellt. Da das Messprinzip und dementsprechend die zugehörigen Angaben der Messunsicherheiten auf Polarkoordinaten basieren, wird im Folgenden die äquivalente Messung  $\mathbf{z}_I = (r, \phi, \varphi, l)$  in Polarkoordinaten verwendet. Hierbei bezeichnen  $r$  und  $\phi$  die Polarkoordinaten der Position  $\mathbf{c}$ .

### L-Shape

Wurde ein L-Shape klassifiziert, so wird in einem ersten Schritt die grobe Position der Ecke  $\hat{\mathbf{c}}$  bestimmt. Hierzu wird die Annahme getroffen, dass der Punkt mit dem geringsten Abstand zu dem Sensor höchstwahrscheinlich einen Eckpunkt repräsentiert. Im weiteren wird dieser Punkt verwendet, um alle Rohpunkte des Segments in zwei Mengen  $\mathcal{M}_1$  und  $\mathcal{M}_2$  zu unterteilen: Die erste Menge enthält nur Punkte, die zu dem ersten Schenkel des L-Shapes gehören. Respektive enthält  $\mathcal{M}_2$  nur Punkte des zweiten Schenkels. Analog zur Berechnung eines I-Shapes werden durch die beiden Mengen die

Regressionsgeraden  $g$  und  $h$  gelegt. Durch den Schnittpunkt der beiden Geraden wird die Positionsbestimmung der Ecke  $\mathbf{c}$  verbessert. Der Anfangspunkt der beiden Schenkel ist somit durch den Punkt  $\mathbf{c}$  definiert. Um die Endpunkte  $\mathbf{a}$  bzw.  $\mathbf{b}$  des L-Shapes zu bestimmen werden alle Punkte auf die Gerade  $g$  bzw.  $h$  projiziert und der Punkt mit dem größten Abstand zu  $\mathbf{c}$  als Endpunkt gewählt. Für jeden der beiden Schenkel lässt sich jetzt die Länge  $l_1$  und  $l_2$  bestimmen. Durch die Winkelhalbierende von  $g$  und  $h$  ist die Orientierung  $\varphi$  des L-Shapes bestimmt.

Die kritischste Stelle in der Berechnung eines L-Shapes ist die Wahl des ersten Eckpunktes  $\hat{\mathbf{c}}$ . In einigen Ausnahmefällen ist durch Messfehler die Annahme falsch, dass die Ecke des L-Shapes der Punkt ist, der am nächsten zum Sensor gelegen ist. Dies führt dazu, dass eine der beiden Mengen  $\mathcal{M}_1$  bzw.  $\mathcal{M}_2$  Punkte des anderen Schenkels enthält und die zugehörige Ausgleichsgerade verfälscht wird. Um diesen Effekt zu kompensieren, wird für die Berechnung der Regressionsgeraden eine Implementierung nach dem RANSAC-Verfahren verwendet. Dadurch können falsch zugeordnete Punkte ignoriert werden. In Extremfällen können die falsch zugeordneten Punkte eine bessere Ausgleichsgerade bilden, als die eigentlichen Punkte des Schenkels. Dieser Effekt ist zum Beispiel zu beobachten, wenn einer der beiden Schenkel aus nur wenigen Punkten besteht. Dies kann jedoch leicht erkannt werden, da die resultierenden Schenkel annähernd parallel sind. Für eine quantitative Evaluierung unterschiedlicher Verfahren zur Erzeugung einer Ausgleichsgeraden sei auf Abschnitt 4.2.6 verwiesen.

In Abbildung 4.8 sind alle Messwerte schematisch dargestellt. Die Messung  $\mathbf{z}_L = (\mathbf{c}, \varphi, l_1, l_2)$  für ein L-Shape setzt sich aus der Position  $\mathbf{c}$ , der Orientierung  $\varphi$  und den Längen  $l_1$  bzw.  $l_2$  der beiden Schenkeln zusammen. Die einzelnen Komponenten der Messung lassen sich aus der Ausgleichsgeraden  $g$  bzw.  $h$  und den Rohpunkten  $\mathbf{x}_i$  wie folgt berechnen:

$$g : y = m_1x + b_1 \quad (4.20)$$

$$h : y = m_2x + b_2 \quad (4.21)$$

$$\mathbf{c} = \left( \frac{b_2 - b_1}{m_1 - m_2}, \frac{b_2m_1 - b_1m_2}{m_1 - m_2} \right)^T \quad (4.22)$$

$$\varphi = \sphericalangle \left( \frac{\mathbf{a} - \mathbf{c}}{l_1} + \frac{\mathbf{b} - \mathbf{c}}{l_2} \right) \quad (4.23)$$

$$l_1 = \|\mathbf{a} - \mathbf{c}\| \quad (4.24)$$

$$l_2 = \|\mathbf{b} - \mathbf{c}\| \quad (4.25)$$

$$\mathbf{a} = \underset{\mathbf{x}_i}{\operatorname{argmax}} \langle \mathbf{x}_i, (1, m_1)^T \rangle \quad (4.26)$$

$$\mathbf{b} = \underset{\mathbf{x}_i}{\operatorname{argmax}} \langle \mathbf{x}_i, (1, m_2)^T \rangle \quad (4.27)$$

Analog zu den I-Shapes wird im weiteren Verlauf dieser Arbeit die Messung  $\mathbf{z}_L = (r, \phi, \varphi, l_1, l_2)$  für ein L-Shape in Polarkoordinaten betrachtet.

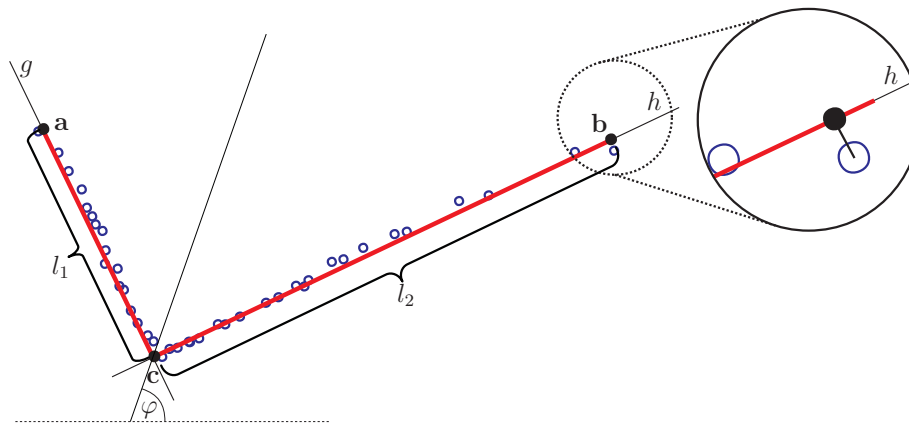


Abbildung 4.8: Schematische Darstellung eines L-Shapes (rote Linie). Durch alle Rohpunkte (blaue Kreise) werden zwei Ausgleichsgeraden  $g$  und  $h$  gebildet, die den beiden Schenkeln entsprechen. Durch den Schnittpunkt der Geraden ist der Punkt  $c$  definiert. Die Projektion der Rohpunkte auf diese Geraden bestimmt die beiden Endpunkte  $a$  bzw.  $b$  des L-Shapes und somit die Längen  $l_1$  bzw.  $l_2$  der beiden Schenkel. Durch die Winkelhalbierende zwischen  $g$  und  $h$  ist die Orientierung  $\varphi$  des L-Shapes gegeben.

### 4.2.3 Modell für Fußgänger

In dem vorherigen Abschnitt wurde beschrieben, wie sich aus den Rohdaten des Laserscanners Merkmale in Form eines I-Shapes bzw. L-Shapes erzeugen lassen. Im Folgenden wird beschrieben, wie mit diesen Messwerten Fußgänger beobachtet werden können.

Für Fußgänger wird die Annahme getroffen, dass sie durch einen Zylinder approximiert werden können. Durch diese Vereinfachung ist das Erscheinungsbild für einen Fußgänger unabhängig von dem Blickwinkel, aus dem er beobachtet wird und kann somit immer durch ein I-Shape beschrieben werden (siehe auch Abschnitt 4.2.2 bzw. Abbildung 4.5f). Für die Bestimmung der Position wird der Schwerpunkt  $c$  der Rohpunkte eines I-Shapes verwendet. Die Breite  $w$  eines Fußgängers ist durch die Länge  $l$  des I-Shapes bestimmt. In Abbildung 4.9 ist das Beobachtungsmodell für einen Fußgänger schematisch dargestellt.

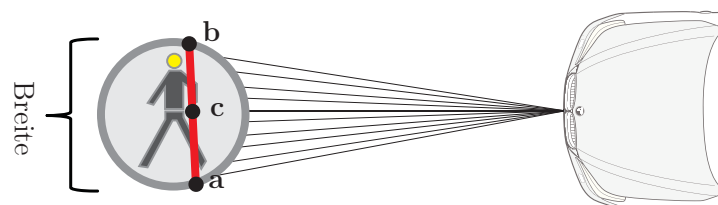


Abbildung 4.9: Messmodell für Fußgänger im Laserscanner. Unter der Annahme, dass sich Fußgänger durch einen Zylinder approximieren lassen, kann seine Breite und Position unabhängig von dem Blickwinkel, unter dem er beobachtet wird, mit einem I-Shape (rote Linie) gemessen werden.

Für das Modell der Fußgänger ergibt sich die Messung  $z_{I-Shape} = (r, \phi, \alpha, l)^T$  wie folgt aus dem Zustand  $\mathbf{x}_t^{ped} = (x, y, v_x, v_y, w, h)^T$ :

$$h_{I-Shape}(\mathbf{x}_t^{ped}) = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arctan\left(\frac{y}{x}\right) \\ \arctan\left(\frac{y}{x}\right) + \frac{\pi}{2} \\ w \end{pmatrix} \quad (4.28)$$

#### 4.2.4 Modell für Fahrzeuge

In dem Messmodell für Fahrzeuge wird die Vereinfachung getroffen, dass sich ein Fahrzeug durch einen Quader beschreiben lässt. Somit ergeben sich je nach Blickwinkel auf das Fahrzeug zwei unterschiedliche Erscheinungsbilder: Ist nur eine Seite des Quaders sichtbar, kann ein I-Shape als Messung erwartet werden (siehe auch Abschnitt 4.2.2 bzw. Abbildung 4.10). Als Messwert für die Position wird die Mitte  $c$  der zugehörigen Fläche angenommen. In Abhängigkeit davon, welche Seite des Fahrzeuges sichtbar ist, kann entweder die Länge oder die Breite beobachtet werden. Die Orientierung  $\alpha$  des I-Shapes leitet sich direkt aus der Orientierung  $\phi$  des Fahrzeuges ab.

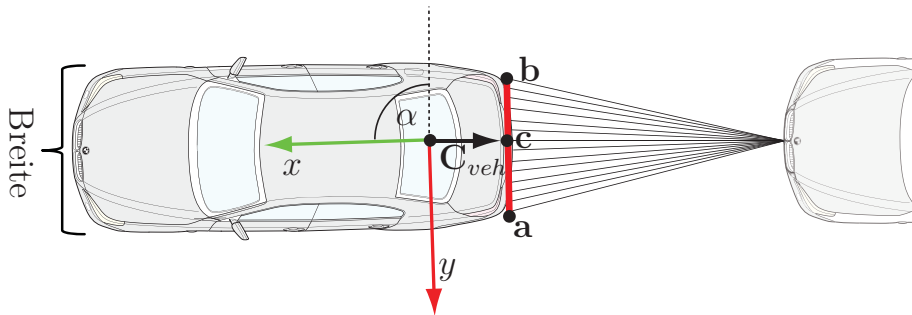


Abbildung 4.10: Messmodell für ein Fahrzeug im Laserscanner. Ist nur eine Seite des Fahrzeuges sichtbar, kann entweder nur die Länge  $l$  oder nur die Breite  $w$  gemessen werden. Die Position des Fahrzeuges ist durch den Mittelpunkt des I-Shapes bestimmt. Die Orientierung des Fahrzeuges  $\phi$  kann direkt durch die Orientierung  $\alpha$  der Laserscanner Merkmale gemessen werden. Der Ursprung des lokalen Fahrzeugkoordinatensystems ist mit  $C_{veh}$  bezeichnet.

Sind zwei Seitenflächen des Fahrzeuges sichtbar, wird ein L-Shape als Messung erwartet (siehe Abschnitt 4.2.2 bzw. Abbildung 4.11). In diesem Fall sind die Fahrzeuglänge bzw. -breite direkt messbar, da sie den Längen der beiden Schenkel des L-Shapes entsprechen. Als Messwert für die Position des Fahrzeuges wird nicht die Mitte einer Fläche, sondern der durch die beiden Flächen bestimmte Eckpunkt  $c$  verwendet, da dieser robuster gegenüber Messfehlern ist. Die Orientierung des L-Shapes  $\alpha$  lässt sich je nach sichtbaren Seiten aus der Orientierung  $\phi$  des Fahrzeuges ableiten.

In diesem Messmodell ist der grundlegende Schritt die Sichtbarkeitsentscheidung, die mit Standardverfahren aus der Computergrafik gelöst werden kann [BGZ02]. Durch die begrenzte Auflösung des

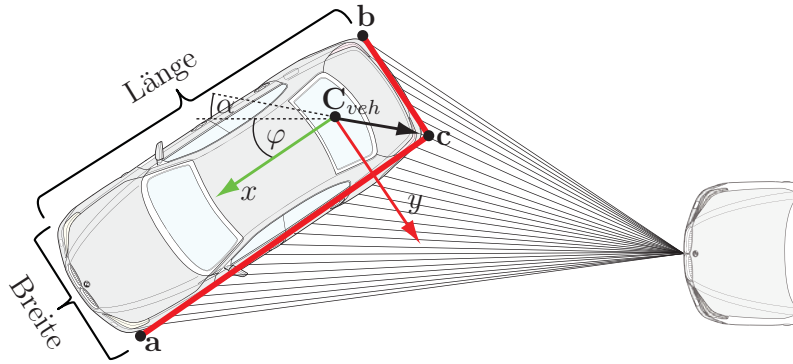


Abbildung 4.11: Messmodell für ein Fahrzeug im Laserscanner. Sind zwei Seiten des Fahrzeuges sichtbar, können Länge und Breite gleichzeitig vermessen werden. Die Position des Fahrzeuges ist durch den Eckpunkt des L-Shapes bestimmt. Die Orientierung  $\varphi$  kann direkt durch die Orientierung  $\alpha$  der Laserscanner Merkmale gemessen werden. Der Ursprung des lokalen Fahrzeugkoordinatensystems ist mit  $\mathbf{C}_{veh}$  bezeichnet.

Laserscanners und unterschiedliche Reflektanzen von Fahrzeugen gibt es hierfür allerdings keine harte Grenze. Vielmehr existiert ein Übergangsbereich, in dem sowohl ein L-Shape als auch ein I-Shape gemessen werden kann. In dieser Situation wird, je nachdem was gemessen wurde, die Entscheidung rückwirkend getroffen.

Für ein L-Shape lässt sich die Messung  $z_{L-Shape} = (r, \phi, \alpha, l_1, l_2)$  wie folgt aus dem Zustand  $\mathbf{x}_t^{veh} = (x, y, v, a, \varphi, \dot{\varphi}, w, l)^T$  berechnen:

$$h_{L-Shape}(\mathbf{x}_t^{veh}) = \begin{pmatrix} \sqrt{(c_x \cos(\varphi) - c_y \sin(\varphi) + x)^2 + (c_x \sin(\varphi) + c_y \cos(\varphi) + y)^2} \\ \arctan\left(\frac{c_x \sin(\varphi) + c_y \cos(\varphi) + y}{c_x \cos(\varphi) - c_y \sin(\varphi) + x}\right) \\ \varphi \pm \frac{\pi}{4} \\ w \\ l \end{pmatrix} \quad (4.29)$$

Der Punkt  $\mathbf{c} = (c_x, c_y)^T$  für die zu messende Ecke des L-Shapes ist im lokalen Fahrzeugkoordinatensystem  $\mathbf{C}_{veh}$  definiert und von der Länge und Breite des Fahrzeuges abhängig. Für das in Abbildung 4.11 dargestellte Beispiel entspricht  $c_x = -1$  und  $c_y = \frac{w}{2}$ .

Die Messung  $z_{I-Shape} = (r, \phi, \alpha, l_1)$  für ein I-Shape ergibt sich analog zu der Messung für ein L-Shape. Einziger Unterschied ist, dass abhängig von der sichtbaren Fläche entweder die Länge oder die Breite gemessen werden kann:

$$h_{I-Shape}(\mathbf{x}_t^{veh}) = \begin{pmatrix} \sqrt{(c_x \cos(\varphi) - c_y \sin(\varphi) + x)^2 + (c_x \sin(\varphi) + c_y \cos(\varphi) + y)^2} \\ \arctan\left(\frac{c_x \sin(\varphi) + c_y \cos(\varphi) + y}{c_x \cos(\varphi) - c_y \sin(\varphi) + x}\right) \\ f(\mathbf{x}_t^{veh}) \\ g(\mathbf{x}_t^{veh}) \end{pmatrix} \quad (4.30)$$



Der Mittelpunkt der sichtbaren Seitenfläche  $\mathbf{c} = (c_x, c_y)^T$  ist im lokalen Fahrzeugkoordinatensystem  $\mathbf{C}_{veh}$  definiert. Für das in Abbildung 4.10 dargestellte Beispiel entspricht  $c_x = -1$  und  $c_y = 0$ . Die Funktionen  $f$  und  $g$  aus Gleichung 4.30 sind wie folgt definiert:

$$f(\mathbf{x}_t^{veh}) = \begin{cases} \varphi, & \text{falls die Seitenfläche sichtbar ist} \\ \varphi - \frac{\pi}{2}, & \text{sonst} \end{cases} \quad (4.31)$$

$$g(\mathbf{x}_t^{veh}) = \begin{cases} w, & \text{falls die Breite gemessen werden kann} \\ l, & \text{sonst} \end{cases} \quad (4.32)$$

#### 4.2.5 Anpassungen an den Partikel-Filter

Falls für das Tracking ein Partikel-Filter (siehe Kapitel 5) zum Einsatz kommt, muss neben den Messfunktionen  $h_{L-Shape}$  bzw.  $h_{I-Shape}$  noch eine Gewichtung der einzelnen Partikel definiert werden. Unter der Annahme eines normalverteilten Messrauschens ergibt sich für die einzelnen Partikel-Filter-Hypothesen  $x^{(i)}$  bei einer vorliegenden L-Shape Messung  $\mathbf{z}_{L-Shape}$  die zugehörige Gewichtung  $w^{(i)}$  des Partikels durch eine Normalverteilung:

$$w^{(i)} = \frac{1}{\sqrt{(2\pi)^5 \cdot \det(\mathbf{R}_{LIDAR})}} \cdot e^{-\frac{1}{2}(\mathbf{h}_{L-Shape}(x^{(i)}) - \mathbf{z}_{L-Shape})^T \mathbf{R}_{LIDAR}^{-1} (\mathbf{h}_{L-Shape}(x^{(i)}) - \mathbf{z}_{L-Shape})} \quad (4.33)$$

Für eine I-Shape Messung, bei der Breite und Länge nicht gleichzeitig messbar sind, verändert sich die Gewichtung der Messung  $\mathbf{z}_{L-Shape}$  bzw. des  $i$ -ten Partikels  $x^{(i)}$  zu:

$$w^{(i)} = \frac{1}{\sqrt{(2\pi)^4 \cdot \det(\mathbf{R}_{LIDAR})}} \cdot e^{-\frac{1}{2}(\mathbf{h}_{I-Shape}(x^{(i)}) - \mathbf{z}_{I-Shape})^T \mathbf{R}_{LIDAR}^{-1} (\mathbf{h}_{I-Shape}(x^{(i)}) - \mathbf{z}_{I-Shape})} \quad (4.34)$$

#### 4.2.6 Bewertung

Die Erzeugung einer Ausgleichsgeraden ist der entscheidende Schritt für die in Abschnitt 4.2.2 erläuterten Messverfahren. Im Folgenden soll deshalb der Einfluss verschiedener Verfahren zur Bildung einer Regressionsgeraden auf die Qualität der Merkmalsgewinnung untersucht werden. Hierzu wurden zwei unterschiedliche Szenarien untersucht. In dem ersten Versuchsaufbau werden mehrere Fahrzeuge auf einer Autobahn betrachtet. Somit ist ein möglichst allgemeines Szenario mit unterschiedlichen Entfernungen, Fahrzeugtypen und Geschwindigkeiten sichergestellt. Für das zweite Szenario wurde ein Datensatz gewählt, der besonders viele Messfehler enthält und dementsprechend als besonders schwierig zu betrachten ist. Für die Evaluierung der Güte der Verfahren wurde die Genauigkeit (RMSE) des Trackings (siehe Kapitel 5) verwendet, da diese im Wesentlichen direkt von der Qualität der Merkmalsgewinnung abhängt. Untersucht wurden die einfache lineare Regression, wie sie auch in Abschnitt 4.2.1 beschrieben ist, sowie eine Implementierung nach dem RANSAC-Verfahren [FB81]. Hierbei wurde der Einfluss der beiden Ansätze für L-Shapes bzw. I-Shapes gesondert betrachtet. In den Tabellen 4.1a bzw. 4.1b ist die Qualität des Trackings für beide Szenarien dargestellt.



	$r$ (RMSE)	$\varphi$ (RMSE)		$r$ (RMSE)	$\varphi$ (RMSE)
Ohne RANSAC	0.17 m	0.44 °	Ohne RANSAC	0.16 m	0.73 °
RANSAC L-Shape	0.17 m	0.38 °	RANSAC L-Shape	0.16 m	0.62 °
RANSAC I-Shape	0.17 m	0.39 °	RANSAC I-Shape	0.16 m	0.64 °
RANSAC beides	0.17 m	0.32 °	RANSAC beides	0.15 m	0.50 °

(a) Szenario 1

(b) Szenario 2

Tabelle 4.1: Einfluss verschiedener Verfahren zur Berechnung von Regressionsgeraden auf die Fehler in der Zustandsschätzung für die Entfernung  $r$  und der lateralen Ablage  $\varphi$ . Es wurden zwei Szenarien mit Fahrzeugen verglichen. Das Szenario 1 beschreibt eine typische Autobahnscene mit unterschiedlichen Fahrzeugmodellen auf mehreren Spuren. Für Szenario 2 wurde ein Datensatz gewählt, der höheres Messrauschen enthält. Der Einfluss einer Implementierung nach RANSAC wurde für L-Shapes und I-Shape gesondert betrachtet.

Insbesondere in dem zweiten Szenario, das viele Messfehler enthält, sind mit dem RANSAC Verfahren deutlich bessere Ergebnisse zu erreichen, da dieser Ansatz Ausreißer sehr gut unterdrücken kann. Messfehler wirken sich häufig direkt auf die Orientierung der L-Shapes und I-Shapes aus. Da die Orientierungsmessung an die Gierrate des Fahrzeugmodells gekoppelt ist, wird dessen laterale Ablage  $\varphi$  beeinflusst. Die Entfernungsmessung  $r$  ist hingegen nur sehr gering von Fehlern in der Orientierung betroffen. Beide Effekte sind sehr gut in den Ergebnissen aus Tabelle 4.1 zu erkennen: Die Schätzung der Entfernung ist weitgehend unabhängig von Fehlern in der Messung der Orientierung. Die Schätzung der lateralen Ablage kann hingegen durch ein robustes Verfahren zur Bestimmung von Regressionsgeraden optimiert werden. In dem allgemeinen Szenario ist der Unterschied der beiden Verfahren nicht mehr so deutlich, aber immer noch vorhanden. Auf eine Evaluierung in Bezug auf Fußgänger wurde verzichtet, da für die Positionsbestimmung von Fußgängern der Schwerpunkt der Rohpunkte als Messung verwendet und somit keine Verbesserung zu erwarten ist.

#### 4.2.7 Zusammenfassung

Für eine Fahrumfelderfassung ist ein Laserscanner im Augenblick einer der vielseitigsten und genauesten Sensoren. Mit seiner Hilfe kann sowohl eine Vorklassifizierung vorgenommen (vergleiche [Wal08]), als auch Zustandsgrößen wie die Position, die Länge, die Breite und die Orientierung von Verkehrsteilnehmern bestimmt werden. Aus diesem Grund wurde das Messmodell für diesen Sensor sehr ausführlich behandelt. In einem ersten Vorverarbeitungsschritt werden die Rohdaten in zusammengehörige Bereiche unterteilt. Hierzu wurde eine Erweiterung des Region-Growing Verfahrens vorgestellt, die im Rahmen einer Fahrumfelderfassung eine zuverlässigere und vollständigere Segmentierung von Fahrzeugen und Fußgängern gewährleistet. Basierend auf diesen Segmenten werden Merkmale extrahiert, die die zu erwartenden Messwerte der Beobachtungsmodelle beschreiben. Hierbei wurde die Annahme getroffen, dass sich Fahrzeuge durch Quader und Fußgänger durch Zylinder approximieren lassen. Der wesentliche Bestandteil dieser Merkmalsextraktion wurde durch ein Verfahren erweitert, das sich als sehr robust gegenüber Messfehlern erwies. Abschließend wurden die eigentlichen Messmodelle für Fahrzeuge und Fußgänger eingeführt und beschrieben, wie sich die jeweiligen Modelle für einen Partikel-Filter verwenden lassen.

### 4.3 Radarsensorik

Seit 1998 haben sich frequenzmodulierte Dauerstrichradare im automotiven Umfeld für Fahrerassistenzsysteme wie das ACC etabliert. In [Foe06] findet sich ein guter Überblick wie mittels Radarsensorik ausgedehnte Objekte des Fahrumfeldes erfasst werden können. Für die in dieser Arbeit verwendeten Sensoren ist eine dedizierte Vorverarbeitung der Messwerte nicht sinnvoll, da auf die sensorinterne Bearbeitung der Daten kein Einfluss genommen werden konnte und die Rohdaten deshalb nicht zur Verfügung standen. Darüber hinaus können für die stark vorverarbeiteten Daten weder zusätzliche Merkmale gebildet, noch eine Datenreduktion erreicht werden. Die Messung  $\mathbf{z}_r$  für ein Zielobjekt ist definiert durch dessen Position in Polarkoordinaten ( $r$  und  $\varphi$ ) sowie der relativen radialen Geschwindigkeit  $v_r$ :

$$\mathbf{z}_r = (r, \varphi, v_r)^T \quad (4.35)$$

In Abbildung 4.12 ist der Zusammenhang zwischen realer Geschwindigkeit  $\mathbf{v}$  und der radialen Geschwindigkeit eines Zielobjektes dargestellt. Die Berechnung des radialen Anteils  $v^{\parallel}$  der Geschwindigkeit ist gegeben durch die Position  $\mathbf{x}$  und dem Geschwindigkeitsvektor  $\mathbf{v}$  des Zielobjektes:

$$v^{\parallel} = \cos(\beta) \|\mathbf{v}\| \quad (4.36)$$

$$= \frac{\mathbf{v}\mathbf{x}}{\|\mathbf{x}\|} \quad (4.37)$$

Aus der Geschwindigkeit des Eigenfahrzeuges  $\mathbf{v}_{ego}$  und der radialen Geschwindigkeit  $v^{\parallel}$  ergibt sich die relative radiale Geschwindigkeit  $v_r$  wie folgt:

$$v_r = v^{\parallel} - \cos(\alpha) \|\mathbf{v}_{ego}\| \quad (4.38)$$

$$= v^{\parallel} - \frac{1}{\sqrt{1 + \left(\frac{y}{x}\right)^2}} \|\mathbf{v}_{ego}\| \quad (4.39)$$

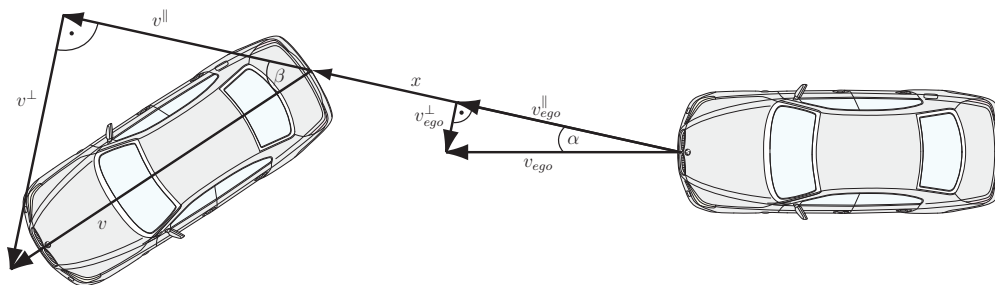


Abbildung 4.12: Radarmessung der radialen Geschwindigkeit. Bei Doppler-Radaren wird die relative radiale Geschwindigkeit  $v_r$  gemessen. Sie ist abhängig von der Position  $\mathbf{x}$ , dem Geschwindigkeitsvektor  $\mathbf{v}$  des Zielobjektes und der Geschwindigkeit des Eigenfahrzeuges  $\mathbf{v}_{ego}$ .

### 4.3.1 Bestimmung der Messunsicherheit

Filterverfahren, wie zum Beispiel der Kalman-Filter (siehe Kapitel 5), setzen voraus, dass die Messunsicherheit bekannt ist. Darüber hinaus wird oft angenommen, dass die Messfehler normalverteilt sind. Die Angaben der Sensorhersteller basieren im Allgemeinen auf genormten Standardzielen und einer Normalverteilung der Messwerte, was unter Umständen eine zu idealisierte Annahme ist. Deshalb wird im Folgenden versucht, diese Werte empirisch zu ermitteln. Hierzu wird das in Kapitel 3.2 beschriebene Verfahren zur Referenzdatengewinnung verwendet.

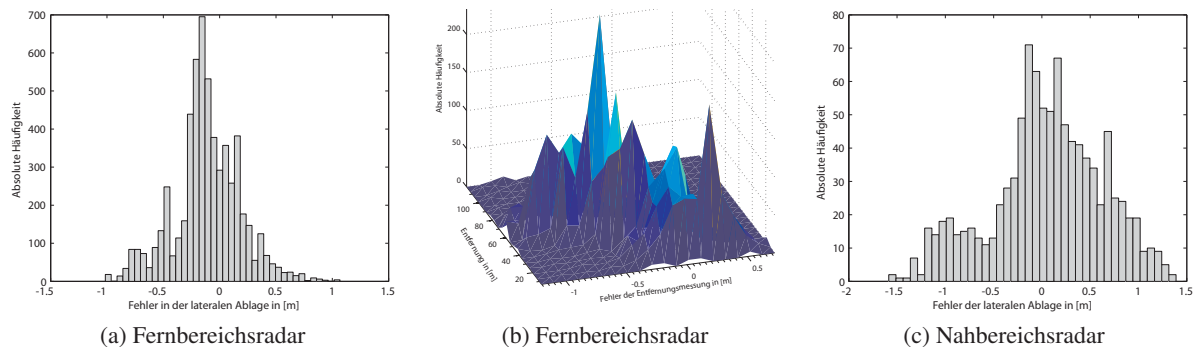


Abbildung 4.13: Messwertverteilung der Fern- und Nahbereichsradare. In den Abbildung (a) bzw. (c) sind die Histogramme der Messfehler für die laterale Ablage dargestellt. Während diese Fehler in (a) mit einer Normalverteilung approximiert werden können, sind in (c) zwei Maxima (Reflexionsquellen) zu erkennen. In dem Histogramm der Abweichungen der Entfernungsmessung aus Abbildung (b) ist eine Abhängigkeit von der Entfernung zu dem Zielobjekt zu erkennen.

In Abbildung 4.13 sind für Nah- und Fernbereichsradare die Histogramme der Fehler aufgetragen. Für die Auswertung diente ein PKW als Referenzziel, wobei die Mitte des Fahrzeughecks den Referenzpunkt darstellt. Für Fernbereichsradare kann näherungsweise von einer Normalverteilung der Fehler für die lateralen Ablage ausgegangen werden (siehe Abbildung 4.13a). Die Verteilung der Abweichung in der Entfernungsschätzung ist allerdings von dem Abstand des Zielobjektes abhängig, was deutlich in Abbildung 4.13b zu erkennen ist: Mit größer werdender Entfernung verschiebt sich die Verteilung nach links. Dies ist dadurch zu erklären, dass bei einem PKW mehrere mögliche Reflexionsquellen existieren. So kann ein Echo zum Beispiel von der Rückseite, den Rückleuchten, der Hinterachse oder dem Motorblock erzeugt worden sein. Insbesondere bei den Nahbereichsradaren ist dieser Effekt gut zu erkennen. In dem Histogramm für den Fehler in der lateralen Ablage aus Abbildung 4.13c ist deutlich zu erkennen, dass die linke Rückleuchte bzw. die linke hintere Ecke des Fahrzeuges für ein zweites lokales Maximum bei -1 m verantwortlich ist.

Im weiteren Verlauf dieser Arbeit wird die Messunsicherheit der Radarsensorik mit der Kovarianzmatrix  $\mathbf{R}_{RADAR}$  bezeichnet.

### 4.3.2 Fahrzeuge

Aufgrund der multiplen Reflexionsquellen kann bei einem Fahrzeug nur bedingt von einer Normalverteilung der Messwerte ausgegangen werden (siehe Abschnitt 4.3.1). Der Effekt dieser systematischen

Fehler kann zum Einen durch eine ausreichend große Varianz minimiert werden. Dies führt zu einer künstlichen Verschlechterung der Messwerte. Zum Anderen können die möglichen Reflexionsquellen für Fahrzeuge explizit modelliert werden. Dies ist besonders dann sinnvoll, wenn durch einen zweiten Sensor, wie zum Beispiel einen Laserscanner, bereits Vorwissen in Bezug auf die Position, Orientierung und die Länge bzw. Breite eines Fahrzeuges vorhanden ist und somit Mehrdeutigkeiten ausgeschlossen werden können. Im Folgenden wird ein derartiges erweitertes Messmodell vorgestellt.

### Multiple Reflexionsquellen

Wie in Abschnitt 4.1 beschrieben ist der Zustand eines Fahrzeuges  $\mathbf{x}_t^{veh}$  definiert als:

$$\mathbf{x}_t^{veh} = (x, y, v, a, \varphi, \dot{\varphi}, w, l)^T \quad (4.40)$$

In dem lokalen Fahrzeugkoordinatensystem  $C_{veh}$  werden vier typische Reflexionsquellen definiert: Die Mitte der Hinterachse  $\mathbf{x}$  bzw. der Heckansicht  $\mathbf{c}$  sowie die rechte und linke hintere Ecke des Fahrzeuges ( $\mathbf{r}$  bzw.  $\mathbf{l}$ ). In Abbildung 4.14 sind diese Punkte schematisch dargestellt. Das Eintreten einer Reflexion an einem dieser Punkte ist allerdings nicht gleich wahrscheinlich. Vielmehr ist die Reflexion abhängig von der Sichtbarkeit bzw. dem Winkel  $\alpha$  und  $\beta$  (siehe Abbildung 4.12). Für die in Abbildung 4.12 dargestellte Konstellation ist es zum Beispiel unwahrscheinlich, dass die hintere rechte Fahrzeugecke für eine Reflexion verantwortlich ist. In dem Messmodell sollte diese Abhängigkeit berücksichtigt werden, um Mehrdeutigkeiten bei der Zuordnung von realen Messwerten zu vorhergesagten Messwerten zu minimieren. Für die Reflexionspunkte an den Ecken eines Fahrzeuges kann dies durch einen einfachen Sichtbarkeitsentscheid der zugehörigen Flächen erfolgen. Hierfür können Standardverfahren aus der Computergrafik verwendet werden, wie sie zum Beispiel in [BGZ02] beschrieben sind.

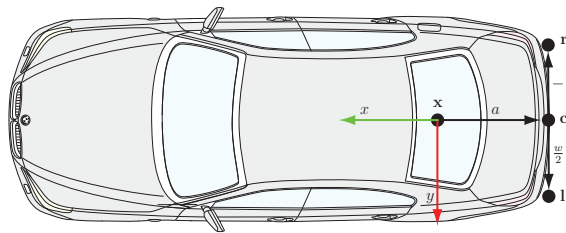


Abbildung 4.14: Modell von vier Reflexionsquellen für ein Fahrzeug: Die Mitte der Hinterachse  $\mathbf{x}$  bzw. der Heckansicht  $\mathbf{c}$  sowie die rechte und linke hintere Ecke des Fahrzeuges ( $\mathbf{r}$  bzw.  $\mathbf{l}$ ).

Für den Reflexionspunkt  $\mathbf{r}$  ist die Messfunktion  $h_r(\mathbf{x}_t^{veh})$  wie folgt definiert:

$$h_r(\mathbf{x}_t^{veh}) = \begin{pmatrix} \sqrt{\hat{x}_r^2 + \hat{y}_r^2} \\ \arctan\left(\frac{\hat{y}_r}{\hat{x}_r}\right) \\ \frac{(v_x, v_y)(\hat{x}_r, \hat{y}_r)}{\|(\hat{x}_r, \hat{y}_r)\|} - \frac{1}{\sqrt{1 + \left(\frac{\hat{y}_r}{\hat{x}_r}\right)^2}} \|\mathbf{v}_{ego}\| \end{pmatrix} \quad (4.41)$$

Die einzelnen Komponenten der Messfunktion lassen sich wie folgt aus den jeweiligen Zustandsgrößen ableiten:

$$\hat{\mathbf{x}}_r = \mathbf{C}_{veh} \mathbf{r} \quad (4.42)$$

$$v_x = \cos(\varphi)v \quad (4.43)$$

$$v_y = \sin(\varphi)v \quad (4.44)$$

Die restlichen Reflexionspunkte werden analog zu 4.41 bestimmt.

### Bewertung

Das Modell der multiplen Reflexionspunkte bietet die Möglichkeit, die Messwertverteilung der Radarsensorik besser zu approximieren, indem objektspezifische Rückstrahl-Charakteristika berücksichtigt werden. Dieses Verfahren ist allerdings auf das Zusatzwissen eines weiteren Sensors angewiesen, der die Mehrdeutigkeiten auflöst, die an den Übergängen zwischen den einzelnen Reflexionspunkten entstehen können. Liegt ein Messwert zum Beispiel zwischen der linken hinteren Ecke und dem Mittelpunkt des Fahrzeughecks, so kann sich das Fahrzeug, je nachdem welchem der Reflexionspunkte der Messwert zugeordnet wird, entweder nach links oder nach rechts bewegt haben. Ein möglicher zweiter Sensor könnte eine Kamera oder ein Laserscanner sein.

In Tabelle 4.2 ist dargestellt, wie sich das Modell der multiplen Reflexionsquellen auf die Genauigkeit der Zustandsschätzung der Entfernung  $r$  und der lateralen Ablage  $\varphi$  auswirkt. Grundlage der Evaluierung war eine Fusion aus Laserscanner und Nahbereichsradaren. Als Qualitätsmaß wurde der RMSE für die Schätzgröße der Entfernung  $r$  sowie der lateralen Ablage  $\varphi$  gewählt (vergleiche die Metriken aus Abschnitt 7.2.1). Die hierfür nötigen Referenzdaten wurden durch einen dedizierten Sensor (siehe Abschnitt 7.1.3) ermittelt.

Ausgangspunkt der Evaluierung war ein Testszenario, in dem ein vorausfahrendes Fahrzeug mehrere Spurwechsel durchführt. Mit diesem Szenario lässt sich der Effekt von multiplen Reflexionsquellen an den Fahrzeugecken und dem Heck simulieren. Es konnte gezeigt werden, dass mit diesem Modell eine Steigerung in der Schätzung der lateralen Ablage erreicht werden kann. Der Einfluss auf die Entfernungsschätzung ist hingegen nur sehr gering.

	$r$ (RMSE)	$\varphi$ (RMSE)
Modell mit einem Echo	0.148 m	0.69 °
Modell mit multiplen Reflexionsquellen	0.143 m	0.57 °

Tabelle 4.2: Vergleich der Fehler in der Zustandsschätzung mit multiplen Reflexionsquellen und mit einem Echo.

### 4.3.3 Fußgänger

Mit den in dieser Arbeit verwendeten Radarsensoren können Fußgänger nur sehr bedingt erfasst werden. Insbesondere die Fernbereichsradare sind auf Fahrzeuge und hohe Reichweiten optimiert und haben einen sehr eingeschränkten Öffnungswinkel. Dementsprechend werden Fußgänger nur sehr selten und nur direkt vor dem eigenen Fahrzeug detektiert. Aktuelle Weiterentwicklungen im automotiven Umfeld, wie scannende Radare oder spezielle Nahbereichsradare, sind in dieser Hinsicht weitaus sensitiver, werden aber im Rahmen dieser Arbeit nicht weiter betrachtet.

Wie in Abschnitt 4.1 beschrieben ist der Zustand eines Fußgänger  $\mathbf{x}_t^{ped}$  wie folgt definiert:

$$\mathbf{x}_t^{ped} = (x, y, v_x, v_y, w, h)^T \quad (4.45)$$

Fußgänger werden als einfaches Punktziel modelliert. Dementsprechend ergibt sich folgende Messfunktion  $h_r(\mathbf{x}_t^{ped})$ :

$$h_r(\mathbf{x}_t^{ped}) = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arctan(y/x) \\ \frac{(v_x, v_y)(x, y)}{\|(x, y)\|} - \frac{1}{\sqrt{1 + (\frac{y}{x})^2}} \|\mathbf{v}_{ego}\| \end{pmatrix} \quad (4.46)$$

### 4.3.4 Anpassungen an den Partikel-Filter

Falls für das Tracking ein Partikel-Filter zum Einsatz kommt, muss neben der Messfunktion  $h_r$  noch eine Gewichtung der einzelnen Partikel definiert werden. Für die Hypothese  $x^{(i)}$  ist die zugehörige Gewichtung  $w^{(i)}$  des Partikels durch folgende Gleichung bestimmt:

$$w^{(i)} = \frac{1}{\sqrt{(2\pi)^3 \cdot \det(\mathbf{R}_{RADAR})}} \cdot e^{-\frac{1}{2}(h_r(x^{(i)}) - \mathbf{z}_r)^T \mathbf{R}_{RADAR}^{-1} (h_r(x^{(i)}) - \mathbf{z}_r)} \quad (4.47)$$

### 4.3.5 Zusammenfassung

Radarsensoren haben sich im automotiven Umfeld etabliert und kommen bisher vornehmlich in Fahrerassistenzsystemen wie dem ACC zum Einsatz. In diesem Abschnitt wurde ein Messmodell für die Radarsensorik vorgestellt, das die charakteristischen Rückstrahleigenschaften von Fahrzeugen explizit berücksichtigt. Es konnte gezeigt werden, dass sich damit im Vergleich zu einem einfachen Punktziel eine höhere laterale Genauigkeit erreichen läßt. Allerdings ist hierfür ein zweiter Sensor (Laserscanner) notwendig, der Mehrdeutigkeiten verringert. Fußgänger werden hingegen als Punktziel modelliert. Es ist dabei zu beachten, dass diese mit der in dieser Arbeit verwendeten Radarsensorik nur sehr eingeschränkt erfasst werden können, da insbesondere die Fernbereichsradare auf Fahrzeuge und hohe Reichweiten optimiert sind und zusätzlich einen sehr eingeschränkten Öffnungswinkel besitzen. Abschließend wurde beschrieben, wie sich die einzelnen Modelle für einen Partikel-Filter verwenden lassen.

## 4.4 Konturmodell Video

Im Vergleich zu den bisher beschriebenen Sensoren besitzen Kameras im Allgemeinen eine sehr hohe Winkelauflösung. Dadurch bieten sie die Möglichkeit, die laterale Ablage von Verkehrsteilnehmern sehr genau zu messen. Allerdings können nur bedingt Angaben über die Entfernung von einzelnen Objekten getroffen werden. Im Folgenden wird ein Messmodell für Fußgänger und Fahrzeuge vorgestellt, das auf Konturverläufen basiert und somit sowohl für Ferninfrarot- als auch normale Videokameras geeignet ist.

Eine Möglichkeit, möglichst beliebige Objektmodelle schnell und einfach zu modellieren und im Videobild zu erkennen, ist ein Konturerkenner welcher anhand einer Konturvorgabe das Erkennen eines Objektes ermöglicht [BH94]. Die Vorlage der Kontur ist dabei durch eine Menge von Liniensegmenten gegeben, welche die markanten Umrisse eines Objektes darstellen. Für jedes Liniensegment werden eine oder mehrere orthogonale Linien definiert, entlang derer nach starken Grauwertdifferenzen gesucht wird. Anhand der Entfernung starker Grauwertdifferenzen zum betrachteten Liniensegment kann das Zutreffen einer bestimmten Objekthypothese überprüft werden.

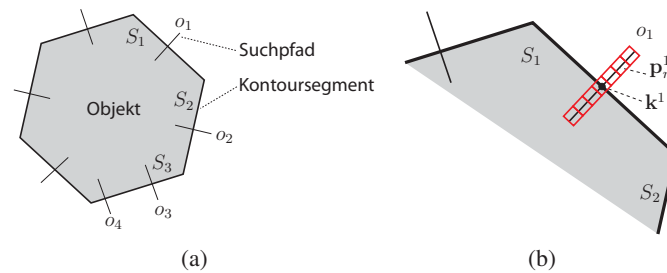


Abbildung 4.15: Allgemeines Konturmodell: Eine Menge an Kontursegmenten beschreibt die markanten Umrisse eines Objektes. Auf den Suchpfaden  $o_i$ , die orthogonal zu den einzelnen Kontursegmenten  $S_j$  definiert sind, wird nach starken Grauwertdifferenzen gesucht. In Abbildung (b) sind die einzelnen Bildpunkte  $\mathbf{p}_m^1$  auf dem Suchpfad  $o_1$  dargestellt. Der Punkt  $\mathbf{k}^1$  beschreibt den Schnittpunkt von  $o_1$  mit  $S_1$ .

Sei  $\mathcal{N} = \{o_1, \dots, o_n\}$  die Menge der zu den Kontursegmenten  $\mathcal{S} = \{S_1, \dots, S_m\}$  orthogonalen Suchpfade  $o_i$  der Länge  $l$ , so wird entlang jeder Normalen  $o_i$  an den zugehörigen Bildpunkten  $\mathbf{p}_m^i$  nach Grauwertdifferenzen über einem bestimmten Schwellwert  $\theta$  gesucht, wobei  $(\mathbf{p}_1^i, \dots, \mathbf{p}_{l-1}^i)$  den Punkten auf dem Suchpfad  $o_i$  entsprechen. Im Folgenden bezeichnet  $\mathbf{k}^i$  den Schnittpunkt der Normalen  $o_i$  mit dem zugehörigen Kontursegment  $S_j$ . Für jeden Suchpfad  $o_i \in \mathcal{N}$  ist folgendes Ähnlichkeitsmaß  $\Psi(i)$  definiert:

$$\Psi(i) = \begin{cases} \min_{\mathbf{p}_m^i} |\mathbf{k}^i - \mathbf{p}_m^i| & , \text{ falls } \exists m : |f(\mathbf{p}_m^i) - f(\mathbf{p}_{m+1}^i)| > \theta \\ \xi & , \text{ sonst} \end{cases} \quad (4.48)$$

Hierbei entspricht  $f(\mathbf{p})$  dem Grauwert im Bild an der Position  $\mathbf{p}$ . Der Parameter  $\xi$  dient als Indikator, dass kein passender Grauwertübergang gefunden wurde. Im Folgenden wird der durchschnittliche Abstand aller Suchpfade als Ähnlichkeitsmaß für die Übereinstimmung des Konturmodells mit dem Bildausschnitt verwendet.



#### 4.4.1 Kopfkontur für Fußgänger

Der allgemeine Ansatz eines Konturmodells lässt sich sehr gut auf das typische Erscheinungsbild eines Fußgängers im Infrarotbild übertragen. Bis zu einer Umgebungstemperatur  $\leq 25^\circ$  heben sich alle nicht mit stark isolierender Kleidung bedeckten Körperteile sehr gut von dem im Allgemeinen kälteren Hintergrund ab (siehe auch Abbildung 4.17a). Im Vergleich zum normalen Videobild sind Fußgänger im Infrarotbild weniger texturiert und im Wesentlichen durch ihre Kontur definiert. Da die Kopfregion, insbesondere das Gesicht, im Allgemeinen nicht mit Kleidung bedeckt ist, hat sich dieser Bereich als sehr charakteristische Region erwiesen. Darüber hinaus bietet das Erscheinungsbild des Kopfes den Vorteil annähernd invariant gegenüber dem Blickwinkel zu sein. Aus diesen Gründen verwenden zum Beispiel Xu et al. [XLF05] ebenfalls das Erscheinungsbild der Kopfes in einer Wärmebildkamera. Hierbei wird als Beobachtungsmodell das Meanshift-Verfahren [FH75] verwendet. In der Arbeit von Yasuno [YRYA05] wird die Methode des Template-Matchings verwendet, um die Position des Kopfes mit einer FIR-Kamera zu verfolgen. Für eine detaillierte Übersicht über aktuelle videobasierte Verfahren zur Erkennung und Verfolgung von Fußgängern im Kontext einer Fahrumfelderfassung sei auf Gavrilla und Munder [GM07] verwiesen.

Im Folgenden wird für die Wärmebildkamera das Beobachtungsmodell einer Kopfkontur eingeführt. Die Region des Kopfes wird durch Kontursegmente, wie sie in Abbildung 4.16 dargestellt sind, approximiert.

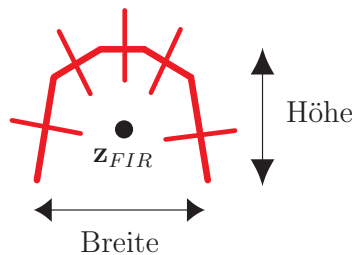


Abbildung 4.16: Konturmodell für den Kopf eines Fußgängers im Wärmebild.

Die Breite und Höhe des Modells ist hierbei nicht in Bild-, sondern in Weltkoordinaten definiert, wodurch eine entfernungsinvariante Darstellung gegeben ist. Für die Messung wird das Konturmodell an die 3D Position des Fußgängerkopfes platziert und mit der Matrix  $\mathbf{P}_{fir}$  in das Kamerabild projiziert. Mit diesem Schritt wird das Modell zum Einen in Bildkoordinaten transformiert und zum Anderen entsprechend der Entfernung des Fußgängers skaliert. Schließlich werden die Grauwertdifferenzen orthogonal zu dieser Kontur gemessen und für jede Normale  $o_i$  das Abstandsmaß  $\psi(i)$  berechnet. In Abbildung 4.17 ist die in das Wärmebild projizierte Kopfkontur dargestellt.

Die allgemeine Messfunktion  $h_{FIR}$  des Messwertes  $\mathbf{z}_{FIR} = (u, v)^T$  für den Kopf ergibt sich aus dem Zustand  $\mathbf{x}_t^{ped} = (x, y, v_x, v_y, w, h)^T$  des Fußgängers wie folgt:

$$h_{FIR}(\mathbf{x}_t^{ped}) = \mathbf{P}_{FIR}(x, y, h, 1)^T \quad (4.49)$$



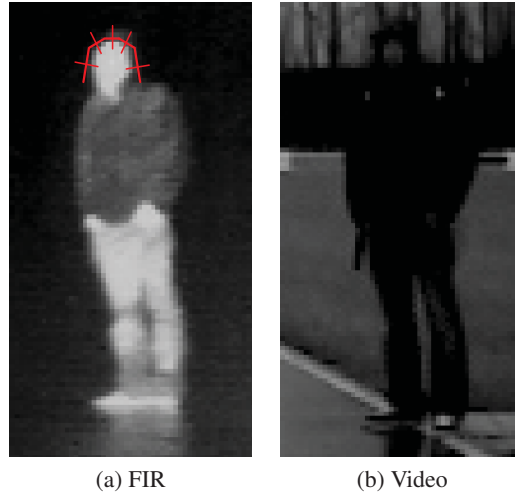


Abbildung 4.17: Abbildung (a) zeigt die Projektion der Kopfkontur in das Wärmebild. Zum Vergleich ist in Abbildung (b) dieselbe Szene aus der Sicht der Videokamera dargestellt. In dem Bild der Ferninfrarotkamera lässt sich der Vordergrund deutlich einfacher von Hintergrund separieren.

#### Anpassungen für den Partikel-Filter

Für den Partikel-Filter wird das Konturmodell verwendet, um eine Hypothese (Partikel) im Bild zu verifizieren. Hierzu wird das durchschnittliche Abstandsmaß aller Suchpfade verwendet. Die Gewichtung  $w^{(i)}$  des zugehörigen Partikels  $x^{(i)}$  einer Objekthypothese erfolgt schließlich mittels einer Normalverteilung:

$$w^{(i)} = \frac{1}{\sqrt{(2\pi)\sigma_{FIR}}} e^{-\frac{1}{2\sigma_{FIR}^2} \left( \frac{1}{n} \sum_{j=1}^n \Psi^{(i)}(j)^2 \right)} \quad (4.50)$$

#### Anpassungen für den Kalman-Filter

Um einen Messwert  $\mathbf{z}_{fir}$  für den Kalman-Filter zu erzeugen, wird iterativ in einem Suchfenster  $\mathcal{F}$ , das abhängig vom Zustand  $\mathbf{x}_t^{ped}$  definiert ist, mit der Gewichtsfunktion  $w^{(K_i)}$  aus 4.50 das Zutreffen des Konturmodells  $K_i$  überprüft. Die Kontur  $K_i \in \mathcal{F}$ , die das größte Gewicht zugewiesen bekommen hat, wird als Messung verwendet. Um Fehlmessungen zu vermeiden, muss das Gewicht zusätzlich über einem Schwellwert  $\Theta_{FIR}$  liegen.

$$\mathbf{z}_{fir} = \begin{cases} \operatorname{argmax}_{K_i} w^{(K_i)} & , \text{ falls } \exists K_i : w^{(K_i)} \geq \Theta_{FIR} \quad \text{wobei } K_i \in \mathcal{F} \\ \text{nicht definiert} & , \text{ sonst} \end{cases} \quad (4.51)$$

#### 4.4.2 Fahrzeugmodell

Fahrzeuge besitzen im Allgemeinen sehr viele vertikale oder horizontale Strukturen, wie zum Beispiel die Reifen, die Stoßstange, Türen oder Fenster. Im Videobild führt dies zu ausgeprägten Kanten, die sich sehr gut für eine Hypothesengenerierung für Fahrzeuge eignen [SBM06]. Dies ist im Wesentlichen auf die klar definierten Formen und Farben und die damit verbundene charakteristische Kontur zurückzuführen. Im Folgenden wird deshalb das Beobachtungsmodell einer Fahrzeugkontur eingeführt, das diese charakteristischen Bildmerkmale verwendet. Für einen sehr guten Überblick über aktuelle bildbasierte Verfahren zur Detektion von Fahrzeugen sei auf die Zusammenfassung von Sun et al. [SBM06] oder auf die Arbeit von Leonhard Walchshäusl [Wal08] verwiesen.

Für Fahrzeuge wird die Annahme getroffen, dass sie sich durch einen Quader approximieren lassen, der durch die Zustandskomponenten Position, Orientierung, Breite, Länge sowie die Höhe eindeutig definiert ist. Für diesen Quader werden die vom Sensor aus sichtbaren Kanten ermittelt und in das Videobild projiziert (siehe Abbildung 4.18). Diese Kanten bilden die Segmente des Konturmodells. Die horizontalen Oberkanten des Quaders werden allerdings nicht berücksichtigt, da keine charakteristische Grauwertkanten am Fahrzeugoberbau erwartet werden können. Von den restlichen horizontalen Kanten werden nur die des Hecks berücksichtigt, da sich das Fahrzeug an dieser Stelle durch seinen eigenen Schatten deutlich von der Straße abhebt. Sind, wie in Abbildung 4.18b dargestellt, zwei Flächen des Quaders sichtbar, so wird für die durch diese Flächen bestimmte vertikale Kante ebenfalls keine Messung durchgeführt, da auch hier keine ausgeprägten Kanten im Bild vorhanden sind. Somit werden nur diejenigen Kontur-Kanten berücksichtigt, die das Fahrzeug deutlich vom Hintergrund (der Straße) abgrenzen. In Abbildung 4.18b ist das Konturmodell für ein Fahrzeug dargestellt. Rote Kanten werden für die Messung herangezogen. Graue Kanten sind zwar sichtbar, werden aber nicht verwendet. Alle nicht sichtbaren Segmente sind gestrichelt dargestellt.

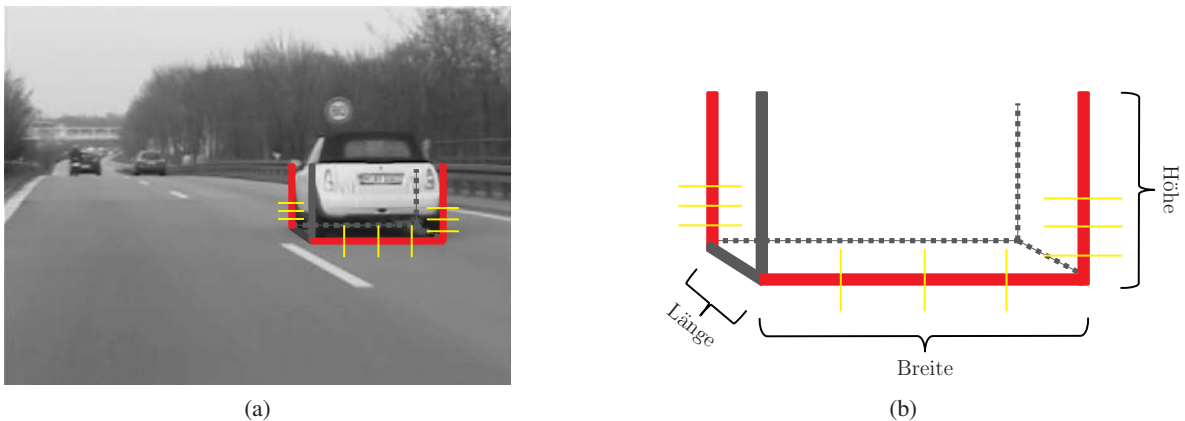


Abbildung 4.18: Konturmodell eines PKW: Links die Projektion der quaderförmigen Objekthypothese in das Videobild. Kanten, für die Messungen erfolgen, sind rot eingezeichnet. Die kürzeren gelben Linien repräsentieren die erwähnten Normalen, entlang derer nach starken Grauwertdifferenzen gesucht wird. In der rechten Abbildung ist das Konturmodell dargestellt. Es ist durch die Zustandskomponenten Position, Orientierung, Breite, Länge sowie die Höhe eindeutig definiert.

Die allgemeine Messfunktion  $h_{Video}$  des Messwertes  $\mathbf{z}_{Video} = (u, v)^T$  für ein Fahrzeug ergibt sich aus dem Zustand  $\mathbf{x}_t^{veh} = (x, y, v, a, \varphi, \dot{\varphi}, w, l)^T$  des Fahrzeuges wie folgt:

$$h_{Video}(\mathbf{x}_t^{veh}) = \mathbf{P}_{Video} \overbrace{\begin{bmatrix} \cos(\varphi) & -\sin(\varphi) & 0 & x \\ \sin(\varphi) & \cos(\varphi) & 0 & y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}^{\mathbf{C}_{veh}} \mathbf{b} \quad (4.52)$$

Der Punkt  $\mathbf{b}$  bezeichnet hierbei den Mittelpunkt der unteren sichtbaren horizontalen Kante des Fahrzeuges in dem Koordinatensystem  $\mathbf{C}_{veh}$ . Sind zwei Flächen sichtbar, so wird diejenige Kante gewählt, die besser einer horizontalen Kante im Bild entspricht. Für diesen Sichtbarkeitsentscheid können Standardverfahren aus der Computergrafik verwendet werden, wie sie zum Beispiel in [BGZ02] beschrieben sind.

#### Anpassungen für den Partikel-Filter

Für den Partikel-Filter wird das Konturmodell verwendet, um eine Hypothese (Partikel) im Bild zu verifizieren. Hierzu wird das durchschnittliche Abstandsmaß aller Suchpfade verwendet. Die Gewichtung  $w^{(i)}$  des zugehörigen Partikels  $x^{(i)}$  einer Objekthypothese erfolgt schließlich mittels einer Normalverteilung:

$$w^{(i)} = \frac{1}{\sqrt{(2\pi)\sigma_{Video}}} e^{-\frac{1}{2\sigma_{Video}^2} \left( \frac{1}{n} \sum_{j=1}^n \Psi^{(i)}(j)^2 \right)} \quad (4.53)$$

#### Anpassungen für den Kalman-Filter

Um einen Messwert  $\mathbf{z}_{fir}$  für den Kalman-Filter zu erzeugen, wird iterativ in einem Suchfenster  $\mathcal{F}$ , das abhängig vom Zustand  $\mathbf{x}_t^{veh}$  definiert ist, mit der Gewichtsfunktion  $w^{(K_i)}$  aus 4.53 das Zutreffen des Konturmodells  $K_i$  überprüft. Die Kontur  $K_i \in \mathcal{F}$ , die das größte Gewicht zugewiesen bekommen hat, wird als Messung verwendet. Um Fehlmessungen zu vermeiden muss das Gewicht zusätzlich über einem Schwellwert  $\Theta_{Video}$  liegen.

$$\mathbf{z}_{video} = \begin{cases} \operatorname{argmax}_{K_i} w^{(K_i)} & , \text{ falls } \exists K_i : w^{(K_i)} \geq \Theta_{Video} \quad \text{wobei } K_i \in \mathcal{F} \\ \text{nicht definiert} & , \text{ sonst} \end{cases} \quad (4.54)$$

## 4.5 Zusammenfassung

Für eine Fahrumfelderfassung ist ein Laserscanner im Augenblick einer der vielseitigsten und genauesten Sensoren. Aus diesem Grund wurde das Messmodell für diesen Sensor sehr ausführlich behandelt. In einem ersten Vorverarbeitungsschritt werden die Rohdaten des Laserscanners in zusammengehörige Bereiche unterteilt. Hierzu wurde eine Erweiterung des Region-Growing Verfahrens vorgestellt, die im Rahmen einer Fahrumfelderfassung eine zuverlässigere und vollständigere Segmentierung von Fahrzeugen und Fußgängern gewährleistet. Basierend auf diesen Segmenten werden Merkmale extrahiert, die die zu erwartenden Messwerte der Beobachtungsmodelle beschreiben. Der wesentliche Bestandteil dieser Merkmalsextraktion wurde durch ein Verfahren erweitert, das sich als sehr robust gegenüber Messfehlern erwies. Abschließend wurden die eigentlichen Messmodelle für Fahrzeuge und Fußgänger eingeführt.

Seit 1998 haben sich Radarsensoren im automotiven Einsatz etabliert und kommen vornehmlich in Fahrerassistenzsystemen wie das ACC zum Einsatz. In diesem Kapitel wurde ein Messmodell vorgestellt, das im Vergleich zu einem einfachen Punktziel eine höhere laterale Genauigkeiten ermöglicht. Das eingeführte Modell der multiplen Reflexionsquellen ist allerdings auf einen zweiten Sensor (Laserscanner) angewiesen, um Mehrdeutigkeiten zu vermeiden bzw. aufzulösen. Fußgänger werden hingegen als einfache Punktziele modelliert. Hierbei ist allerdings zu beachten, dass Fußgänger mit der in dieser Arbeit verwendeten Radarsensorik nur sehr eingeschränkt erfasst werden können, da insbesondere die Fernbereichsradare auf Fahrzeuge und hohe Reichweiten optimiert sind und zusätzlich einen sehr eingeschränkten Öffnungswinkel besitzen.

Im Vergleich zu den entfernungsgebenden Sensoren, wie Laserscanner oder Radar, besitzen Kameras eine sehr hohe Winkelauflösung. Dadurch bieten sie die Möglichkeit, die laterale Ablage von Verkehrsteilnehmer sehr präzise zu messen. Für Fußgänger wurde ein Beobachtungsmodell vorgestellt, das das charakteristische Erscheinungsbild von Fußgängern in einer Ferninfrarotkamera berücksichtigt. Hierbei hat sich die Kontur des Fußgängerkopfes als sehr gutes Merkmal erwiesen, da dieser im Allgemeinen nicht mit Kleidung bedeckt und seine Kontur weitgehend invariant gegenüber dem Blickwinkel auf den Fußgänger ist. Ab einer Umgebungstemperatur  $\geq 25^\circ$  nimmt jedoch der Kontrast zwischen Fußgänger und Hintergrund deutlich ab und die Position des Kopfes kann nur noch sehr bedingt beobachtet werden. Fahrzeuge besitzen sehr viele vertikale oder horizontale Strukturen und heben sich durch ihre klaren Flächen und Farben deutlich von Hintergrund ab. Aus diesem Grund wurde ein Konturmodell beschrieben, das den Umriss eines Fahrzeuges durch einen Quader approximiert.

Alle beschriebenen Messmodelle - insbesondere die videobasierten - sind sehr einfach gehalten, da im Rahmen einer Fahrumfelderfassung und der Verwendung von Partikel-Filtern die Echtzeitfähigkeit von besonderer Bedeutung ist. Für alle Modelle wurde beschrieben, wie sie sich für einen Partikel-Filter verwenden lassen.

# TRACKING

---

Der Begriff Tracking<sup>1</sup> beschreibt den Prozess, mit dem Objekte über mehrere Zeitschritte hinweg verfolgt werden. Dabei wird versucht, aus Messungen bzw. Beobachtungen des Objektes Zustandsgrößen wie zum Beispiel Position, Geschwindigkeit, Breite oder Orientierung abzuleiten bzw. zu schätzen. Ein weiterer wichtiger Aspekt des Trackings ist die Einbeziehung des zeitlichen Verlaufes bzw. der letzten Beobachtungen des Objektes. Dementsprechend wird ein Objekt im Kontext seiner Historie betrachtet. Auf diese Weise können zum Einen Störeinflüsse wie zum Beispiel Messfehler kompensiert und zum Anderen Zustandsgrößen wie die Beschleunigung abgeleitet werden, die sich einer direkten Beobachtung entziehen. In der Literatur wird der Prozess des Trackings auch als Filterung bezeichnet, da die Auswirkungen von Messfehlern abgeschwächt werden.

Für Fahrerassistenzsysteme wie die aktive Gefahrenbremsung oder den Spurwechsellassistenten ist ein präzises und zuverlässiges Tracking der Positionen, Geschwindigkeiten, Beschleunigungen und Raummaße von anderen Verkehrsteilnehmern wie Fußgängern und Fahrzeugen von entscheidender Bedeutung. Aus diesem Grund werden in diesem Kapitel unterschiedliche Trackingverfahren wie Kalman-Filter und Partikel-Filter diskutiert und Möglichkeiten aufgezeigt, wie die Messungen mehrerer Sensoren in die dargestellten Verfahren integriert werden können. Dabei werden zuerst die theoretischen Grundlagen der Objektverfolgung behandelt. Anschließend werden die einzelnen Verfahren unter automotiven Aspekten und in Hinblick auf eine Fahrumfelderfassung evaluiert.

## 5.1 Grundlagen der Objektverfolgung

Für den Menschen ist sein umfangreiches Vorwissen entscheidend, um Objekte visuell zu verfolgen. Die Kenntnisse über die Szene und die zu verfolgenden Objekte ermöglichen es ihm, genaue Angaben zum Aufenthaltsort eines Objektes, sowie der zu erwartenden Bewegung zu treffen. Somit können selbst bei Verdeckungen noch Aussagen oder Annahmen getroffen werden: Ein Ball, der plötzlich hinter einem parkenden Fahrzeug auf die Straße rollt, impliziert zum Beispiel, dass sehr wahrscheinlich ein Kind dem Ball folgen wird. Für eine maschinelle Erfassung ist eine derartig umfangreiche Wissensbasis zur Zeit praktisch nicht realisierbar, da eine komplette Modellierung noch zu aufwendig ist. Vielmehr wird versucht, durch einfache Modelle die Realität anzunähern.

---

<sup>1</sup> engl. für das Verfolgen bzw. das Nachführen

Das Vorwissen des Computers wird dabei im Allgemeinen in folgende zwei Modelle unterteilt:

- **Objektmodell (a priori Wissen):**  
Charakteristische Merkmale, welche das zu verfolgende Objekt möglichst eindeutig festlegen. Dies können zum Beispiel 3D-Modelle, markante Konturen, die Objektfarbe oder Farb-Histogramme sein. Häufig sind in dem Objektmodell auch weitere Zustandsinformationen über Geschwindigkeit und die Bewegungsrichtung enthalten. In einem zusätzlichen Bewegungsmodell wird definiert, in welcher Art und Weise sich der Objektzustand über die Zeit verändert.
- **Beobachtungsmodell (likelihood):**  
Beschreibt, wie die Sensorinformationen zu behandeln und auszuwerten sind. So werden zum Beispiel Messungenauigkeiten berücksichtigt und die Rohdaten auf eine Form gebracht, in welcher sie mit dem aktuellen Objektzustand verglichen werden können (z.B. Kantenextraktion).

Zur Lösung des Problems der Objektverfolgung werden im Rahmen dieser Arbeit nur Prädiktionsfilter betrachtet. Bei einem Prädiktionsfilter läuft das Eingangssignal durch zwei parallele Zweige. Der erste Zweig liefert aus den vergangenen Eingangswerten einen Schätzwert (Prädiktionsschritt) für den aktuellen Eingangswert. Das Eingangssignal auf dem zweiten Zweig wird mit der Prädiktion verglichen. Schließlich lässt sich mit der Differenz der beiden Signale das Eingangssignal korrigieren (Korrekturschritt).

### 5.1.1 Mathematische Formulierung

Im Folgenden wird für die Objektverfolgung eine mathematische Formulierung gegeben. Diese fungiert als gemeinsame theoretische Grundlage für die Verfahren (Kalman-Filter und Partikel-Filter), die im weiteren Verlauf dieser Arbeit näher betrachtet werden (siehe auch Abschnitt 5.2 sowie 5.3). Die Formulierung ist im Wesentlichen aus [RA04] und [DGA00] entnommen.

Basierend auf einem diskreten Zeitbegriff wird der Objektzustand zur Zeit  $t$  als  $x_t \in \mathbb{R}^n$  bezeichnet. Das Ziel der Objektverfolgung ist es, den Objektzustand  $x_t$  aus Messungen bzw. Beobachtungen (vergleiche Kapitel 4)  $z_{1:t} = \{z_1, \dots, z_t\}$  mit  $z_t \in \mathbb{R}^m$  zu bestimmen. Der zeitliche Verlauf der Zustände bis zur Zeit  $t$  wird mit  $x_{0:t} = \{x_0, \dots, x_t\}$  bezeichnet und als Markov-Kette mit initialer Verteilung  $p(x_0)$  modelliert. Zudem wird die Vereinfachung getroffen, dass ein beliebiger Objektzustand nur vom vorhergehenden Zustand abhängt. Die Objektdynamik stellt also eine zeitliche Markov-Kette erster Ordnung dar:

$$\forall t : p(x_t | x_{0:t-1}) = p(x_t | x_{t-1}) \quad (5.1)$$

Die Übergangswahrscheinlichkeit  $p(x_t | x_{t-1})$  ist durch das Bewegungsmodell (siehe Abschnitt 4.1) gegeben. Dieses entspricht der Anwendung einer nicht notwendigerweise linearen oder zeit-invarianten Funktion  $f$  auf den Objektzustand  $x$ :

$$x_t = f_t(x_{t-1}, u_t, w_t) \quad (5.2)$$

Fehlerhafte Modellierungen des Bewegungsmodells oder unvorhergesehene Störungen werden durch das Prozessrauschen  $w_t$  ausgeglichen bzw. berücksichtigt. Mittels  $u_t$  wird eine optionale zusätzliche Systemeingabe modelliert, wie sie zum Beispiel bei Bewegungen des Beobachters auftritt.

Das Ziel der Objektverfolgung ist es, den Objektzustand  $x_t$  rekursiv aus Messungen  $z_{1:t} = \{z_1, \dots, z_t\}$  mit  $z_t \in \mathbb{R}^{d_z}$  zu schätzen. Diese Messungen werden aus einem geschätzten Zustand mittels der Messfunktion  $h$  gewonnen (siehe Kapitel 4):

$$z_t = h_t(x_t, v_t) \quad (5.3)$$

Die Unsicherheit in der Messung wird durch  $v_t$  modelliert. Analog zu  $f_t$  muss die Funktion  $h_t$  nicht notwendigerweise linear in  $x_t$  oder  $v_t$  sein. Im weiteren Verlauf werden die Rauschgrößen  $w_t$  und  $v_t$  als voneinander unabhängig angenommen, sowie als weißes Rauschen modelliert. Sie stellen also einen diskreten stochastischen Prozess von unkorrelierten Zufallsvariablen mit Erwartungswert Null und konstanter Varianz dar. Der Initialzustand  $x_0$  wird als rauschfrei angenommen. Um den Objektzustand  $x_t$  basierend auf den Messungen  $z_{1:t}$  zu schätzen, muss folgende a posteriori Verteilung berechnet werden:

$$p(x_{0:t} | z_{1:t}) \quad (5.4)$$

Da im vorliegenden Fall in jedem Zeitschritt nur die a posteriori Schätzung des aktuellen Zeitschrittes von Interesse ist, wird nicht die gesamte Zustandshistorie betrachtet, sondern nur der aktuelle Zeitpunkt. Damit vereinfacht sich die a posteriori Verteilung auf nachfolgende Randverteilung:

$$p(x_t | z_{1:t}) \quad (5.5)$$

Ausgehend von einer initialen Verteilung  $p(x_0) \hat{=} p(x_0 | z_0)$ , mit  $z_0 = \emptyset$ , erfolgt die Berechnung dieser Verteilung durch rekursive Anwendung der zwei bereits erwähnten Schritte Prädiktion und Korrektur. Angenommen, die benötigte Verteilung  $p(x_{t-1} | z_{1:t-1})$  zum Zeitpunkt  $t-1$  ist bekannt, dann ist es möglich, die Prädiktion durch Anwendung des Bewegungsmodells mittels der Chapman-Kolmogorow-Gleichung zu berechnen:

$$\begin{aligned} p(x_t | z_{1:t-1}) &= \int p(x_t, x_{t-1} | z_{1:t-1}) dx_{t-1} \\ &= \int p(x_t | x_{t-1}, z_{1:t-1}) \cdot p(x_{t-1} | z_{1:t-1}) dx_{t-1} \\ &= \int p(x_t | x_{t-1}) \cdot p(x_{t-1} | z_{1:t-1}) dx_{t-1} \end{aligned} \quad (5.6)$$

Hierbei wurde ausgenutzt, dass  $p(x_t | x_{t-1}, z_{1:t-1}) = p(x_t | x_{t-1})$  gilt, da die Gleichung (5.2) eine Markov-Kette erster Ordnung darstellt. Die Wahrscheinlichkeitsverteilung  $p(x_t | x_{t-1})$  des Bewegungsmodells ist ebenfalls durch Gleichung (5.2) definiert.

Erhält man zum Zeitpunkt  $t$  eine Messung  $z_t$ , so lässt sich der Korrektur-Schritt mittels der Prädiktion durch Anwendung der Regel von Bayes berechnen:

$$\begin{aligned} p(x_t | z_{1:t}) &= p(x_t | z_t, z_{1:t-1}) \\ &= \frac{p(z_t | x_t, z_{1:t-1}) \cdot p(x_t | z_{1:t-1})}{p(z_t | z_{1:t-1})} \\ &= \frac{p(z_t | x_t) \cdot p(x_t | z_{1:t-1})}{p(z_t | z_{1:t-1})} \\ &= k_t \cdot p(z_t | x_t) \cdot p(x_t | z_{1:t-1}) \end{aligned} \quad (5.7)$$

Der Faktor  $k_t$  stellt hierbei eine Normalisierungskonstante dar:

$$\frac{1}{k_t} = p(z_t | z_{1:t-1}) = \int p(z_t | x_t) \cdot p(x_t | z_{1:t-1}) dx_t \quad (5.8)$$



Die Wahrscheinlichkeitsverteilung  $p(z_t|x_t)$  des Beobachtungsmodells (die *Likelihood-Funktion*) wird durch Gleichung (5.3) definiert. Zu beachten ist, dass für die Herleitung der Gleichung (5.7) die Vereinfachung  $p(z_t|x_t, z_{1:t-1}) = p(z_t|x_t)$  verwendet wurde. Für die Beobachtungen  $z_{1:t} \triangleq \{z_1, \dots, z_t\}$  wird also angenommen, dass diese sowohl untereinander stochastisch unabhängig sind (siehe Gleichung (5.9)), als auch im Hinblick auf das Bewegungsmodell (siehe Gleichung (5.10)).

$$p(z_{1:t}|x_{1:t}) = \prod_{i=1}^t p(z_i|x_i) \quad (5.9)$$

$$p(x_t, z_{1:t-1} | x_{1:t-1}) = p(x_t|x_{t-1}) \cdot \prod_{i=1}^{t-1} p(z_i|x_i). \quad (5.10)$$

Die Herleitung der zwei Rekursionsgleichungen (5.6) und (5.7) findet sich unter anderem in [DGA00] und [IB98].

Mit diesen beiden Rekursionsgleichungen lässt sich ein Filter konstruieren, welcher das Schätzproblem optimal löst. Allerdings ist dies nur im theoretischen Sinne möglich, da eine analytische Lösung im allgemeinen Fall nicht bestimmt werden kann. Nur in stark eingeschränkten Problemklassen, welche in den nächsten Abschnitten kurz erläutert werden, kann die a posteriori Verteilung exakt und komplett durch eine endliche Menge von Parametern bestimmt werden.

### 5.1.2 Optimale Algorithmen

In den folgenden Spezialfällen ist es möglich, optimale Algorithmen mit endlichem Rechenaufwand zur rekursiven Zustandsschätzung anzugeben:

1. Für lineare Zustandstransformationen, lineare Messfunktionen, sowie initiale, unimodale und gauß-verteilte Wahrscheinlichkeitsverteilungen erhält man aus den Gleichungen (5.6) und (5.7) den häufig verwendeten Kalman-Filter [Kal60].
2. Wenn der Zustandsraum diskret ist, und nur endlich viele Zustände möglich sind, so liefern gitterbasierte Verfahren eine optimale Lösung [RA04].
3. Für sehr spezielle nichtlineare Bewegungsgleichungen und lineare Beobachtungsgleichungen ist es möglich, exakte Lösungen anzugeben [RA04, Ben81, Dau86].

Als ein Vertreter dieser optimalen Algorithmen wird im Folgenden die Klasse der Kalman-Filter näher betrachtet.

## 5.2 Klasse der Kalman-Filter

Der Kalman-Filter und seine Varianten sind sehr verbreitete Verfahren zur Zustandsschätzung linearer und nichtlinearer Systeme. Mit ihnen kann das Problem der Objektverfolgung nicht nur sehr effizient, sondern unter bestimmten Voraussetzungen auch optimal gelöst werden. Eine gute Einführung in den Kalman-Filter und seine Erweiterungen ist in [WB01, GA93, JU04] zu finden.

Insbesondere in einer Fahrumfelderfassung wird die Dynamik der zu verfolgenden Objekte häufig durch nichtlineare Systeme beschrieben (siehe Abschnitt 4.1). Deshalb werden im Folgenden Erweiterungen des Kalman-Filters beschrieben, die eine Verwendung nichtlinearer Objektmodelle gestatten.



Abschließend werden Möglichkeiten aufgezeigt, wie die Messungen mehrerer Sensoren in die beschriebenen Verfahren integriert werden können. Abschließend werden der Extended-Kalman-Filter und der Unscented-Kalman-Filter gegenüber gestellt. Für einen konkreten Vergleich zwischen dem Kalman-Filter und dem Partikel-Filter sei auf das Ende des Abschnittes über die Partikel-Filter verwiesen (siehe 5.3.4).

### 5.2.1 Standard Kalman-Filter

Der Kalman-Filter (KF) setzt voraus, dass die a posteriori Verteilung zu jedem Zeitpunkt durch eine gaußsche Normalverteilung beschrieben werden kann und somit durch den Mittelwert und die Kovarianz vollständig charakterisiert wird. Wenn die Verteilung  $p(x_{t-1}|z_{1:t-1})$  eine Normalverteilung ist, so ist die darauffolgende a posteriori Verteilung  $p(x_t|z_{1:t})$  im nächsten Zeitschritt ebenfalls normalverteilt, vorausgesetzt folgende Bedingungen sind erfüllt:

- das dynamische Modell  $f_t(x_{t-1}, u_t, w_t)$  aus Gleichung (5.2) ist linear in  $x_{t-1}$ ,  $w_t$  und  $u_t$
- die Rauschgrößen  $w_t$  und  $v_t$  sind komponentenweise normalverteilt mit bekanntem Mittelwert und Kovarianz
- die Messfunktion  $h_t(x_t, v_t)$  aus Gleichung (5.3) ist linear in  $x_t$  und  $v_t$

Unter diesen Voraussetzungen können die Gleichung (5.2) und (5.3) wie folgt geschrieben werden:

$$x_t = \mathbf{F}_t x_{t-1} + \mathbf{B}_t u_t + w_t \quad (5.11)$$

$$z_t = \mathbf{H}_t x_t + v_t \quad (5.12)$$

Die beiden Matrizen  $\mathbf{F}_t$  und  $\mathbf{H}_t$  beschreiben das lineare dynamische Modell  $f_t$  bzw. die lineare Messfunktion  $h_t$ . Durch  $u_t$  werden zusätzliche Systemeingaben (z.B. Eigenbewegungen des Beobachters) modelliert und durch die Matrix  $\mathbf{B}_t$  auf den Zustandsvektor abgebildet. Die Rauschgrößen  $w_t$  und  $v_t$  sind mittelwertsfreie, multivariate normalverteilte Zufallsvariablen der Kovarianzen  $\mathbf{Q}_t$  und  $\mathbf{R}_t$ .

Für die Zustandsschätzungen werden folgende Notationen verwendet:

- $\hat{x}_t$  bezeichnet den Schätzwert des Zustandes  $x_t$  zum Zeitpunkt  $t$
- $\hat{x}_t^-$  bezeichnet den prädierten (vorhergesagten) Schätzwert zum Zeitpunkt  $t$

Damit ergeben sich für den Kalman-Filter folgende Gleichungen für die Prädiktion und Korrektur:

$$\hat{x}_t^- = \mathbf{F}_t \hat{x}_{t-1} + \mathbf{B}_t u_t \quad (5.13)$$

$$\hat{x}_t = \hat{x}_t^- + \mathbf{K}_t \cdot (z_t - \mathbf{H}_t \hat{x}_t^-) \quad (5.14)$$

Der sogenannte *Kalman-Faktor* bzw. *Kalman-Gain*  $\mathbf{K}_t$  berechnet sich wie folgt:

$$\mathbf{K}_t = \mathbf{P}_t^- \mathbf{H}_t^T [\mathbf{H}_t \mathbf{P}_t^- \mathbf{H}_t^T + \mathbf{R}_t]^{-1} \quad (5.15)$$

Der Kalman-Faktor übernimmt hierbei im Wesentlichen die Aufgabe die Messung in Abhängigkeit der aktuellen Fehlerkovarianz der Zustandsschätzung bzw. der Messunsicherheit zu gewichten und

auf die einzelnen Zustandsgrößen abzubilden. Schließlich lassen sich die zugehörigen prädizierten und korrigierten Fehlerkovarianzmatrizen der Zustandsschätzung bestimmen:

$$\mathbf{P}_t^- = \mathbf{F}_t \mathbf{P}_{t-1} \mathbf{F}_t^T + \mathbf{Q}_t \quad (5.16)$$

$$\mathbf{P}_t = \mathbf{P}_t^- [\mathbf{I} - \mathbf{K}_t \mathbf{H}_t] \quad (5.17)$$

Sind die oben genannten (stark restriktiven) Bedingungen erfüllt, so liefert der Kalman-Filter die optimale Lösung für das Tracking Problem aus Abschnitt 5.1.1. Die Fehlerkovarianzmatrix  $\mathbf{P}_t$  spiegelt die Unsicherheit der Schätzung wider. Liegen zu einem Zeitpunkt keine Messungen des Zustandes vor, so wird nur der Prädiktionsschritt ausgeführt und die Zustandsunsicherheit erhöht sich um das Prozessrauschen. Treffen neue Messungen ein, so sinkt die Unsicherheit. In Abbildung 5.1 ist die Funktionsweise des Kalman-Filters zusammengefasst dargestellt.

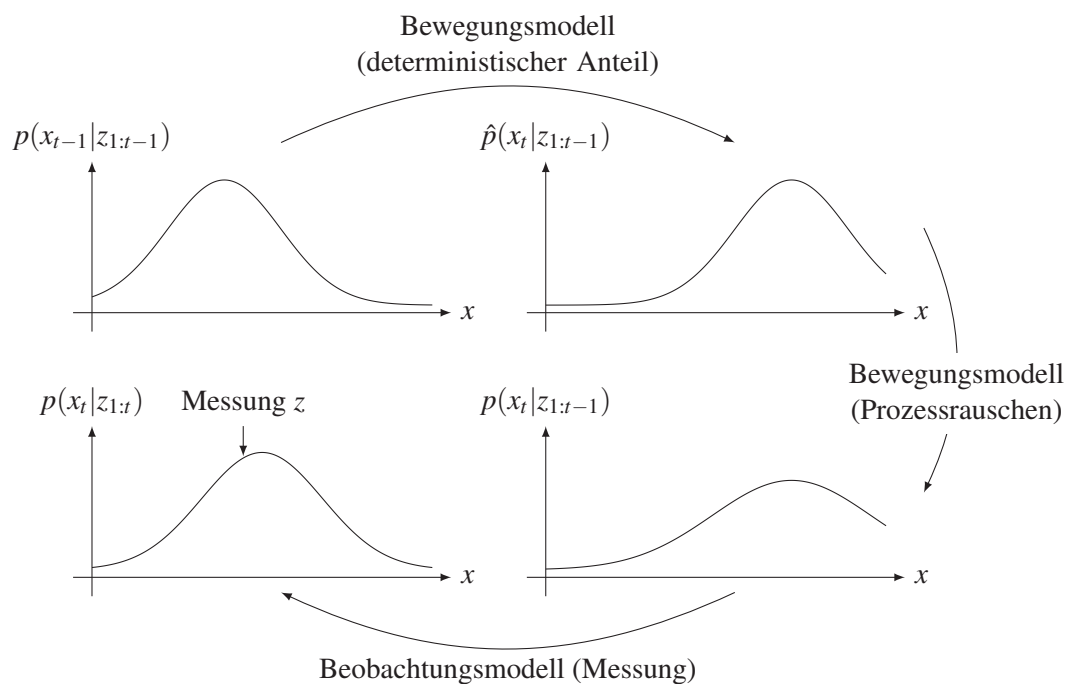


Abbildung 5.1: Ein Iterationszyklus des Kalman-Filters: Dargestellt ist die Wahrscheinlichkeitsverteilung eines zu schätzenden, eindimensionalen Zustandes. Zuerst wird das Bewegungsmodell angewendet und somit eine Vorhersage des Zustandes im nächsten Zeitschritt getroffen. Um die Unsicherheit der Vorhersage zu modellieren, wird diese anschließend mit einem Rauschen überlagert. Anhand der so vorhergesagten Objektposition werden Messungen vorgenommen, mit welchen die Vorhersage korrigiert wird.

### 5.2.2 Extended-Kalman-Filter

Um die starken Einschränkungen des Kalman-Filters aufzulockern, wurden Erweiterungen entwickelt, die es gestatten, das Grundprinzip des Kalman-Filters auch bei nichtlinearen Problemen anzuwenden. Ein weitverbreiteter Vertreter ist der Extended-Kalman-Filter (EKF) [WB01, GA93]. Sein Grundprinzip besteht darin, das nichtlineare Problem in einer lokalen Umgebung des zu schätzenden Zustandes durch eine Linearisierung zu approximieren. Dies erfolgt nach dem Prinzip der Taylorreihenapproximation, welche nach der ersten Ableitung abgeschnitten wird. Die Bewegungs- und Beobachtungsmodelle  $x_t$  und  $z_t$  mit den zugehörigen nichtlinearen Funktionen  $f$  bzw.  $h$  werden wie folgt approximiert:

$$x_t \approx f(\hat{x}_{t-1}, u_t, 0) + \mathbf{A} \cdot (x_{t-1} - \hat{x}_{t-1}) + \mathbf{W} \cdot w_t \quad (5.18)$$

$$z_t \approx h(\hat{x}_t^-) + \mathbf{H} \cdot (x_t - \hat{x}_t^-) + \mathbf{V} \cdot v_t \quad (5.19)$$

Die für die Taylorreihe benötigten Jacobi-Matrizen berechnen sich wie folgt:

$$\begin{aligned} \mathbf{A}_{[ij]} &= \frac{\partial f_i}{\partial x_j}(\hat{x}_{t-1}, u_t, 0) \quad , \quad \mathbf{W}_{[ij]} = \frac{\partial f_i}{\partial w_j}(\hat{x}_{t-1}, u_t, 0) \\ \mathbf{H}_{[ij]} &= \frac{\partial h_i}{\partial x_j}(\hat{x}_t^-, 0) \quad , \quad \mathbf{V}_{[ij]} = \frac{\partial h_i}{\partial v_j}(\hat{x}_t^-, 0) \end{aligned} \quad (5.20)$$

Damit ergeben sich für den Extended-Kalman-Filter folgende Gleichungen für die Prädiktion und Korrektur:

$$\hat{x}_t^- = f(\hat{x}_{t-1}, u_t, 0) \quad (5.21)$$

$$\hat{x}_t = \hat{x}_t^- + \mathbf{K}_t \cdot (z_t - h(\hat{x}_t^-, 0)) \quad (5.22)$$

Der *Kalman-Gain* für den Extended-Kalman-Filter lässt sich wie folgt berechnen:

$$\mathbf{K}_t = \mathbf{P}_t^- \mathbf{H}^T [\mathbf{H} \mathbf{P}_t^- \mathbf{H}^T + \mathbf{V} \mathbf{R}_t \mathbf{V}^T]^{-1} \quad (5.23)$$

Schließlich lassen sich die zugehörigen prädierten und korrigierten Fehlerkovarianzmatrizen der Zustandsschätzung bestimmen:

$$\mathbf{P}_t^- = \mathbf{A} \mathbf{P}_{t-1} \mathbf{A}^T + \mathbf{W} \mathbf{Q} \mathbf{W}^T \quad (5.24)$$

$$\mathbf{P}_t = \mathbf{P}_t^- [\mathbf{I} - \mathbf{K}_t \mathbf{H}] \quad (5.25)$$

Zu beachten ist, dass es sich bei dem Extended-Kalman-Filter nicht mehr um einen optimalen Schätzer handelt. Die lineare Approximation ist abhängig von der Problemstellung und nur in einer kleinen Umgebung des tatsächlichen Systemzustandes zulässig. Bei stark nichtlinearen Zustandsänderungen divergiert der Extended-Kalman-Filter. In diesem Zusammenhang ist unter divergieren zu verstehen, dass sich das Ergebnis der Objektverfolgung immer weiter von der Realität entfernt. Im Allgemeinen führt dies zu einem Abbruch der Verfolgung (Objektverlust), da sich ab einem gewissen Zeitpunkt die Messwerte zu stark von der Prädiktion unterscheiden. Eine weitere notwendige Forderung ist die Differenzierbarkeit der Funktionen  $f$  und  $h$ .

### 5.2.3 Unscented-Kalman-Filter

Bei stark nichtlinearen Systemen ist die Approximation des Extended-Kalman-Filters nicht mehr zulässig. So können zum Beispiel systematische Fehler und nicht mehr normalverteilte Wahrscheinlichkeitsverteilungen entstehen. Beides kann die Qualität der Schätzung negativ beeinflussen [JU04]. Aus diesem Grund verwendet der Unscented-Kalman-Filter (UKF) [JUDW95, JUDW00] die Unscented Transformation als Modell der Fehlerfortpflanzung. Es kann gezeigt werden, dass dieses Modell den Mittelwert und die Kovarianz bis zur dritten Stufe der Taylorreihen-Approximation einer beliebigen nichtlinearen Funktion korrekt abbildet [WCD<sup>+</sup>06]. Das Konzept der Transformation ist der Monte-Carlo Simulation sehr ähnlich: Es werden Beispielpunkte bestimmt, die mit der nichtlinearen Funktion  $f$  transformiert werden. Im Gegensatz zur Monte-Carlo Simulation werden diese Punkte nicht zufällig gezogen, sondern so gewählt, dass die Kovarianzmatrix und der Mittelwert der Punktmenge genau der originalen Verteilung entsprechen. In der Literatur werden diese Punkte als Sigma-Punkte bezeichnet. Schließlich werden die Kovarianz und der Mittelwert der transformierten Punktmenge bestimmt.

Bei dieser Variante bleibt die grundlegende Struktur des Kalman-Filters (Prädiktion und Korrektur) erhalten. Für die Wahl der Sigma-Punkte wird der ursprüngliche Zustandsvektor  $\hat{x}$  wie folgt um die Rauschgrößen  $w_t$  und  $v_t$  erweitert:

$$\hat{x}_{t-1}^a = [\hat{x}_{t-1}^T \quad w_{t-1}^T \quad v_{t-1}^T]^T \quad (5.26)$$

Die Wahl der Sigma-Punkte  $\chi_{t-1}$  und der zugehörigen Gewichte  $W_i$  aus dem erweiterten Zustand  $\hat{x}_{t-1}^a$  ist in [JU04] beschrieben und wird als gegeben angenommen. Im weiteren Verlauf entspricht  $\chi^x$  den Punkten ohne die Erweiterung aus 5.26.

Prädiktion:

$$\chi_t^x = f(\chi_{t-1}^x, u_t) \quad (5.27)$$

$$\hat{x}_t^- = \sum_{i=0}^{2L} W_i \chi_{i,t}^x \quad (5.28)$$

$$\mathbf{P}_t^- = \sum_{i=0}^{2L} W_i [\chi_{i,t}^x - \hat{x}_t^-] [\chi_{i,t}^x - \hat{x}_t^-]^T \quad (5.29)$$

$$\Upsilon_t = h(\chi_t^x, u_t) \quad (5.30)$$

$$\hat{z}_t^- = \sum_{i=0}^{2L} W_i \Upsilon_{i,t} \quad (5.31)$$

Korrekturschritt:

$$\mathbf{S} = \sum_{i=0}^{2L} W_i^{(c)} [\Upsilon_{i,t} - \hat{y}_t^-] [\Upsilon_{i,t} - \hat{y}_t^-]^T \quad (5.32)$$

$$\mathbf{P}_{xy} = \sum_{i=0}^{2L} W_i^{(c)} [\chi_{i,t}^x - \hat{x}_t^-] [\Upsilon_{i,t} - \hat{y}_t^-]^T \quad (5.33)$$

$$\mathbf{K}_t = \mathbf{P}_{xy} \mathbf{S}^{-1} \quad (5.34)$$

$$\hat{x}_t = \hat{x}_t^- + \mathbf{K}_t (z_t - \hat{z}_t^-) \quad (5.35)$$

$$\mathbf{P}_t = \mathbf{P}_t^- - \mathbf{H} \mathbf{P}_{xy} \mathbf{H}^T \quad (5.36)$$

Eine analytische und teilweise aufwendige Berechnung der Jacobimatrizen, wie bei dem Extended-Kalman-Filter, ist nicht notwendig. Die Wahl und Berechnung der Sigma-Punkte ist nicht trivial, aber unabhängig von den jeweiligen System- und Messmodellen und kann somit als „black box“ Verfahren betrachtet werden. In Bezug auf die Komplexität ist der Unscented-Kalman-Filter bis auf einen konstanten Faktor vergleichbar mit dem Extended-Kalman-Filter. Wie in [JUDW95] beschrieben kann der UKF den Zustand und die Fehler-Kovarianz bis zum vierten Grad korrekt vorhersagen. Dementsprechend ist er dem EKF überlegen, der in der Zustandsschätzung nur bis zum zweiten Grad korrekt approximiert. Allerdings liefert der UKF in der Realität nur dann bessere Ergebnisse, wenn die Wölbung und die höheren Momente der Zustandsfehlerverteilung signifikant sind [LJ03].

### 5.2.4 Fusion

Unabhängig von der gewählten Erweiterung existieren bei einem Kalman-Filter verschiedene Möglichkeiten die Messungen mehrerer Sensoren zu fusionieren. Im Allgemeinen wird zwischen zwei Ansätzen unterschieden (siehe auch Abbildung 5.2): Fusion der Zustände [BSC86, GH02] und Fusion der Messwerte [RM88, MDW94, GH01].

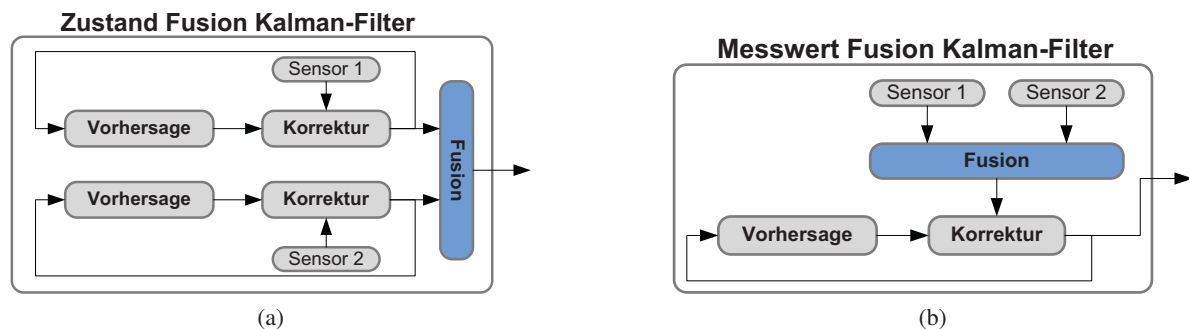


Abbildung 5.2: Fusionsmöglichkeiten beim Kalman-Filter. In Abbildung (a) wird das Ergebnis von mehreren unabhängigen Kalman-Filtern fusioniert. In (b) werden von einem einzigen Kalman-Filter die fusionierten Messwerte verarbeitet.

**Zustandsfusion** Bei der Fusion der Zustände wird für jeden Sensor ein unabhängiger Kalman-Filter verwendet. Nach dem Korrekturschritt werden die einzelnen Zustände zu einem Zustand fusioniert. Dieses Verfahren wird auch oft als Track-Level Fusion bezeichnet. Gao und Harris [GH02] haben diese Art der Fusion durch eine Rückkopplung erweitert. Hierbei wird das Ergebnis der Fusion in die einzelnen Kalman-Filter integriert. Für einen Vergleich der verschiedenen Varianten der Zustandsfusion sei auf [GH02] verwiesen.

**Messwertfusion** Die zweite Möglichkeit der Fusion ist die Kombination von Messwerten, die sich in mehrere Varianten untergliedert. So können zum Beispiel alle Sensoren sequentiell in den Kalman-Filter integriert werden, wobei die Korrektur für die Messung  $n$  auf den Ergebnissen des Schrittes  $n - 1$  operiert. Für nichtlineare Messwerte muss allerdings beachtet werden, dass der Verarbeitungsprozess mit der genauesten Messung beginnt.

Eine weitere Variante besteht darin, die einzelnen Messwerte  $\mathbf{z}_i$  gemäß ihrer jeweiligen Kovarianzen  $\mathbf{R}_i$  zu gewichten und zu einem neuen gemeinsamen Messwert  $\mathbf{z}$  zu kombinieren:

$$\mathbf{z} = \left( \sum_{i=1}^n \mathbf{R}_i^{-1} \right)^{-1} \sum_{i=1}^n \mathbf{R}_i^{-1} \mathbf{z}_i \quad (5.37)$$

Entsprechend werden die gemeinsamen Matrizen  $\mathbf{R}$  und  $\mathbf{H}$  gebildet:

$$\mathbf{H} = \left( \sum_{i=1}^n \mathbf{R}_i^{-1} \right)^{-1} \sum_{i=1}^n \mathbf{R}_i^{-1} \mathbf{H}_i \quad (5.38)$$

$$\mathbf{R} = \left( \sum_{i=1}^n \mathbf{R}_i^{-1} \right)^{-1} \quad (5.39)$$

Diese Art der Fusion ist sehr effizient durchführbar. Allerdings können hiermit nur gleichartige Messwerte wie zum Beispiel von zwei Radaren fusioniert werden.

Eine dritte Variante konkateniert alle Messungen  $\mathbf{z}_i$  zu einer großen virtuellen Messung  $\mathbf{z}$ :

$$\mathbf{z} = [\mathbf{z}_1, \dots, \mathbf{z}_n]^T \quad (5.40)$$

Analog dazu werden die Matrizen  $\mathbf{R}$  und  $\mathbf{H}$  gebildet:

$$\mathbf{H} = (\mathbf{H}_1, \dots, \mathbf{H}_n)^T \quad (5.41)$$

$$\mathbf{R} = \text{diag}(\mathbf{R}_1, \dots, \mathbf{R}_n) \quad (5.42)$$

Diese Variante stellt weder Bedingungen hinsichtlich der Art der Messwerte noch in Bezug auf die Reihenfolge der Abarbeitung. Sie kann somit als sehr allgemeine Art der Messwertfusion betrachtet werden. Nach [GH01] ist sie auch in Bezug auf die zu erwartenden Ergebnisse zu bevorzugen oder zumindest äquivalent. Allerdings wächst mit der Zahl der Sensoren auch die Dimension der Matrizen  $\mathbf{R}$  und  $\mathbf{H}$ , was zu einem höheren Rechenaufwand im Vergleich zu den anderen Verfahren führt.

Im weiteren Verlauf dieser Arbeit wird die Konkatenation von Messwerten verwendet, da zum Einen die Messwertfusion nach [RM88] der Zustandsfusion vorzuziehen ist und zum Anderen sehr unterschiedliche Sensoren zur Fahrumfelderfassung verwendet werden sollen.

### 5.2.5 Kompensation von Latenzzeiten

Für zukünftige Fahrerassistenzsysteme, wie zum Beispiel die aktive Gefahrenbremsung, ist eine unverfälschte Zustandsschätzung von sehr großer Bedeutung. Insbesondere bei hohen Relativgeschwindigkeiten zwischen dem eigenen Fahrzeug und anderen Objekten kann der Einfluss von Latenzzeiten in der Sensorik nicht vernachlässigt werden. So ist zum Beispiel in einem typischen Stadtszenario ( $v = 14 \frac{\text{m}}{\text{s}}$ ) eine Latenzzeit von 50 ms für einen Fehler von 70 cm in der Entfernungsmessung von Fußgängern verantwortlich. In dem Kapitel 3.3 wurde bereits beschrieben, wie sich die Latenzzeiten von Sensoren experimentell bestimmen lassen. Im Folgenden soll diskutiert werden, wie bestehende Latenzen in den Vorgang der Filterung integriert und somit kompensiert werden können. Ein guter Überblick über verschiedene Ansätze findet sich in [BLN<sup>+</sup>98, vdMWJ04].

Eine Möglichkeit, Latenzzeiten der Länge  $m$  zu kompensieren, besteht darin, alle Messungen und Zustandsgrößen aus den vergangenen  $n$  Arbeitsschritten zu speichern. Trifft die verspätete Messung ein, wird sie zeitlich korrekt einsortiert und alle Messungen werden in der zeitlich korrekten Reihenfolge abgearbeitet [BLN<sup>+</sup>98]. Für lineare Systeme ist diese Vorgehensweise die optimale Lösung. Allerdings sind der Rechenaufwand und der Bedarf an zusätzlichem Speicherplatz sehr hoch. Im Rahmen dieser Arbeit müssten zum Beispiel alle Bilder der letzten  $m$  Zeitschritte gespeichert werden. Aus diesem Grund wird oftmals nur der vorhergesagte Zustand  $\hat{x}_{t-m}^-$  zum Zeitpunkt  $t - m$  gespeichert. Für den Extended-Kalman-Filter ergibt sich dann folgender geänderter Korrekturschritt, wobei  $z_{t-m}$  der um  $m$  Zeitschritte verzögerten Messung entspricht:

$$\hat{x}_t = \hat{x}_t^- + \mathbf{K}_t \cdot (z_{t-m} - h_t(\hat{x}_{t-m}^-, 0)) \quad (5.43)$$

Diese Vorgehensweise ist nur eine Approximation und nicht die optimale Lösung. Dennoch hat sie sich im Rahmen einer Fahrumfelderfassung als ausreichend erwiesen, um den Einfluss von Latenzzeiten deutlich zu minimieren (siehe Tabelle 5.1). Um die Auswirkungen dieser Art der Kompensation zu untersuchen, wurden zwei unterschiedliche Szenarien evaluiert. Das erste Szenario beinhaltet nur wenige Abschnitte mit hohen Relativgeschwindigkeiten und sollte deshalb nur sehr bedingt von Latenzzeiten beeinflusst werden. In dem zweiten Versuchsaufbau nähert sich ein Fahrzeug mit  $70 \frac{km}{h}$  einem Fußgänger. In Tabelle 5.1 sind für beide Szenarien die Fehler der Zustandsschätzung mit und ohne Kompensation der Latenzzeit angegeben. Es ist gut zu erkennen, dass ohne zeitliche Kompensation ein deutlicher Fehler in der Entfernungsmessung entsteht. Durch den modifizierten Korrekturschritt aus 5.43 kann in beiden Szenarien eine deutliche Steigerung der Genauigkeit der Schätzung erreicht werden. Einzig die Schätzung der lateralen Ablage wird minimal verschlechtert.

	$RMSE_r$	$RMSE_\varphi$	$RMSE_v$		$RMSE_r$	$RMSE_\varphi$	$RMSE_v$
ohne Komp.	0.22 m	0.69 °	0.76 $\frac{m}{s}$	ohne Komp.	0.96 m	0.24 °	0.54 $\frac{m}{s}$
mit Komp.	0.16 m	0.75 °	0.60 $\frac{m}{s}$	mit Komp.	0.19 m	0.18 °	0.34 $\frac{m}{s}$

(a) Szenario 1 (b) Szenario 2

Tabelle 5.1: Für zwei Szenarien wurde der Fehler in der Zustandsschätzung mit und ohne Kompensation von Latenzzeiten untersucht. Im ersten Szenario sind nur wenige Abschnitte mit hoher Relativgeschwindigkeit vorhanden. Im zweiten Szenario nähert sich ein Fahrzeug mit  $70 \frac{km}{h}$  einem Fußgänger. In diesem Szenario ist der Einfluss von Latenzzeiten besonders ausgeprägt.

### 5.2.6 Vergleich EKF und UKF

Nachdem in den vorherigen Abschnitten die theoretischen Grundlagen der einzelnen Kalman-Filter Varianten beschrieben wurden, sollen im Folgenden der Extended- und der Unscented-Kalman-Filter verglichen werden. In dieser Gegenüberstellung werden sowohl die besonderen Anforderungen im automotiven Umfeld, als auch die konkreten Mess- und Objektmodelle aus dem vorherigen Kapitel berücksichtigt. Zunächst wird anhand von drei Testszenarien die Qualität der Schätzung einzelner Zustandsgrößen evaluiert. Anschließend werden die Laufzeit der beiden Verfahren analysiert und unterschiedliche Aspekte der Implementierung und der Praxistauglichkeit bewertet.



### Qualität der Schätzung

Der Unscented-Kalman-Filter wurde entwickelt, um die Zustands- und Fehlerkovarianz für stark nicht-lineare Systeme besser zu approximieren, als es mit dem Extended-Kalman-Filter möglich ist. Hierdurch kann eine genauere Schätzung der Zustandsgrößen erwartet werden. Im Folgenden soll am Beispiel der nichtlinearen Messfunktion des Laserscanners (siehe Kapitel 4.2.4) sowie des nichtlinearen Bewegungsmodells für Fahrzeuge (siehe Kapitel 4.1.1) der Einfluss des Unscented-Kalman-Filters auf die Genauigkeit der Zustandsschätzung untersucht werden.

Zur Bewertung der Qualität der Zustandsschätzung wurden drei repräsentative Testszenarien ausgewählt, für die hochgenaue Referenzdaten zur Verfügung standen. Die einzelnen Szenarien können hinsichtlich der enthaltenen Fahrdynamik untergliedert werden. So enthält das erste Szenario einige Passagen mit engen Kurven und starken Beschleunigungen und somit einen hohen Anteil an Längs- und Querdynamik. In der zweiten Testszene wird ausschließlich die Längsdynamik betrachtet. Es sind weder Kurven noch Spurwechsel enthalten. In dem dritten Szenario führt das vorausfahrende Fahrzeug acht sehr schnelle Spurwechselmanöver durch. Hiermit soll der Einfluss der Querdynamik evaluiert werden.

In Tabelle 5.2 sind die ermittelten Ergebnisse zusammengefasst. Die Tabellenspalten enthalten die untersuchten Szenen. Jeweils zwei Tabellenzeilen beinhalten die Werte einer Metrik für den Extended sowie den Unscented-Kalman-Filter. Als Qualitätsmaß der Evaluierung wurde der RMSE für die Schätzgröße der Entfernung  $r$ , der lateralen Ablage  $\varphi$  sowie der Geschwindigkeit  $v$  gewählt. Die restlichen Zustandsgrößen wie Länge, Breite oder die Orientierung werden in diesem Vergleich nicht berücksichtigt, da sowohl ihr zeitliches Verhalten als auch die zugehörige Messfunktion linear sind und somit kein Unterschied zu erwarten ist.

	Einheit	Szenario 1 (Längs- und Querdynamik)	Szenario 2 (Längsdynamik)	Szenario 3 (Querdynamik)
$RMSE_r$ EKF	[m]	0.22	0.16	0.10
$RMSE_r$ UKF	[m]	0.22	0.16	0.10
$RMSE_\varphi$ EKF	[°]	0.59	0.47	0.59
$RMSE_\varphi$ UKF	[°]	0.54	0.47	0.53
$RMSE_v$ EKF	$[\frac{m}{s}]$	0.73	0.55	0.29
$RMSE_v$ UKF	$[\frac{m}{s}]$	0.70	0.47	0.23

Tabelle 5.2: Für drei Szenarien und unterschiedliche Varianten des Kalman-Filters wurde die Qualität der Zustandsschätzung für die Entfernung  $r$ , die laterale Ablage  $\varphi$  sowie die Geschwindigkeit  $v$  eines vorausfahrenden Fahrzeuges untersucht. Jeweils zwei Tabellenzeilen beinhalten die Werte dieser Metriken für den Extended sowie den Unscented-Kalman-Filter. Die Szenarien unterscheiden sich hinsichtlich der enthaltenen Fahrdynamik des zu verfolgenden Fahrzeuges. Szenario 1 enthält Längs- und Querdynamik. Szenario 2 beinhaltet nur ein Beschleunigungsmanöver. Im dritten Szenario werden schnelle Spurwechsel vollzogen.

Aus dem Vergleich der beiden Filtervarianten können folgende Ergebnisse abgeleitet werden: Der Einfluss der Längsdynamik auf die Qualität der Entfernungsmessung lässt sich sehr gut an der ersten Zeile



der Tabelle 5.2 ablesen. So konnte für Szenario 3, das den geringsten Anteil an Beschleunigungs- und Bremsmanövern enthält, die Entfernung am Besten bestimmt werden. Allerdings ist kein signifikanter Unterschied zwischen dem EKF und dem UKF zu beobachten. Dies ist dadurch zu erklären, dass die Entfernung nur unwesentlich von dem nichtlinearen dynamischen Modell beeinflusst wird.

Mit dem UKF kann nur in den Szenen eine bessere Schätzung der lateralen Ablage  $\phi$  erreicht werden, die einen großen Anteil an Querdynamik beinhalten (vergleiche die zweite Zeile in Tabelle 5.2). Dies ist darin begründet, dass der nichtlineare Anteil in dem dynamischen Modell für Fahrzeuge im Wesentlichen von der Gierrate bzw. der Orientierung bestimmt ist und sich deutlich stärker auf die Prädiktion der lateralen Ablage als auf die Vorhersage der Entfernung auswirkt.

Mit einem Laserscanner kann die Geschwindigkeit  $v$  nicht direkt gemessen werden. Deshalb wird sie indirekt über das dynamische Modell und die Änderung der Position geschätzt. Somit kann sich der nichtlineare Anteil auch auf die Schätzung der Geschwindigkeit auswirken. Für alle drei Testszene wird durch die Verwendung des UKF eine bessere Schätzung der Geschwindigkeit erreicht.

### Laufzeitanalyse

In einer Fahrumfelderfassung sind Echtzeitanprüche von besonderer Bedeutung. In einer Stadt oder auf einer Autobahn sind 10 bis 15 Objekte, die gleichzeitig verfolgt werden müssen, keine Seltenheit. In Tabelle 5.3 ist für den Kalman-Filter und den Unscented-Kalman-Filter jeweils die durchschnittliche Rechenzeit angegeben, die für einen Vorhersage- bzw. Korrekturschritt benötigt wird. Grundlage der Berechnung war das Objektmodell für Fahrzeuge aus Abschnitt 4.1.1 und das Messmodell für den Laserscanner aus Abschnitt 4.2.4. Die Ausführung erfolgte auf einem Intel Core 2 Duo mit 2.4 GHz, wobei keine Optimierung für moderne Mehrkernprozessoren vorgenommen wurde. Für eine ausführliche Betrachtung derartiger Optimierungen sei auf Kapitel 7.2.5 verwiesen.

Es war zu erwarten, dass der UKF für beide Schritte mehr Rechenzeit benötigt, da sowohl das dynamische Modell, als auch das Messmodell auf alle Sigma-Punkte angewendet werden müssen. Aus Sicht der Laufzeitanalyse ist dies der einzige Unterschied der beiden Varianten. Dementsprechend ist der Unterschied in der Laufzeit hauptsächlich von der Anzahl der Sigma-Punkte und der Komplexität der jeweiligen Modelle abhängig. Unter der Annahme von 20 Objekten, die gleichzeitig verfolgt werden müssen, kann mit beiden Varianten eine echtzeitfähige<sup>2</sup> Fahrumfelderfassung realisiert werden.

Filter	Vorhersage	Korrektur
EKF	0.120 ms	0.025 ms
UKF	0.350 ms	0.100 ms

Tabelle 5.3: Vergleich der durchschnittlichen Rechenzeit für EKF und UKF. Grundlage ist das Objektmodell für Fahrzeuge und das Messmodell für den Laserscanner. Es ist jeweils die Rechenzeit für eine Ausführung des Vorhersage- bzw. Korrekturschrittes angegeben.

<sup>2</sup> In Rahmen dieser Arbeit wird ab einer Frequenz von 15 Hz von Echtzeitfähigkeit ausgegangen.

### Aspekte der Implementierung

Neben der Genauigkeit und der Echtzeitfähigkeit ist der Implementierungsaufwand ein weiterer wichtiger Aspekt für den Vergleich von Filterverfahren. Für bestimmte Systemmodellierungen kann es zum Beispiel erhebliche Unterschiede in der technischen Umsetzung geben. Wie bereits in den vorherigen Abschnitten erwähnt wurde, existieren für beliebige Systemmodelle sowohl für den UKF, als auch den EKF generische Implementierungen. Der einzige Unterschied der beiden Verfahren in Bezug auf die technische Umsetzung besteht in der Formulierung der Eingangsgrößen. So werden für den EKF die Jacobi-Matrix für die Funktionen des Objekt- und Messmodells ( $f_t(x_{t-1}, u_t, w_t)$  bzw.  $h_t(x_t, v_t)$ ) benötigt. Die analytische Berechnung dieser Matrizen kann unter Umständen nicht trivial sein. Bei den UKF ist dieser Schritt nicht notwendig. Hier werden nur die konkreten Funktionen  $f_t$  und  $h_t$  benötigt.

### 5.2.7 Zusammenfassung

Der Kalman-Filter und seine Varianten sind sehr verbreitete Verfahren zur Zustandsschätzung linearer und nichtlinearer Systeme. Mit ihnen kann das Problem der Objektverfolgung nicht nur sehr effizient, sondern unter bestimmten Voraussetzungen auch optimal gelöst werden. Selbst in sehr komplexen Stadtszenarien mit bis zu 50 Objekten kann mit aktueller Hardware eine echtzeitfähige Fahrumfelderfassung realisiert werden. Für stark nichtlineare Systeme, wie sie zum Beispiel in dem dynamischen Modell für Fahrzeuge zum Einsatz kommen, bietet der Unscented-Kalman-Filter einige Vorteile gegenüber dem Extended-Kalman-Filter. Zum Einen gestaltet sich die Implementierung deutlich einfacher, da eine analytische und teilweise sehr aufwendige Berechnung der Jacobimatrizen nicht notwendig ist. Zum Anderen profitiert die Qualität der Schätzung der lateralen Ablage deutlich von der besseren Behandlung der nichtlinearen Systeme. Allerdings ist dieser Effekt nur in Szenen mit einem großen Anteil an Querdynamik zu beobachten und mit einer deutlich erhöhten Rechenzeit verbunden. Dennoch ist, bei ausreichender Rechenkapazität, der Unscented-Kalman-Filter für die Realisierung einer Fahrumfelderfassung zu bevorzugen, da er in der Umsetzung einfacher zu handhaben ist und mindestens dieselbe Qualität liefert.

### 5.3 Partikel-Filter

Bei den Partikel-Filtern wird die a posteriori Verteilung des zu verfolgenden Objekts nicht wie bei den Kalman-Filtern durch eine geschlossene Gleichung mittels Normalverteilungen angegeben, sondern durch eine endliche Menge diskreter Einzelhypothesen (*Partikel* bzw. *Samples*) genähert. Sowohl das dynamische Modell als auch die Messfunktion können von beliebiger (nichtlinearer) Form sein. Die Forderung nach der Differenzierbarkeit entfällt ebenso wie die Annahme, dass das Prozess- und Messrauschen normalverteilt sind. Allerdings handelt es sich bei Partikel-Filtern, im Vergleich zu den in Abschnitt 5.1.2 beschriebenen Algorithmen, nicht mehr um optimale bzw. analytische Algorithmen. In der Literatur sind sie auch unter den Bezeichnungen *sequential Monte Carlo methods* [DGA00], *condensation-algorithm* [IB98] und *bootstrap-filter* [CS86] zu finden.

Der stochastische Ansatz, in dem sich ein Partikel-Filter vom Kalman-Filter unterscheidet, ist die Modellierung des zeitlichen Verlaufs einer multimodalen Wahrscheinlichkeitsverteilung anstatt der Schätzung eines einzigen Zustands. Diese Modellierung entspricht einer gleichzeitigen Berücksichtigung mehrerer alternativer Hypothesen über den tatsächlichen Zustand des modellierten Systems, was mit einem Kalman-Filter grundsätzlich nicht möglich ist. Dies wird erreicht, indem die a posteriori Wahrscheinlichkeitsverteilung des Objektzustandes durch eine zufällig generierte Menge von Partikeln repräsentiert wird. Diese Menge wird auf der Grundlage eines dynamischen Modells propagiert und durch die Anwendung eines Beobachtungsmodells neu bewertet. In Abbildung 5.3 ist ein Iterationszyklus eines Partikel-Filters dargestellt.

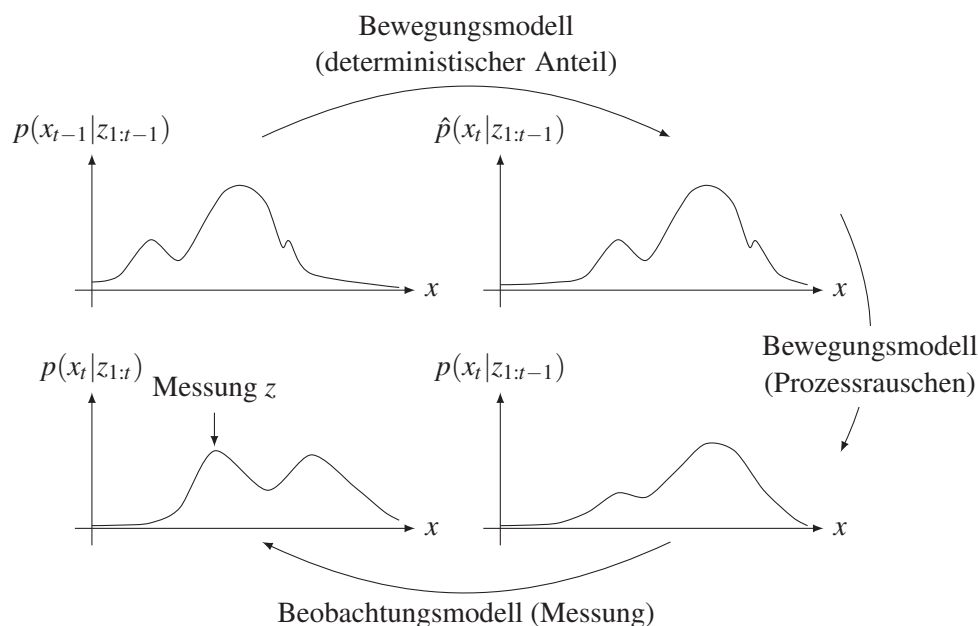


Abbildung 5.3: Darstellung eines Partikel-Filter-Iterationsschrittes. Die Schritte sind analog zum Kalman-Filter (vgl. Abbildung 5.1). Anhand des Bewegungsmodells erfolgt die Prädiktion der einzelnen Hypothesen, gefolgt von einer Überlagerung durch das Prozessrauschen. Schließlich werden die prädizierten Hypothesen mittels der Messungen des Beobachtungsmodells korrigiert.

Der Basis-Algorithmus (*sequential sampling*) wurde bereits in den 50er-Jahren von Hammersley und Morton[HM54] eingeführt, jedoch aufgrund des Rechenaufwandes lange Zeit weitestgehend ignoriert. Erst durch die gestiegene Rechenkapazitäten wurden Partikel-Filter praxistauglich. Seitdem hat sich die Forschungsaktivität erheblich erhöht und diverse Erweiterungen und Anwendungen wurden entwickelt [DGA00, Che03]. Im Folgenden werden einige dieser Erweiterungen vorgestellt, die im Hinblick auf eine Fahrumfelderfassung von besonderem Interesse sind. Sie verfolgen im Allgemeinen das Ziel, die Genauigkeit zu steigern bzw. den Rechenaufwand zu verringern.

### 5.3.1 Kernel-Partikel-Filter

Die Motivation des Kernel-Partikel-Filters [CA03] besteht darin, die Partikel dichter um die Maxima der zu schätzenden Verteilungsdichte zu positionieren und somit in diesen relevanten Bereichen eine bessere Abtastung zu erreichen. Hierzu wird das Mean-Shift-Verfahren [FH75] verwendet, um die Maxima der Verteilungsdichte zu finden und die Partikel in deren Richtung zu verschieben. Für eine detaillierte Einführung und Anwendungsbeispiele aus der Bildverarbeitung sei auf [SS06, CA03, CRM03] verwiesen.

Beim Kernel-Partikel-Filter wird das Verfahren der *Kernel Density Estimation* (auch *Parzen window method* [Par62]) angewandt, um eine unbekannte kontinuierliche Verteilungsdichte mit Hilfe von endlich vielen Partikeln abzuschätzen. Jedem Partikel wird dazu ein so genannter Kernel  $K$  zugewiesen, der zum Beispiel einer Gaußfunktion entsprechen kann (siehe auch Abbildung 5.4a). Anschließend werden die Partikel, ausgehend von einem Startwert, mittels des Mean-Shift-Verfahrens iterativ entlang der Gradienten zu lokalen Maxima der kontinuierlichen Verteilungsdichte verschoben. Auf diese Weise häufen sich Partikel in der Nähe von lokalen Maxima. In Abbildung 5.4b ist das Ergebnis dieses Prozesses nach mehreren Iterationsstufen dargestellt.

Bei einer Anwendung des Mean-Shift-Verfahrens im Kontext einer Fahrumfelderfassung muss beachtet werden, dass die einzelnen Komponenten der mehrdimensionalen Partikel unterschiedlich skaliert sind und in verschiedenen Einheiten vorliegen (z.B. Position in  $m$ , Beschleunigung in  $m/s^2$ ). Deshalb werden sie vor Anwendung des *Mean Shift-Verfahrens* auf  $N(0, 1)$  normalisiert und nach der Verschiebung wieder zurück transformiert.

Die Wahl der Kernelbreite  $h$  hat ebenfalls großen Einfluss auf das Ergebnis, denn je größer dieser Wert ist, desto weniger Cluster werden gebildet. Im Extremfall werden alle Partikel hin zu einem Punkt, d.h. einer Hypothese, verschoben. Dadurch geht die Eigenschaft des Partikel-Filters verloren, mehrere verschiedene Hypothesen gleichzeitig zu verfolgen. Wie in [Sil86] gezeigt wurde, existiert bezüglich des mittleren quadratischen Fehlers zwischen der a posteriori Verteilung und der *Kernel Density Estimation* eine optimale Kernelbreite:

$$h_{opt} = \left( \frac{4}{(d+2) \cdot N} \right)^{\frac{1}{d+4}} \quad (5.44)$$

Obwohl Gleichung (5.44) nur für gleichgewichtete Partikel, einer gaußverteilten a posteriori Verteilung und der Einheitsmatrix als Kovarianzmatrix optimal ist, kann diese Kernelbreite auch bei multimodalen Verteilungen verwendet werden, um einen suboptimalen Filter zu erlangen [DGA00].

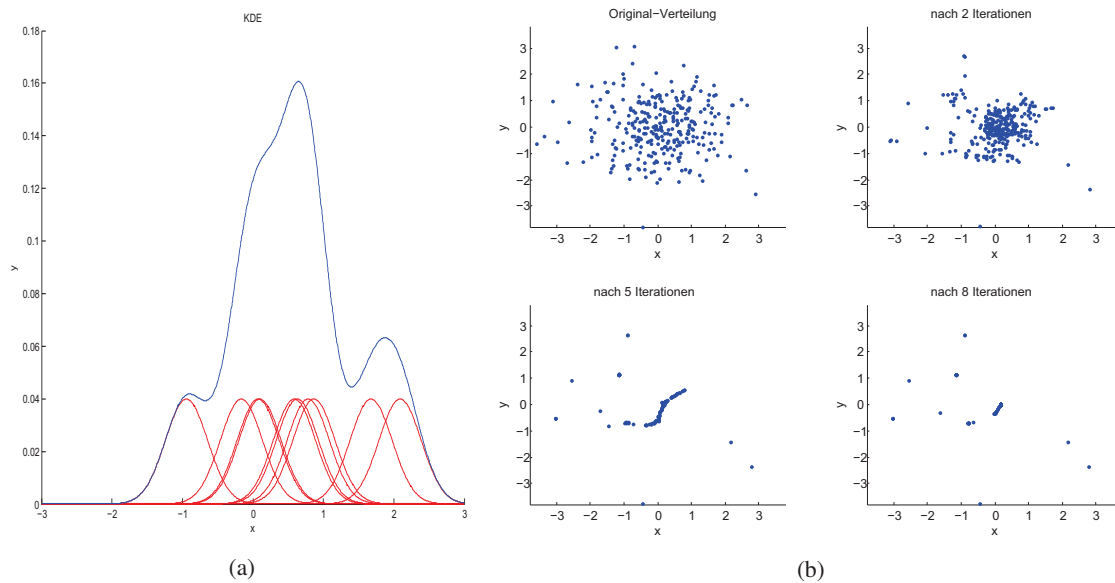


Abbildung 5.4: Beispiel für die *Kernel Density Estimation* (KDE) und das Mean-Shift-Verfahren. Gegeben sind  $N = 10$  gleichgewichtete Partikel in den Mittelpunkten ihrer jeweiligen Kernfenster (rot) mit  $h = 0.3$ . Die durch diese Partikel repräsentierte Funktion wird mittels der KDE geschätzt (blau). Abbildung (b) zeigt das Mean-Shift-Verfahren nach mehreren Iterationen, angewendet auf eine Verteilung von  $N = 300$  gleichgewichteten Partikel der Dimension  $d = 2$ . Aufgrund der gewählten Kernelbreite  $h = 0.5$  bilden sich mehrere Cluster von Partikeln heraus.

### 5.3.2 Annealed-Partikel-Filter

Ein Standard-Partikel-Filter ist schlecht geeignet, um hochdimensionale Verteilungsdichten zu verfolgen, da die Anzahl der benötigten Partikel exponentiell mit der Zahl der Dimensionen steigt. Es können, aufgrund der notwendigen Rechenzeit, nicht genug mögliche Systemzustände  $x_t$  nachgemessen werden, um die Beobachtungsdichte  $p(z_t|x_t)$  hinreichend zu approximieren. Deutscher et. al schlugen deshalb in ihrer Arbeit [DBR00] den Annealed-Partikel-Filter (APF) vor. Ihr Objektmodell, das menschliche Körperbewegungen beschreibt, umfasste dabei 29 Freiheitsgrade. Es konnte gezeigt werden, dass 400 Partikel ausreichend sind, um dieses komplexe Modell zu verfolgen. Die Idee des Annealed-Partikel-Filter besteht darin, durch *simulated annealing*<sup>3</sup> [KGJV83], die Partikel im hochdimensionalen Zustandsraum zu den relevanten Stellen (Maxima) zu lenken und dabei lokale Maxima zu umgehen.

Das Verfahren arbeitet in mehreren Schichten, wobei für jede Schicht  $m$  ein Glättungsparameter  $\beta_m \in \{\beta_M, \beta_{M-1}, \dots, \beta_0\}$  mit  $\beta_M < \beta_{M-1} < \dots < \beta_0 = 1$  definiert wird. Mit diesem Parameter wird der Einfluss der Beobachtungsdichte gesteuert. Der Prozess beginnt in der Schicht  $M$  mit einer stark geglätteten Beobachtungsdichte mit nur wenigen lokalen Maxima. In der letzten Schicht wird die ursprüngliche Beobachtungsdichte verwendet. In jeder Schicht wird die modifizierte Beobachtungsdichte durch

<sup>3</sup> Die simulierte Abkühlung (engl. simulated annealing) ist ein Verfahren zum Auffinden einer approximativen Lösung von Optimierungsproblemen und ist an den Abkühlungsprozess von Metallen angelehnt. Die Grundidee besteht darin, dass bei hohen Temperaturen lokale optimale Lösungen wieder verlassen werden können. Mit abnehmender Temperatur wird diese Möglichkeit sukzessive eingeschränkt.

eine Menge  $S_{t,m}$  an Partikeln abgetastet. Anschließend folgt ein *resampling* Schritt durch Ziehen mit Zurücklegen anhand der Gewichte der Partikel. Dadurch werden niedrig gewichtete Partikel verworfen. Die so erhaltene Menge  $S_{t,m}$  wird schließlich mit einem multivariaten, gaußverteilten Rauschen  $P_m$  überlagert und der nächsten Schicht übergeben. Das Rauschen  $P_M$  ist dabei derart gewählt, dass sich die Varianz in jeder Schicht verringert  $P_M > \dots > P_0$ . Dadurch wird sichergestellt, dass die Partikel dichter beieinander liegen, während der Algorithmus gegen Maxima konvergiert. In Abbildung 5.5 ist die Funktionsweise des APF schematisch für vier Schichten dargestellt.

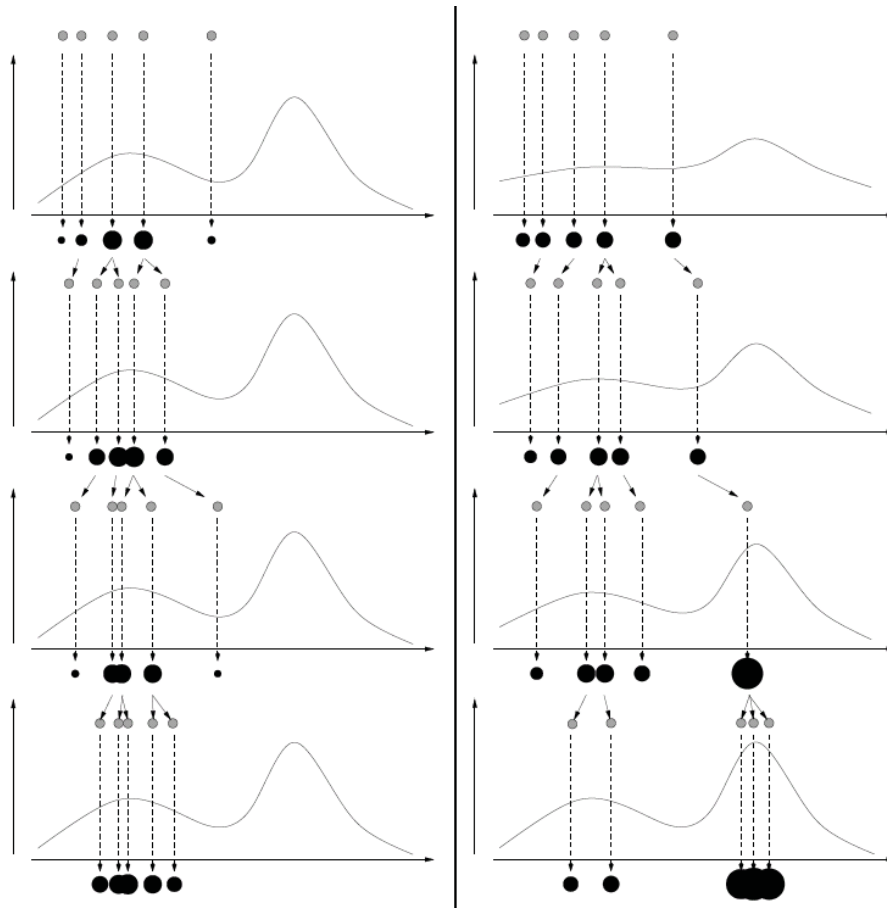


Abbildung 5.5: Schematische Funktionsweise des Annealed-Partikel-Filters (APF).

Links: Ohne den APF bleiben die Partikel in einem lokalen Maximum hängen. Rechts: Der APF bewirkt, dass die Partikel aus diesem lokalen Maximum entkommen können. (Abbildung wurde aus [GPS<sup>+</sup>07] entnommen)

### 5.3.3 Fusion

Nachdem in den vorherigen Abschnitten unterschiedliche Varianten des Partikel-Filters aufgezeigt wurden, soll im Folgenden geklärt werden, wie sich mehrere Sensoren in das Konzept des Partikel-Filters integrieren lassen. Analog zu der Klasse der Kalman-Filter lassen sich drei grundlegende Konzepte unterscheiden [IB05] (siehe auch Abbildung 5.6):

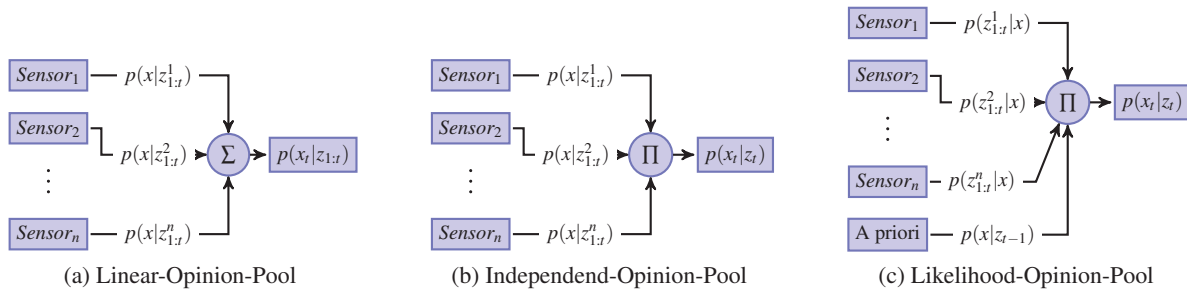


Abbildung 5.6: Unterschiedliche Modelle der Bayesianen Fusion. In (a) ist der Linear-Opinion-Pool dargestellt. Die a posteriori Verteilungen der einzelnen Sensoren werden linear kombiniert. Der Independend-Opinion-Pool (b) ist ein Spezialfall von (a). Hier wird angenommen, dass die Beobachtungen der Sensoren unabhängig sind. Der in (c) dargestellt Likelihood-Opinion-Pool ist ein rekursiver Prozess, in dem die a priori Verteilung in die Fusion einbezogen wird.

**Linear-Opinion-Pool:** In der als Linear-Opinion-Pool bezeichneten Methode werden die a posteriori Verteilungen  $p(x|z_{1:t}^i)$  der einzelnen Sensoren zu einer gesamten a posteriori Verteilung kombiniert (siehe auch Abbildung 5.6a):

$$p(x_t|z_{1:t}) = \sum_1^n \omega_i p(x_t|z_{1:t}^i) \quad \text{mit} \quad \sum_1^n \omega_i = 1 \quad (5.45)$$

**Independend-Opinion-Pool:** Der sogenannte Independend-Opinion-Pool ist ein Spezialfall des bereits genannten Linear-Opinion-Pool (siehe auch Abbildung 5.6b). Hier wird angenommen, dass die Beobachtungen der einzelnen Sensoren unabhängig sind. Somit berechnet sich die fusionierte a posteriori Verteilung wie folgt:

$$p(x_t|z_{1:t}) = \prod_1^n p(x_t|z_{1:t}^i) \quad (5.46)$$

Die beiden zuerst genannten Arten der Fusion sind auch unter dem Begriff der Track-Level Fusion bekannt.

**Likelihood-Opinion-Pool:** Die dritte Möglichkeit der Fusion (Likelihood-Opinion-Pool) arbeitet auf Messwert- bzw. Partikelebene und setzt ebenfalls voraus, dass die einzelnen Messungen voneinander unabhängig sind. Zusätzlich wird die a priori Verteilung  $p(x_t|z_{1:t-1})$  mit in den Fusionsprozess einbezogen. Dies erfolgt implizit durch die Anwendung des Bewegungsmodells. Dadurch werden Messungen, die aus unwahrscheinlichen Bewegungen resultieren würden, entsprechend niedrig gewichtet.

Im weiteren Verlauf dieser Arbeit wird die Methode des Likelihood-Opinion-Pools verwendet, da diese auf Messwert- bzw. Partikelebene fusioniert und somit bessere Ergebnisse erwarten lässt (siehe frühe Fusion Abschnitt 2.3.3).



### 5.3.4 Bewertung der Partikel-Filter

Im folgenden Abschnitt wird die Schätzgenauigkeit der verschiedenen Partikel-Filter-Verfahren unter vordefinierten Szenarien für Fahrzeuge und Fußgänger evaluiert. Um die gleichen Randbedingungen zu schaffen, wurden für alle Partikel-Filter dieselben Werte für das Messrauschen der Sensoren und das Prozessrauschen der Bewegungsmodelle verwendet. Für diese Evaluierung wird nur der Laserscanner verwendet. Eine Evaluierung in Bezug auf eine Fusionierung verschiedener Sensorinformationen findet am Ende dieses Abschnittes statt. Aus Übersichtsgründen werden im Folgenden für die zu vergleichenden Filterverfahren folgende Abkürzungen verwendet: PF = Standard-Partikel-Filter, KPF = Kernel-Partikel-Filter bzw. APF = Annealed-Partikel-Filter.

Da es sich bei Partikel-Filtern um nichtdeterministische Verfahren handelt, können die Ergebnisse der verschiedenen Durchläufe schwanken. Um diesen probabilistischen Einfluss auf das Ergebnis der Evaluierung zu minimieren, wird für jedes Szenario das Ergebnis von 25 Durchläufen gemittelt. Zusätzlich wird in Abhängigkeit der Partikelzahl die Standardabweichung der Ergebnisse der einzelnen Durchläufe bestimmt. Dadurch ist eine Abschätzung möglich, wie stark das Ergebnis pro Durchlauf variiert. Die Anzahl der verwendeten Partikel ist im Wesentlichen für die Qualität der Schätzung verantwortlich. Allerdings ist die Laufzeit hierzu direkt proportional. Um für diesen Zielkonflikt einen geeigneten Kompromiss zu finden, wird für jedes Szenario der Einfluss der Partikelzahl auf die Qualität der Schätzung evaluiert.



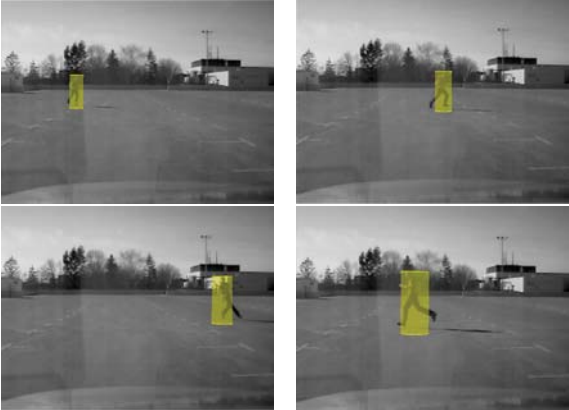
#### Beschreibung der Testszenerien

Die Tabelle 5.4 gibt einen Überblick über die verwendeten Szenarien. Alle Szenen wurden unter optimalen Umweltbedingungen aufgezeichnet und können hinsichtlich der Objektklassen in Fahrzeug- und Fußgängerszenarien untergliedert werden. Innerhalb der jeweiligen Klassen erfolgt eine weitere Unterteilung in Bezug auf die Objektdynamik. Die für die Evaluierung nötigen Referenzdaten wurden durch einen dedizierten Sensor (siehe auch Abschnitt 7.1.3) ermittelt. Als Qualitätsmaß wurde der RMSE für die Schätzgröße der Entfernung  $r$ , der Geschwindigkeit  $v$  sowie der lateralen Ablage  $\varphi$  gewählt (siehe auch Abschnitt 7.2.1)

Im ersten Szenario führt das zu verfolgende Fahrzeug eine Vollbremsung mit einer Verzögerung von  $9 \frac{m}{s^2}$  durch. Dieses extreme Manöver ist für ein Filterverfahren besonders anspruchsvoll, da sich Zustandsgrößen, die nicht direkt beobachtet werden können, in kurzer Zeit sehr stark ändern. Im zweiten Szenario wird eine typische Situation in der Stadt oder auf der Landstraße adressiert, in der ein vorausfahrendes Fahrzeug verfolgt wird. Um den Anspruch zu steigern, sind einige Passagen mit engen Kurven und starken Beschleunigungen enthalten. Das dritte Szenario adressiert einen sehr dynamischen Fußgänger, der sehr oft und zufällig die Richtung wechselt. Dieses Szenario ist von besonderem Interesse, da sich wie in dem zweiten Experiment Zustandsgrößen sehr stark und unvorhersagbar ändern können.



Tabelle 5.4: Szenenüberblick.

Szenenbeschreibung	Szenenausschnitte
<p>Szenario 1: Vollbremsung eines vorausfahrenden Referenzfahrzeuges bei 50 km/h. Dauer: 8.7s (288 Frames).</p>	 <p>The images show a first-person view from a vehicle on a road. In the top-left frame, a small green bounding box is centered on a car in the distance. In the top-right frame, a purple bounding box is centered on the same car. In the bottom-left frame, a blue bounding box is centered on the car. In the bottom-right frame, a green bounding box is centered on the car. The car's position relative to the camera changes slightly between frames, indicating movement.</p>
<p>Szenario 2: Dynamische Verfolgungsfahrt eines anderen PKW auf einer kurvenreichen Teststrecke. Dauer: 54.0s (1783 Frames).</p>	 <p>The images show a first-person view from a vehicle on a winding road. In the top-left frame, a purple bounding box is centered on a car ahead. In the top-right frame, a green bounding box is centered on the car. In the bottom-left frame, a blue bounding box is centered on the car. In the bottom-right frame, a green bounding box is centered on the car. The road curves to the right, and the car's position changes accordingly.</p>
<p>Szenario 3: Rennender und abrupt die Richtung wechselnder Fußgänger. Dauer: 15.5s (521 Frames).</p>	 <p>The images show a first-person view from a vehicle in a parking lot. In the top-left frame, a yellow bounding box is centered on a pedestrian. In the top-right frame, a yellow bounding box is centered on the pedestrian. In the bottom-left frame, a yellow bounding box is centered on the pedestrian. In the bottom-right frame, a yellow bounding box is centered on the pedestrian. The pedestrian is moving across the frame.</p>

### Senario 1 „Vollbremsung“

Die Schätzfehler der Filter für das erste Szenario sind in Abbildung 5.7 dargestellt. Es ist deutlich zu erkennen, dass der Annealed-Partikel-Filter durchgehend höhere Genauigkeiten liefert, als die restlichen Verfahren. Allerdings nimmt mit steigender Anzahl an Partikeln der Unterschied deutlich ab. Dieser Effekt war zu erwarten, da es sich bei den Varianten APF und KPF um Optimierungen hinsichtlich der Anzahl der benötigten Partikel bzw. der Geschwindigkeit der Konvergenz handelt. Der Einfluss der Partikelanzahl auf die Qualität der Schätzung ist deutlich ersichtlich. Bei der Verwendung von mehr als 1600 Partikeln konnte keine signifikante Steigerung der Ergebnisse mehr erreicht werden. Bei dem Annealed-Partikel-Filter ist eine Konvergenz sogar schon bei 400 Partikeln erkennbar.

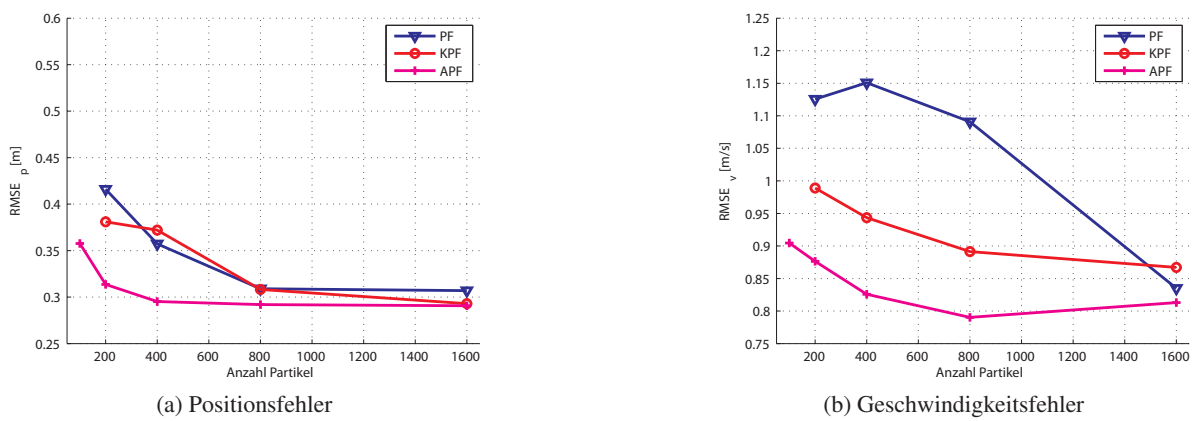


Abbildung 5.7: Schätzfehler der verschiedenen Partikel-Filter bei verwendeter Partikelzahl  $N = 100, 200, 400, 800, 1600$  für das Szenario 1. Nicht eingetragene Werte bedeuten, dass mit dem jeweiligen Filter in mehr als 50% der Durchläufe das Objekt nicht durchgehend verfolgt werden konnte. Hier betrifft das den Fall für den PF mit  $N = 100$  und den KPF mit  $N = 100$ .

In Tabelle 5.5 ist prozentual angegeben, in wie vielen der 25 Durchläufe das Objekt nicht durch die gesamte Testszene hindurch kontinuierlich verfolgt werden konnte. Hieraus ist zum Einen ersichtlich, dass in diesem Testszenario die Verwendung der Partikel-Filter erst ab einer Partikelzahl von  $N \geq 400$  sinnvoll ist und zum Anderen, dass der Annealed-Partikel-Filter deutlich weniger Partikel benötigt, um ein Fahrzeug robust verfolgen zu können.

Filter \ Anzahl Partikel	-	100	200	400	800	1600
PF	-	×	24	0	0	0
KPF	-	×	28	0	0	0
APF	-	16	12	0	0	0

Tabelle 5.5: Prozentsatz der Durchläufe, bei denen der verwendete Filter das Objekt nicht über die gesamte Testszene durchgehend verfolgen konnte. Mit dem Symbol „×“ werden die Fälle gekennzeichnet, bei denen mehr als 50% der Durchläufe einen Objektverlust vorwies.

Bei Partikel-Filtern handelt es sich, wie bereits erwähnt, um nichtdeterministische Algorithmen. Die Ergebnisse aus verschiedenen Durchläufen können dementsprechend selbst bei identischer Parametrisierung (Partikelanzahl, Prozess- und Messrauschen) schwanken. Um den Einfluss dieses Verhaltens auf die Schätzfehler zu evaluieren, bietet sich eine Betrachtung der Standardabweichungen der Fehlerwerte innerhalb der erfolgreichen Durchläufe an. In Abbildung 5.8 ist das Ergebnis dieser Evaluierung dargestellt. Es ist zu erkennen, dass die Schwankung der Fehlerwerte in den verschiedenen Durchläufen derselben Testszene mit zunehmender Partikelzahl sinkt und ab 800 Partikel nur noch sehr gering ist. Darüber hinaus ist eine bessere Konvergenz des Annealed-Partikel-Filters im Vergleich zum Standard- oder Kernel-Partikel-Filter ersichtlicher.

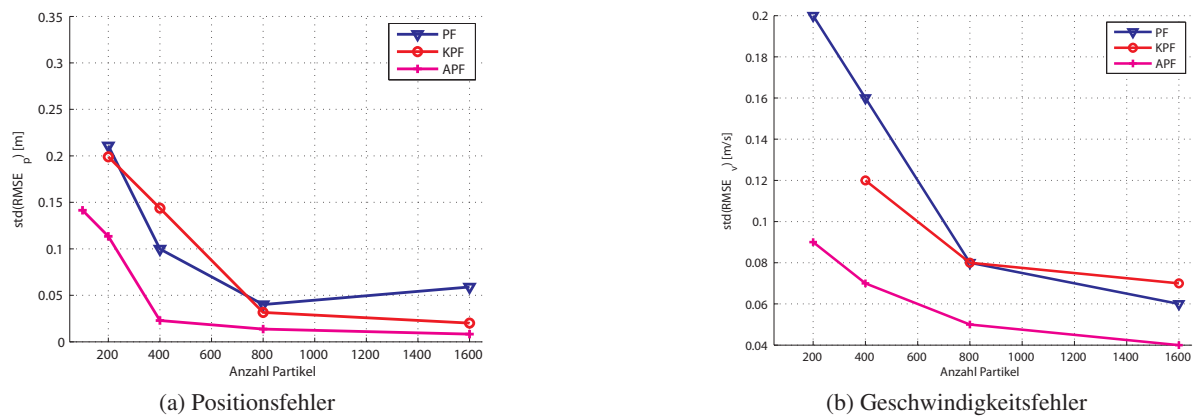


Abbildung 5.8: Experiment „Vollbremsung“: Standardabweichungen der Schätzfehler innerhalb der erfolgreichen Durchläufe der verschiedenen Partikel-Filter bei verwendeter Partikelzahl  $N = 100, 200, 400, 800, 1600$ . Je mehr Partikel verwendet werden, desto mehr gleichen sich die Resultate bei verschiedenen Durchläufen an.

## Szenario 2 „Dynamische Fahrt“

Für das zweite Szenario sind in Abbildung 5.9 die Schätzfehler der einzelnen Filtervarianten dargestellt. Dieses Szenario beinhaltet, abgesehen von einigen Passagen mit engen Kurven oder starken Beschleunigungen, nur wenige abrupte Zustandsänderungen. Da in Kurvenfahrten die Positionsänderungen durch die direkte Messung der Orientierung und der Position gut beobachtet werden kann, sind für den Fehler in der Positionsschätzung  $RMSE_p$  keine signifikanten Unterschiede zwischen den verschiedenen Partikel-Filter Varianten zu erkennen. Lediglich in der Schätzung der Geschwindigkeit ist der Annealed-Partikel-Filter den beiden anderen Verfahren überlegen, da sich der APF in den Phasen der starken Beschleunigung schneller an die neue Beobachtungsdichte adaptiert.

In Tabelle 5.6 ist angegeben, in wie vielen der 25 Durchläufe das Objekt nicht durch die gesamte Testszene hindurch kontinuierlich verfolgt werden konnte. Hier fällt auf, dass die Objektverluste im Vergleich zum ersten Szenario häufiger sind, und dementsprechend die Verwendung der Partikel-Filter erst ab einer Partikelzahl von mindestens 400 bis 800 Sinn macht.

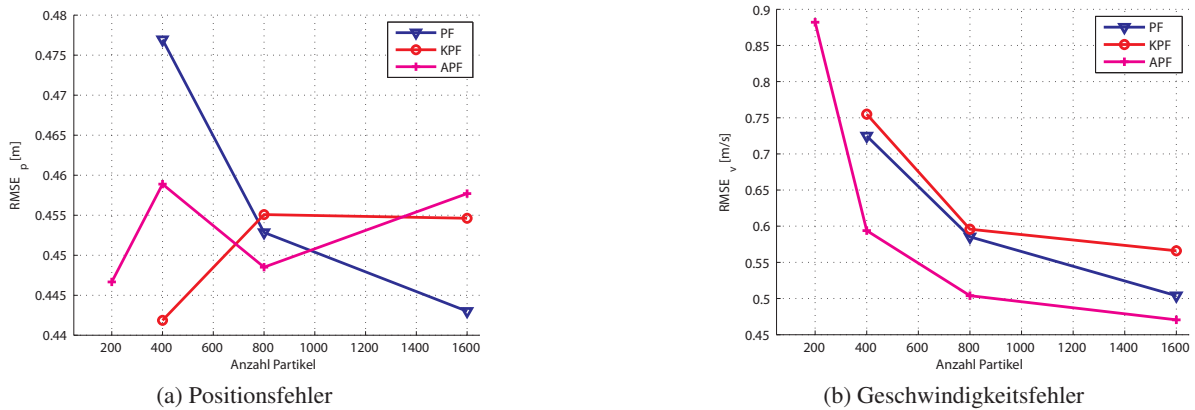


Abbildung 5.9: Schätzfehler der verschiedenen Partikel-Filter bei verwendeter Partikelzahl  $N = 200, 400, 800, 1600$  für das Szenario 2. Nicht eingetragene Werte bedeuten, dass mit dem jeweiligen Filter in mehr als 50% der Durchläufe das Objekt nicht durchgehend verfolgt werden konnte.

Filter \ Anzahl Partikel	-	100	200	400	800	1600
PF	-	×	×	12	4	0
KPF	-	×	×	28	0	0
APF	-	×	10	0	0	0

Tabelle 5.6: Prozentsatz der Durchläufe, bei denen der verwendete Filter das Objekt nicht über die gesamte Testszene durchgehend verfolgen konnte. Mit dem Symbol „×“ werden die Fälle gekennzeichnet, bei denen mehr als 50% der Durchläufe einen Objektverlust vorwies.

### Szenario 3 „Fußgänger“

Die Werte des Positionsfehlers  $RMSE_p$ , sowie dessen Standardabweichung über die einzelnen Durchläufe sind für das dritte Szenario in Abbildung 5.10 dargestellt. Da für die Objektgeschwindigkeit keine Referenzdaten vorlagen, konnten diese nicht evaluiert werden. Wie bereits in den anderen beiden Szenarien liefert der Annealed-Partikel-Filter durchgehend genauere Ergebnisse als der Standard-Partikel-Filter. Bereits ab 200 Partikeln ist bei dem APF eine deutliche Konvergenz zu erkennen. Wie aus Abbildung 5.10b ersichtlich ist, spielt der Faktor Zufall bereits ab 200 Partikeln nur noch eine untergeordnete Rolle.

In Tabelle 5.7 ist wieder angegeben, in wie vielen der 25 Durchläufe das Objekt nicht durch die gesamte Testszene hindurch kontinuierlich verfolgt werden konnte. Aufgrund der abrupten Richtungswechsel des Objekts versagt der Kernel-Partikel-Filter nahezu vollständig, da sich hier die Verschiebung der Partikel in Richtung lokaler Maxima der Beobachtungsdichte (Abschnitt 5.3.1) negativ auswirkt. Diese Verschiebung betrifft auch die Zustandsgrößen, die nicht gemessen werden können. Rennt der Fußgänger zum Beispiel in eine Richtung, adaptiert sich der Kernel-Partikel-Filter in stärkerem Maße als die anderen Partikel-Filter an den zugehörigen Geschwindigkeitsvektor. Der Kernel-Partikel-Filter misst dem bisherigen Geschwindigkeitsvektor also eine sehr große Gewichtung bei, so dass die einzel-

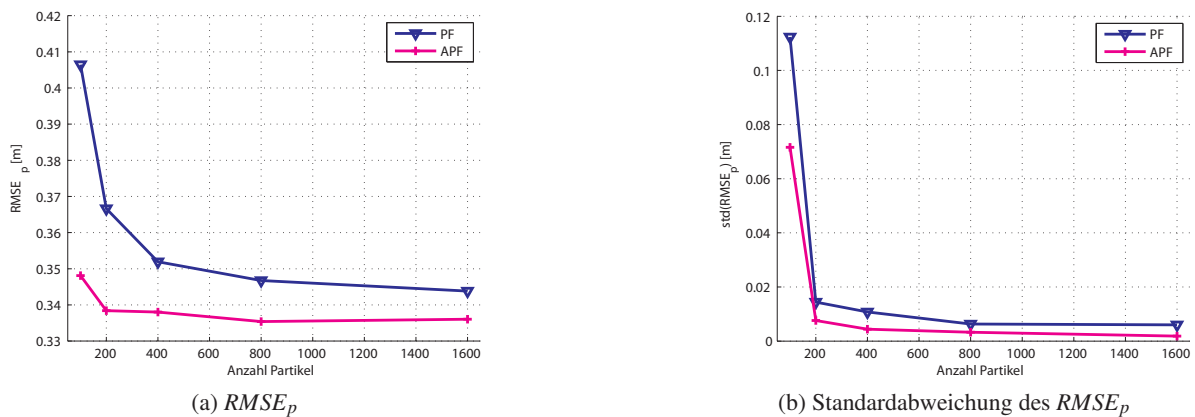


Abbildung 5.10: Experiment „Fußgänger“: Positionsfehler  $RMSE_p$  der verschiedenen Partikel-Filter bei verwendeter Partikelzahl  $N = 100, 200, 400, 800, 1600$ .

nen Partikel im Zustandsraum zu dicht um diesen Geschwindigkeitsvektor herum verteilt sind. Erfolgt eine plötzliche und starke Änderung des Geschwindigkeitsvektors, so kann diese von den Partikeln nicht mehr erfasst werden.

Filter \ Anzahl Partikel	-	100	200	400	800	1600
PF	-	4	0	0	0	0
KPF	-	×	×	×	×	×
APF	-	4	0	0	0	0

Tabelle 5.7: Prozentsatz der Durchläufe, bei denen der verwendete Filter das Objekt nicht über die gesamte Testszene durchgehend verfolgen konnte. Mit dem Symbol „×“ werden die Fälle gekennzeichnet, bei denen mehr als 50% der Durchläufe einen Objektverlust vorwiesen.

Im Gegensatz zu den ersten beiden Szenarien ist mit dem Standard-Partikel-Filter und dem Annealed-Partikel-Filter eine Verfolgung von Fußgängern, selbst bei extremen Bewegungsabläufen, bereits ab einer Partikelzahl von  $N = 200$  möglich. Dies ist dadurch zu erklären, dass das Objektmodell von Fußgängern in Bezug auf die Freiheitsgrade wesentlich einfacher ist, als das von Fahrzeugen. Somit kann durch die vorhandenen Partikel eine effektivere und konzentriertere Abtastung der Beobachtungsdichte erreicht werden.

### Ergebnisse der Sensorfusion

Um die Auswirkungen zusätzlicher Sensorinformationen auf die Genauigkeit der Schätzung zu überprüfen, wurden die verschiedenen Beobachtungsmodelle aus dem Kapitel 4 kombiniert und unter Verwendung des Standard-Partikel-Filters evaluiert. Hierzu wurden jeweils für Fahrzeuge und Fußgänger eines der in Abschnitt 5.3.4 eingeführten Szenarien verwendet. Die Fusion wurde nach dem Prinzip des Likelihood-Opinion-Pools realisiert (siehe Abschnitt 5.3.3). Vorgreifend sei an dieser Stelle erwähnt, dass sich der Laserscanner als bester Sensor herausstellte was die Schätzgenauigkeit betrifft.

Ohne diesen konnte keine Fusionierung der restlichen verfügbaren Sensoren annähernd gleichwertige Resultate liefern. Somit ist in der folgenden Evaluierung keine Sensorkombination vertreten, welche den Laserscanner nicht verwendet.

Für Fahrzeuge wurden anhand des zweiten Szenarios vier unterschiedliche Kombinationen von Beobachtungsmodellen evaluiert (Laserscanner, Laserscanner + Radar, Laserscanner + Kontur sowie Laserscanner + Radar + Kontur). In Abbildung 5.11 sind die Schätzfehler der einzelnen Sensorkonfigurationen dargestellt. In diesem Szenario konnte durch keine Kombination eine signifikante Steigerung der Positionsschätzung erreicht werden. Die Schätzung der Geschwindigkeit profitiert hingegen von zusätzlichen Messwerten. So führt zum Beispiel die direkte Messung der relativen radialen Geschwindigkeit des Radarsensors zu einer schnelleren Anpassung bei starken Beschleunigungen. Obwohl mit den Videosensoren keine direkte Messung der Geschwindigkeit möglich ist, lässt sich bei ihnen ein ähnlicher Effekt beobachten, der durch die größere Abtastrate der Sensoren von 30 Hz zu erklären ist. Durch die Kombination von drei Sensoren war keine weitere Steigerung zu erreichen.

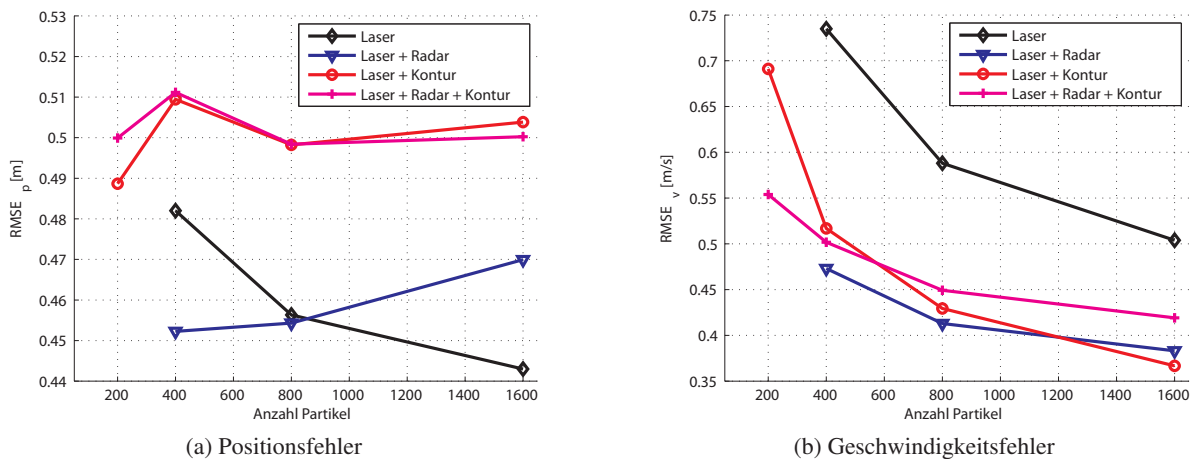


Abbildung 5.11: Schätzfehler für vier Kombinationen von Beobachtungsmodellen in dem zweiten Szenario.

Für Fußgänger sind in Abbildung 5.12 die Werte des Positionsfehlers  $RMSE_p$  für zwei verschiedene Sensorkombinationen (Laser und Laser + Kopfkontur) dargestellt. Auf eine Untersuchung der Radarmodelle wurde verzichtet, da Fußgänger nur sehr bedingt von Radarsensoren beobachtet werden können. Darüber hinaus war eine Betrachtung der Geschwindigkeit nicht möglich, da hierfür keine Referenzdaten verfügbar waren.

In diesem Szenario konnte durch die Fusion eine signifikante Steigerung von über 20 % in der Positionsschätzung erreicht werden. Dies ist zum Einen auf die dreifach höhere Abtastrate der Wärmebildkamera im Vergleich zum Laserscanner zurückzuführen. Somit können dynamische Bewegungen schneller erfasst werden. Zum Anderen ist die Winkelauflösung der Kamera dafür verantwortlich, dass in den Bereichen, in denen sich der Fußgänger quer zum Fahrzeug bewegt, dessen Position präziser geschätzt werden kann.

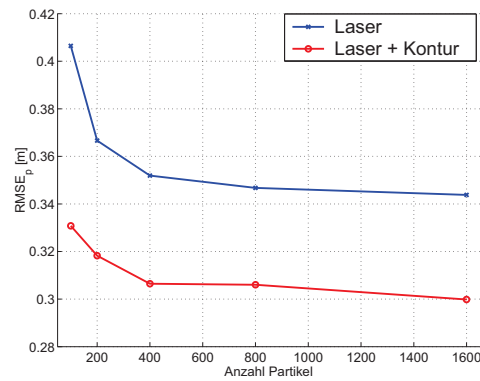


Abbildung 5.12: Experiment „Fußgänger“: Positionsfehler  $RMSE_p$  der verschiedenen Partikel-Filter bei verwendeter Partikelzahl  $N = 100, 200, 400, 800, 1600$ .

### Analyse der Rechenzeit

Aktuelle Fahrerassistenzsysteme, wie zum Beispiel das ACC, sind auf ein echtzeitfähiges Fahrumfeld-erfassungssystem angewiesen. Deshalb ist die benötigte Rechenzeit ein wichtiger Aspekt für die Wahl eines geeigneten Filterverfahrens. Im Allgemeinen müssen mehrere Objekte gleichzeitig verfolgt werden, was zusätzlich zu einer Vervielfältigung des Rechenaufwandes führt. In Abbildung 5.13a ist für die vorgestellten Partikel-Filter Varianten die durchschnittliche Zeit dargestellt, die zur Verfolgung eines Fahrzeuges bei unterschiedlicher Partikelanzahl benötigt wird. Es ist zum Einen deutlich zu erkennen, dass die Rechenzeit nahezu linear mit der Anzahl der Partikel steigt. Zum Anderen ist zu beobachten, dass sowohl der Annealed- als auch der Kernel-Partikel-Filter rechenintensiver sind als der Standard-Partikel-Filter. Der Kernel-Partikel-Filter benötigt um ein vielfaches mehr Zeit, da das verwendete Mean Shift-Verfahren aus Abschnitt 5.3.1 eine quadratische Laufzeit besitzt. Für den APF ist der zusätzliche Aufwand im Wesentlichen von der Anzahl der internen Iterationen abhängig.

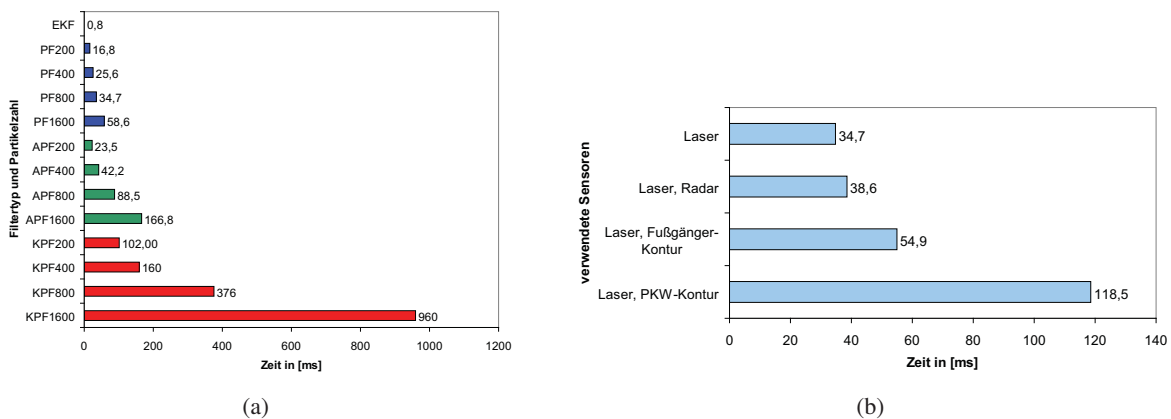


Abbildung 5.13: Laufzeitanalyse der Partikel-Filter: In (a) ist die Mittlere Rechenzeit der verschiedenen Filterverfahren bei der Verfolgung eines Fahrzeuges und unterschiedlicher Partikelanzahl angegeben. In dem Diagramm (b) ist die durchschnittliche Rechenzeit des Standard-Partikel-Filters mit  $N = 800$  Partikel bei Verwendung verschiedener Sensoren dargestellt.



In Abbildung 5.13b ist ein Rechenzeitvergleich für die Fusionierung verschiedener Sensorinformationen angegeben. Hieraus ist erkennbar, dass sich eine Hinzunahme der Radarmesswerte zu den Laserscannerdaten nur unmerklich auf die Laufzeit auswirkt. Die Berechnungen der Fahrzeug- bzw. der Kopfkontur hat hingegen einen deutlichen Einfluss auf die Rechenzeit, da für jeden Partikel explizit im Bild nachgemessen werden muss. Die Laufzeitanalyse erfolgte mit einem Pentium 4 der Taktfrequenz 3.2 GHz.

Um Fahrzeuge robust verfolgen zu können, sind je nach Szenario mindestens 400 bis 800 Partikel notwendig (siehe vorheriger Abschnitt). Unter diesen Umständen ist mit der vorliegenden Implementierung keine echtzeitfähige Fahrumfelderfassung realisierbar. Insbesondere die Fusion von Video Sensoren führt zu einer zusätzlichen Steigerung des Rechenaufwandes. Für Fußgänger sind hingegen nur 100 bis 200 Partikel notwendig, um eine robuste Verfolgung zu gewährleisten. Somit können bei einer Fusion von Laserscanner und Video bis zu zwei Fußgänger in Echtzeit verfolgt werden. Es ist allerdings anzumerken, dass die verwendete Implementierung nicht optimiert ist, und dass sich Partikel-Filter einfach parallelisieren lassen und deshalb sehr gut mit aktuellen Multiprozessor Architekturen skalieren.

### Vergleich mit dem Extended-Kalman-Filter

Nachdem in den vorherigen Abschnitten die einzelnen Partikel-Filter untereinander verglichen wurden, soll im Folgenden die Qualität der Zustandsschätzung des Extended-Kalman-Filters gegenüber dem Annealed-Partikel-Filter untersucht werden. Für die Evaluierung wurden die Testszenen aus Abschnitt 5.3.4 verwendet. In Tabelle 5.8 ist das Ergebnis dieser Analyse dargestellt. Die Tabellenspalten enthalten die untersuchten Szenen. Jeweils zwei Tabellenzeilen beinhalten die Werte einer Metrik für den Extended-Kalman-Filter bzw. den Annealed-Partikel-Filter.

	Einheit	Szenario 1	Szenario 2	Szenario 3
$RMSE_p$ EKF	[m]	0.50	0.45	0.40
$RMSE_p$ APF (1600)	[m]	0.30	0.42	0.34
$RMSE_v$ EKF	$[\frac{m}{s}]$	1.99	0.47	–
$RMSE_v$ APF (1600)	$[\frac{m}{s}]$	0.82	0.46	–

Tabelle 5.8: Für drei Szenarien wurde die Qualität der Zustandsschätzung für die Position  $p$  sowie die Geschwindigkeit  $v$  untersucht. Jeweils zwei Tabellenzeilen beinhalten die Werte dieser Metriken für den Extended-Kalman-Filter (EKF) sowie den Annealed-Partikel-Filter (APF) bei  $n = 1600$  Partikeln. Mit dem Szenario 1 und 2 wird das Tracking von Fahrzeugen evaluiert. Szenario 3 adressiert hingegen Fußgänger.

Aus dem Vergleich der beiden Filtervarianten können folgende Ergebnisse abgeleitet werden: Für das Szenario 1, das eine „Vollbremsung“ enthält, können mit dem APF deutlich geringere Fehler für die Schätzung der Geschwindigkeit erreicht werden. Dies ist darin begründet, dass die Geschwindigkeit nicht direkt beobachtet werden kann. Durch die zufällige Streuung der Partikel adaptiert sich insbesondere der APF deutlich schneller an eine veränderte Geschwindigkeit als der EKF. Eine bessere Schätzung der Geschwindigkeit hat einen direkten Einfluss auf die Qualität der Positionsbestimmung. Deshalb ist in diesem Szenario für den APF im Vergleich zum EKF ein geringerer Fehler für die Schätzung der Position zu beobachten.



In dem kurvenreichen zweiten Szenario ist kein signifikanter Unterschied in den beiden Verfahren zu erkennen. Zum Einen enthält dieses Szenario nur wenige Abschnitte mit Längsdynamik. Zum Anderen kann durch die direkte Messung der Fahrzeugorientierung eine Änderung in der Querdynamik gut beobachtet werden. Somit sind im Gegensatz zum ersten Szenario nur Passagen enthalten, die von beiden Filtervarianten äquivalent behandelt werden können.

Das sehr dynamische Verhalten des Fußgängers in dem dritten Szenario kann durch den APF besser geschätzt werden. Dies ist dadurch zu erklären, dass der Partikel-Filter mehrere Hypothesen gleichzeitig verfolgen kann und somit schneller auf die dynamischen Bewegungen des Fußgängers reagiert.

### 5.3.5 Fazit

In diesem Abschnitt wurden verschiedene Varianten des Partikel-Filters beschrieben und Möglichkeiten aufgezeigt, wie die Messungen mehrerer Sensoren in die dargestellten Verfahren integriert werden können. Anschließend wurden die Ansätze für Fahrzeuge und Fußgänger unter verschiedenen Szenarien in Bezug auf Genauigkeit und Rechenaufwand evaluiert. Hierbei erwies sich der Annealed-Partikel-Filter als bestes Verfahren, da er schon bei kleiner Partikelanzahl konvergiert und somit bei geringerem Rechenaufwand genauere Ergebnisse liefert. Zudem war ein Zusammenhang zwischen der Komplexität der Objektmodelle und der benötigten Partikelanzahl zu beobachten. So ließen sich Fußgänger deutlich genauer verfolgen als Fahrzeuge.

Neben der Evaluierung der Verfahren mit Hinblick auf einen Laserscanner wurden die Auswirkungen zusätzlicher Sensorinformationen auf die Genauigkeit der Schätzung und den Rechenaufwand untersucht. Für Fahrzeuge wurden hierfür vier unterschiedliche Kombinationen aus Laserscanner, Radar und Videosensorik betrachtet. Lediglich für die Schätzung der Geschwindigkeit konnte durch eine Fusion aus Laserscanner und den Radar Sensoren eine gesteigerte Genauigkeit erzielt werden. Die restlichen Zustandsgrößen profitieren nur unwesentlich von der zusätzlichen Sensorik. Für Fußgänger konnte hingegen bei einer Fusion von Laserscanner und Ferninfrarotkamera eine deutliche Steigerung der Positionsgenauigkeit erreicht werden.

Abschließend wurde die Qualität der Zustandsschätzung des Extended-Kalman-Filters gegenüber dem Annealed-Partikel-Filter untersucht. In stark dynamischen Szenarien kann mit dem Annealed-Partikel-Filter eine bessere Schätzgenauigkeit erreicht werden. Allerdings sind für Fahrzeuge hierfür je nach Szenario mindestens 400 bis 800 Partikel notwendig. Unter diesen Umständen ist mit keiner der untersuchten Partikel-Filter Varianten eine Fahrumfelderfassung in Echtzeit realisierbar. Für Fußgänger sind hingegen nur 100 bis 200 Partikel notwendig, um eine robuste Verfolgung zu gewährleisten. Somit können selbst bei einer Fusion von Laserscanner und Ferninfrarotkamera bis zu zwei Fußgänger in Echtzeit verfolgt werden.

## 5.4 Zusammenfassung

Im ersten Teil des Kapitels wurden für die Klasse der Kalman-Filter Erweiterungen beschrieben, die zum Einen die Verwendung nichtlinearer Objekt- und Messmodelle gestatten und zum Anderen Messungen mehrerer Sensoren fusionieren. Zudem wurden Möglichkeiten aufgezeigt, wie vorhandene Latenzzeiten in den Vorgang der Filterung integriert und somit kompensiert werden können. Abschließend wurden die unterschiedlichen Filterverfahren in Bezug auf die Anforderungen einer Fahrumfel-

derfassung untereinander verglichen. Bei ausreichender Rechenkapazität, ist der Unscented-Kalman-Filter zu bevorzugen, da er in der Umsetzung einfacher zu handhaben ist und für nichtlineare Systeme die besten Ergebnisse hinsichtlich der Genauigkeit der Zustandsschätzung erreicht werden können.

Im zweiten Teil des Kapitels wurden nichtdeterministische Verfahren der Objektverfolgung betrachtet. Als Hauptvertreter dieser Gruppe wurden verschiedene Erweiterungen der Partikel-Filter evaluiert, die im Hinblick auf eine Fahrumfelderfassung eine gesteigerte Genauigkeit versprechen bzw. den Rechenaufwand im Vergleich zum Standard-Partikel-Filter verringern. Der Annealed-Partikel-Filter erwies sich hierbei als bestes Verfahren aus dieser Gruppe, da er schon bei kleiner Partikelanzahl konvergiert und somit bei geringerem Rechenaufwand genauere Ergebnisse liefert. Zudem war ein Zusammenhang zwischen der Komplexität der Objektmodelle und der benötigten Partikelanzahl zu beobachten. So ließen sich Fußgänger deutlich effizienter verfolgen als Fahrzeuge. In einem Vergleich mit dem Kalman-Filter konnten in stark dynamischen Szenarien bessere Schätzgenauigkeiten erreicht werden. Allerdings ist der Annealed-Partikel-Filter noch nicht für alle Objektklassen echtzeitfähig und somit für heutige Fahrerassistenzsysteme nur sehr bedingt geeignet.

# VISUALISIERUNG VON HETEROGENEN SENSORDATEN

---

Die Visualisierung von Daten kann als eine Berechnungsmethode zur Überführung von symbolischer in geometrische Information definiert werden [MDB87]. Damit werden relevante Aspekte abstrakter Datenmengen in eine bildlich veranschaulichte Form gewandelt, um einen umfassenden Überblick zu gewähren, verborgene Datenzusammenhänge aufzuschlüsseln und sowohl Erkenntnis als auch Kommunikation zu erleichtern [SM00]. Shneiderman sieht die Essenz von Informationsvisualisierung in der Beschleunigung menschlichen Denkens durch Werkzeuge, die die menschliche Intelligenz verstärken. [Che04]

Speziell für die Entwicklung von maschinellen Umfelderkassungssystemen, die aufgrund zahlreicher Sensoren ein hohes Datenaufkommen aufweisen, spielen geeignete Darstellungsstrategien eine große Rolle. Für die Umsetzung eines Perzeptionssystems benötigt der Entwickler eine Veranschaulichung der Sensordaten in allen Ausprägungsstufen. Angefangen bei Rohdaten über aggregierte Daten bis hin zu fusionierten und interpretierten Daten sollten adäquate Visualisierungsmechanismen verfügbar sein. Diese erleichtern sowohl das Auffinden, als auch das Beseitigen von Programmfehlern und Falschparametrisierungen. Fehlerbehaftete Sensorkalibrierungen oder ungenaue zeitliche Synchronisierungen beispielsweise können gewissermaßen auf einen Blick erkannt und behoben werden.

Eine adäquate Sensordatenvisualisierung bietet des Weiteren eine stete und unmittelbare Performanzüberwachung des Systems. Zum Einen lassen sich aus dieser sinnvolle Verbesserungsmöglichkeiten und Erweiterungen erkennen und ableiten, zum Anderen kann diese der Ausgangspunkt für gänzlich neuartige Einblicke und Ideen sein.

Eine weitere wichtige Zielsetzung einer Datenveranschaulichung ist die Vereinfachung von Präsentation und Kommunikation. Eine bildliche Aufbereitung von Sensordaten kann Experten als Diskussionsgrundlage dienen, aber auch fachfremden Entwicklern eine gemeinsame Schnittstelle bereitstellen. Kommunikationslücken zwischen Sensorikern und automotiven HMI-Designern<sup>1</sup> können dadurch überbrückt, und der Ideen- und Informationsaustausch vereinfacht und gefördert werden.

In den folgenden Abschnitten werden Visualisierungskonzepte für das im Rahmen dieser Arbeit umgesetzte Perzeptionssystem vorgestellt und diskutiert. Dabei wird in Abschnitt 6.1 zunächst ein Überblick über den Datenfluss innerhalb eines automotiven Umfelderkassungssystems gegeben, woraus auf grundsätzliche Darstellungsrichtlinien basierende allgemeine Visualisierungsanforderungen abgeleitet

---

<sup>1</sup> HMI ist die Abkürzung für Human Machine Interface und bezeichnet die Schnittstelle zwischen Mensch und Maschine.

werden. Speziell für die Anforderungen während des Entwicklungsprozesses werden im nachfolgenden Abschnitt 6.2 Darstellungslösungen vorgestellt und exemplarisch illustriert. Abschnitt 6.3 widmet sich Darstellungsformen die sowohl die Systempräsentation vor unbedarften Dritten als auch die Kommunikation zwischen Entwicklern erleichtern. Abschließend wird in Abschnitt 6.4 als Beispiel für eine visualisierungsgetriebene Konzeptidee eine Fahrbahnerkennung mittels der Flottenbewegung vorausfahrender Fahrzeuge vorgestellt.

## 6.1 Grundlagen der Visualisierung

Visualisierungsmechanismen beziehen sich stets auf eine Datenbasis. Dieser Abschnitt gibt somit zunächst einen Überblick über den Datenbestand und -fluss eines Umfelderkassungssystems und behandelt anschließend grundlegende Visualisierungsrichtlinien, die beispielhaft auf die Belange von Perzeptionssystemen umgesetzt werden.

### 6.1.1 Daten und Informationsfluss

In einem Umfelderkassungssystem liegen Daten in unterschiedlichen Verarbeitungszuständen vor, welche durch geeignete Algorithmik ineinander überführt werden. Den Anfang des Informationsflusses (siehe Abbildung 6.1) bilden Rohdaten, die unmittelbar von der verwendeten Sensorik stammen. Durch Merkmalsextraktion und Zusammenführung werden diese Basisdaten in vorverarbeitete Daten umgewandelt. In einem letzten Schritt werden die Daten durch wissensbasierte und modellgetriebene Verfahren interpretiert. Dabei nimmt das Datenvolumen von links nach rechts ab, wobei zugleich der Abstraktionsgrad ansteigt. Die drei Ausprägungsstufen der Daten liegen in der Regel zeitgleich vor und werden meist zyklisch im Takt der Sensoraktualisierungsraten aufgefrischt. Eine geeignete dynamische Visualisierung sollte Schritt halten mit den Datenaktualisierungsraten der Sensorik und dem Entwickler Darstellungsmechanismen für alle Datenausprägungen und -überführungen zur Verfügung stellen.

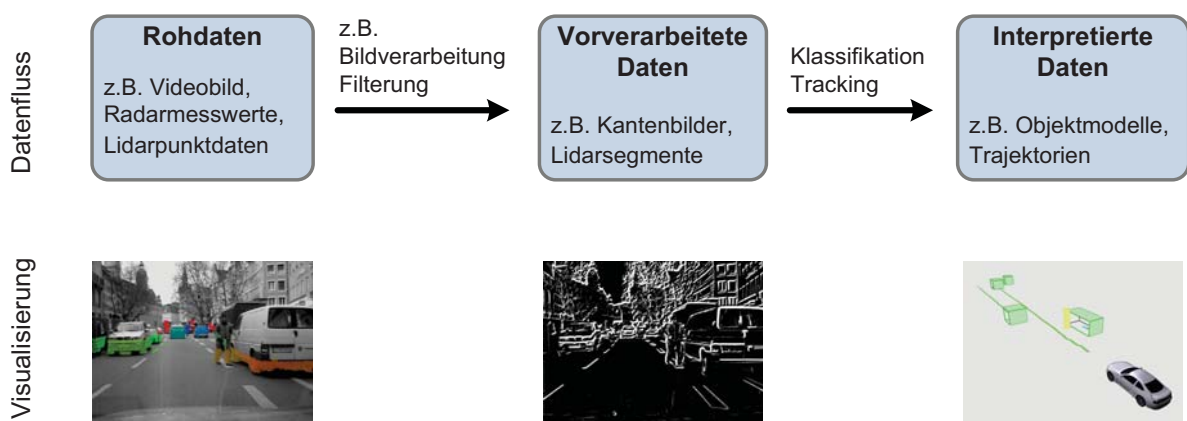


Abbildung 6.1: Datenausprägungen und Informationsfluss innerhalb eines Umfelderkassungssystems mit Visualisierungsbeispielen.

### 6.1.2 Designrichtlinien

Im vorangegangenen Abschnitt wurden die Daten einer Fahrumfelderfassung sowie deren Entstehung betrachtet. Neben der Analyse dieser Daten in Bezug auf Art und Struktur ist die Wahl geeigneter Darstellungsformen ein wichtiger Aspekt während des Designs einer Visualisierung wissenschaftlicher Daten. Schumann [SM00] beschreibt folgende drei grundsätzlichen Anforderungen, die jede Visualisierung erfüllen sollte:

**Expressivität:** Eine gute Visualisierung darf ausschließlich die in den Daten enthaltenen bzw. daraus abgeleiteten Informationen verwenden und muss sie möglichst unverfälscht wiedergeben. Auf diese Weise werden Fehlentscheidungen bzw. Falschinterpretationen vermieden, die darauf zurückzuführen sind, dass die Visualisierung dem Anwender Fakten suggeriert, die nicht in den Daten enthalten sind.

**Effektivität:** Um dem Anwender Informationen in intuitiver und ergonomischer Form zu präsentieren, versucht eine effektive Visualisierung die Möglichkeiten des Ausgabemediums sowie die Fähigkeiten der menschlichen Wahrnehmung optimal auszunützen (vgl. [Mac86]). Informationen können zum Beispiel durch Farben effektiv kodiert werden, da sie ähnlich zu Texturen in den frühen Phasen des visuellen Wahrnehmungsprozesses ausgewertet werden und somit zu einer geringen kognitiven Belastung des Anwenders führen [SM00].

**Angemessenheit:** Wenn Nutzen und Aufwand einer Visualisierung in einem vertretbaren Verhältnis zueinander stehen, spricht man von einer angemessenen Darstellung. Unter Aufwand sind in diesem Zusammenhang nicht nur die Kosten zur Erstellung der Visualisierung, sondern auch der zeitliche Aufwand zur Generierung der grafischen Darstellung zu verstehen. Eine zu komplexe und somit langsame Visualisierung ist nicht mehr ergonomisch und verliert folglich ihre Effektivität.

Eine expressive und effektive Darstellung ist eine notwendige, aber in vielen Fällen nicht ausreichende Voraussetzung für eine adäquate Visualisierung von wissenschaftlichen Daten. Oft können erst durch Manipulation bzw. Interaktion mit den Daten neue Erkenntnisse gewonnen oder Unstimmigkeiten geklärt werden. In der Theorie existieren unzählige Interaktionsmöglichkeiten für den Anwender. In der Praxis lassen sich hierfür allerdings grundlegende Designrichtlinien, die Shneidermann [Shn96] als „Information Seeking Mantra“ bezeichnet, finden. Im Folgenden werden diese Richtlinien am Beispiel der Visualisierung eines Fahrumfelderfassungssystems weiter erläutert:

**Kontext:** Der Anwender benötigt einen Überblick über die zu visualisierenden Daten, um grobe Zusammenhänge zu erfassen bzw. interessante Bereiche identifizieren zu können. Während der Anwender in Details vertieft ist, trägt eine Übersichtsdarstellung zusätzlich dazu bei, den Überblick zu behalten und Daten in ihrem globalen Kontext zu betrachten.

Für ein Fahrumfelderfassungssystem sind räumliche Daten aus dem Bereich von wenigen Zentimetern bis zu mehreren hundert Metern typisch. Vordefinierte Kamerapositionen wie beispielsweise die Vogelperspektive helfen dem Anwender, den räumlichen Überblick innerhalb der virtuellen Darstellung des Fahrumfeldes zu behalten. Darüber hinaus können ein fahrzeugfestes Koordinatensystem oder die Visualisierung der Position des Eigenfahrzeuges als zusätzliche Orientierungshilfe dienen.

Die Grundlage für die Visualisierung eines Fahrumfelderfassungssystems bilden Szenenkataloge, die sich aus hunderten Aufnahmen aus mehreren Jahren zusammensetzen können. Um in

diesen riesigen Zeitspannen den Überblick nicht zu verlieren, werden die einzelnen Aufnahmesequenzen nach Aufzeichnungsdatum und -zeit gruppiert bzw. sortiert. Je nach Anforderung können andere Sortierungs- bzw. Gruppierungskriterien wie z.B. Kreuzungsszenarien oder Aufzeichnungen mit Gegenverkehr von Interesse sein. Ein Navigationsbalken informiert über die aktuelle Position innerhalb einer Aufnahmesequenz.

**Fokus:** Nachdem sich der Anwender einen groben Überblick über die Daten verschafft hat, bietet eine vergrößerte Darstellung die Möglichkeit, interessante Daten bzw. Bereiche genauer zu untersuchen. Erfolgt die Vergrößerung in mehreren Stufen, verliert der Benutzer weder die Orientierung noch den Blick für Zusammenhänge.

In einer Fahrumfelderfassung lassen sich einige Vorgänge, wie z.B. die Vorverarbeitung oder bestimmte Messwerte, nur aus der Nähe oder unter einem bestimmten Blickwinkel untersuchen oder beobachten. Mit Hilfe einer virtuellen Kamera hat der Benutzer die Möglichkeit jede beliebige Position und jeden Blickwinkel innerhalb einer virtuellen dreidimensionalen Repräsentation des Fahrumfeldes einzunehmen (siehe Abschnitt. 6.2.5). Für die Untersuchung von stationären Objekten wie zum Beispiel Ampeln, Verkehrsschildern oder sonstiger Randbebauung kann der Wechsel zu einer ortsfesten Kameraeinstellung, die nicht an die Bewegung des Eigenfahrzeuges gekoppelt ist, hilfreich sein.

**Filterung:** Neben der vergrößerten Darstellung bietet das Ausblenden bzw. Herausfiltern von irrelevanten oder störenden Daten dem Anwender die Möglichkeit sich auf ausgewählte Daten zu konzentrieren oder bestimmte Aspekte hervorzuheben.

Insbesondere in einer multisensoriellen Fahrumfelderfassung kann der Blick für das Wesentliche von Messwerten verschiedener Sensoren oder Daten unterschiedlichen Abstraktionsgrades verdeckt oder beeinträchtigt sein. Durch das Ein- bzw. Ausblenden von Informationen, wie zum Beispiel bestimmte Sensoren oder Objektklassen, hat der Anwender die Kontrolle darüber, welche Daten im Augenblick relevant sind.

Es gibt allerdings Situationen, in denen weder das Ein- noch das Ausblenden ein zufriedenstellendes Ergebnis liefert. Beispielsweise können Messwerte im Augenblick von Interesse sein, obwohl sie andere Informationen verdecken und somit als störend empfunden werden. Dieser Zielkonflikt kann durch verschiedene Detail- oder Transluzenzstufen<sup>2</sup> entschärft werden. vermeiden)

**Optionale Details (details-on-demand):** Sobald die zu visualisierende Datenmenge durch Zoom- und Filteroperationen auf ein überschaubares Maß reduziert worden ist, sollten Detailinformationen zu einzelnen Datensätzen oder Gruppen bei Bedarf eingeblendet werden können.

In einer Fahrumfelderfassung gibt es viele Informationen wie z.B. Beschleunigungen, Gierraten oder Kovarianzmatrizen, die entweder selten benötigt oder schwer grafisch darzustellen sind. Um die Visualisierung nicht unnötig zu überfrachten, werden diese Daten erst bei Bedarf und meist in textueller Form präsentiert.

**In Beziehung setzen (relate):** Im Allgemeinen existieren Beziehungen und Zusammenhänge zwischen den zu visualisierenden Daten. Dem Anwender muss die Möglichkeit gegeben werden, bekannte Abhängigkeiten hervorzuheben und unbekannte Zusammenhänge neu zu erschließen.

---

<sup>2</sup> Transluzenz ist die partielle Lichtdurchlässigkeit eines Körpers.



Insbesondere bei entfernungsgebender Sensorik (Laserscanner oder Radar) ist es schwierig, einen Bezug zwischen einzelnen räumlichen Messwerten und realen Objekten herzustellen: Wurde ein Radarecho von einem Baum oder einem Menschen generiert? Diese Frage ist ohne zusätzliches Kontextwissen in Form weiterer Verarbeitungsergebnisse oder ergänzender Sensorik meistens nicht zu beantworten. Eine einfache Art den Anwender in diesem Prozess zu unterstützen, ist die Projektion von 3D Daten in ein Videobild. Abstrakte dreidimensionale Daten können so mit „Bildausschnitten“ attribuiert werden, was sowohl zu einem besseren Verständnis der Daten, als auch zu einer besseren Orientierung innerhalb der virtuellen Umgebung führt.

Jede Stufe im Datenflussdiagramm (siehe Abbildung 6.1) erhöht den Abstraktionsgrad der Daten und erzeugt somit neue Abhängigkeiten. So kann es zum Beispiel von Interesse sein, welche Rohdaten zu welchem Ergebnis der Vorverarbeitung gehören oder welche Messwerte in der Datenassoziation welchen Objekten zugeordnet wurden. Derartige Beziehungen können zum Beispiel durch Verbindungslinien zwischen den beteiligten Objekten hervorgehoben werden.

## 6.2 Visualisierung des Informationsflusses

Ein Softwareentwickler, der ein automotives Perzeptionssystem umsetzt, arbeitet an komplexen Algorithmen, die Sensordaten in weiterverarbeitete und diese in interpretierte Daten wandeln. Um solch eine anspruchsvolle Tätigkeit ausführen zu können, benötigt ein Experte Einblick in die Abläufe und Gespür für die Daten und Werkzeuge zur Fehlererkennung und -suche. Eine rein textuelle Darstellung der internen Mechanismen leistet diese Anforderungen nur unzulänglich und würde schnell die kognitiven Fähigkeiten eines Menschen übersteigen. Wesentlich effektiver und somit hilfreicher ist eine grafische Aufbereitung der Daten und Strukturen, wodurch der Entwickler Zusammenhänge schneller erschließen kann und ein Kontrollinstrument an die Hand bekommt. Durch eine geeignete visuelle Aufbereitung steht dem Entwickler des Weiteren eine Plattform zur Verfügung, die es erlaubt schwierige Sachverhalte zu veranschaulichen und einem fachfremden Dritten darzulegen.

In den folgenden Abschnitten werden die wesentlichen Daten und Strukturen des Informationsflusses eines Perzeptionssystems herausgegriffen und mögliche Darstellungskonzepte beschrieben. Zudem werden diese Konzepte, die zum Teil auf pragmatischen Visualisierungsansätzen aus der Sensor-domäne beruhen, einer systematischen Ordnung zugeführt. Im praktischen Umsetzungsprozess, der im Rahmen dieser Arbeit entstandenen Fahrumfelderfassung, haben sich die aufgezeigten Visualisierungsmethoden als unentbehrliches Entwicklungswerkzeug erwiesen. Eine fundierte Aussage bezüglich des quantitativen Nutzens kann erst durch aufwendige Benutzerstudien getroffen werden. Gleichwohl ist, durch die konsequente Berücksichtigung der in Abschnitt 6.1.2 eingeführten Designrichtlinien sowie durch die gewonnenen empirischen Werte, ein positiver Effekt auf den Entwicklungsprozess sehr wahrscheinlich.

### 6.2.1 Sensorverbauorte und -erfassungsbereiche

Die Sensoren eines multisensoriellen Perzeptionssystems sind in der Regel an unterschiedlichen Orten verbaut. Um die Messwerte korrekt in ein gemeinsames Koordinatensystem zu transformieren, ist eine exakte Kalibrierung notwendig, die den Verbauort, die Orientierung sowie interne Abbildungseigenschaften jedes Sensors bereitstellt. Die Güte einer durchgeführten Kalibrierung wird indirekt an den

Messwertüberlagerungen sichtbar. Sie kann aber auch direkt über die Visualisierung der Sensorverbauorte und -orientierungen plausibilisiert werden. Im Falle eines automotiven Umfelderkassungssystems wird das Sensorsystem direkt am Fahrzeug befestigt. Steht nun ein virtuelles dreidimensionales Modell des Fahrzeugs zur Verfügung, können die Verbauorte der Sensoren direkt an diesem Modell augmentiert werden, was einen unmittelbaren Rückschluss auf die Qualität der Kalibrierung zulässt (siehe Abbildung 6.2a).

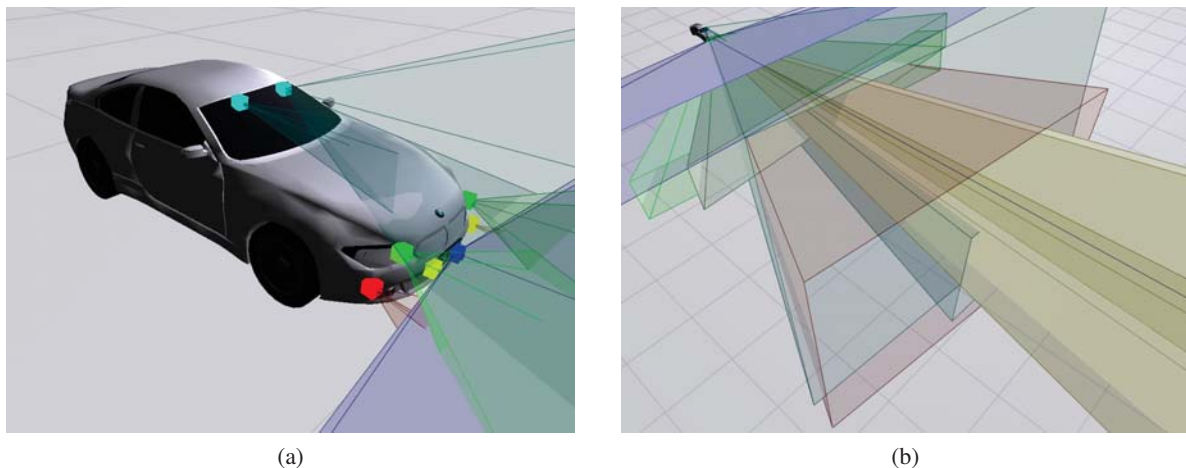


Abbildung 6.2: Visualisierte Sensorverbauorte und -erfassungsbereiche. Abbildung (a) zeigt die Verbauorte von acht Sensoren in einem Versuchsfahrzeug anhand unterschiedlich eingefärbter Würfel. Die jeweiligen approximierten Sichtbereiche sind in Abbildung (b) anhand von transluzenten Pyramiden dargestellt.

Von großer Relevanz für die sensorbasierten Anwendungen aber auch für die Algorithmik sind des Weiteren die räumlichen Erfassungsbereiche der einzelnen Sensoren und deren gegenseitige Überlagerungen. Ein Sensorerfassungsbereich definiert sich über Öffnungswinkel und Sensorreichweite. In der virtuellen Umgebung eignen sich somit entsprechend ausgerichtete Pyramiden oder Kegel, um die jeweiligen Sensorsichtbereiche approximiert darzustellen (siehe Abbildung 6.2b). Die exakte Nachbildung vieler Sensorsichtbereiche ist jedoch nicht trivial, da die Erfassungsbereiche nur ungenau bestimmbar sind, die Grenzen nicht scharf definiert sind oder die Struktur der Bereiche komplexe Formen annehmen. Der Sichtbereich von Mehrstrahlradaren beispielsweise setzt sich messprinzipbedingt aus einzelnen sich überlagernden Radarkeulen zusammen, wobei deren Ränder nicht scharf begrenzt sind.

### 6.2.2 Rohdaten und vorverarbeitete Sensordaten

Neben Rohdaten, die unmittelbar von der Sensorik geliefert werden, sollte der Entwickler visuellen Zugriff auf weiterverarbeitete Daten und möglichst alle relevanten Zwischenzustände haben. Grundsätzlich unterteilen sich Daten in Werte und Relationen. Zu Ersteren zählen beispielsweise bildgebende Datenquellen und metrische Messdaten. Letztere sind logische Verknüpfungen zwischen Daten wie beispielsweise räumliche Messwertüberlagerungen von unterschiedlichen Sensoren an identischen Objekten.



### Mehrdimensionale Messdaten

Zweidimensionale Messdaten werden häufig durch Streudiagramme dargestellt. Dabei werden Wertepaare, die durch grafische Symbole gekennzeichnet sind, in ein kartesisches Koordinatensystem eingetragen. Ein Beispiel dafür sind aus einer Vogelperspektive betrachtete Radardaten mit den Attributen Entfernung und Winkel, die in kartesische Koordinaten transformiert wurden (siehe Abbildung 6.3). Die unterschiedliche Farb- und Formgebung der Symbole kodiert die Sensorzugehörigkeit des Messwertes. Eine symbolische Darstellung des Versuchsfahrzeuges als Verbauort der Sensoren gibt einen Anhaltspunkt für die Orientierung des Streudiagramms. Das Streudiagramm erlaubt Distanzschätzungen und gibt Aufschluss über Datenhäufungen.

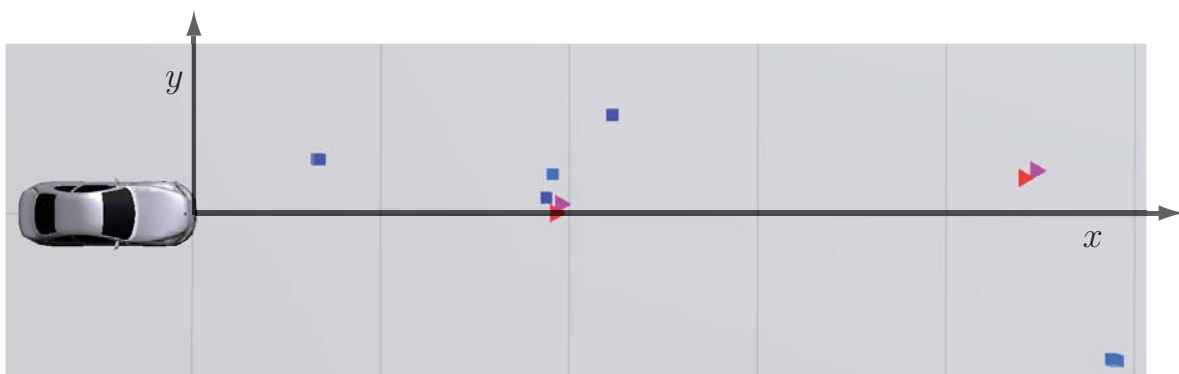


Abbildung 6.3: Zweidimensionale Entfernungsdaten von vier Radarsensoren aus einer Vogelperspektive betrachtet und als Streudiagramm dargestellt. Die unterschiedlichen Sensorzugehörigkeiten sind durch Farb- und Formkodierungen der Symbole gekennzeichnet.

Aktuelle Sensoren, die ein immer detaillierteres Abbild der Realität erstellen, generieren zunehmend dreidimensionale oder sogar mehrdimensionale Messwerte. Eine mögliche Lösung [Spe00], diese Daten zu visualisieren, sind multiple Streudiagramme, die jeweils zwei Dimensionen in einem Standardstreudiagramm darstellen. Damit ergeben sich für  $r$  Dimensionen insgesamt  $r(r-1)$  Streudiagramme. Dadurch dass jeder Messwert in jedem Streudiagramm existiert, ergeben sich  $r$  analoge Repräsentationen des selben Datenpunktes. Dies erhöht die kognitive Last des Benutzers und verringert zudem den freien Platz für Beschriftungen.

Eine isolierte dreidimensionale Darstellung von räumlichen Sensordaten bietet eine platzsparende Datenvisualisierung, da ein einzelnes Visualisierungsdiagramm ausreichend ist. Zudem ist die Verknüpfung zwischen Realität und deren Abbild für den Betrachter meist intuitiver. Im Gegensatz zu multiplen Streudiagrammen sind mit der Projektion von dreidimensionalen Umgebungen Verdeckungseffekte und Zuordnungsschwierigkeiten verbunden. Diese inhärenten Nachteile einer räumlichen Datenvisualisierung können durch eine interaktive Raumnavigation (siehe auch Abschnitt 6.2.5) aufgelöst werden [Shn03]. Höherwertige Dimensionen von räumlichen Messwerten lassen sich über Farbe, Gestalt und Größe der Datenpunkte kodieren (siehe Abbildungen 6.4 und 6.5).

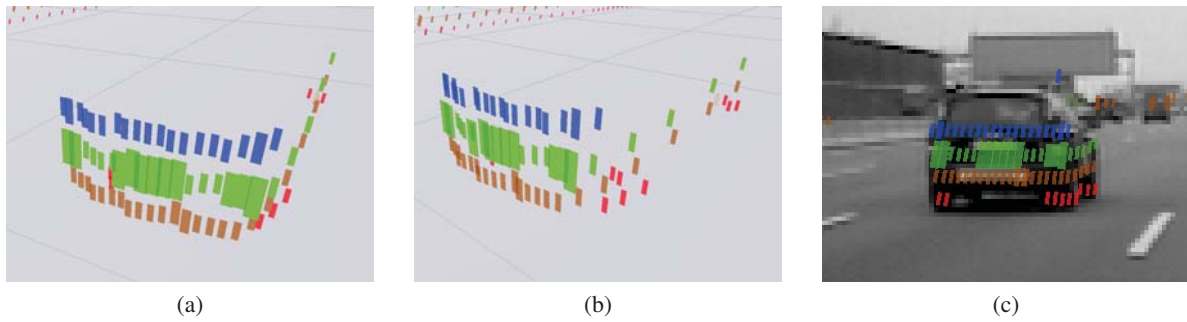


Abbildung 6.4: Multidimensionale Laserscannerdaten unter zwei verschiedenen räumlichen Ansichten (a) und (b) sowie als Projektion in das Videobild (c). Neben der Position im Raum wird durch die Visualisierung die Reflektionsstärke über die Größe, die Abtastebene über die Farbe, und der Kanal über den Grad der Transluzenz dargestellt. Die typische L-Form des Fahrzeugs wird ebenso sichtbar wie die starken Reflektionen des Nummernschildes und der Rückstrahler.

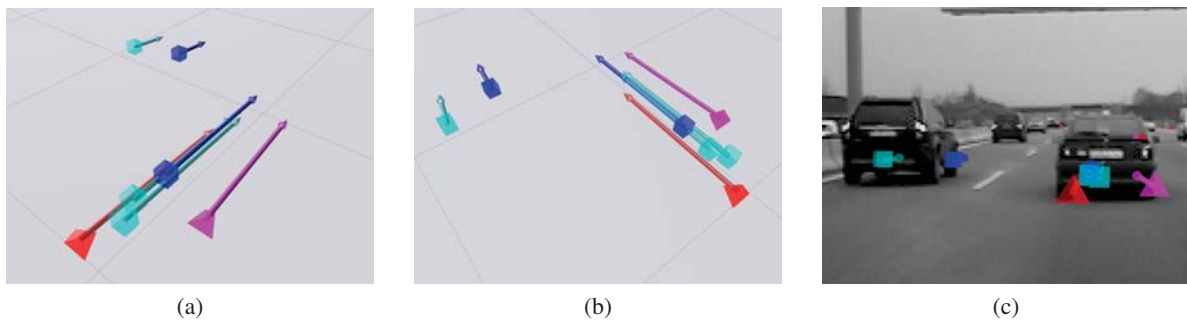


Abbildung 6.5: Multidimensionale Radardaten (a) und (b) sowie deren Projektion in das korrespondierende Videobild (c). Neben der Position des Radarechos im Raum wird die, über den Dopplereffekt gemessene, radiale Relativgeschwindigkeit über die Lage und Länge eines Pfeils sichtbar gemacht.

### Dynamische Bilddaten

Ein digitalisiertes Bild wird in der Regel als Matrix abgelegt, deren Elemente diskrete Werte (z.B. schwarz-weiß Kamera), dreidimensionale oder mehrdimensionale Vektoren (z.B. RGB Farbkamera bzw. hyperspektrale Sensoren) oder kontinuierliche Werte (z.B. Spektrometer) annehmen können [FP03]. Gewöhnlich werden diese Matrixeinträge direkt auf die Bildschirmspeichermatrix abgebildet und somit dargestellt. Die Abbildungsfunktion ist abhängig von der bildgebenden Sensorik und der Anwendung. Während z.B. RGB-Farbkamerabilder unverändert übertragen und dargestellt werden können, muss bei mehrdimensionalen Bilddaten eine Beschneidung erfolgen oder eine Zuordnungsvorschrift (z.B. Falschfarbdarstellung) angewendet werden.

Von der dreidimensionalen Welt können zweidimensionale Bilder erzeugt werden, indem sie durch ein optisches System auf eine Bildebene projiziert wird. Perspektivische Kameras setzen dabei diese Abbildung um, wobei sich deren Projektionseigenschaften über mathematische Modelle beschreiben

lassen. Eine detaillierte Einführung in Sensormodelle findet sich in Abschnitt 2.1. Hier sei lediglich das grobe Prinzip erklärt: Ein dreidimensionaler Punkt  $\mathbf{x}^W$ , der in einem Koordinatensystem  $W$  vorliegt, wird zunächst über eine Transformationsmatrix  $\mathbf{T}_C^W \in \mathbb{R}^{4 \times 4}$  in einen Punkt  $\tilde{\mathbf{x}}^C$  in Kamerakoordinaten  $C$  überführt und anschließend über eine Projektionsmatrix  $\mathbf{K} \in \mathbb{R}^{4 \times 4}$  auf einen zweidimensionalen Punkt  $\mathbf{x}$  in der Bildebene projiziert:

$$\mathbf{x} = \mathbf{K} \mathbf{T}_C^W \mathbf{x}^W \quad (6.1)$$

Über diese Rechenvorschrift können Messwerte von entfernungsgebenden Sensoren, aber auch interpretierte Daten, direkt in das Kamerabild projiziert und damit Bildinhalte überlagert werden. Dadurch wird es dem Betrachter möglich, Zusammenhänge zwischen Daten unterschiedlicher Sensoren herzustellen und auf den Ursprung der Messungen zu schließen (vgl. Abbildungen 6.4c, 6.5c und 6.8b). Eine weitere Möglichkeit, Bilddaten in Verbindung mit anderen Sensordaten darzustellen, ist eine texturbasierte Anzeige in einer virtuellen dreidimensionalen Umgebung. Dazu werden die Bilder über ein Musterabbildungsverfahren<sup>3</sup> als Textur auf eine virtuelle Ebene im Raum gelegt [Fol95]. Um ein ansprechendes Ergebnis zu erhalten, kann die Darstellung durch ein geeignetes Interpolationsverfahren geglättet werden (vgl. Abbildung 6.6a). Liegt diese Ebene nun parallel zur Bildebene der entsprechenden Kamera und wird zusätzlich das Modell der virtuellen Szenenkamera (siehe auch Abschnitt 6.2.5) an das der Kamera angeglichen, sind analog zur obigen Bildprojektion passgenaue Überlagerung erzielbar (vgl. Abbildung 6.6b). Mit Hilfe dieser Vorgehensweise können die Bilder mehrerer Kameras zur Überlagerung gebracht werden [HZ03]. Über den Grad der Transluzenz und über die Ausdehnung der Ebenen wird eine Aufmerksamkeitssteuerung realisierbar.

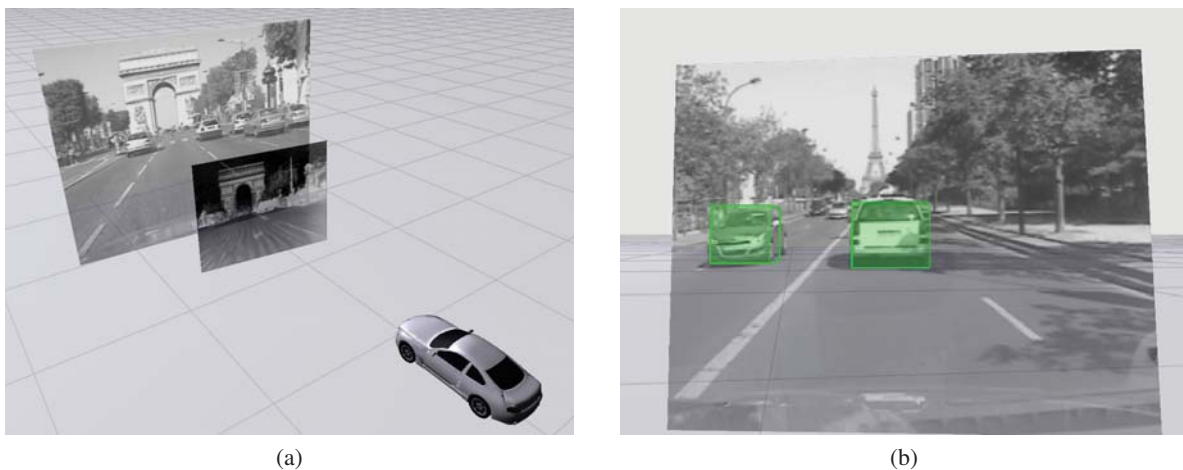


Abbildung 6.6: Visualisierung von Bilddaten in einer dreidimensionalen Umgebung. Abbildung (a) zeigt das Umgebungsbild von einer Videokamera und einer Wärmebildkamera in einer virtuellen Umgebung als Texturen auf zwei Ebenen parallel zu den jeweiligen Bildebenen der Kameramodelle. In Abbildung (b) wurde die Lage der virtuellen Szenenkamera mit der Lage der realen Videokamera in Einklang gebracht. Dadurch ist eine passgenaue Projektion der erkannten Objekte in das Bild möglich.

<sup>3</sup> engl. texture mapping

### Vorverarbeitete Daten und Relationen

Die Datenvorverarbeitung bündelt Daten mit ähnlichen Eigenschaften (Aggregation), extrahiert durch algorithmische Umformungen markante Merkmale (Merkmalsextraktion) oder filtert die Datenströme anhand ausgewählter Attribute (Filterung). Alle relevanten Vorverarbeitungsergebnisse sollten sichtbar gemacht werden.

Verbindungslinien sind die einfachste Art, Beziehungen bzw. Relationen von aggregierten Datenverbänden darzustellen. Über die Farbgebung und die Linienstärke kann die Relation mit weiteren Eigenschaften attribuiert werden. Die Abbildung 6.7a zeigt das Ergebnis einer Aggregation von Laserscannerdaten, die mittels Linienzügen visualisiert wird. Fehler in der Segmentierung in Form von unzugeordneten Rohdaten werden unmittelbar sichtbar.

Die Visualisierung von Filterungsvorgängen lässt sich über eine vollständige Darstellungsunterdrückung oder durch eine adäquate Transluzenz- bzw. Farbgebung erreichen. In Abbildung 6.7b werden Daten, die eine bestimmte Eigenschaft nicht erfüllen, stark durchscheinend gezeichnet und damit aus dem Wahrnehmungsfokus entfernt. Gleichzeitig sind sie noch sichtbar, um Rückschlüsse auf die Arbeitsweise der Filterungsroutine zuzulassen.

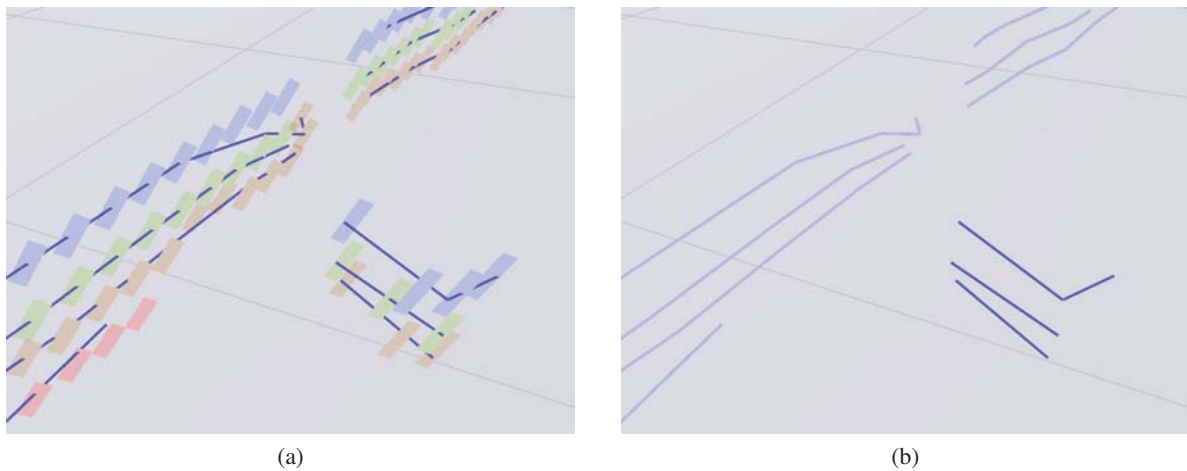


Abbildung 6.7: Abbildung (a) zeigt durch Linienzüge visualisierte Relationen zwischen Rohdaten. In Abbildung (b) wird die Transluzenz von Linienzügen erhöht, die eine Länge von 4.0 m überschreiten. Dadurch wird ein möglicher Filtervorgang zwischen Fahrzeughypothesen und Randbebauung veranschaulicht.

### 6.2.3 Interpretierte Daten

Am Ende des Datenflusses einer Fahrumfelderfassung liegen interpretierte Daten in Form von Objektlisten vor. Dazu zählen alle für die nachfolgende Applikationsschicht relevanten Objekte wie zum Beispiel Fahrbahnmarkierungen und -begrenzungen, Lichtsignalanlagen und Verkehrsteilnehmer. Jedes dieser Objekte wird über Attribute wie Objekttyp, Ausdehnung, Orientierung, Konfidenzen, zeitlicher Verlauf, Position, Geschwindigkeit und Beschleunigung definiert, die zu einem Zustandsvektor

zusammengefasst werden. Aufgabe der Visualisierung ist es, diese Objekte in Verbindung mit ihrem Zustand darzustellen.

Die Veranschaulichung von Position, Ausdehnung, Orientierung und Typ der Objekte kann dabei über die Platzierung von primitiven geometrischen Modellen erreicht werden. Fahrzeuge können durch Quader, Fußgänger und Lichtsignalanlagen durch Zylinder und Fahrbahnmarkierungen durch Linien- oder Bezierkurven repräsentiert werden (siehe auch Abbildung 6.8). Sofern eine detailliertere Perception notwendig wird, wie beispielsweise der Radstand von Fahrzeugen, die Fußstellung von Fußgängern oder eine exakte Geometrie der Signalanlagen, kann eine sukzessive geometrische Verfeinerung dieser groben Basismodelle vorgenommen werden. Weitere Zustandsinformationen der Objekte, wie Geschwindigkeiten, aber auch Beschleunigungswerte, können anhand von Pfeilen wiedergegeben werden. Mittels der Zeichenfarbe und -transluzenz lassen sich sowohl kontinuierliche als auch diskrete Zustandsattribute wie beispielsweise Konfidenz und Objekttyp sichtbar machen.

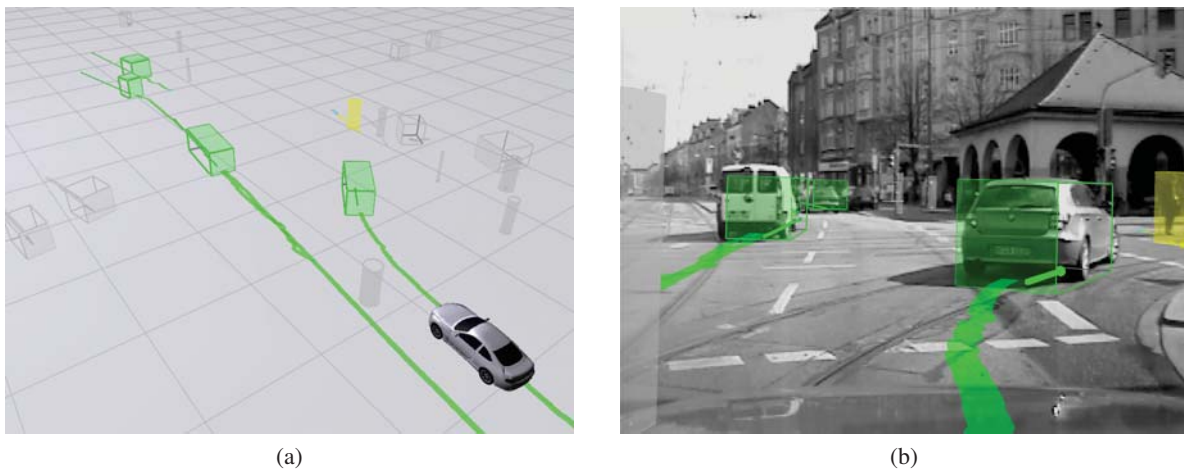


Abbildung 6.8: Visualisierte Objektdaten, deren Trajektorien und Bewegungsrichtungen. Abbildung (a) zeigt eine dreidimensionale Schrägansicht auf eine Straßenszene mit eingeblendeten Objekten. Grau eingefärbte Objekte sind Objekthypothesen, grüne Quader stellen erkannte Fahrzeuge dar und gelbe Zylinder Fußgänger. Das eigene Fahrzeug ist durch ein Modell visualisiert. Abbildung (b) zeigt die Projektion der Szenerie in das Bild einer Videokamera.

#### 6.2.4 Zeitlicher Verlauf der Messungen

Erst vielfache Messungen pro Sekunde machen Sensoren tauglich für eine automotiv Umfelderfassung. Diese zeitliche Komponente oder vierte Dimension eines Messwertes kann in der Regel durch eine mit der Messung Schritt haltende Visualisierung dargestellt werden. Eine benutzergetriebene Unterbrechung des zeitlichen Ablaufs durch Einzelschritt-, Zeitlupen- und Pausenmodus, gewährt dem Anwender die volle Kontrolle über die Datensequenzen. Der Benutzer kann durch diese Vorgehensweise jedoch lediglich eine Momentaufnahme der Datenaufnahme einsehen, was gegebenenfalls von Nachteil ist.



Eine weitere Möglichkeit, den zeitlichen Verlauf von Messwerten darzustellen, ist in Abbildung 6.9a zu sehen. Vergleichbar zu den Trajektorienverläufen in Abbildung 6.8 werden zeitgleich mit den aktuellen Daten alte Daten an ihren jeweiligen absoluten Messorten angezeigt. Wenn neuere von älteren Messwerten unterscheidbar sein sollen, kann eine adäquate Farb- oder Transluzenzgebung (siehe auch Abschnitt 6.2.2) der Messwerte von Nutzen sein. Diese verschränkte Darstellung ist durchführbar, solange sich die Messwerte nicht gegenseitig überlagern und damit eine zeitliche Differenzierung erschweren. Vor Allem bei Bilddaten ist es dann sinnvoller, für jeden Zeitschritt eine gesonderte Anzeige zu verwenden.

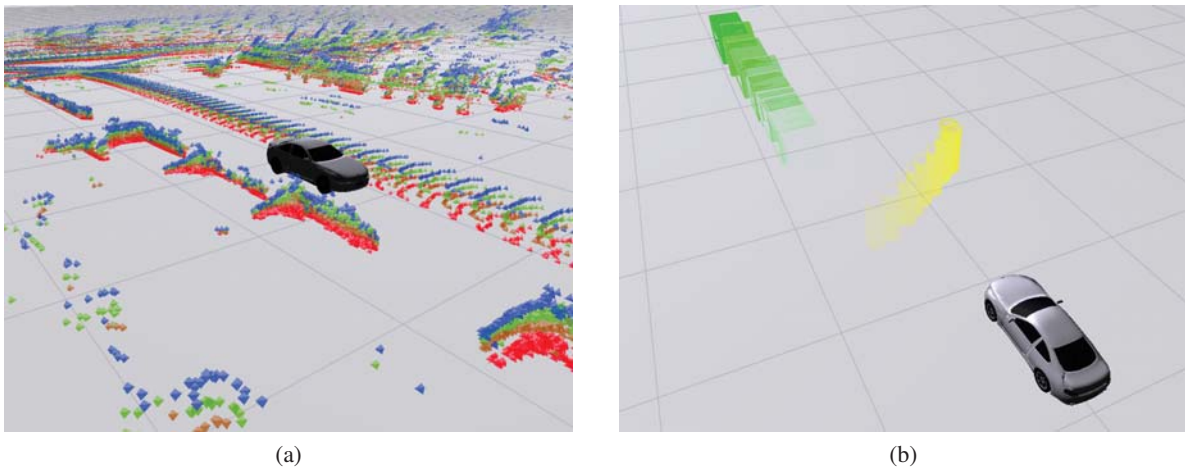


Abbildung 6.9: Visualisierung von zeitlichen Messwertverläufen. Abbildung (a) zeigt Laserscannerdaten in einem Weltkoordinatensystem, die während eines Abbiegevorgangs an einer Kreuzung aufgezeichnet wurden. Am Straßenrand sind parkende Fahrzeuge zu erkennen, und auch die Datenspur des vorausfahrenden Fahrzeugs ist sichtbar. Abbildung (b) zeigt Objektdaten von einem Fahrzeug (grün) und einem Fußgänger (gelb). Weiter in der Zeit zurückliegende Objektpositionen werden dabei durchscheinender gezeichnet.

### 6.2.5 Navigation und Interaktion

Mehrdimensionale Messwerte sowie weiterverarbeitete Daten werden, wie in Abschnitt 6.2.2 beschrieben, in einer dreidimensionalen Darstellung (virtuellen Welt) visualisiert. Hierbei können bedingt durch den Informationsverlust der Projektion von dreidimensionalen Daten auf einen zweidimensionalen Bildschirm Verdeckungen bzw. Mehrdeutigkeiten entstehen. Eine Möglichkeit, dieses Problem zu lösen, sind drei vordefinierte orthogonale Projektionen der Szene, wie sie häufig in Programmen zur 3D Modellierung verwendet werden. Diese Darstellungsart ist besonders dann vorteilhaft, wenn Objekte sehr genau manipuliert werden müssen, da keine perspektivischen Verzerrungen auftreten. Allerdings muss der Anwender gleichzeitig drei unterschiedliche Repräsentationen von einem Objekt verarbeiten, was sich in einer erhöhten kognitiven Belastung niederschlägt. [Spe00]

Eine weitere Möglichkeit Verdeckungen und Mehrdeutigkeiten aufzulösen, ist die Navigation, da der Anwender selbst bestimmt unter welchem Blickwinkel Daten untersucht werden sollen<sup>4</sup>. Insbesondere in einer Fahrumfelderfassung, in der die Manipulation von Objekten eine untergeordnete Rolle spielt, stellt eine einzige dreidimensionale Ansicht der virtuellen Welt eine platzsparende, und den Anwender kognitiv wenig belastende, Alternative zu multiplen Streudiagrammen dar. Allerdings muss hierfür eine intuitive und effektive Navigationsmöglichkeit gefunden werden, die im Folgenden näher beschrieben wird.

## Navigation

Je nach verfügbaren Eingabegeräten bzw. Aufgabenstellung existieren mehrere Möglichkeiten in einer virtuellen Welt zu navigieren: Ware und Osborne [WO90] benötigen zum Beispiel für die von ihnen beschriebenen Metaphern der „Szene in der Hand“, des „Augapfels in der Hand“ oder des „fliegenden Fahrzeugs“ ein Eingabegerät mit sechs Freiheitsgraden. Hierbei steuert das Eingabegerät direkt eine virtuelle Kamera, manipuliert die gesamte Szene bei fixiertem Blickwinkel oder kontrolliert die Bewegung eines virtuellen Fahrzeuges. Während der Entwicklung einer Fahrumfelderfassung sind allerdings konventionelle Eingabegeräte wie Maus oder Tastatur vorzuziehen, da sie weit verbreitet und sowohl am Arbeitsplatz, als auch im Fahrzeug problemlos verwendet werden können. Für die Maus wird von Nielson [NO87] eine Technik beschrieben, die zwei- in dreidimensionale Bewegungen überführt, indem der Bewegungsvektor der Maus mit den Projektionen der Koordinatenachsen verglichen wird, und die resultierende Bewegung entlang der Achsen mit der größten Korrelation durchgeführt wird. Diese Art der Navigation ist nicht sehr intuitiv, da bestimmte Bewegungsrichtungen nur über Umwege realisiert werden können. Im Arcball Modell von Shoemaker [Sho92] werden zweidimensionale Mausbewegungen als Kreisbögen auf einer Kugeloberfläche interpretiert. Dadurch können beliebige dreidimensionale rotatorische Eingaben vorgenommen werden. Zu Navigationszwecken im automotiven Einsatz ist dieses Interaktionsmodell zu mächtig, da nicht alle rotatorischen Freiheitsgrade benötigt werden. Eine ähnliche Navigationsmöglichkeit, die zudem translatorische Eingaben ermöglicht, ist die Metapher eines Orbits von Tan [TRC01] bzw. Phillips und Badler [PB88], die wie folgt an die Bedürfnisse einer Fahrumfelderfassung angepasst wurde (vgl. auch Abbildung 6.10):

Der Betrachter befindet sich in dem Punkt  $P$  auf der Oberfläche einer Kugel mit Radius  $r$  und blickt in Richtung des Kugelmittelpunktes  $C$ , der sich immer auf der  $xy$ -Ebene des Weltkoordinatensystems befindet (siehe Abschnitt 2.4.2). Ausgehend von dem Punkt  $P$  hat der Betrachter nur zwei Möglichkeiten seine Position auf der Oberfläche zu verändern: Er kann sich entweder auf einer Kreisbahn bewegen, die parallel zur  $xy$ -Ebene liegt und durch  $P$  verläuft, oder er läuft die Umlaufbahn entlang, die orthogonal zur  $xy$ -Ebene ist und durch  $P$  verläuft. Dadurch können Objekte, die sich innerhalb der Kugel befinden, von allen Seiten begutachtet werden. Beide Varianten entsprechen in Kugelkoordinaten<sup>5</sup> einer Änderung des Azimutwinkels  $\phi$  bzw. des Polarwinkels  $\theta$ . Um ein „Überschlagen“ der Kamera an den Polen zu vermeiden, ist  $\theta$  auf das Intervall  $[-90^\circ; +90^\circ]$  beschränkt. Sollen Details vergrößert betrachtet werden, muss der Benutzer bedingt durch sein starres Sichtfeld näher an das Objekt gelangen. Dies lässt sich dadurch erreichen, dass der Radius  $r$  und somit die Höhe der Umlaufbahn variiert wird. Letztendlich erhält der Betrachter die Möglichkeit, sich frei in der virtuellen Welt zu bewegen, indem der Kugelmittelpunkt translatiert wird.

<sup>4</sup> ‘for 3D to be useful, you’ve got to be able to move it’ [Spe00]

<sup>5</sup> In Kugelkoordinaten ist ein Punkt  $P$  über seinen Abstand  $r$  zum Kugelmittelpunkt  $C$ , den Azimutwinkel  $\phi$  sowie den Polarwinkel  $\theta$  eindeutig bestimmt.



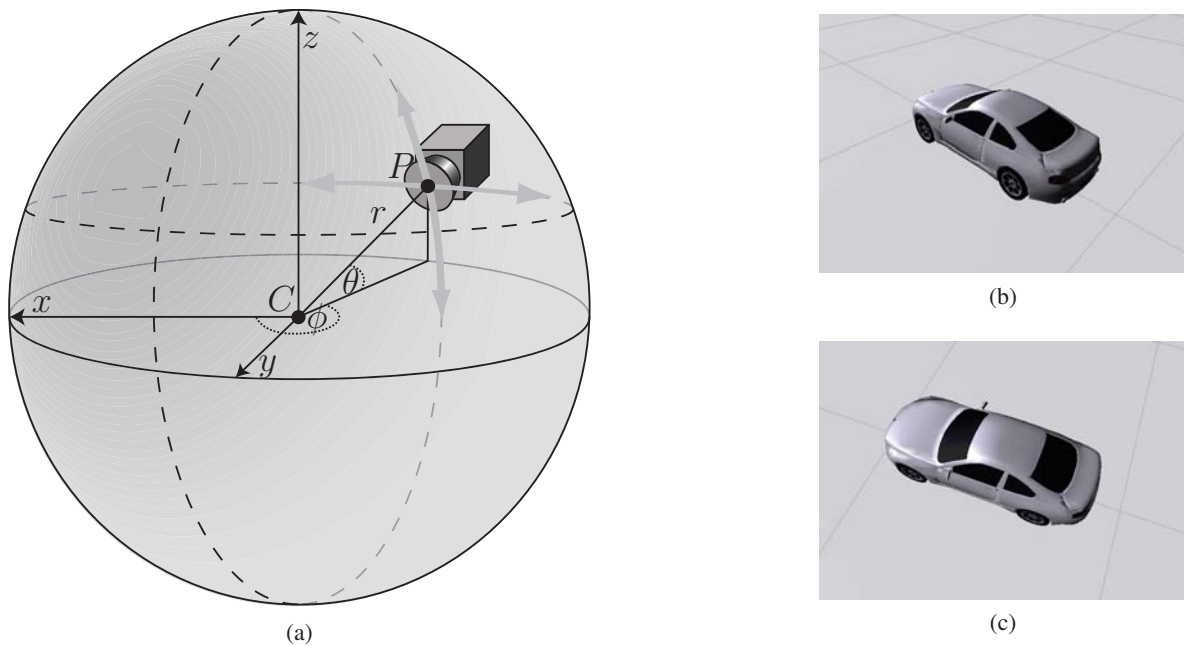


Abbildung 6.10: Orbit-Navigation innerhalb einer virtuellen Umgebung: Der Radius  $r$ , der Azimutwinkel  $\phi$  sowie der Polarwinkel  $\theta$  bestimmen eindeutig die Position  $P$  einer virtuellen Kamera auf einer Kugeloberfläche (a). Bewegt sich die Kamera auf einer Umlaufbahn, können Objekte, die sich innerhalb der Kugel befinden, von allen Seiten betrachtet werden (siehe (b) und (c)).

Im Folgenden muss geklärt werden, wie konventionelle Eingabemedien wie Maus oder Tastatur auf die oben beschriebene Navigationsmetapher abgebildet werden können. Da diese weniger Freiheitsgrade als die Orbit-Metapher besitzen, können nicht alle Parameter gleichzeitig verändert werden. Deshalb wird die Navigation in folgende drei Abschnitte unterteilt, die einzeln gesteuert werden:

**Translation** Um eine intuitive Bedienung zu gewährleisten, erfolgt die Translation in der egozentrischen Sicht des Betrachters. Eine Bewegung der Maus um  $\Delta u$  nach links bzw. rechts hat eine laterale Verschiebung der Szene aus Sicht des Anwenders zur Folge. Analog wird eine Vor bzw. Zurück-Bewegung um  $\Delta v$  auf eine longitudinale Änderung abgebildet. Die Granularität der Translation lässt sich über den Skalierungsfaktor  $\tau$  steuern. Schließlich berechnet sich der neue Kugelmittelpunkt  $\tilde{\mathbf{c}}$  aus dem Azimutwinkel  $\phi$ ,  $\Delta u$ ,  $\Delta v$ ,  $\tau$  sowie dem alten Kugelmittelpunkt  $\mathbf{c}$  wie folgt:

$$\tilde{\mathbf{c}} = \mathbf{c} + \begin{bmatrix} \cos \phi & -\sin \phi & 0 & 0 \\ \sin \phi & \cos \phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tau \Delta u + \begin{bmatrix} \cos \phi & -\sin \phi & 0 & 0 \\ \sin \phi & \cos \phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \tau \Delta v \quad (6.2)$$

**Änderung des Blickwinkels** Ausgehend vom Punkt  $P$  hat der Betrachter nur zwei Möglichkeiten, seine Position auf der Kugeloberfläche und damit seinen Blickwinkel zu verändern: Er kann sich entweder auf einer Umlaufbahn bewegen, die parallel zur  $xy$ -Ebene liegt und durch  $P$  verläuft, oder er verwendet die Umlaufbahn, die orthogonal zur  $xy$ -Ebene ist und ebenfalls durch  $P$  verläuft. Jede der beiden Varianten wird eigenständig mit der Maus kontrolliert. Im ersten Fall ergibt sich der neue Azimutwinkel  $\tilde{\phi}$  aus der Links- bzw. Rechtsbewegung der Maus um  $\Delta u$  sowie dem alten Azimutwinkel  $\phi$  wie folgt:

$$\tilde{\phi} = \phi + \frac{\pi}{180} \Delta u \quad (6.3)$$

Im zweiten Fall berechnet sich der neue Polarwinkel  $\tilde{\theta}$  aus der Bewegung der Maus um  $\Delta u$  und dem alten Polarwinkel  $\theta$ :

$$\tilde{\theta} = \theta + \frac{\pi}{180} \Delta u \quad (6.4)$$

**Zoom** In vielen Anwendungen wird das „Scroll-Rad“ der Maus dazu verwendet, um den Zoomfaktor zu bestimmen. Analog wird der Radius  $r$  direkt durch das „Scroll-Rad“ variiert.

Die Tastatur lässt sich analog zur Maus auf die Orbitnavigation abbilden, da Bewegungen der Maus mit unterschiedlichen Tasten sehr gut simuliert werden können. Der neue Punkt  $P$  auf der Kugeloberfläche berechnet sich als

$$\mathbf{p} = \mathbf{c} + \begin{pmatrix} r \sin \theta \cos \phi \\ r \sin \theta \sin \phi \\ r \cos \theta \\ 1 \end{pmatrix} \quad (6.5)$$

Das Referenzkoordinatensystem, in dem die Navigation durchgeführt wird kann frei gewählt werden. Mit einer virtuellen Kamera in Weltkoordinaten (siehe auch Abschnitt 2.4.2) beispielsweise lassen sich stationäre Objekte wie Ampeln, Verkehrsschilder oder sonstige Randbebauung gut beobachten. Bewegte Objekte hingegen verschwinden sehr schnell aus dem Blickfeld des Betrachters. Um dennoch den Fokus nicht zu verlieren, müsste die Position der Kamera in Weltkoordinaten ständig manuell korrigiert werden. In dieser Situation ist die Koppelung der Position und Orientierung der virtuellen Kamera an andere Objekte, wie zum Beispiel dem Eigenfahrzeug, deutlich effektiver. In diesem Betrachtungsmodus ist die Navigation auf das Koordinatensystem des Zielobjektes bezogen. Dadurch befindet sich das fokussierte Objekt fortwährend im Blickfeld der virtuellen Kamera, da jede Positions- und Orientierungsänderung des Zielobjektes automatisch übernommen wird.

## Interaktion

Sobald die zu visualisierende Datenmenge durch Navigation und Zoom sowie optionaler Filteroperationen auf ein überschaubares Maß reduziert worden ist, benötigt der Anwender weitere Interaktionsmöglichkeiten, um Beziehungen zwischen Daten zu definieren oder Detailinformationen zu einzelnen Datensätzen zu erhalten. Der Anwender hat, wie in Abbildung 6.11a dargestellt, die Möglichkeit die Entfernung zwischen ausgewählten Objekten oder Messwerten zu ermitteln. Diese Art der Beziehung

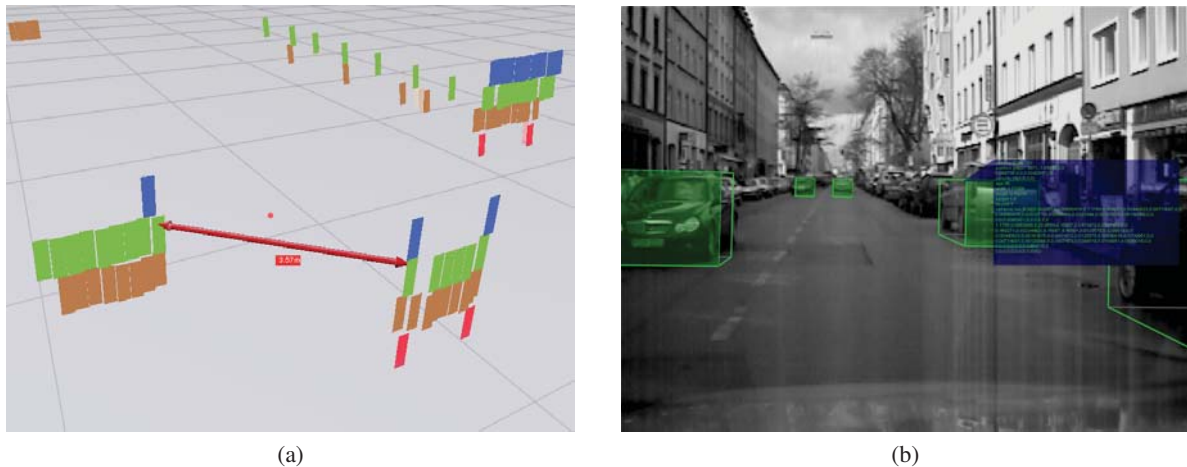


Abbildung 6.11: Unterschiedliche Interaktionsmöglichkeiten: (a) Bestimmung der Entfernung zwischen zwei Messwerten des Laserscanners. (b) Verfügbare Detailinformation zu einem ausgewählten Objekt werden in textueller Form angezeigt.

ist beispielsweise hilfreich, um Fehlfunktionen in der Vorverarbeitung oder Datenassoziation zu untersuchen.

In einer Fahrumfelderfassung sind einzelne Objekte wie Fahrzeuge oder Fußgänger mit zusätzlichen Informationen wie z.B. Klassifikationsgüten, Merkmalsvektoren, Beschleunigungen oder Kovarianzmatrizen attribuiert, die selten benötigt oder schwer darzustellen sind. Um die Visualisierung nicht unnötig zu überfrachten, kann der Anwender entscheiden, zu welchem Objekt diese Daten in textueller Form präsentiert werden sollen (siehe Abbildung 6.11b).

### 6.2.6 Werkzeuge der Implementierung

Eine aufwändige Visualisierung von mehrdimensionalen Daten mit Transluzenzdarstellungen, Kantenglättung, Lichtberechnungen und Interaktionsmöglichkeiten benötigt ein hohes Maß an Rechenleistung. Um dennoch den Echtzeitanforderungen gerecht zu werden, ist es sinnvoll alle Darstellungsrechnungen auf einem dedizierten Grafikprozessor auszuführen. Damit kümmert sich ein spezialisierter Rechenkern um die Visualisierung, und der Hauptprozessor steht vollständig für die Sensordatenverarbeitung zur Verfügung. Die Spezifikation OpenGL [SWND05] bietet hierfür eine plattform- und programmiersprachenunabhängige Schnittstelle zur Anwendungsprogrammierung von 3D-Grafik.

## 6.3 Präsentation und Kommunikation

Die Präsentation und Kommunikation von erzielten Ergebnissen ist eine wichtige Visualisierungsanforderung. Sie schafft eine Diskussionsgrundlage für Experten und veranschaulicht Fakten und Zusammenhänge, so dass sie auch für fachfremde Entwickler leicht verständlich werden. Kommunikationslücken zwischen Sensorikern und automotiven HMI-Designern können dadurch überbrückt, und

der Ideen- und Informationsaustausch vereinfacht bzw. gefördert werden. Selbst Laien können durch bildstarke Darstellungen einen Einblick in die Arbeitsweise einer Fahrumfelderfassung gewinnen.

Im Folgenden werden unterschiedliche Präsentationsmedien und -szenarien beschrieben, um Informationen einer Fahrumfelderfassung zu visualisieren. Eine Live-Darstellung des Perzeptionssystems während des Fahrbetriebs stellt dabei andere Anforderungen an die Wahl des Ausgabemediums und des Visualisierungskonzeptes als eine expertengeführte Diskussionsrunde über interne Systemabläufe. Zunächst wird auf die konventionelle Visualisierungsumsetzung im Fahrzeug eingegangen, die eine Systemüberwachung zur Laufzeit ermöglicht. Anschließend werden zukunftsweisende Darstellungskonzepte vorgestellt, die mittels erweiterter Realität gänzlich neuartige Einblicke in die Daten erlauben.

### 6.3.1 Konventionelle Ausgabemedien

Die Erkennungsgüte und Performanz von automotiven Umfelderkennungssystemen kann besonders eindrucksvoll im laufenden Betrieb im Fahrzeug vermittelt werden. Voraussetzung für eine intuitive Wahrnehmung ist eine echtzeitfähige Visualisierung, die dem Betrachter den aktuellen Zustand des Perzeptionssystems widerspiegelt. Dabei sollte die Art und Weise der Darstellung den Betrachter in der Zuordnung von maschinell erkannten Objekten (virtuelle Objekte) zu realen Umgebungsobjekten unterstützen. Um diese Anforderungen umzusetzen, haben wir in unserem Versuchsträger die Anzeige in eine Kamera- und Übersichtsansicht aufgeteilt (siehe Abbildung 6.12).

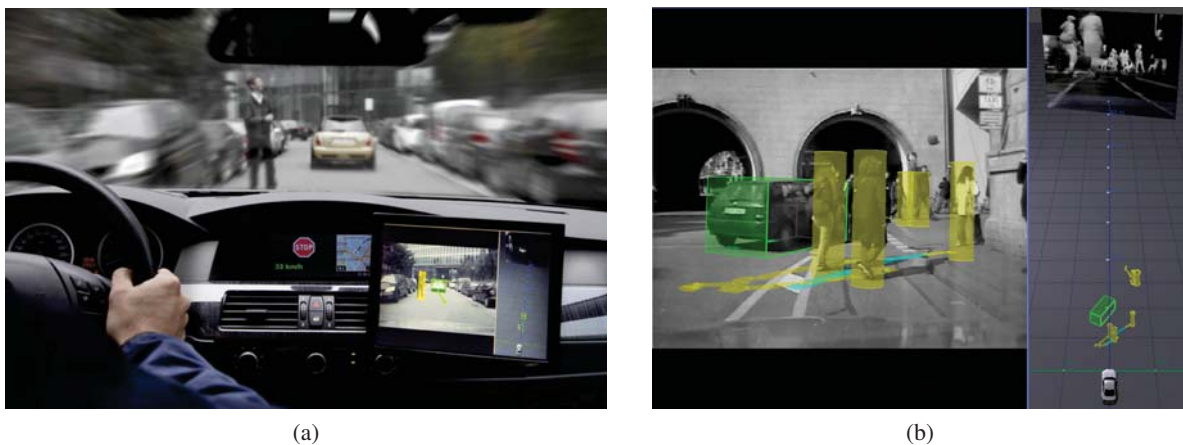


Abbildung 6.12: Visualisierung der Fahrumfelderfassung im Fahrzeug. Abbildung (a) zeigt rechts unten den zu Präsentations- und Überwachungszwecken integrierten Bordmonitor mit der Anzeige der aktuellen Fahrszene. In Abbildung (b) ist die Anzeigeaufteilung des Bordmonitors gesondert illustriert. Links ist das aktuelle Videokamerabild mit einprojizierten maschinell erkannten Umfeldobjekten dargestellt. Die rechte Hälfte des Monitorbildes zeigt eine Vogelperspektivenansicht der erfassten Szenerie mit einem weiteren räumlich eingebetteten Sensorbild (in der Abbildung ein Wärmebild).

Auf der linken Bildschirmhälfte wird die vertraute Sicht auf die Straße durch eine den Perzeptionsbereich überblickende Videokamera dargestellt, die mit allen virtuellen Objekten halbtransparent überlagert wird. Durch die ähnlichen Blickrichtungen von Kamera und Betrachter wird die Zuordnung von

realem zu virtuellem Objekt erleichtert. Die Vogelperspektivenansicht auf der rechten Bildschirmhälfte erlaubt einen Szenerieüberblick und gibt exakten Aufschluss über Positionierung und Ausrichtung der virtuellen Objekte.

### 6.3.2 Erweiterte Realität

Unter Erweiterter Realität, engl. Augmented Reality (AR), versteht man die Erweiterung der realen Welt durch virtuelle Informationen. So kann zum Beispiel mit Hilfe der AR einem Chirurgen während einer Operation die optimale Position einer Prothese visualisiert oder ein Pilot bei schlechter Sicht sicher zum Flughafen geleitet werden. Im Allgemeinen ist die Erweiterung der Realität nicht nur auf eine Sinnesmodalität beschränkt. Vielmehr sind neben optischer Überlagerungen auch akkustische Anreicherungen denkbar. Für eine detaillierte Einführung in die erweiterte Realität sei auf Azuma [Azu97, ABB<sup>+</sup>01] verwiesen.

Eine ökonomische und realitätsnahe Evaluierung neuer Konzepte ist ein wichtiger Bestandteil in der Entwicklung zukünftiger Fahrerassistenzsysteme. Auf diesem Gebiet gewinnen Methoden der erweiterten Realität neben klassischen Fahrsimulatoren immer mehr an Bedeutung. So können Benutzeroberflächendesigner zum Beispiel neuartige Fahrerassistenzsysteme, bereits bevor sie technisch realisierbar sind, erleben und evaluieren. Bock et al. [BM06, BMF07] bilden mit Hilfe der AR gefährliche Situationen in einer realen Umgebung nach, indem dem Fahrer auf einem Testgelände virtueller Straßenverkehr simuliert wird. Ein ähnliches Prinzip wird von Regenbrecht et al. [RBW05] für ein erweitertes Fahrsicherheitstraining verwendet: Die Realität wird hierbei sowohl für den Schüler als auch den Lehrer mit virtuellen Verkehrsteilnehmern und künstlichen Wetterbedingungen überlagert. Auf diese Weise lernt der Schüler in einer sorgfältig definierten Umgebung, kritische Situationen wie Unfälle, Hindernisse oder schlechte Wetter- bzw. Straßenbedingungen zu meistern.

Im Vergleich zu den starren vordefinierten Anzeigekonzepten, wie sie im vorherigen Abschnitt 6.3.1 beschrieben wurden, kann durch die Verwendung der erweiterten Realität ein sehr individueller Eindruck von den Ergebnissen einer Fahrumfelderfassung vermittelt werden. Dies vermittelt fachfremden Entwicklern mehr Verständnis für Messwerte bzw. Daten und die daraus resultierende Zusammenhänge und Probleme. Je nach verwendetem Anzeigegerät kann die AR für Experten auch als Arbeits- bzw. Diskussionsgrundlage dienen und somit die Zusammenarbeit während der Entwicklung der Fahrumfelderfassung fördern.

In den folgenden Abschnitten wird exemplarisch an zwei unterschiedlichen Szenarien (Labor und Fahrzeug) beschrieben, wie sich Konzepte der erweiterten Realität als Präsentationsmittel in einer Fahrumfelderfassung nutzen lassen. Neben technischen Aspekten in der Realisierung wird in einer abschließenden Diskussion über Erfahrungen berichtet, die während der Erprobung im Labor bzw. Fahrzeug gesammelt wurden. Die Umsetzung und die Auswertung der Ergebnisse erfolgte in Kooperation mit dem Lehrstuhl für Informatikanwendungen in der Medizin & Augmented Reality der TU München. Eine detaillierte Betrachtung findet sich in der gemeinsamen Veröffentlichung [TLWK07].

### Trackingsystem

In der erweiterten Realität können virtuelle Informationen nur dann korrekt visualisiert werden, wenn der aktuelle Blickwinkel und die Position des Betrachters in Bezug auf die reale Welt bekannt sind.

Sobald sich der Benutzer allerdings frei bewegen kann, ist diese Beziehung nicht mehr statisch und die Visualisierung muss jeder Positions- und Blickwinkeländerung folgen.

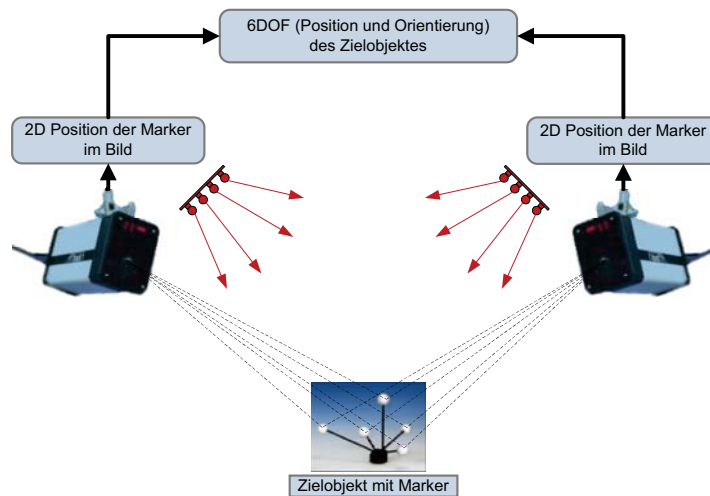


Abbildung 6.13: Schematischer Aufbau eines optische Trackingsystems: Zwei Infrarotkameras bestimmen die Position der retroreflektierenden Kugeln des Zielobjektes im Bild. Über Epipolargeometrie und Triangulierung wird aus der Bildposition in beiden Kameras die Position der Kugeln im Raum ermittelt. Schließlich kann aus der Lage der einzelnen Marker zueinander die Orientierung des Zielobjektes bestimmt werden.

Für das automatische Verfolgen dieser Parameter werden in der erweiterten Realität optische, magnetische oder akkustische Trackingverfahren verwendet. Hierbei bieten optische Verfahren im Allgemeinen die höchste Genauigkeit [Sch04]. Allerdings ist eine Sichtverbindung zwischen den Kameras und den zu verfolgenden Objekten erforderlich. In der technischen Umsetzung wurde ein kommerzielles optisches Trackingsystem<sup>6</sup> verwendet. In diesem System werden die zu verfolgenden Objekte mit mehreren retroreflektierenden<sup>7</sup> Kugeln (Marker) versehen. Infrarotkameras erkennen die Marker im Bild und bestimmen mittels Epipolargeometrie und Triangulierung deren Position im Raum. Aus der bekannten räumlichen Anordnung der Marker zueinander kann die Position und Orientierung der Zielobjekte im Raum bestimmt werden. Um die zugrundeliegende maschinelle Bildverarbeitung zu vereinfachen und optimale Lichtbedingungen zu gewährleisten, wird die komplette Szene mit zeitlich synchronisierten Infrarotblitzen beleuchtet. Der schematische Aufbau des verwendeten optischen Trackingsystems ist in Abbildung 6.13 dargestellt. Eine detaillierte Einführung sowie eine Beschreibung weiterer Trackingverfahren für die erweiterte Realität wird in [RDB01] gegeben.

## Anzeigegeräte

Neben dem Tracking sind geeignete Anzeigegeräte (siehe Abbildung 6.14) ein wichtiger Aspekt in der erweiterten Realität. Sie sind dafür verantwortlich, dass Betrachter künstliche Objekte als Teil

<sup>6</sup> DTrack von Advanced Realtime Tracking GmbH (A. R. T.).

<sup>7</sup> Als Retroreflektoren werden Reflektoren bezeichnet, die weitgehend unabhängig von ihrer Ausrichtung, den größten Teil der eingehenden Strahlung zurück in Richtung der Strahlungsquelle reflektieren.



der realen Welt wahrnehmen. Im Folgenden werden die Anzeigeräte erläutert, die in der auf AR basierenden Visualisierung eines automotiven Perzeptionssystems verwendet wurden.

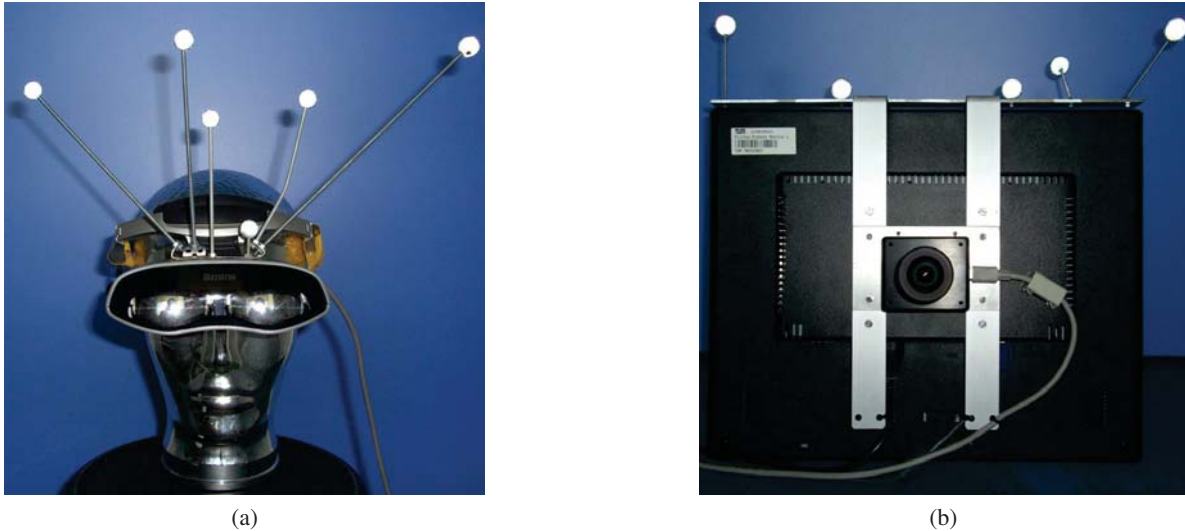


Abbildung 6.14: Unterschiedliche Ausgabegeräte für die erweiterte Realität: Abbildung (a) zeigt ein HMD (Sony Glasstron PLM-S700). Um einen durchsichtigen Bildschirm zu simulieren wurde an die Rückseite eines TFT Monitors (b) eine Kamera installiert. Die zusätzlich angebrachten Markierungskugeln in Abbildung (a) und (b) dienen der robusten Verfolgung der Position und Orientierung der jeweiligen Anzeigeräte.

**Head-Mounted Display** Ein Head-Mounted Display (HMD) ist ein auf dem Kopf getragenes visuelles Ausgabemedium, das am Computer erzeugte Bilder auf augennahen Bildschirmen darstellt oder direkt auf die Netzhaut projiziert [Wik07]. In der AR werden HMDs dazu verwendet, um die Realität mit computererzeugten Bildern zu überlagern [RR91]. Dieser Effekt lässt sich mit einem halbtransparenten Bildschirm oder mit zusätzlichen Kameras erreichen. Bei letzteren wird das Videobild der Kamera als Hintergrund für die Augmentierung verwendet und auf einem undurchsichtigen Bildschirm dargestellt.

Für die Visualisierung wurde ein „Sony Glasstron PLM-S700“ verwendet, das über zwei semitransparente Bildschirme mit einer Auflösung von  $800 \times 600$  Bildpunkten verfügt. Mit diesem HMD kann allerdings kein dreidimensionaler Seheindruck generiert werden, da beide Bildschirme nicht unabhängig voneinander angesteuert werden können. Sechs optische Marker, die, wie in Abbildung 6.14a dargestellt, an der Oberseite angebracht sind, ermöglichen ein robustes und präzises Verfolgen der Position und Orientierung [Bau07].

**Tragbarer „Video-see-through“ Bildschirm** Neben halbtransparenten Bildschirmen, wie sie in vielen HMD's zum Einsatz kommen, kann das von einer Kamera erzeugte Abbild der Umwelt dazu verwendet werden, um die Realität mit virtuellen Informationen zu überlagern und dadurch einen transparenten Bildschirm zu simulieren. Hierfür wird eine Kamera an der Rückseite eines Bildschirms angebracht und das von ihr generierte Bild auf der Vorderseite dargestellt (siehe Abbildung 6.14b). Somit lässt sich für den Benutzer der Eindruck eines durchsichtigen Bildschirms erzeugen. Für die



Visualisierung wird ein auf diese Weise modifizierter 19" Bildschirm verwendet, auf dessen Oberseite zusätzlich fünf optische Marker angebracht wurden, um ein robustes Verfolgen der Position und Orientierung zu ermöglichen.

### Visualisierung im Labor

Für die Realisierung des Systems wurde im ersten Schritt eine Laborumgebung gewählt, da hier eine Evaluierung und Optimierung unter vordefinierten Bedingungen und unter dem Ausschluss von externen Störgrößen erfolgen kann. Der gesamte Versuchsaufbau ist in Abbildung 6.15a schematisch dargestellt.

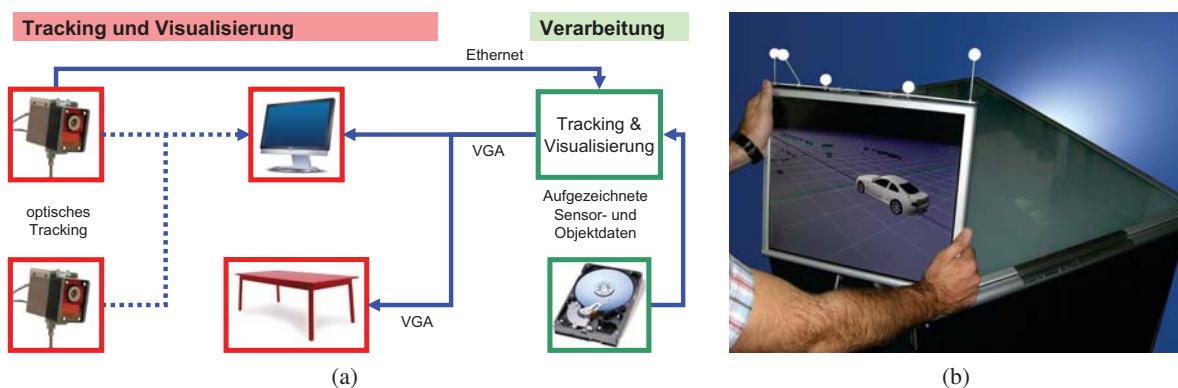


Abbildung 6.15: Visualisierung im Labor: Abbildung (a) zeigt einer Übersichtsdarstellung für das Visualisierungssystem in der Laborumgebung. Der tragbare TFT Monitor in Abbildung (b) dient in der Laborumgebung als Lupen- bzw. Detailvisualisierung für die Daten, die auf dem Rückprojektionstisch in einer Vogelperspektive dargestellt werden.

In dieser Umgebung werden aufgezeichnete Daten gleichzeitig auf einem Rückprojektionstisch<sup>8</sup> und einem tragbaren Bildschirm (siehe Abbildungen 6.14b und 6.15b) dargestellt. Der Tisch dient zum Einen als Übersichtsdarstellung und trägt somit zu einer besseren Orientierung innerhalb der Szene bei. Zum Anderen bietet er eine Diskussionsgrundlage, wenn mehrere Benutzer die Daten bzw. Ergebnisse begutachten. Nachdem mit Hilfe des Tisches interessante Bereiche ermittelt wurden, ermöglicht der tragbare Bildschirm eine Lupen- bzw. Detailvisualisierung beliebiger Szenenausschnitte (siehe Abbildung 6.15b). Der Anwender kann somit durch die Bewegung des Bildschirms intuitiv den Blickwinkel auf die Daten bestimmen. Hierzu wird durch ein Trackingsystem fortlaufend die Position und Orientierung der Anzeigeräte bestimmt. Die „video-see-through“ Funktionalität des tragbaren Bildschirmes ist in diesem Szenario nicht sinnvoll und wurde deaktiviert, da die zu visualisierenden Daten außerhalb des Labors aufgezeichnet wurden.

### Kontaktanaloge Visualisierung im Fahrzeug

Im Fahrzeug werden die Daten der Fahrfelderfassung sowohl auf einem HMD als auch auf einem tragbaren videobasierten Bildschirm visualisiert. Besonderes Interesse gilt hierbei innerstädtischen

<sup>8</sup> Ein Rückprojektionstisch ist ein Tisch, bei dem das Bild von unten auf eine halbtransparente Tischplatte projiziert wird.

Szenen bzw. Fahrten auf der Autobahn. Um den Fahrer nicht unnötig von seiner primären Aufgabe abzulenken, wurde der Versuchsaufbau (siehe Abbildung 6.16a) so gewählt, dass beide Anzeigemodalitäten ausschließlich dem Beifahrer zur Verfügung stehen. Durch die kopfabhängig Perspektivenwahl soll mit beiden Geräten eine möglichst individuelle Visualisierung erreicht werden. Um eine derartige AR realisieren zu können, wurde ein Tracking-System in das Fahrzeug integriert, das die Position und Orientierung der Anzeigeräte robust verfolgt und darüber hinaus jede Ablenkung und Störung des Fahrers und Anwenders vermeidet. Aus diesem Grund wurden die beiden Kameras des Trackingsystems, wie in Abbildung 6.16b zu sehen, auf der Rücksitzbank mit Blick auf den Beifahrer verbaut. Der maximal mögliche Interaktionsbereich beschränkt sich somit auf den Beifahrer.

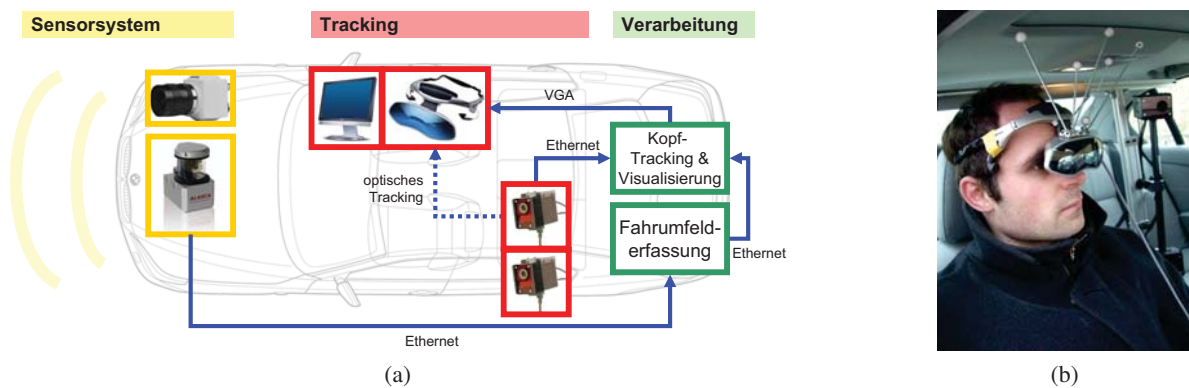


Abbildung 6.16: Kontaktnaloge Visualisierung im Fahrzeug: (a) Übersichtsdarstellung für die AR-Visualisierung im Fahrzeug. Ein HMD (c) und ein tragbarer Bildschirm werden verwendet, um Daten der Fahrumfelderfassung zu visualisieren. Für eine korrekte Überlagerung der realen Welt mit Sensor- und Objektdaten verfolgt ein Trackingsystem (b) die Position und Orientierung der Anzeigeräte.

### Spatial Relationship Graph

Das Visualisierungssystem muss kontinuierlich die 3D Positionen und Orientierungen (Pose) von Sensordaten, Objekten und Anzeigeräten zueinander in Bezug setzen. Diese Beziehungen, die sich aus Translationen, Rotationen, Skalierungen oder Projektionen zusammensetzen, lassen sich durch eine Transformationsmatrix  $\mathbf{T}$  beschreiben. In Abbildung 6.17 sind die Beziehungen zwischen allen in der AR-Visualisierung beteiligten Komponenten in einem so genannten „Spatial Relationship Graph“ (SRG) [NWB04, PHBK06] dargestellt. Hierbei werden Komponenten durch Knoten und räumlichen Beziehungen zwischen den einzelnen Knoten durch gerichtete Kanten symbolisiert.

Der linke Teil des SRGs beschreibt den internen Aufbau der Fahrumfelderfassung, die aus mehreren Sensoren besteht. Messwerte werden im Allgemeinen im Koordinatensystem des zugehörigen Sensors beschrieben (Kanten  $\mathbf{O}$  und  $\mathbf{P}$ ), wohingegen die Position und Orientierung der Sensoren selbst relativ zum Fahrzeugkoordinatensystem angegeben werden (Kanten  $\mathbf{M}$  und  $\mathbf{N}$ ). Innerhalb des SRGs bezeichnet der „VP Render“ Knoten den virtuellen Viewport, bezüglich dessen die Szene dargestellt wird. In einer eigenständigen Visualisierung wird dieser Viewport, wie in Abschnitt 6.2.5 beschrieben, mit der Maus oder Tastatur kontrolliert (Kanten  $\mathbf{Q}$  und  $\mathbf{R}$ ). In der AR-Visualisierung hingegen bestimmen die getrackten Anzeigeräte über die Kanten  $\mathbf{I}$ ,  $\mathbf{J}$  bzw.  $\mathbf{K}$  wie die Szene dargestellt werden soll.

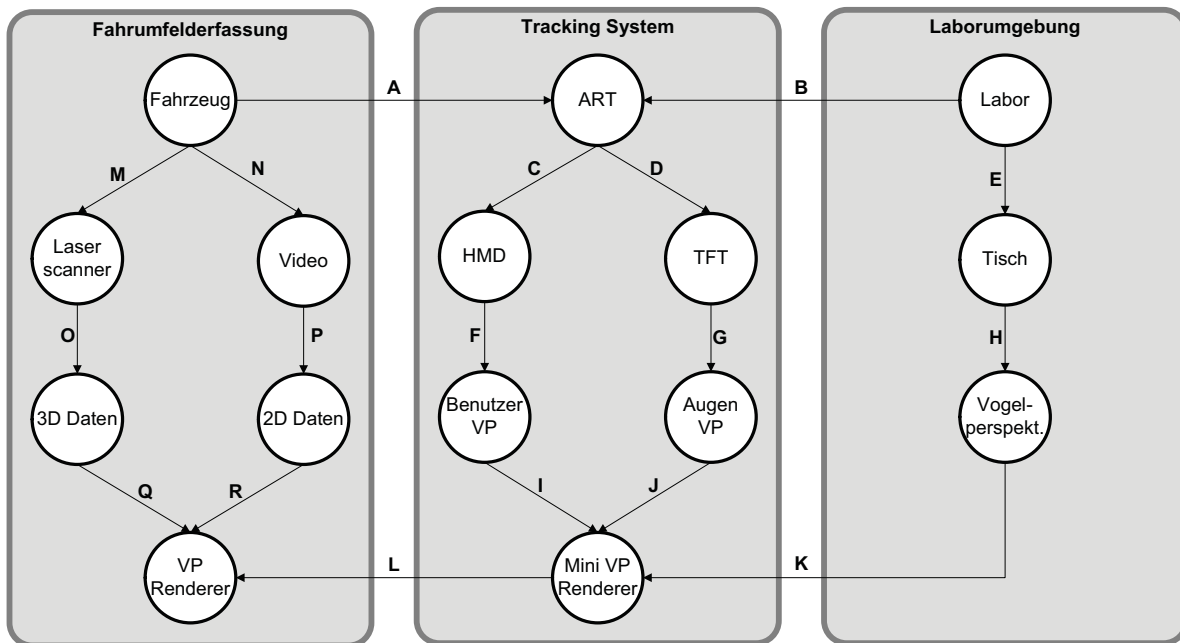


Abbildung 6.17: Der „Spatial Relationship Graph“ des Visualisierungssystem für die Labor- und Fahrzeugumgebung. Er beschreibt die Beziehung zwischen Sensordaten, Objekten und Anzeigegeräten in Bezug auf ihre Position und Orientierung und ihren Einfluss auf die Darstellung der Daten.

Der mittlere Abschnitt des SRGs beschreibt den Aufbau des Trackings, das benötigt wird, um innerhalb der Fahrzeugumgebung die Position und Orientierung der mobilen Anzeigegeräte (TFT und HMD) zu ermitteln. Als Trackingsystem wird ein optischer Tracker verwendet (siehe Abschnitt 6.3.2). Der Ursprung des Trackerkoordinatensystems ist abhängig vom Fahrzeugkoordinatensystem (siehe Kante **A**). Für beide Anzeigegeräte ermittelt das Trackingsystem die Position und Orientierung der optischen Marker (Kante **C** bzw. **D**). Im Allgemeinen stimmt die Pose der Marker nicht mit der vordefinierten inhärenten Blickrichtung der Anzeigegeräte überein, was mit den zusätzlichen Kanten **F** und **G** kompensiert wird. Das HMD bzw. TFT kontrollieren letztendlich über die Kanten **J** und **I** die eigentliche Darstellung der Szene. Für die Laborumgebung wird die Darstellung um den Faktor  $\frac{1}{10}$  skaliert (Kante **L**).

Im rechten Teil des SRG ist die Pose des Rückprojektionstisches innerhalb der Laborumgebung sowie die Position der Vogelperspektive relativ zum Tisch durch die Kanten **E** bzw. **H** definiert. Über die Kante **K** wird bestimmt, wie die Daten auf dem Tisch dargestellt werden soll.

### Kalibrierung der Anzeigegeräte

Um Daten auf dem HMD, TFT oder dem Rückprojektionstisch darstellen zu können, muss der Blickwinkel des Anwenders relativ zu den Anzeigegeräten beschrieben werden (Kanten **F**, **G** und **H** aus Abbildung 6.17). Das Trackingsystem definiert anhand der Position der optischen Marker zueinander

die Position und Lage des Zielkoordinatensystems. Im Allgemeinen befindet sich dieses Koordinatensystem innerhalb der Marker und nicht auf oder innerhalb des Anzeigegerätes. Deshalb sind zusätzliche Kalibrierungsschritte notwendig, um die Transformation zwischen dem Zielkoordinatensystem und dem Ausgabegerät bzw. dem Betrachter zu ermitteln. Diese Schritte werden im Folgenden näher erläutert.

Für das TFT (Kante **F** aus Abbildung 6.17) muss die Transformation zwischen Zielkoordinatensystem und Mittelpunkt des Bildschirms gefunden werden. Mit einem getrackten Zeigestab werden hierzu alle vier Ecken ( $P_1, P_2, P_3, P_4$ ) der sichtbaren Anzeigefläche des TFTs markiert. Durch diese vier Punkte ist der Mittelpunkt  $\mathbf{c} = \frac{\mathbf{p}_3 + \mathbf{p}_1}{2}$  und das Koordinatensystem des Bildschirms mit den Basisvektoren  $\mathbf{b}_1 = \mathbf{p}_1 - \mathbf{p}_2$ ,  $\mathbf{b}_2 = \mathbf{p}_2 - \mathbf{p}_3$  und  $\mathbf{b}_3 = \mathbf{b}_1 \times \mathbf{b}_2$  eindeutig bestimmt.

Für die Kalibrierung des HMDs (Kante **G** aus Abbildung 6.17) wurde die „Single Point Active Alignment Method“ (SPAAM) [TN00] verwendet. Hierbei muss der Anwender ein Fadenkreuz, das auf dem Bildschirm des HMDs dargestellt ist, mit mehreren vordefinierten Punkten im dreidimensionalen Raum zur Deckung bringen. Der Algorithmus berechnet anschließend die Projektionseigenschaften sowie die Transformation zwischen dem Ziel- und Anzeigekoordinatensystem.

Der Rückprojektionstisch befindet sich an einer vordefinierten Position innerhalb der Laborumgebung. Eine Kalibrierung für die Kante **H** aus Abbildung 6.17 ist deshalb nicht erforderlich.

### Aspekte der Umsetzung

Die AR Visualisierung wurde im ersten Schritt in der Laborumgebung entwickelt bzw. getestet und anschließend ins Fahrzeug integriert. Die in diesem Prozess gesammelten Erfahrungen sowie geschlossenen Kompromisse werden im Folgenden beschrieben und diskutiert.

**Laborumgebung** Um aufgezeichnete Sensordaten in einer Laborumgebung zu visualisieren, mussten folgende Anpassungen vorgenommen werden:

In einer Fahrumfelderfassung sind Sensordaten mit einer Entfernung von über 100m keine Seltenheit. Damit dennoch eine vollständige Visualisierung in der räumlich beschränkten Laborumgebung erfolgen kann, müssen die original Sensordaten um den Faktor  $\frac{1}{10}$  skaliert werden. Darüber hinaus, lässt sich ein skaliertes Abbild der Umgebung einfacher aus verschiedenen Blickwinkeln außer- oder innerhalb des Fahrzeuges untersuchen.

Im Allgemeinen bewegt sich das Fahrzeug in der realen Welt. Eine Darstellung der Sensordaten in Weltkoordinaten erscheint deshalb nicht sinnvoll, da das Fahrzeug während einer Testfahrt schnell die Laborumgebung verlassen würde. Stattdessen, wird das Koordinatensystem des Fahrzeuges auf den Ursprung des Trackingkoordinatensystems fixiert (siehe Kanten **B** und **E** in Abbildung 6.17). Die Daten werden somit relativ zum Fahrzeugkoordinatensystem visualisiert.

Die Laborumgebung besteht aus einer vollständigen virtuellen Umgebung, in der keine realen Objekte wie das Eigenfahrzeug oder Sensordaten als Orientierungshilfen existieren. Für einen intuitiven Umgang mit den Daten, muss der Benutzer jedoch wissen, wo und aus welchem Blickwinkel die Szene betrachtet werden soll. Hierfür hat sich die Übersichtsdarstellung auf dem Rückprojektionstisch als hilfreich erwiesen. Darüber hinaus ist die Position des Tisches an den Ursprung des Trackingkoordinatensystems ausgerichtet, was als weitere Orientierungshilfe für den Anwender dient.

**Tracking im Fahrzeug** Das nötige Trackingvolumen innerhalb des Fahrzeuges kann durch einen Quader beschrieben werden (40 cm vertikal, 80 cm horizontal und 60 cm vor den Beifahrer). In diesem Bereich wird gewährleistet, dass dem Anwender ausreichend Raum für die Interaktion mit den beiden Anzeigegeräten zur Verfügung steht. Um dieses Areal überwachen zu können, wurde ein optisches Trackingsystem verwendet, das über zwei Weitwinkelkameras mit einer Basisbreite von 50 cm verfügt. Die beiden optischen Achsen treffen sich in einer Entfernung von 160 cm, und das Trackingvolumen beginnt in einem Abstand von 30 cm, wo sich die Grenzen der Sichtbereiche nahezu orthogonal schneiden. Mit diesem Aufbau kann für ein Standardziel, abhängig von dessen Entfernung (ca. 50-100 cm) zu den beiden Kameras, eine Auflösung von 0.25 mm bis zu 0.5 mm erreicht werden [Bau07].

Bedingt durch die Größe des Trackingsystems und den minimalen Abstand des Trackingvolumens, war eine Installation des Systems vor dem Beifahrer über dem Armaturenbrett nicht möglich. Die Kameras konnten auch nicht weiter auf die Seite des Fahrers verschoben werden, da jede Ablenkung und Störung des Fahrers vermieden werden sollte. Aus diesem Grund mussten die Kameras hinter den Vordersitzen angebracht werden. Der Versuch das System horizontal hinter der Kopfstütze des Beifahrers oder des Fahrers anzubringen scheiterte, da entweder der Kopf des Anwenders oder eine der beiden Kopfstützen die Sicht der Kameras zu stark beeinträchtigten. Schließlich ermöglichte eine vertikale Installation des Trackingsystems, wie sie in Abbildung 6.18c dargestellt ist, eine ausreichende Sicht auf das geforderte Trackingvolumen. Allerdings können unter bestimmten Voraussetzungen immer noch Verdeckungen entstehen. Zum Einen kann der rechte Ellbogen des Fahrers, wenn er auf der Armlehne zu weit nach hinten bewegt wird, die Sicht der Kameras auf den unteren Teil des Volumens beeinträchtigen. Zum Anderen werden die Marker des tragbaren „Video-see-through“-Bildschirm, wenn er zu weit nach rechts bewegt wird, teilweise vom linken Arm oder der linken Schulter des Beifahrers verdeckt.



Abbildung 6.18: Installation und Ausrichtung des optischen Trackingsystems im Fahrzeug. In Abbildung (a) ist die Befestigung an den Handgriffen zu sehen. Die vertikale Verbauposition über der Rücksitzbank zwischen den Vordersitzen und die Ausrichtung der Kameras auf das Trackingvolumen (Beifahrer) ist in Abbildung (b) und (c) zu sehen.

Nachdem ein geeigneter Verbauort für das Trackingsystem gefunden wurde, musste sichergestellt werden, dass keine Infrarotlichtquelle innerhalb des Fahrzeuges die Qualität des Trackings negativ beeinflusst. Untersuchungen haben ergeben, dass der Rückspiegel je nach Einstellung unerwünschte Re-



flektionen der infraroten Beleuchtung erzeugt. Als weitere Störquelle konnte die Nachtbeleuchtung, die im Himmel des Fahrzeuges über dem Schalthebel integriert ist, identifiziert werden. Diese statischen Lichtquellen können allerdings vernachlässigt werden, indem die zugehörigen Bereiche auf dem Kamerabild ignoriert werden. Dieses Verfahren, das auch unter dem Namen „blind spot detection“ bekannt ist, lässt sich jedoch nicht auf externe Störeinflüsse, wie zum Beispiel Sonneneinstrahlung anwenden. Die Thermoverglasung der Windschutzscheibe absorbiert den infraroten Anteil des Sonnenlichts nicht zu 100%, deshalb kann es je nach Einfallswinkel zum Teil zu erheblichen Einschränkungen der Trackingqualität kommen.

Ein letzter Punkt in Bezug auf das Tracking betrifft die stabile und vibrationsfreie Befestigung des Trackingsystems im Fahrzeug. Um sicherzustellen, dass sich das Trackingkoordinatensystem durch Vibrationen des Fahrzeuges während der Fahrt nicht verändert, muss eine starre Verbindung zwischen dem Fahrzeug und dem Trackingsystem erreicht werden. Eine Konstruktion, die die Isofix Kindersitzhalterung und die beiden Handgriffe über den Rücksitztüren verwendet, gewährt ausreichend Stabilität, um den Einfluss lateraler und longitudinaler Vibrationen zu minimieren. Vertikale Schwingungen wurden mit einer zusätzlichen Stütze im Fußraum gedämpft. Die komplette Konstruktion ist in Abbildung 6.18 zu sehen. Um den verbleibenden Einfluss von Schwingungen auf die Trackingqualität zu ermitteln, wurde ein Zielobjekt in einer Entfernung von 120 cm zu den Kameras starr mit der Karosserie verbunden. Das Schwingungsverhalten der Marker ist hierbei zu vernachlässigen, da sie im Vergleich zu den Kameras sehr leicht sind. Basierend auf aufgezeichneten Daten wurde für ein stillstehendes Fahrzeug eine Standardabweichung für die Position von unter 0.1 mm und für die Orientierung von unter 0.2 Grad ermittelt. Während der Fahrt erhöhte sich die Positionsungenauigkeit für alle drei Achsen auf 4 mm. Die Standardabweichung der Orientierung blieb unter einem Grad.

**Systemperformanz** Für eine korrekte Überlagerung der Realität mit virtuellen Informationen sind kurze Systemantwortzeiten außerordentlich wichtig. Deshalb wurde besonderes Augenmerk auf eine echtzeitfähige Implementierung gelegt. Die Systemantwortzeit setzt sich aus den Latenzen für das Tracking, für die Kamera des tragbaren Bildschirms, für die Fahrumfelderfassung, für die Darstellung der Szene und für die Kommunikation der Subsysteme untereinander zusammen. Das Visualisierungssystem wurde in OpenGL implementiert und wird von einer 3D Grafikkarte beschleunigt ausgeführt. Dies führt zu einer Darstellungszeit von ca. 20 ms. Die Fahrumfelderfassung benötigt weitere 20 bis 30 ms. Zusammen mit der durchschnittlichen Latenz des Trackingsystems von ca. 20 ms ergibt sich eine durchschnittliche Systemantwortzeit von ca. 70 ms. Der Beitrag der Kommunikationswege ist vernachlässigbar. In Tabelle 6.1 sind alle Latenzzeiten die innerhalb der Visualisierung entstehen aufgelistet.

	Fahrumfelderfassung	Tracking	Darstellung	Kommunikation
Latenz in ms	30 ms bis 40 ms	20 ms	20 ms	≤ 1 ms

Tabelle 6.1: Latenzzeiten der einzelnen Systemkomponenten für die AR basierte Visualisierung im Fahrzeug.

**Anzeigeegeräte** Während der Fahrt können unterschiedlichste Lichtverhältnisse vorherrschen: Von annähernder Dunkelheit in der Nacht auf Landstraßen bis zu direkter Sonneneinstrahlung am Tag. Unter derartigen Bedingungen kann die Wahrnehmung der erweiterten Realität geschmälert werden, wenn Anzeigeegeräte nicht in der Lage sind einen ausreichenden Kontrast im Vergleich zur realen Welt zu erzeugen. Darüber hinaus erhöht sich bei starken Helligkeitsschwankungen die Fokussierungszeit der Augen, was den Anwender schneller ermüden lässt. Insbesondere bei optisch durchsichtigen Anzeigeegeräten wie dem HMD kann der Kontrast in hellen Umgebungen derart abnehmen, dass die Visualisierung für den Menschen nicht mehr sichtbar ist.



Abbildung 6.19: Im Fahrzeug können Sensor- und Objektdaten auf einem tragbaren Bildschirm dargestellt. Abbildung (a) zeigt die Daten des Laserscanners auf einem Parkplatz. Die Beine des Fußgängers werden von den Sensormesswerten überlagert. In Abbildung (b) wird das Ergebnis der Fahrumfelderfassung während der Fahrt auf dem tragbaren TFT visualisiert. Das Bild der Kamera wurde mit drei erkannten Fahrzeugen (grüne Quader) überlagert.

Die Darstellung von Sensor- und Objektdaten erscheint auf dem tragbaren Bildschirm in einem guten Kontrastverhältnis zur realen Welt, da das TFT eine Helligkeit von  $1000 \frac{cd}{m^2}$  besitzt. Zusätzlich regelt die Treibersoftware der Kamera über die Blende die Helligkeit des Bildes. Abbildung 6.19 zeigt die Visualisierung an einem sonnigen Tag. Allerdings existieren bei der Verwendung des TFTs einige Einschränkungen: Aufgrund der zentralen Position der Kamera muss der Bildschirm sehr hoch gehalten werden, um einen ausreichenden Blick auf die Szene zu erlangen. So belegt das Armaturenbrett selbst unter maximaler Ausnutzung des zur Verfügung stehenden Raumes innerhalb des Fahrzeuges noch immer ein Drittel des Bildes. Darüber hinaus ist der Bildschirm verhältnismäßig schwer und lässt sich deshalb nur sehr unkomfortabel über längere Zeit verwenden. Ein letzter Punkt betrifft die Position und Anordnung der optischen Marker. Der Anwender muss stets darauf achten, dass die Marker nicht den Fahrzeughimmel berühren und deformiert werden. Da alle Marker auf der Oberseite des Bildschirms annähernd auf einer Achse angebracht sind, kann der Tracker Rotationen um diese Achse schlecht verfolgen.

Beim HMD (Sony Glasstron) wird die reale Welt durch einen halbtransparenten Bildschirm mit virtuellen Objekten überlagert. Der Kontrast der Darstellung ist hier deshalb von besonders großer Bedeutung. An einem sonnigen Tag ist der Kontrast des HMDs im Vergleich zum dominanten Umgebungslicht zu gering, und die Visualisierung ist kaum noch sichtbar. An bedeckten Tagen oder nachts kann hingegen ein gutes Ergebnis, wie in Abbildung 6.20 dargestellt, erreicht werden.



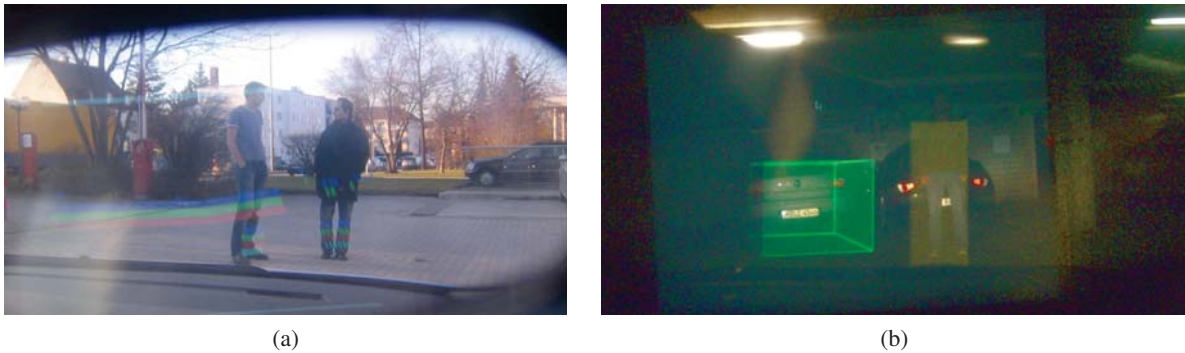


Abbildung 6.20: Im Fahrzeug werden Sensor- und Objektdaten auf einem HMD dargestellt. Abbildung (a) zeigt die Daten des Laserscanners auf einem Parkplatz. Die Beine der Fußgänger werden von den Messwerten überlagert. Um einen guten Kontrast zu erreichen, wurde in Abbildung (b) das Ergebnis der Fahrumfelderfassung in einer Tiefgarage auf dem HMD visualisiert. Die Realität wird mit dem detektierten Fußgänger (gelber Zylinder) und dem erkannten Fahrzeug (grüner Quader) überlagert.

### 6.4 Fahrbahnerkennung mittels Flottenbewegung

Eine geeignete Visualisierung eines Fahrumfelderfassungssystems macht Zusammenhänge zwischen den Daten transparenter und liefert gänzlich neue Einblicke in die Sensordaten. Somit ist sie ein Ideenlieferant für Verbesserungsmöglichkeiten oder gänzlich neuartige Einfälle. Am Beispiel einer umgesetzten Anwendung zur Ermittlung der befahrbaren Fläche durch die Flottenbewegung vorausfahrender Fahrzeuge soll dieses Potential verdeutlicht werden. Die Anregung für die Umsetzung dieses Systems stammt von der Visualisierung der erkannten Fahrzeugflotte und deren Trajektorien (vgl. Abbildung 6.8).

Viele Automobilhersteller haben schon Systeme zur Spurverlassenswarnung auf dem Markt. Bisher funktioniert dieses Fahrerassistenzsystem jedoch lediglich auf gut ausgebauten Straßen mit entsprechend vorhandenen Begrenzungsmarkierungen. Fehlen diese Markierungen jedoch, ist eine Lokalisierung der Straße durch diese meist kamerabasierten Ansätze nicht mehr möglich. Neuere Vorgehensweisen versuchen dieses Defizit durch eine direkte meist texturbasierte Fahrbahnerkennung auszugleichen.

Einen anderen Ansatz verfolgt das im Folgenden detailliert beschriebene System. Dieses ist allerdings nicht als Konkurrenz zu bereits verfügbaren Fahrspurerkennungssystemen zu sehen. Vielmehr liegt das Potential dieser Vorgehensweise in der Informationsfusion, indem die Qualität und Zuverlässigkeit durch zusätzliche Evidenz erhöht wird.

#### Umsetzung

Die einzelnen Trajektorienverläufe der vorausfahrenden Fahrzeugflotte werden gemäß ihrer geschätzten Querausdehnung verbreitert und zu einer aus Dreiecken bestehenden Trajektorienfläche zusammengefügt. Anschließend werden alle Trajektorienflächen zu einem oder mehreren Bereichspolygo-

nen aggregiert. Diese Polygone stellen den befahrbaren Bereich dar. Durch den Flottenverlauf können zusätzlich Straßenverlauf und -krümmung abgeleitet werden (siehe Abbildung 6.21).

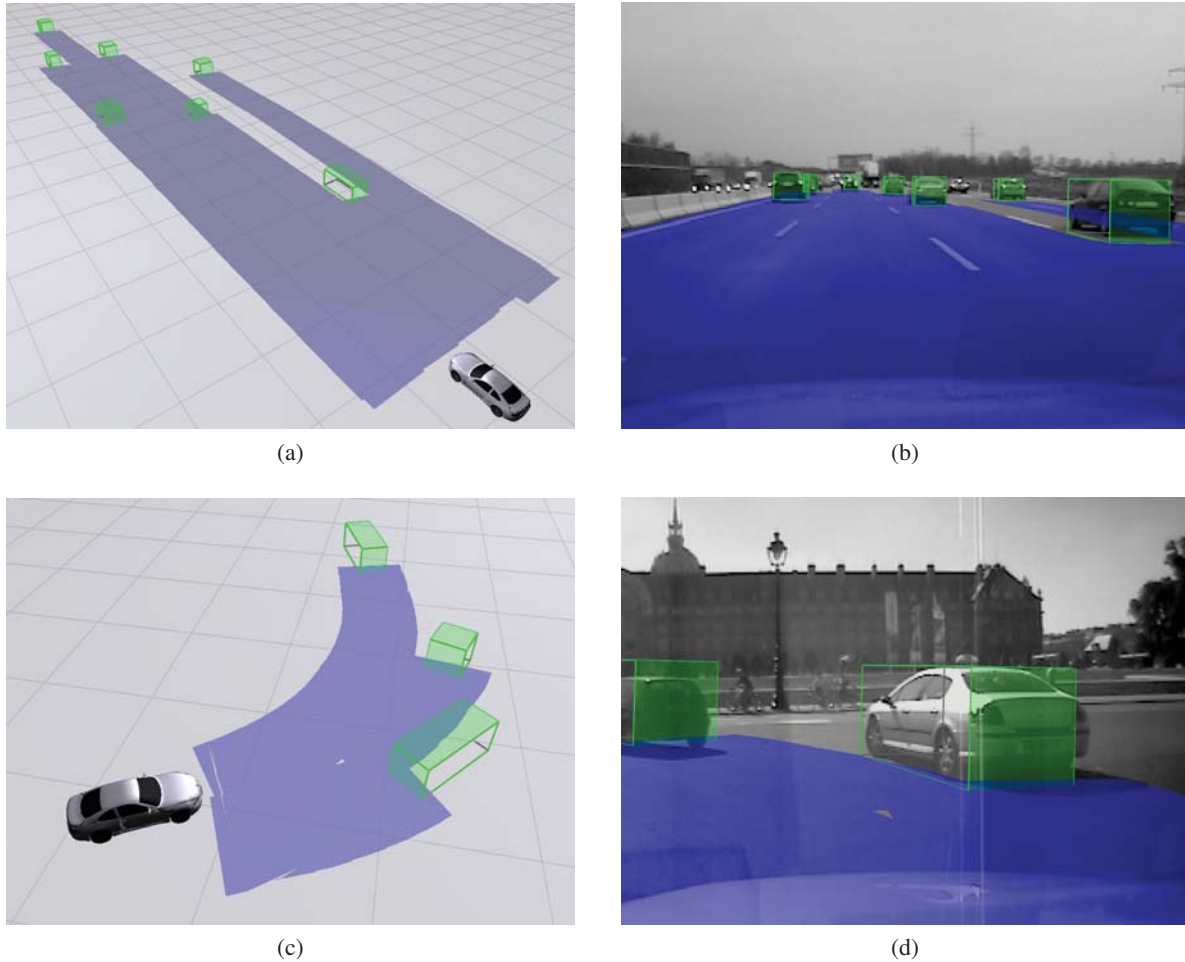


Abbildung 6.21: Befahrbarer Bereich (blaues Polygon), der durch die Flottenbewegung vorausfahrender Fahrzeuge (grüne Quader) ermittelt wurde. Die Abbildungen (a) und (b) bzw. (c) und (d) zeigen dabei jeweils die gleiche Szenerie (Autobahn bzw. Kreisverkehr) aus unterschiedlichen Ansichten (Vogelperspektive und Projektion in fahrzeugintegrierte Videokamera).

Der Trajektorienverlauf  $T_f = (V_1, V_2, \dots, V_n)$  eines einzelnen Fahrzeugs  $f$  in der Flotte  $\mathcal{F}$  (siehe Abbildung 6.22a) ist das Grundgerüst für das Dreiecksnetz. Es ergibt sich ein schlauchähnliches Konstrukt indem durch jeden zweidimensionalen Verlaufsvektor  $V_i$  mit  $i \in \{1, \dots, n-1\}$  eine zur Strecke  $\overline{V_i V_{i+1}}$  orthogonale Begrenzungslinie  $\overline{V_i^l V_i^r}$  gelegt wird. Durch den aktuellen Fahrzeugstandort  $v_n$  wird dabei eine Begrenzungslinie parallel zu  $\overline{V_{n-1}^l V_{n-1}^r}$  gelegt. Dabei entspricht die Länge von  $\overline{V_i^l V_i^r}$  bzw. von  $\overline{V_i^r V_i^l}$  der Hälfte der gemessenen Fahrzeugbreite  $\gamma_f$  vergrößert um eine durchschnittliche Seitenabstandsweite  $\delta$  zu anderen Objekten. Erfahrungsgemäß hat sich ein Wert von  $\delta = 0.5$  m als praktikabel erwiesen. Besser wäre allerdings eine geschwindigkeitsabhängige bzw. kontextsensitive Seitenabstandsweite-

daption. Das Dreiecksnetz eines einzelnen Trajektorienverlaufs besteht nun aus der Dreiecksmenge

$$\mathcal{D}_f = \{[V_i V_{i+1} V_i^l], [V_i V_{i+1} V_i^r], [V_i^l V_{i+1} V_{i+1}^l], [V_i^r V_{i+1} V_{i+1}^r]\} \quad \forall i \in \{1, \dots, n-1\} \quad (6.6)$$

Nachdem dieses Dreiecksnetz  $\mathcal{D}_f$  für jedes Fahrzeug  $f \in \mathcal{F}$  erzeugt wurde, werden alle diese Netze zu einem Gesamtpolygon  $\mathcal{G} = \bigcup_{f \in \mathcal{F}} \mathcal{D}_f$  aggregiert. Überlappende Bereiche werden mittels Tessellierung<sup>9</sup> in eine flache Dreieckstruktur umgewandelt. Die Begrenzung des befahrbaren Bereichs bildet die äußere Kontur von  $\mathcal{G}$  (siehe Abbildung 6.22b für ein Beispiel mit zwei Flottenfahrzeugen).

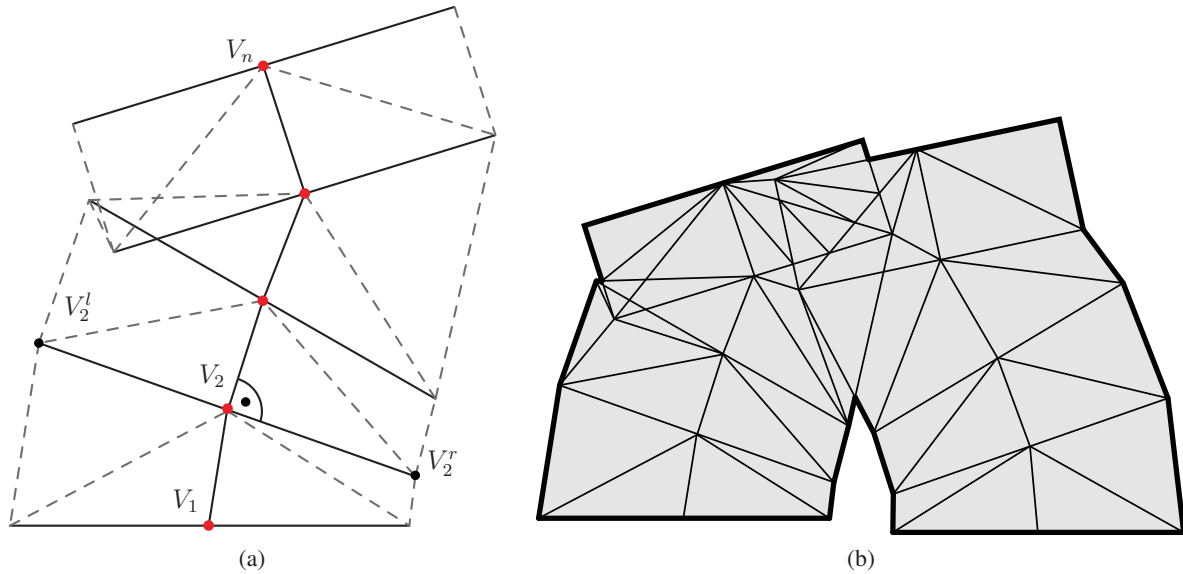


Abbildung 6.22: Polygonerzeugung aus den Flottentrajektorien. Abbildung (a) zeigt die diskrete Trajektorie eines einzelnen Fahrzeugs (rote Punkte) und das daraus erzeugte Dreiecksnetz. In Abbildung (b) werden die Dreiecksnetze von zwei Fahrzeugen mittels Tessellierung zu einem gemeinsamen Polygon aggregiert.

## Bewertung

Eine kontinuierliche Ermittlung der befahrbaren Bereiche ist mit dem vorgestellten Verfahren nicht möglich. Im Wesentlichen müssen hier zwei Vorbedingungen erfüllt sein, damit die Erkennung funktioniert: Es müssen vorausfahrende Fahrzeuge vorhanden sein, und diese müssen über das Perzeptionssystem auch erkannt werden. Erst dann kann über die Fahrzeugtrajektorien und die Breitenschätzung der Fahrzeuge ein befahrbarer Bereich abgeleitet werden. Selbst dann deckt jedoch die ermittelte Region, abhängig von der Flottengröße und -verteilung, selten die gesamte zur Verfügung stehende Fahrbahn ab. Meist wird lediglich eine Untermenge detektiert (siehe Abbildung 6.21b). Des Weiteren ist die Qualität der Bereichserkennung direkt an die Qualität der Fahrumfelderfassung gekoppelt. Ist beispielsweise die Breitenschätzung der Flottenfahrzeuge ungenau, wirkt sich das unmittelbar auf die Ausdehnung der einzelnen Spuren aus. Eine Krümmungsschätzung der Fahrbahn aufgrund der Flottenbewegung ist ebenfalls mit Einschränkungen verbunden. Nur wenn die detektierte Flotte eine ausreichend große Zahl an Fahrzeugen beinhaltet, können Ausreißer, die den globalen Krümmungswert

<sup>9</sup> Tessellierung ist eine Technik, die sich mit der Zerlegung von Polygonen beschäftigt.

verfälschen würden (z.B. durch einen Spurwechsel), ausfindig gemacht werden. Ist die beobachtbare Flotte sehr klein oder besteht lediglich aus einem einzelnen Fahrzeug, sind die generierten Spurdaten, selbst bei einer fehlerlosen Fahrumfelderfassung, mit einer verminderten Verlässlichkeit zu betrachten. Das Vorderfahrzeug kann sich beispielsweise regelwidrig verhalten, oder durch einen Fahrfehler oder äußere Einwirkungen von der Fahrbahn abkommen.

Trotz dieser signifikanten Nachteile hat das umgesetzte Fahrbahnerkennungssystem mittels der Flottenbewegung einige entscheidende Vorteile. Ist eine Fahrumfelderfassung im Fahrzeug, wie sie für eine steigende Anzahl an Fahrerassistenzsysteme (z.B. Abstandshaltung, Gefahrenbremsung, Spurwechselassistent) benötigt wird, ohnehin vorhanden, verursacht diese Art der Fahrbahnerkennung keine zusätzlichen Hardwarekosten, sofern die nötige Software mit den verfügbaren Ressourcen auskommt. Selbst wenn dieses Verfahren aufgrund der signifikanten Nachteile nicht als exklusives Spurerkennungssystem betreibbar ist, liegt das Potential in der Informationsfusion. Eine herkömmliche markierungsbasierte Spurerkennung kann nämlich in vielen Fällen von zusätzlichen Informationen profitieren. Sind beispielsweise die Spurmarkierungen durch Verdeckung, schlechte Wetterverhältnisse oder an Baustellen nicht oder nur unzulänglich erkennbar, ist eine robuste Detektion der Markierungsverläufe nicht gewährleistet. Zudem ist die Entfernungreichweite der visuellen Spurerkennung durch die Auflösung der bildgebenden Sensorik begrenzt und könnte über die Flottenbewegung, die mittels entfernungsgebender Sensorik erfasst wird, vergrößert werden.

Zukünftige Spurerkennungssysteme werden von den Informationen aus markierungsbasierter Spurerkennung, Straßenkartendaten, Fahrzeug-zu-Fahrzeug-Kommunikation und Flottenbewegung profitieren. Wie eine solche Informationsfusion gewinnbringend realisiert werden kann, ist Gegenstand kommander Forschungsaktivitäten.

## 6.5 Zusammenfassung und Diskussion

In diesem Kapitel wurden Visualisierungsmöglichkeiten von heterogenen Sensordaten eines automotiven Umfelderkennungssystems vorgestellt. Die Gründe, warum eine geeignete Informationsrepräsentation von Vorteil ist, sind vielfältig: Der Systementwickler wird in der Umsetzung unterstützt, indem ihm durch eine adäquate Visualisierung ein Instrument an die Hand gegeben wird, das interne Abläufe und Zusammenhänge transparenter macht, die Fehlererkennung- und -suche erleichtert und ein Gespür für die Daten und Datenbeziehungen fördert. Durch eine geeignete visuelle Aufbereitung steht dem Entwickler des Weiteren eine Plattform zur Verfügung, die es erlaubt schwierige Sachverhalte zu veranschaulichen und einem unbedarften Dritten darzulegen. Dadurch werden Kommunikationslücken zwischen fachfremden Experten geschlossen und der Ideen- und Informationsaustausch vereinfacht und gefördert. Die Präsentations- und Kommunikationswirkung ist ein weiteres wichtiges Ziel einer Visualisierung. Durch bildstarke Informationsrepräsentationen können verborgene Prozesse sichtbar gemacht und medienwirksam dargestellt werden.

Um dieses Visualisierungspotential ausschöpfen zu können, wurde im Rahmen dieser Arbeit eine dreidimensionale virtuelle Fahrumgebungsdarstellung entwickelt, in der alle relevanten Daten des Informationsflusses eines automotiven Perceptionssystems repräsentiert werden können. Zudem wurde eine Navigations- und Interaktionsmöglichkeit in die virtuelle Umgebung integriert, die dem Benutzer einen Überblick über die Daten, eine Auflösung von Verdeckungen und Mehrdeutigkeiten und eine Betrachtung von Details ermöglicht. Die verwendeten Konzepte, die zum Teil auf pragmatischen

Visualisierungsansätzen aus der Sensordomäne beruhen, wurden einer systematischen Ordnung zugeführt.

Das entwickelte Darstellungssystem wird mittels Zeichenfilter und frei konfigurierbarer Ansichtsmodi vielfältig und modular eingesetzt. Zum Einen dient es als Offline-Entwicklungswerkzeug am Arbeitsplatz. Zum Anderen demonstriert es die aktuelle Perzeptionsleistung während der Fahrt im Versuchsträger. Zusätzlich wurde die gleiche Visualisierung verwendet, um über Anzeigegeräte wie „head-mounted-displays“, „see-through TFT’s“ und Rückprojektionstische die Sensordaten in einer erweiterten Realität erlebbar zu machen. Das Fallbeispiel „Fahrbahnerkennung mittels Flottenbewegung“ am Ende des Kapitels macht deutlich, wie Visualisierungen den Ideenentstehungsprozess anstoßen können.

Als Darstellungsmöglichkeiten für heterogene Sensordaten wurden Streudiagramme, eine virtuelle dreidimensionale Fahrzeugumgebung und Messwertprojektionen in das Videobild vorgestellt. Eine von diesen Möglichkeiten alleine genügt jedoch nicht vollständig den Visualisierungsanforderungen wie Expressivität, Effektivität und Angemessenheit. Streudiagramme erlauben ähnlich wie Vogelperspektivendarstellungen eine gute Repräsentation von Entfernungen und Datenhäufungen. Lagebeziehungen von räumlichen Daten können aber nur unzureichend veranschaulicht werden. Einen besseren Einblick und eine flexiblere Darstellung gewährt eine interaktiv dreidimensionale Umgebungsmodellierung des Fahrzeugs und der entfernungsbeinhaltenden Sensormesswerte. Eine Benutzernavigation ist dabei unabdingbar, damit sowohl Verdeckungen der Daten untereinander als auch Mehrdeutigkeiten durch den fehlenden stereoskopischen Raumeindruck aufgelöst werden können. Gleichzeitig erhöht eine adäquate Navigationsmöglichkeit die Flexibilität der Darstellung. Sowohl vogelperspektivische als auch isometrische Ansichten sind dadurch denkbar. Für Bilddaten ist eine räumliche Visualisierungsumgebung jedoch weniger geeignet. Es können zwar Kamerabilder als Textur auf Polygone aufgebracht und dadurch dargestellt werden, eine Zuordnung von räumlichen Messwerten zu zweidimensionalen Bilddaten ist dadurch aber nicht unmittelbar möglich. Erst durch eine Übertragung der intrinsischen und extrinsischen Kamerageometrie auf eine virtuelle Szenenkamera ist eine korrekte Anreicherung der Bilddaten mit Entfernungsdaten durchführbar. Der Gesamtüberblick über die räumlichen Daten geht aber meist verloren, da der Kameraerfassungsbereich oft nur einen Ausschnitt der Szene überdeckt. Aufgrund der unterschiedlichen Stärken und Schwächen dieser Darstellungsmöglichkeiten hat sich eine Verbunddarstellung aus „Projektion in das Bild“ und dreidimensionaler virtueller Szenenvisualisierung sowohl im Fahrzeug als auch in der Entwicklungsumgebung bewährt. Nachteile der Darstellungsformen wie der mangelhafte Gesamtüberblick in der Bildprojektion und die unzulängliche Bilddarstellung in der virtuellen Umgebung werden dadurch wechselseitig kompensiert.

Ein möglicher Ansatz, die starren Visualisierungsmethoden aufzulockern, bietet der Einsatz von Erweiterter Realität (AR). Durch kontaktanaloge Visualisierungen mittels getrackter Anzeigemedien können Messwerte in Verbindung mit der Realität erlebbar gemacht werden. Dazu wurden optische Trackingsysteme genutzt, um sowohl in der Labor- als auch in der Fahrzeugumgebung „head-mounted-displays“ und „see-through TFT’s“ als AR-Darstellungsgeräte nutzbar zu machen. Die verwendeten optischen markerbasierten Trackingsysteme liefern zwar eine hohe Positionsgenauigkeit, sind aber relativ anfällig für Verdeckungen und ungünstige Beleuchtungsbedingungen. Speziell in der Fahrzeugumgebung wird die Trackingleistung durch starke Sonneneinstrahlung negativ beeinflusst. Ein weiterer Nachteil bedingt sich durch die beschränkten Raumverhältnisse im Fahrzeug. Der dadurch reduzierte Interaktionsradius wird zusätzlich durch die erhabenen Markerbäume begrenzt. Vibrationen der

Trackingkameras konnten durch einen robusten Einbau minimiert werden. Latenzzeiten wurden mittels einer mit einem 3D-Prozessor beschleunigten Anzeige und kurzen Kommunikationspfaden gering gehalten. Trotz der bestehenden technischen Probleme, die möglicherweise mit alternativen Trackingansätzen (magnetisch) beseitigt werden könnten, fördert die aktuelle Umsetzung von AR im Kontext von Perzeptionssystemen das Sensordatenverständnis von fachfremden Entwicklern und die Gruppenarbeit während der Entwicklungsphase eines Umfelderkennungssystems.

Zudem bieten sowohl die beschriebenen herkömmlichen Visualisierungsansätze als auch die Ansätze der Erweiterten Realität nicht nur eine Plattform zur Fehlererkennung und -beseitigung, sondern auch eine Inspirationsquelle für kreatives Denken und innovative Ideen. Dabei ist ein quantitativer Leistungsvergleich der einzelnen Visualisierungsmöglichkeiten für automotiv Perzeptionssysteme nicht einfach. Selbst wenn sinnvolle Vergleichsmetriken zur Verfügung stünden, könnten erst aufwendige Nutzerstudien eine abgesicherte Aussage treffen. Die in diesem Kapitel aufgestellten Thesen stützen sich deshalb vorwiegend auf wissenschaftlich fundierten Visualisierungsrichtlinien und die im Rahmen dieser Arbeit gewonnene Erfahrung bei der Umsetzung der Fahrumfelderfassung.





# SYSTEMBEWERTUNG

Eine Systembewertung bzw. -evaluierung generiert Gütemaße, die zur Systembeurteilung herangezogen werden können. Wesentliche Gütemaße im Kontext von Umfelderkennungssystemen sind Genauigkeiten, Objekterkennungsraten, Latenzen und Laufzeiten. Dadurch kann zum Einen ein Vergleich zwischen unterschiedlichen Sensoren, Algorithmen und Parameterkonfigurationen durchgeführt und zum Anderen eine grundsätzliche Eignung der Fahrumfelderfassung für diverse Fahrerassistenzfunktionen überprüft werden. Damit in einem Fahrumfelderfassungssystem das Zusammenspiel zwischen Klassifikation und Tracking beurteilt werden kann, sind Evaluierungs- und Referenzdaten erforderlich. Evaluierungsdaten werden mittels Perzeptionsalgorithmen erzeugt, die auf Sensordaten arbeiten. Referenzdaten fungieren als Bewertungsgrundlage und ermöglichen durch einen Abgleich mit der Systemausgabe eine metrikgestützte Gewinnung der Gütemaße (siehe auch Abbildung 7.1).

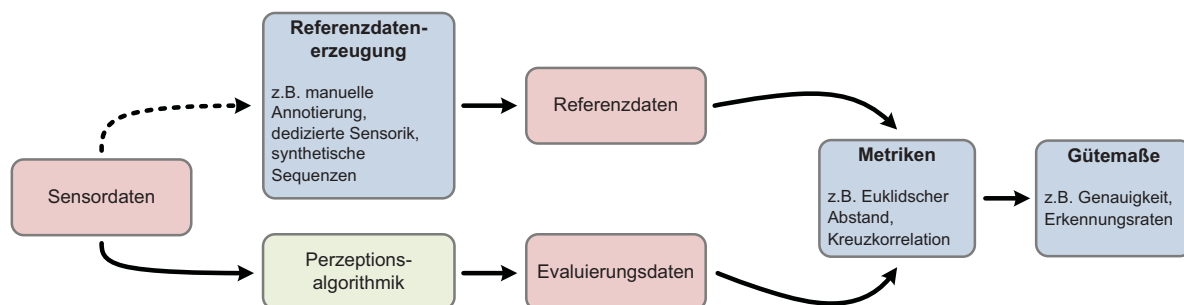


Abbildung 7.1: Übersichtsschaubild zur Systembewertung. Sensordaten sind die Ausgangsdaten zur Referenzdatenerzeugung und zur Evaluierungsdatengewinnung durch Perzeptionsalgorithmen. Durch den Abgleich der Referenz- und Evaluierungsdaten mit Hilfe von adäquaten Metriken sind Gütemaße ableitbar.

Das vorliegende Kapitel ist zweigeteilt. Zunächst werden semiautomatische (Abschnitt 7.1.1), algorithmusgetriebene (Abschnitt 7.1.2), sensorgestützte (Abschnitt 7.1.3) und hybride (Abschnitt 7.1.4) Verfahren zur Referenzdatengewinnung vorgestellt, die für eine Evaluierung von Fahrumfelderfassungssystemen geeignet sind. Aufgrund der speziellen Anforderungen mussten zum Teil neue Vorgehensweisen entwickelt werden. Anschließend wird auf einige Metriken eingegangen, mittels derer die Systemperformanz in Zahlen gefasst werden kann. Im zweiten Teil des Kapitels wird eine Systembewertung durchgeführt. Hierbei werden sowohl die Trackingleistung für Fußgänger und Fahrzeuge evaluiert als auch das Gesamtsystem der umgesetzten Fahrumfelderfassung bestehend aus Tracking

und Klassifikation betrachtet. Die Gesamtsystemleistung wird dabei unter realistischen Fahrsituationen und in verschiedenen Umgebungen, wie Stadt, Land und Autobahn, sowie unter verschiedenen Witterungsbedingungen ermittelt. Zudem werden die Synergieeffekte des multisensoriellen Sensorsystems bewertet.

### 7.1 Referenzdatengewinnung

Damit eine Systemevaluierung durchgeführt werden kann, sind Referenzdaten zwingend erforderlich. Diese ermöglichen, durch einen Abgleich mit der Systemausgabe, eine metrikgestützte Gewinnung der Gütemaße. Zur Erzeugung von Referenzdaten zur Evaluierung von Fahrumfelderfassungssystemen gibt es verschiedene Vorgehensweisen:

- Durch die rechnergestützte Simulation von Fahrszenen sowie den entsprechenden Sensordaten können nahezu beliebige Testsequenzen synthetisch erzeugt werden. Aufgrund der Erzeugungsmodelle, sind sämtliche Parameter des Fahrumfeldes bekannt. Die Umsetzung solcher Simulationsmodelle kann jedoch, insbesondere bei multisensoriellen Systemen, sehr aufwendig sein, da zusätzlich zu Objektgeometrien und -dynamiken ebenso alle Sensormodelle nachgebildet werden müssen. Darüber hinaus kann durch diese Vorgehensweise eine reale Datenaufnahme lediglich approximiert werden.
- Des Weiteren können Referenzdaten von einem dedizierten Sensor abgeleitet werden. Dieser ist oftmals teuer und nur unter bestimmten Randbedingungen einsetzbar und somit kein permanenter Bestandteil des Fahrumfelderfassungssystems. Meist liefert solch ein Sensor, neben sehr hohen Zustandsgenauigkeiten, Objektklassenzugehörigkeiten und Latenzinformationen. Dadurch ist er, neben der Evaluierung von Klassifikationsalgorithmen, ebenso zur Bewertung von Trackingverfahren geeignet.
- Referenzdaten können zudem mittels einer manuellen Attributierung von Aufnahmesequenzen gewonnen werden. Dabei werden in den Sensordaten die Positionen und Klassenzugehörigkeiten aller relevanten Objekte per Hand markiert. Die dazu erforderlichen Arbeitsschritte können für den Anwender sehr monoton und zeitaufwendig sein. Zudem wird die erzielbare Positionsgenauigkeit durch die Genauigkeit der manuellen Markierungen begrenzt. Der Vorteil einer manuellen Attributierung von Testsequenzen liegt in der hohen Flexibilität hinsichtlich der zu markierenden Bereiche sowie in der uneingeschränkten Szenenauswahl.

In der Literatur finden sich vornehmlich bildgestützte Verfahren zur Referenzdatenerzeugung. Dörmann [DM00] und Jaynes [JWSX02] entwickelten die Systeme Viper und ODViS zur Leistungsbewertung von Erkennungssystemen, die ausschließlich auf Videosequenzen basieren. Diese Werkzeuge bieten Schnittstellen zur Referenzdatenerzeugung, Metriken zur Evaluierung und diverse Visualisierungskomponenten. Im Applikationsbereich der Videoüberwachung stellt Black et al. [BER03] ein System vor, das eine pseudosynthetische Bilddatenerzeugung ermöglicht. Dazu werden Bildsequenzen von Fußgängertrajektorien extrahiert und vor neue wechselnde Hintergrundansichten eingefügt. Dadurch kann eine Vielzahl von Szenen synthetisiert werden. Bertozzi et al. [BBG<sup>+</sup>04] beschreiben ein System zur Evaluierung von bildbasierten Fußgängererkennungsalgorithmen. Das vorgestellte Werkzeug ermöglicht manuelle und algorithmusgetriebene Attributierungen von Bildsequenzen, sowie eine anschließende Analyse der Erkennungsraten. Ein Anwender erzielt etwa 100 Bildannotationen in

der Stunde. Eine manuelle Bild-für-Bild-Annotierung von Sensordaten ist somit entsprechend zeitaufwändig. Im nächsten Abschnitt wird ein Werkzeug vorgestellt, das eine semiautomatische Erzeugung von Referenzdaten ermöglicht. Es ist dabei nicht auf Bilddaten beschränkt, sondern erlaubt zusätzlich eine effiziente Annotierung von räumlichen Sensordaten. Die Konzeption und Umsetzung des Werkzeugs zur Referenzdatenerzeugung erfolgte in Kooperation mit dem FORWISS<sup>1</sup> Institut der Universität Passau. Eine detaillierte Betrachtung findet sich in der gemeinsamen Veröffentlichung [TFF<sup>+</sup>07].

### 7.1.1 Semiautomatische Referenzdatenerzeugung

Zur Ableitung von räumlichen Informationen wird vorausgesetzt, dass an der Erfassungssensorik zumindest zwei Kameras oder alternativ mindestens ein bild- und ein entfernungsgebender Sensor beteiligt sind. Im aktuell verwendeten System werden für die Referenzdatenerzeugung eine Videokamera und ein Laserscanner eingesetzt. Neben den Sensoranforderungen wird von der Annahme ausgegangen, dass sich alle relevanten Objekte des Fahrumfeldes, wie beispielsweise Fußgänger und Fahrzeuge, durch einen umschließenden Quader (siehe Abbildung 7.2a) ausreichend genau approximieren lassen.

Ein Keyframe-Konzept ist, neben der Möglichkeit Daten in drei Dimensionen zu markieren, die wesentliche Neuerung des Annotierungswerkzeugs. Statt Objekte Bild-für-Bild zu markieren, werden die Objekttrajektorien durch einzelne Kontrollquader beschrieben. Eine lineare Interpolation, hinsichtlich der Position und Orientierung, zwischen den Kontrollquadern (siehe Abbildung 7.2b) stellt eine durchgängige Bild-für-Bild-Markierung sicher.

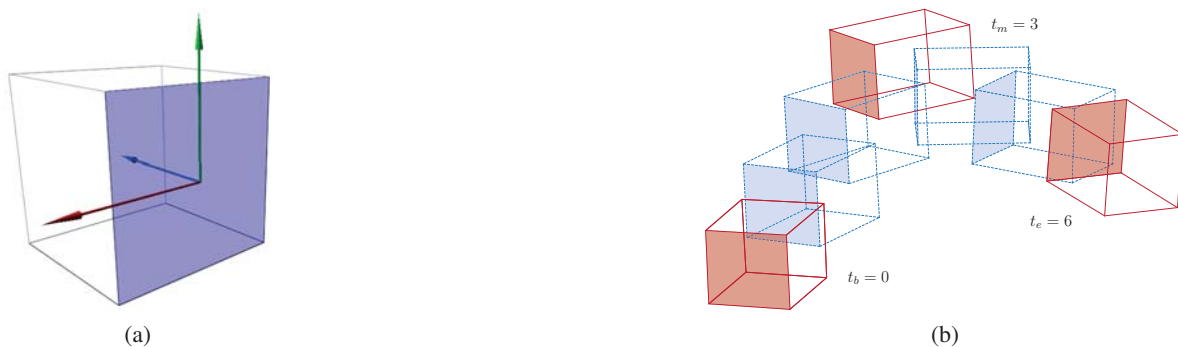


Abbildung 7.2: Quadermodell zur Approximierung von Verkehrsteilnehmern. In Abbildung (a) ist ein Quadermodell mit Ursprung und Ausrichtung des lokalen Koordinatensystems dargestellt. Abbildung (b) zeigt linear interpolierte Quader (gestrichelte Linien) zu den Zykluszeiten  $t = 1, 2, 4, 5$  zwischen manuell platzierten Kontrollquadern (durchgezogene Linien) zu den Zykluszeiten  $t_b = 0$ ,  $t_m = 3$  und  $t_e = 6$ . Die Quaderfläche, in deren Zentrum der Ursprung des lokalen Koordinatensystems liegt, ist farblich hervorgehoben.

<sup>1</sup> Institut für Softwaresysteme in technischen Anwendungen der Informatik.

### Vorgehensweise zur Attributierung einer Szene

Zwei Arbeitsschritte sind nötig, um einen Kontrollquader positionsgenau zu plazieren. Zunächst wird das zu attributierende Objekt im bildgebenden Sensor mit einem umschließenden Rechteck markiert, welches die Rückseite des Kontrollquaders darstellt. Dieser Quader wird auf eine vorgegebene Initialdistanz (z.B. 3 m) zum Kamerasensor gesetzt. Dadurch ist nicht gewährleistet, dass die Abmessungen des Quaders den wirklichen Objektabmessungen entsprechen. Lediglich die Projektion in das Kamerabild ist korrekt (siehe Abbildung 7.3a). Im zweiten Arbeitsschritt wird die Entfernung des Objektes zur Kamera manuell korrigiert. Dazu wird der Kontrollquader entlang der Sichtstrahlen der Kamera verschoben. Zeitgleich wird er in einer Art und Weise skaliert, dass die Projektionsgröße des Quaders im Bildbereich konstant bleibt. Der zweite Arbeitsschritt ist abgeschlossen, wenn die Rückfläche des Kontrollquaders auf den Entfernungsdaten zu liegen kommt (siehe Abbildung 7.3b), bzw. im Fall von zwei Kameras, die korrespondierenden Objektbereiche überdeckt. Anschließend kann noch die Tiefenabmessung und die Orientierung des Quaders angepasst werden.

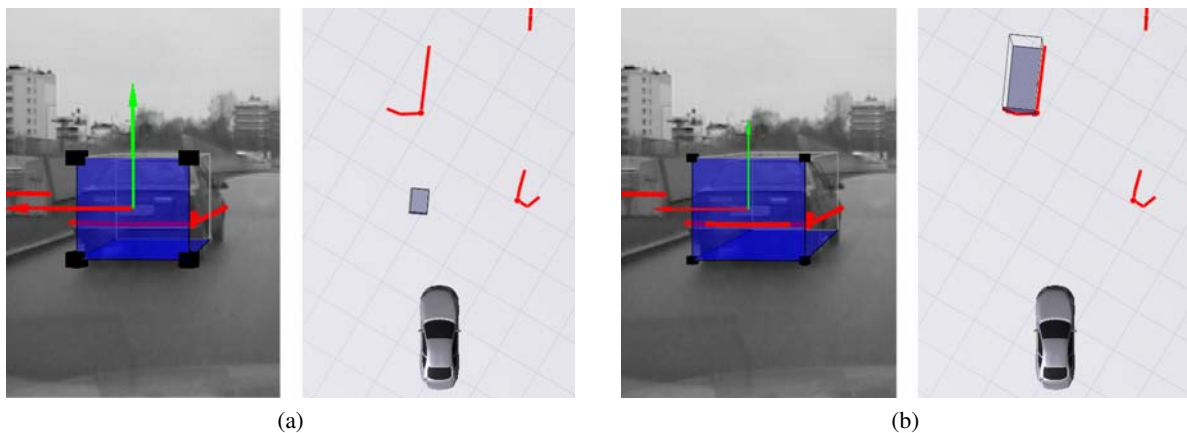


Abbildung 7.3: Arbeitsschritte zur Markierung eines Objektes in drei Dimensionen. Als ersten Schritt wird das Objekt im bildgebenden Sensor markiert (a). Anschließend wird die Entfernung, korrespondierend zu den Entfernungsdaten, justiert (b).

Durch die beiden Arbeitsschritte wird ein einzelner Kontrollquader erzeugt, der anhand der Zustandsgrößen Position, Dimension und Orientierung beschrieben wird. Zur Evaluierung einer gesamten Sequenz ist es jedoch notwendig, für jeden diskreten Zeitschritt Referenzzustandswerte verfügbar zu haben. Dies kann erzielt werden, indem die obigen Arbeitsschritte für jeden Zeitschritt gesondert durchgeführt werden.

Dieses zeitaufwendige Vorgehen kann optimiert werden, indem für ein Objekt lediglich eine Folge von Kontrollquadern plaziert wird und durch eine lineare Interpolation die Objektzustände zwischen den Kontrollquadern zum Zeitpunkt  $t_b$  und  $t_e$  approximiert werden (siehe Abbildung 7.3b). Je nach Fahrmanöver kann die Interpolation dabei sowohl in Fahrzeugkoordinaten als auch in Weltkoordinaten angestoßen werden. Sei  $p : \mathcal{T} \rightarrow \mathbb{R}^3$  eine Funktion, die dem Objekt in jedem diskreten Zeitwert einen Positionswert zuordnet. So gilt für die Positionen der zu interpolierenden Zeitschritte  $t_i$ :

$$\forall t_i \in \mathcal{T}, t_b < t_i < t_e \quad p(t_i) = p(t_b) + \frac{p(t_e) - p(t_b)}{t_e - t_b} \cdot (t_i - t_b) \quad (7.1)$$

Die Objektausdehnungen und -orientierungen der Quader zwischen den Kontrollquadern werden analog zur Position interpoliert. Zur effizienten Annotierung von kompletten Sequenzen hat sich ein objektweises Vorgehen und eine „Divide and Conquer“-Strategie bewährt:

1. Zunächst werden zwei Kontrollquader plaziert. Einer zum Zeitpunkt  $t_b$ , in dem das Objekt zum ersten Mal in der Sensorik sichtbar wird. Der zweite Kontrollquader zum Zeitpunkt  $t_e$ , wenn das Objekt den Sichtbereich der Sensorik verlässt. Das Zeitintervall  $[t_b, t_e]$  wird als eine Sektion bezeichnet.
2. Falls die Interpolationsgenauigkeit in dieser Sektion  $[t_b, t_e]$  nicht ausreichen sollte, kann ein weiterer Kontrollquader in der Sektionsmitte  $t_m$  eingefügt und repositioniert werden. Dadurch werden zwei neue Sektionen  $[t_b, t_m]$  und  $[t_m, t_e]$  erzeugt, für die jeweils eine separate lineare Interpolation angestossen wird. Insgesamt wird dadurch die Objekttrajektorie besser approximiert (siehe auch Abbildung 7.2b).
3. Der zweite Schritt wird mit den Untersektionen solange rekursiv wiederholt, bis die gewünschte Genauigkeit erreicht worden ist.

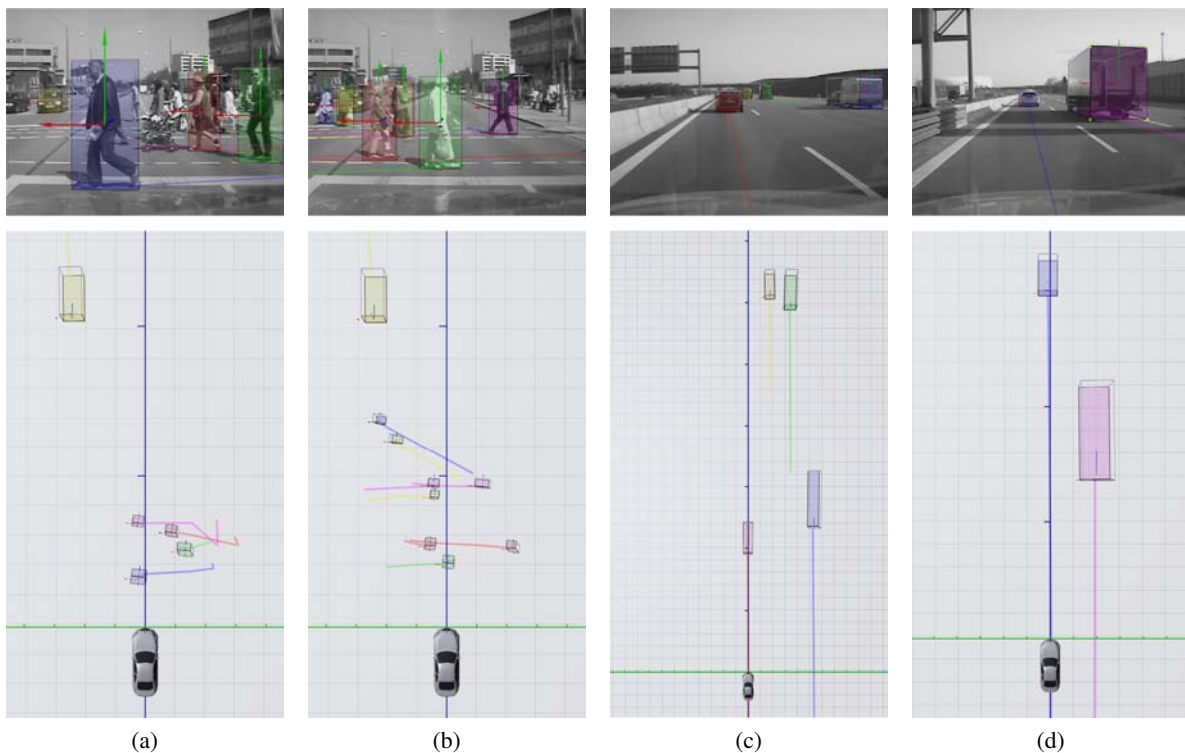


Abbildung 7.4: Zwei semiautomatisch attributierte Szenen: Die Abbildungen (a) und (b) zeigen Szenenschnappschüsse einer Stadtszene (Fußgängerüberweg) mit 12 Fußgängern und einem Fahrzeug zu unterschiedlichen Zeitpunkten. In den Abbildungen (c) und (d) sind Ausschnitte einer Autobahnszene mit 20 Fahrzeugen abgebildet. Die obere Bildreihe gibt die Ansicht auf die jeweilige Szene durch eine Videokamera wieder. Die untere Bildreihe zeigt das Ergebnis der Attributierung aus einer Vogelperspektive.

## Bewertung

Mit der „Divide and Conquer“-Strategie kann eine komplexe Szene, wie jene in Abbildung 7.4a und 7.4b mit 12 Fußgängern und einem Fahrzeug, in weniger als 10 Minuten annotiert werden. Für die Autobahnscene in Abbildung 7.4c und 7.4d, die 20 Fahrzeuge enthält, wurden 15 Minuten Annotierungszeit benötigt. Bei beiden Szenarien bewegen sich die Objekte mit annähernd konstanter Geschwindigkeit bzw. sind weitgehend stationär. Dadurch wird eine geringe Anzahl von Kontrollquadranten zur Approximierung der Trajektorie benötigt. In Szenen mit nichtlinearen Bewegungsphasen müssen entsprechend mehr Kontrollquadranten plaziert werden, wodurch sich der Annotierungsaufwand erhöht. Eine zusätzliche Steigerung der Annotierungseffizienz kann durch nichtlineare (z.B. kubische) Interpolationsverfahren erzielt werden.

Um die zu erwartende Genauigkeit dieses Verfahrens zu bestimmen, wurden mehrere Szenen mit einem hochgenauen dedizierten Referenzsensor (siehe Abschnitt 7.1.3) vermessen und nachträglich mit dem semiautomatischen Keyframe-Verfahren attribuiert. In Tabelle 7.1 ist ein Vergleich der ermittelten Abweichungen zwischen Referenz und manueller Attributierung für Fußgänger sowie Fahrzeuge in zwei unterschiedliche Szenarien dargestellt. In dem ersten Szenario bewegen sich die Zielobjekte auf einfachen linearen Trajektorien. In dem Zweiten verhalten sie sich hingegen sehr dynamisch. Die Qualität der Entfernungs- und Geschwindigkeitsbestimmung wird nur sehr gering von der linearen Interpolation beeinflusst. Für die laterale Ablage (Winkel  $\phi$ ) ist hingegen ein deutlicher Zusammenhang zu erkennen. Die Qualität des Attributs Geschwindigkeit konnte für Fußgänger nicht evaluiert werden, da der dedizierte Referenzsensor hierfür keine Messungen liefert.

	Abstand $r$ (RMSE)	Winkel $\phi$ (RMSE)	Geschwindigkeit $v$ (RMSE)
Fahrzeuge Szenario 1 (1443 Zyklen)	0.09 m	0.3 °	0.37 $\frac{m}{s}$
Fahrzeuge Szenario 2 (1865 Zyklen)	0.10 m	0.5 °	0.36 $\frac{m}{s}$
Fußgänger Szenario 1 (1250 Zyklen)	0.14 m	0.4 °	nicht verfügbar
Fußgänger Szenario 2 (2680 Zyklen)	0.20 m	0.8 °	nicht verfügbar

Tabelle 7.1: Vergleich der Genauigkeiten zwischen semiautomatisch erzeugten Referenzdaten für Fahrzeuge und Fußgänger mit den Genauigkeiten, die mit einem dedizierten Referenzsensor erreicht werden können. Es wurden jeweils zwei Szenarien untersucht: In Szenario 1 bewegen sich die Zielobjekte auf einfachen linearen Trajektorien. In Szenario 2 verhalten sich die Zielobjekte sehr dynamisch.

Die Referenz, die durch diese manuelle Attributierung gewonnen werden kann, ist gut geeignet die Performanz von Hypothesengenerierungs- und Klassifikationsverfahren zu bestimmen. Des Weiteren können aus den markierten Sensordaten Trainingsdaten für einen Klassifikator gewonnen werden. Für die Evaluierung von Trackingverfahren, speziell zur Bestimmung von Zustandsgenauigkeiten, ist ein manuelles Attributieren jedoch nur bedingt geeignet. Die Genauigkeit der Referenz ist stark von der



Markierungsgenauigkeit und den dafür verwendeten Sensoren abhängig. Bei hohen Relativgeschwindigkeiten wird die Qualität der Referenz durch Latenzzeiten, die durch die Sensorik oder das Aufzeichnungsverfahren bedingt sind, negativ beeinflusst.

### 7.1.2 Algorithmusgestützte Referenzdatenerzeugung

Im vorherigen Abschnitt wurde eine Methode erläutert, wie räumliche Sensordaten sehr effizient annotiert werden können. Bei diesem Verfahren besteht die Hauptaufgabe des Anwenders darin, neue Objekte zu erzeugen, diese zu klassifizieren, die Qualität der Interpolation zwischen den einzelnen Keyframes zu überwachen und gegebenenfalls neue Keyframes einzufügen. Insbesondere bei Szenen, mit vielen Objekten, die sich nicht linear bewegen, können diese Arbeitsschritte viel Zeit in Anspruch nehmen. Im Folgenden wird deshalb eine Erweiterung dieses Ansatzes vorgestellt, bei der durch eine Vorverarbeitung automatisch Referenzdaten erzeugt werden. Hierfür können Algorithmen gewählt werden, die nicht den Ansprüchen eines Echtzeitsystems genügen müssen. Bei diesem Ansatz arbeitet der Anwender nicht mehr ausschließlich auf den Rohdaten. Vielmehr bekommt er vom System sowohl für die Klassifikation, als auch für die Wahl der Keyframes Vorschläge unterbreitet. Dadurch lässt sich der Arbeitsaufwand auf ein Überwachen und Korrigieren der Referenzdaten aus der Vorverarbeitung reduzieren. Aus diesem Zusammenhang kann folgende Grundregel abgeleitet werden: Je besser die Vorverarbeitung, desto schneller und effizienter die Attributierung.

Im Folgenden soll ein Fahrumfelderfassungssystem als eine mögliche Art der Vorverarbeitung betrachtet werden. Das Fahrumfelderfassungssystem liefert zu jedem Zeitpunkt  $t$  eine Schätzung des Objektzustandes  $\mathbf{x}$ . Verglichen mit einer manuellen Vorgehensweise entspricht dies einer Bild-für-Bild Annotation der Szene. Im Gegensatz zu dem Keyframe Konzept ist ein nachträgliches manuelles Korrigieren der Daten sehr ineffizient, da statt weniger Stützstellen jeder Zeitpunkt  $t$  betrachtet werden muss. Deshalb ist es sinnvoll, die durch das Fahrumfelderfassungssystem erzeugten Daten  $\mathcal{M} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  durch eine Menge an Keyframes  $\mathcal{K} = \{\mathbf{k}_1, \dots, \mathbf{k}_m\}$  anzunähern. Hierzu wird ein iteratives Verfahren verwendet, dass  $\mathcal{M}$  durch eine Menge an Regressionsgeraden beschreibt. Beginnend bei dem ersten Element von  $\mathcal{M}$  werden solange weitere Elemente  $\mathbf{x}_i \in \mathcal{M}$  zu einer Regressionsgerade hinzugefügt, bis die Standardabweichung dieser Approximation einen Schwellwert  $\tau$  übersteigt. Die Menge  $\mathcal{K}$  an Keyframes wird um den aktuellen Punkt  $\mathbf{x}_i$  erweitert und eine neue Ausgleichsgerade begonnen.

In Abbildung 7.5 ist die Trajektorie eines vorausfahrenden Fahrzeuges auf einer kurvenreichen Strecke dargestellt. Die Trajektorie besteht aus insgesamt 789 einzelnen Punkten. Für diese Szene sind 63 Keyframes ausreichend, um die Daten bei einer Standardabweichung von 3 cm anzunähern. Für Teilabschnitte der Szene, die Kurven enthalten, werden deutlich mehr Keyframes benötigt, als für lineare Abschnitte.

Das semiautomatische Annotieren in Verbindung mit der soeben vorgestellten Vorverarbeitung ermöglicht eine effiziente räumliche Attributierung von vielfältigen Fahrscenen. Für die Evaluierung von Trackingverfahren, speziell zur Bestimmung von Zustandsgenauigkeiten, ist ein manuelles Attributieren allerdings nur bedingt geeignet. So können zum Beispiel nichtlineare Objekttrajektorien, nur ungenau durch einfache Linienzüge approximiert werden. Zum Anderen ist die Genauigkeit von den direkt ableitbaren Attributen wie Position, Ausdehnung oder Orientierung abhängig von der Qualität der verwendeten Sensorik. Mit einer semiautomatischen Referenzdatenerzeugung ist es somit nicht



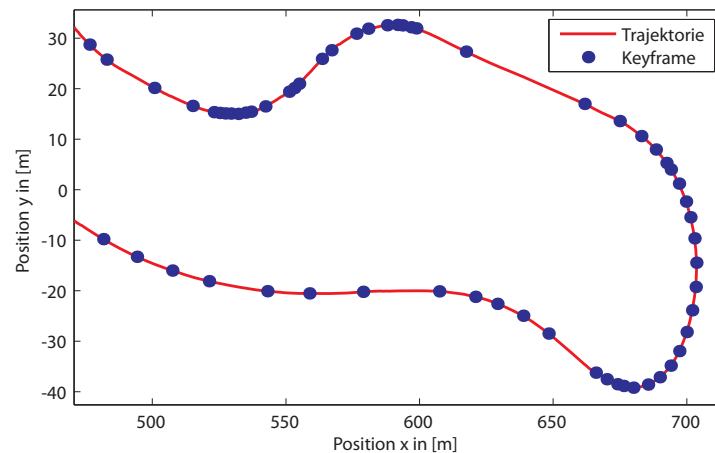


Abbildung 7.5: Die Trajektorie eines vorausfahrenden Fahrzeuges auf einer kurvenreichen Strecke. 63 Keyframes sind ausreichend, um die Szene mit einer durchschnittlichen Abweichung von 3 cm zu approximieren. Dies entspricht einer Reduktion um den Faktor 10 im Vergleich zu allen Punkten der Trajektorie.

möglich, Daten zu erzeugen, die präziser sind als die zugrundeliegende Sensorik. Dies ist im Allgemeinen ausreichend, um Aussagen hinsichtlich der Klassifikationsgüte zu treffen. Für die Evaluierung von Trackingverfahren werden hingegen höhere Anforderungen an die Genauigkeit der Referenzdaten gestellt. Insbesondere wenn abgeleitete oder nicht direkt messbare Größen wie die Beschleunigung oder die Gierrate untersucht werden sollen, ist es notwendig, einen zusätzlichen, von der Fahrumfelderfassung unabhängigen Sensor zu verwenden, der die Attribute des Referenzobjektes präzise ermittelt.

### 7.1.3 Referenzsensor

Im Folgenden wird am Beispiel einer Inertialplattform mit GPS Empfänger erläutert, wie durch zusätzliche Sensorik hochgenaue Referenzdaten erzeugt werden können. Das vorgestellte Verfahren erweitert den Ansatz von Vogel [Vog07], mit dem Fußgänger zentimetergenau lokalisiert und getrackt werden können, um weitere Attribute wie Orientierung, Geschwindigkeit und Beschleunigung von beliebigen Referenzobjekten. Abschließend werden die Vor- und Nachteile dieser Technik diskutiert.

#### Inertialplattform mit Carrier Phase Differential GPS

Grundlage des Verfahrens sind zusätzliche zeitlich synchronisierte Sensoren, die sowohl im Eigenfahrzeug, als auch im Referenzobjekt verbaut sind. Im Eigenfahrzeug befindet sich eine Inertialplattform<sup>2</sup> (RT3003 [Oxf08]), die mittels Carrier Phase Differential GPS<sup>3</sup> (CPD-GPS) die Ausrichtung nach Norden, sowie die Position bezüglich eines globalen Weltkoordinatensystems bestimmt. Eine zweite Inertialplattform mit CPD-GPS Empfänger, die am Zielobjekt angebracht ist, liefert Daten über die Position

<sup>2</sup> Eine Inertialplattform besteht aus mehreren Gyroskopen und Beschleunigungssensoren. Durch eine Kombination mit einem CPD-GPS Empfänger können die Position, Geschwindigkeit, Beschleunigung sowie die Orientierung der Plattform sehr präzise gemessen werden.

<sup>3</sup> Satellitengestütztes System zur Bestimmung der Position auf der Erdoberfläche.

und die Ausrichtung des Referenzobjektes. Darüber hinaus können mit dem zusätzlichen Sensor weitere Attribute wie die Geschwindigkeit  $v_{ref}$ , die Beschleunigung  $a_{ref}$  und die Gierrate  $\dot{\phi}_{ref}$  gemessen werden. Durch die zeitliche Synchronisation zwischen dem Referenzobjekt und dem Eigenfahrzeug ist es letztendlich möglich, aufgezeichnete Sensordaten mit den Messwerten des Referenzsensors zu attributieren. Weitere Details zu diesem Verfahren wurden bereits im Kapitel 3.2 beschrieben.

## Bewertung

Durch den Einsatz von dedizierter Sensorik erfolgt die Annotation von Szenen bereits zum Zeitpunkt der Aufnahme. Deshalb ist mit diesem Ansatz ein zeitaufwendiges manuelles Nachbearbeiten der aufgezeichneten Daten nicht mehr erforderlich. Vielmehr können Attribute wie die Position, die Geschwindigkeit oder die Beschleunigung direkt aus der Messung abgeleitet werden. Weitere Attribute wie zum Beispiel die Objektmaße oder die Klassifikation stehen als a priori Wissen zur Verfügung, wenn die Referenzobjekte mit zusätzlicher Sensorik ausgestattet werden. Darüber hinaus ist diese Art der Referenzdatenerzeugung unabhängig von der zu evaluierenden Sensorik. So können zum Beispiel höhere Reichweiten von mehr als 200 m sowie Erfassungsbereiche von bis zu  $360^\circ$  realisiert werden. Insbesondere bei der Bewertung von Trackingverfahren sind hochgenaue und unabhängige Referenzdaten entscheidend. Durch dedizierte Sensorik können zum Beispiel Attribute wie die Gierrate oder die Beschleunigung direkt gemessen werden, die sonst nur indirekt bestimmt werden könnten. Neben der Messgenauigkeit der dedizierten Sensorik beeinflussen die zeitliche Synchronisation und vorhandene Latenzzeiten die Qualität der Referenzdaten. Für das Fallbeispiel aus dem vorherigen Abschnitt sind in Tabelle 7.2 die durch den Hersteller angegebenen Genauigkeiten der jeweiligen Attribute dargestellt.

	Position $\mathbf{p}$ (RMSE)	Geschwindigkeit $v$ (RMSE)	Beschleunigung $a$ (RMSE)	Orientierung $\phi$ (RMSE)	Gierrate $\dot{\phi}$ (RMSE)
Eigenfahrzeug	0.02 m	$0.02 \frac{m}{s}$	$0.01 \frac{m}{s^2}$	$0.1^\circ$	$0.1 \frac{^\circ}{s}$
Referenz	0.02 m	$0.02 \frac{m}{s}$	$0.01 \frac{m}{s^2}$	$0.1^\circ$	$0.1 \frac{^\circ}{s}$

Tabelle 7.2: Genauigkeiten der verwendeten Inertialplattform *RT3003* für das Referenzobjekt sowie das Eigenfahrzeug.

Mit diesem Verfahren können nur vordefinierte Szenarien mit wenigen Objekten attribuiert werden, da alle Objekte mit zusätzlicher Sensorik ausgestattet werden müssen. Realitätsnahe Aufnahmen wie zum Beispiel in einer Stadt mit mehreren Fußgängern und Fahrzeugen sind, wenn überhaupt, nur mit sehr hohem organisatorischen und finanziellen Aufwand realisierbar. Für derartige Szenarien stellt die Kombination von dedizierter Sensorik und manueller Nachbearbeitung einen guten Kompromiss zwischen Genauigkeit und Kosten dar. Diese hybride Referenzdatenerzeugung wird im nächsten Abschnitt genauer betrachtet.

### 7.1.4 Hybride Referenzdatenerzeugung

In sehr komplexen und realitätsnahen Szenarien mit mehreren Fußgängern und Fahrzeugen, wie zum Beispiel in einer Stadt, ist es nicht praktikabel, jeden Verkehrsteilnehmer mit einer eigenen zusätzlichen Referenzsensorik auszustatten. Die verfügbare Sensorik sollte auf die für diese Szene wichtigsten Teilnehmer beschränkt werden. Die restlichen Objekte können, wie in Abschnitt 7.1.1 beschrieben, nachträglich manuell attribuiert werden. Durch diese Kombination können sehr flexibel und effizient komplexe Szenarien bearbeitet werden. In Abbildung 7.6 trägt ein Fußgänger zum Beispiel einen GPS-Empfänger auf dem Rücken. Für ihn stehen somit hochgenaue Referenzdaten (blauer Zylinder) zur Verfügung. Die restlichen Objekte dieser Szene wurden nachträglich manuell attribuiert.

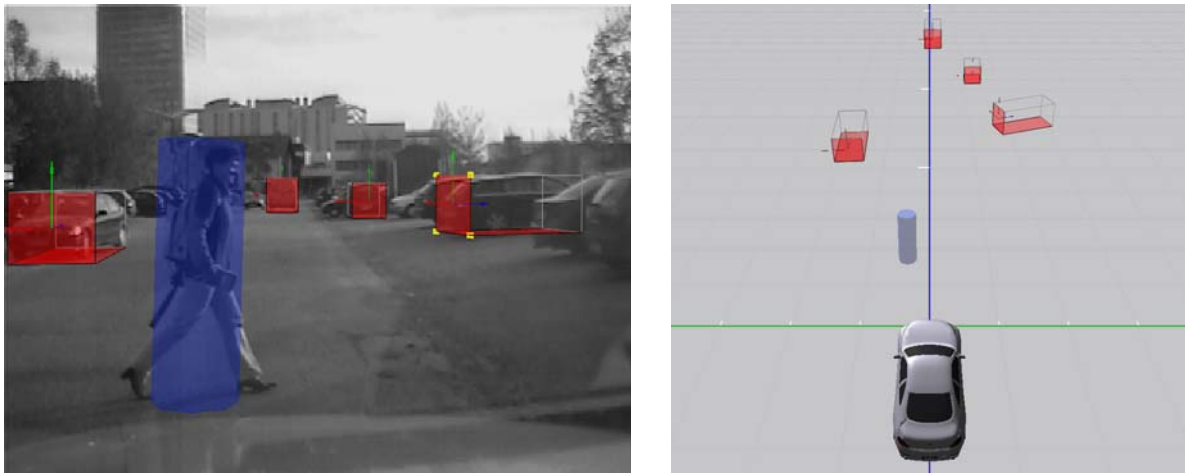


Abbildung 7.6: Kombination aus dedizierter Sensorik und manueller Nachbearbeitung. Der Fußgänger trägt einen GPS-Empfänger auf dem Rücken. Für ihn stehen somit hochgenaue Referenzdaten (blauer Zylinder) zur Verfügung. Die vier Fahrzeuge in der Szene wurden nachträglich manuell attribuiert.

### 7.1.5 Zusammenfassung und Bewertung

In den vorherigen Abschnitten wurden mehrere Möglichkeiten aufgezeigt, wie für ein Fahrumfelderfassungssystem Referenzdaten erzeugt werden können. Die vorgestellten Verfahren adressieren im wesentlichen die zwei größten Anforderungen, die im Allgemeinen an ein System zur Gewinnung von Referenzdaten gestellt werden: Zum Einen sollte der gesamte Prozess so effizient wie möglich sein. Zum Anderen werden insbesondere für die Evaluierung von Trackingverfahren sehr präzise und von der verwendeten Sensorik unabhängige Referenzdaten benötigt. Um aussagekräftige Daten zu erhalten, müssen sehr große Datenmengen unterschiedlichster Szenarien attribuiert werden. Hierzu wurden zwei Ansätze aufgezeigt, die die zu bearbeitende Datenmenge durch ein Keyframe-Konzept um den Faktor 5-10 reduzieren. In Verbindung mit einer „Divide and Conquer“-Strategie können somit große Szenenkataloge sehr effizient attribuiert werden. Wird zusätzlich eine maschinelle Vorverarbeitung (Fahrumfelderfassung) eingesetzt, so muss der Anwender die Vorschläge der Vorverarbeitung nur noch überwachen und gegebenenfalls ergänzen.

Mit den semimanuellen Verfahren können Attribute, wie die Position und die Klassifizierung von Objekten, präzise erfasst werden. Allerdings korreliert die Qualität der Referenzdaten stark mit der Genauigkeit der zugrundeliegenden Sensorik. Darüber hinaus können Attribute wie die Geschwindigkeit, die Beschleunigung oder die Gierrate nur sehr eingeschränkt und in schlechter Qualität ermittelt werden. Aus diesem Grund wurde ein Verfahren vorgestellt, das mit Hilfe eines dedizierten Referenzsensors sehr effizient hochgenaue und unabhängige Referenzdaten erzeugt. Allerdings ist die Attributierung von realitätsnahen Szenarien mit diesem Verfahren nur mit sehr hohem organisatorischen und finanziellen Aufwand möglich, da alle beteiligten Objekte mit zusätzlicher Sensorik ausgestattet werden müssen. Für derartige Szenarien ist die Kombination von dedizierter Sensorik und manueller Nachbearbeitung ein guter Kompromiss zwischen Genauigkeit und Kosten.

Die Gewinnung von Referenzdaten ist trotz der vorgestellten Verbesserungen ein arbeits- bzw. kostenintensiver Prozess. Eine pauschale Aussage, welches der vorgestellten Verfahren bevorzugt werden soll, kann nicht getroffen werden. Vielmehr ist die Wahl sehr stark von den Kriterien und Metriken der Evaluierung abhängig. So sind zum Beispiel für einen Vergleich der Genauigkeit von verschiedenen Trackingverfahren hochgenaue Referenzdaten erforderlich. Um Aussagen über die Erkennungsraten eines Klassifikators oder über die Gesamtperformanz eines Fahrumfelderfassungssystems treffen zu können, sind hingegen manuell attributierte Daten ausreichend.

## 7.2 Evaluierung

In diesem Abschnitt werden zum Einen die Trackingleistung für Fußgänger und Fahrzeuge evaluiert und zum Anderen die Gesamtsystemleistung der umgesetzten Fahrumfelderfassung bestehend aus Tracking und Klassifikation betrachtet. Die Evaluierung wird hinsichtlich Genauigkeiten, Erkennungsraten und Laufzeit anhand von repräsentativen Testszenerarien durchgeführt. Nach einem Überblick über die verwendeten Szenarien und einer Beschreibung der Testsystemparametrisierung werden die Evaluierungsergebnisse und Problemkonstellationen aufgezeigt. Zunächst wird auf die Metriken eingegangen, die für die Leistungsbewertungen herangezogen werden.

### 7.2.1 Metriken

Metriken sind in diesem Zusammenhang Indikatoren, anhand derer die Systemperformanz in Zahlen gefasst werden kann. Sie sollten möglichst intuitiv sein und alle signifikanten Leistungsfaktoren des Systems abdecken. Ein Abgleich der Ergebnisse einer maschinellen Fahrumfelderfassung mit einer, im Vorfeld generierten, Referenz bildet die Grundlage für die Berechnung der Metriken.

Im Rahmen einer Fahrumfelderfassung können Metriken in drei Kategorien untergliedert werden. Die eine Gruppe spiegelt dabei vornehmlich die (Fehl-)Erkennungsrate wider, mit der Objekte klassifiziert werden. Die andere Gruppe beinhaltet die Genauigkeiten, mit der Objekte getrackt werden. Die dritte Kategorie umfasst applikationsspezifische Kriterien wie die benötigte Rechenzeit oder dedizierte Sensorerfassungsbereiche. In den nächsten drei Abschnitten werden die für die Evaluierung verwendeten Metriken beschrieben.

## Metriken für die Klassifikation

**Erkennungsrate (Richtig-Positiv-Rate)** Ein Objekt gilt als erkannt, wenn durch die Klassifikation innerhalb der Objektlebenszeit die korrekte Objektklasse bestimmt wurde, und der Abstand zwischen Referenzobjekt und Evaluierungsobjekt unter einem Wert  $\delta$  liegt. Die Erkennungsrate wird objektspezifisch für Fahrzeuge als  $RPR_V$  und für Fußgänger als  $RPR_P$  notiert. Diese Bewertungskriterien geben das Verhältnis von richtig klassifizierten Objekten zu allen verfügbaren Objekten an. In der Auswertung wurde  $\delta = 1.5\text{ m}$  gewählt.

**Fehlerkennungsrate pro Stunde (FER)** Diese Metrik bildet die Fehlerkennungsrate auf eine zeitliche Basis ab und wird mit  $FER$  bezeichnet. Sie ist durch den Quotienten der Fehlererkennungen pro Stunde definiert. Die Erkennungsrate wird objektspezifisch für Fahrzeuge als  $FER_V$  und für Fußgänger als  $FER_P$  angegeben.

**Matthews Correlation Coefficient (MCC)** Der Matthews Correlation Coefficient [BBC<sup>+</sup>00] kombiniert die Erkennungs- und die Fehlerkennungsrate in einer gemeinsamen Metrik und ermöglicht somit einen direkten Vergleich zwischen zwei Klassifikatoren. Er ist wie folgt definiert:

$$MCC = \frac{t_p t_n - f_p f_n}{\sqrt{(t_p + f_p)(t_p + f_n)(t_n + f_p)(t_n + f_n)}} \quad (7.2)$$

Hierbei bezeichnet  $t_p$  und  $t_n$  die Anzahl der korrekt bzw. nicht erkannten Fahrzeuge und Fußgänger. Analog dazu beschreiben  $f_p$  und  $f_n$  die korrekt klassifizierten Hintergrundobjekte bzw. die Anzahl der Fehlerkennungen. Der Wertebereich dieser Metrik liegt zwischen  $-1$  und  $1$ , wobei ein perfekter Klassifikator durch einen Wert von  $1$  charakterisiert ist. Der  $MCC$  wird objektspezifisch für Fahrzeuge als  $MCC_V$  und für Fußgänger als  $MCC_P$  notiert.

## Metriken für das Tracking

**Zeitliche Abdeckung (ZA)** Die zeitliche Abdeckung beschreibt den Prozentsatz in Bezug auf die Zeit, in der Objekte korrekt klassifiziert und verfolgt werden können. Sei  $t_\lambda$  die Zeit in der ein Objekt  $\lambda$  erfolgreich klassifiziert und verfolgt werden konnte und  $l_\lambda$  die zugehörige Lebensdauer, dann berechnet sich die zeitliche Abdeckung zu:

$$ZA = \frac{\sum_{\forall \lambda \in \mathcal{L}} t_\lambda}{\sum_{\forall \lambda \in \mathcal{L}} l_\lambda} \quad (7.3)$$

Die zeitliche Abdeckung wird für Fahrzeuge als  $ZA_V$  und für Fußgänger als  $ZA_P$  angegeben.

**Mean Square Error (MSE)** Diese Metrik beschreibt die mittlere quadratische Abweichung des gefilterten Wertes  $\hat{x}_t$  von der zu schätzenden Größe  $x_t$  und ist wie folgt definiert:

$$MSE = \frac{1}{T} \sum_{t=1}^T (\hat{x}_t - x_t)^2 \quad (7.4)$$

Für erwartungstreu<sup>4</sup> Schätzverfahren entspricht der MSE der Varianz der Abweichung. Eine geringe mittlere quadratische Abweichung bedeutet, dass gleichzeitig Bias und Varianz des Schätzers klein sind, und das Verfahren unverfälschte Ergebnisse mit einer geringen Streuung liefert.

<sup>4</sup> Weicht ein Schätzer im Mittel von der zu schätzenden Größe ab, so ist er verzerrt (engl. bias) und nicht erwartungstreu.

**RMSE der Entfernung (r)** Der RMSE ist als  $\sqrt{MSE}$  definiert und bewahrt die Einheit der zu schätzenden Größen. Der RMSE der Entfernung (r) beschreibt die Qualität der Entfernungsschätzung und wird objektspezifisch für Fahrzeuge und Fußgänger als  $RMSE_r^V$  bzw.  $RMSE_r^P$  notiert.

**RMSE der lateralen Ablage ( $\varphi$ )** Die laterale Ablage  $\varphi$  ist relativ zum Eigenfahrzeug in Polarkoordinaten definiert und beschreibt den Azimutwinkel der Position eines Verkehrsteilnehmers. Über den RMSE dieser Schätzung können Aussagen bezüglich der Qualität getroffen werden. Für Fahrzeuge und Fußgänger wird die Notation  $RMSE_\varphi^V$  bzw.  $RMSE_\varphi^P$  verwendet.

**RMSE der Geschwindigkeit (v)** Der RMSE der Geschwindigkeit (v) beschreibt die Qualität der Geschwindigkeitsschätzung und wird objektspezifisch für Fahrzeuge und Fußgänger als  $RMSE_v^V$  bzw.  $RMSE_v^P$  bezeichnet.

### Sonstige Evaluierungskriterien

**Rechenzeit** Da die Fahrumfelderfassung als automotiv Anwendung konzipiert ist, sind Echtzeitanforderungen von großer Bedeutung. Die Rechenzeit  $t_{Zyk}$  wird in Millisekunden angegeben und bezeichnet die Ausführungszeit, die für einen Zyklus (siehe auch Abbildung 2.9) benötigt wird. Alternativ dazu kann die Rechenzeit indirekt als Frequenz  $f_{Zyk}$  (Zyklusdurchläufe pro Sekunde) angegeben werden. Überschreitet  $f_{Zyk}$  einen Wert von ca. 15, werden Echtzeitanforderungen erfüllt.

**Erfassungsbereiche** Damit eine aussagekräftige Performanzevaluierung einer Fahrumfelderfassung durchgeführt werden kann, müssen die Erfassungsbereiche der Sensoren mit berücksichtigt werden. Würden die Erfassungsbereiche nicht berücksichtigt werden, fließen Referenzobjekte, die außerhalb der Erfassungsbereiche liegen, in die Erkennungsratenberechnung mit ein. Diese wird vermindert und verfälscht somit die Systemleistung.

Es ist zu beachten, dass sich die vorgestellten Metriken zu einem gewissen Grad gegenseitig beeinflussen. So führt zum Beispiel eine Zunahme der Erkennungsrate zu einer veränderten Berechnungsgrundlage für die Metriken des Trackings. Bei der Bewertung sollte als weiterer Aspekt berücksichtigt werden, dass beim Abgleich mit der Referenz verfälschte Ergebnisse auftreten können. Beinhaltet die Referenz teilverdeckte Objekte, die von der Fahrumfelderfassung nicht erfasst werden, erniedrigt sich die Erkennungsrate (TPR). Objekte die in der Referenz hingegen nicht berücksichtigt werden, können zu erhöhten Fehlerkennungsraten (FER) führen. Die Trackingmetriken werden dadurch nur geringfügig betroffen. Grundsätzlich sollte eine möglichst umfassende Referenz erzeugt werden, um Vergleiche zwischen verschiedenen Algorithmusvarianten, Sensorkonfigurationen und Testszenarien durchführen zu können.

## 7.2.2 Bewertung des Trackingsystems

In diesem Abschnitt wird das Trackingsystem hinsichtlich der Genauigkeiten und dem Mehrwert der Sensorfusion bewertet. Nach einem Überblick über die verwendeten Szenarien und einer Beschreibung der Testsystemparametrisierung erfolgt eine Diskussion der Ergebnisse.

### Beschreibung der Testszenarien

Für die Bewertung des Trackingsystems wurde eine repräsentative Stichprobe aus allen aufgezeichneten Testszenarien ausgewählt. Tabelle 7.3 listet die verwendeten Testszenarien mit einer kurzen Beschreibung sowie einzelnen Szenenausschnitten auf. Die dargestellten Ausschnitte zeigen das Ergebnis der Fahrumfelderfassung zu einem ausgewählten Zeitpunkt. Fußgänger und Fahrzeuge werden durch gelbe Zylinder bzw. grüne Quader symbolisiert.

Alle Szenen wurden unter optimalen Umweltbedingungen aufgezeichnet und können hinsichtlich der Objektklassen in Fahrzeug- und Fußgängerszenarien untergliedert werden. Innerhalb der jeweiligen Klassen erfolgt eine weitere Unterteilung in Bezug auf die Objektdynamik. In sämtlichen Szenen wurden mittels eines dedizierten Sensors (siehe auch Abschnitt 7.1.3) alle relevanten Verkehrsteilnehmer annotiert. Dadurch wird für die nachfolgende Auswertung eine Bewertungsgrundlage geschaffen.

Tabelle 7.3: In dieser Tabelle wird ein Überblick über die Szenen gegeben, die zur Bewertung der Trackingleistung herangezogen wurden. Neben einer kurzen Szenenbeschreibung sind in der rechten Tabellenhälfte Szenenausschnitte (Video- oder Ferninfrarotkamera) mit maschinell erkannten Objekten zu sehen.




Szenenbeschreibung	Szenenausschnitte
Szenarien für Fahrzeuge	
<p>Szene 1: Dynamische Verfolgungsfahrt eines anderen PKW auf einer kurvenreichen Teststrecke. (Dauer 54 Sekunden)</p>	
<p>Szene 2: Fahrgeschwindigkeit auf einer Autobahn ohne dynamische Anteile bis zu einer Entfernung von 120m. (Dauer 1:22 Minuten)</p>	
<p>Szene 3: Mehrere Spurwechselmanöver eines vorausfahrenden PKWs auf einer Teststrecke. (Dauer 1:03 Minuten)</p>	



Tabelle 7.3: In dieser Tabelle wird ein Überblick über die Szenen gegeben, die zur Bewertung der Trackingleistung herangezogen wurden. Neben einer kurzen Szenenbeschreibung sind in der rechten Tabellenhälfte Szenenausschnitte (Video- oder Ferninfrarotkamera) mit maschinell erkannten Objekten zu sehen.





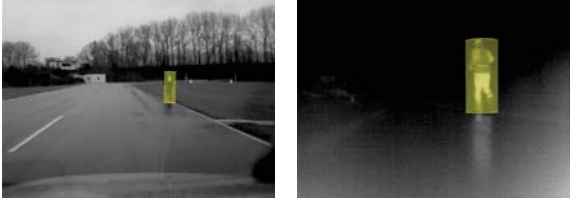

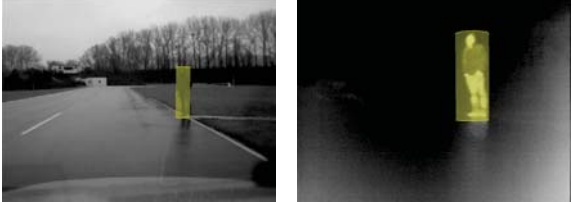
Szenenbeschreibung	Szenenausschnitte
<p>Szene 4: Vollbremsung eines vorausfahrenden Referenzfahrzeuges bei 50 km/h. (Dauer 9 Sekunden)</p>	
<p>Szenarien für Fußgänger</p>	
<p>Szene 5: Rennender und abrupt die Richtung wechselnder Fußgänger. (Dauer 16 Sekunden)</p>	
<p>Szene 6: Ein Fußgänger geht auf der rechten Straßenseite. (Dauer 4 Sekunden)</p>	
<p>Szene 7: Fußgänger quert die Straße von rechts nach links. (Dauer 4 Sekunden)</p>	
<p>Szene 8: Ein Fußgänger rennt auf der rechten Straßenseite. (Dauer 2 Sekunden)</p>	

Tabelle 7.3: In dieser Tabelle wird ein Überblick über die Szenen gegeben, die zur Bewertung der Trackingleistung herangezogen wurden. Neben einer kurzen Szenenbeschreibung sind in der rechten Tabellenhälfte Szenenausschnitte (Video- oder Ferninfrarotkamera) mit maschinell erkannten Objekten zu sehen.

Szenenbeschreibung	Szenenausschnitte
<p>Szene 9: Fußgänger rennt über die Straße von rechts nach links. (Dauer 2 Sekunden)</p>	
<p>Szene 10: Fußgänger steht am Straßenrand. (Dauer 4 Sekunden)</p>	

### Testsystem

Für die Evaluierung der Trackingleistung wird das im Grundlagenkapitel 2 in Abbildung 2.9 illustrierte multisensorielle Fahrumfelderfassungssystem verwendet. Die einzelnen Systemkomponenten sind wie folgt parametrisiert:

**Datenvorverarbeitung** Für die Datenvorverarbeitung der Laserscannerrohdaten wird die inkrementelle Linienaggregation verwendet (siehe auch Abschnitt 4.2.1). Um die Robustheit der Vorverarbeitung gegenüber Messfehlern zu erhöhen, wird zusätzlich ein Verfahren nach RANSAC verwendet (siehe Abschnitt 4.2.2).

**Klassifikation** Die Klassifikation der einzelnen Verkehrsteilnehmer wird manuell durchgeführt. Dadurch kann die Evaluierung unabhängig von der Klassifikation durchgeführt werden.

**Tracking** Für das Verfolgen der Objekte wird ein Extended-Kalman-Filter verwendet (siehe auch Abschnitt 5.2.2), da die Evaluierung in Bezug auf zukünftige Fahrerassistenzsysteme durchgeführt wurde und somit Echtzeitfähigkeit gefordert war. Für Fahrzeuge erfolgt eine Fusion der Messwerte des Laserscanners (siehe Abschnitt 4.2.1) und der Radarsensoren (siehe Abschnitt 4.3). Fußgänger werden durch Laserscannermesswerte und durch die Kopfkontur im Wärmebild (siehe Abschnitt 4.4.1) verfolgt. Die Messwerte mehrerer Sensoren werden nach dem Prinzip der Messwertfusion (siehe Abschnitt 5.2.4) integriert.

### Evaluierung für Fahrzeuge

Die Szenarien, die in Abschnitt 7.2.2 eingeführt wurden, werden in diesem Abschnitt bewertet. Dabei wird der Einfluss von unterschiedlichen Sensorkonfigurationen auf die Systemleistung untersucht. Um den Mehrwert des Fusionssystems in Bezug auf das Tracking von Fahrzeugen zu evaluieren, werden drei Sensorkonfigurationen betrachtet (siehe auch Tabelle 7.4): Die erste Konfiguration ( $K_1$ ) beinhaltet nur den Laserscanner. In der zweiten Konfiguration ( $K_2$ ) wird ausschließlich die Radarsensorik verwendet. Die dritte Konfiguration ( $K_3$ ) fusioniert die Messwerte der Radarsensorik mit dem Laserscanner.

Konfiguration	Laserscanner	Radar
$K_1$	x	
$K_2$		x
$K_3$	x	x

Tabelle 7.4: Übersicht der verwendeten Sensorkonfigurationen.

In Tabelle 7.5 sind die ermittelten Ergebnisse zusammengefasst. Die Tabellenspalten enthalten die untersuchten Szenen. Jeweils drei Tabellenzeilen fassen die Werte einer Metrik für alle Sensorkonfigurationen zusammen. Als Qualitätsmaß der Evaluierung wurde der RMSE für die Schätzgröße der Entfernung  $r$ , der laterale Ablage  $\phi$  sowie der Geschwindigkeit  $v$  gewählt (siehe Abschnitt 7.2.1). In den Szenarien 1 und 3 befand sich das zu verfolgende Fahrzeug teilweise nicht in dem Erfassungsbereich der Radarsensorik. Aufgrund der unvermeidlichen Objektverluste des Trackings und der daraus resultierenden unvollständigen Daten wurde in diesen Szenen auf eine Evaluierung der zweiten Konfiguration verzichtet.

	Einheit	Szenario 1	Szenario 2	Szenario 3	Szenario 4
$RMSE_r^V(K_1)$	[m]	0.22	0.16	0.08	0.47
$RMSE_r^V(K_2)$	[m]	–	0.79	–	0.80
$RMSE_r^V(K_3)$	[m]	0.22	0.22	0.10	0.34
$RMSE_\phi^V(K_1)$	[°]	0.59	0.45	0.57	0.49
$RMSE_\phi^V(K_2)$	[°]	–	0.97	–	1.20
$RMSE_\phi^V(K_3)$	[°]	0.59	0.43	0.58	0.52
$RMSE_v^V(K_1)$	$[\frac{m}{s}]$	0.73	0.25	0.34	1.99
$RMSE_v^V(K_2)$	$[\frac{m}{s}]$	–	0.15	–	0.78
$RMSE_v^V(K_3)$	$[\frac{m}{s}]$	0.63	0.15	0.32	0.76

Tabelle 7.5: Auswertung der Szenen 1 bis 4 (siehe Tabelle 7.3) mithilfe der Metriken aus Abschnitt 7.2.1. Es werden drei Sensorkonfigurationen betrachtet: Die erste Konfiguration verwendet nur den Laserscanner. In der zweiten wird ausschließlich die Radarsensorik verwendet. In der dritten Konfiguration wird eine Fusion aus Radar und Laserscanner evaluiert.

Aus dem Vergleich der Sensorkonfigurationen in den unterschiedlichen Szenarien können folgende Ergebnisse abgeleitet werden:

**Schätzung der Entfernung:** Für Szenarien, die einen geringen Anteil an starken Beschleunigungen enthalten (Szenario 1, 2 und 3), kann die Qualität der Entfernungsschätzung ( $RMSE_r^V$ ) des Laserscanners durch keine der anderen Sensorkonfigurationen  $K_2$  bzw.  $K_3$  gesteigert werden. In diesen Szenarien ist selbst für die Fusion mit der Radarsensorik kein signifikanter Unterschied zu Konfiguration  $K_1$  zu erkennen. Wird ausschließlich die Radarsensorik verwendet, ist ein deutlicher Abfall der Qualität der Entfernungsschätzung zu beobachten. Lediglich in dem hochdynamischen vierten Szenario birgt die Fusion zwischen Radar und Laserscanner einen deutlichen Mehrwert. Dies ist dadurch zu erklären, dass an dieser Stelle, bedingt durch das dynamische Modell, die Schätzung der Entfernung sehr stark von der Geschwindigkeit abhängt und diese durch die Radarsensorik direkt gemessen werden kann. Somit reagiert Konfiguration  $K_3$  deutlich schneller auf die Änderung der Geschwindigkeit, was sich positiv auf die Qualität der Entfernungsschätzung auswirkt.

**Schätzung der lateralen Ablage:** Für die Schätzung der lateralen Ablage  $\varphi$  konnte in keinem der untersuchten Szenarien ein signifikanter Qualitätsunterschied zwischen Konfiguration  $K_1$  und  $K_3$  festgestellt werden. Wird hingegen nur die Radarsensorik ( $K_2$ ) verwendet, ist ein deutlicher Abfall zu beobachten, der in der geringeren Winkelauflösung der verwendeten Radarsensoren begründet ist.

**Schätzung der Geschwindigkeit:** Bei einer näheren Betrachtung der untersuchten Konfigurationen in Bezug auf die Qualität der Geschwindigkeitsschätzung ( $RMSE_v^V$ ) ist in allen Szenarien ein deutlicher Vorteil der Radarsensorik zu beobachten (vergleiche  $K_1$  und  $K_2$ ). Dies ist darin begründet, dass mit einem Radarsensor die Geschwindigkeit direkt gemessen werden kann. Dies hat zwei Vorteile gegenüber der indirekten Messung durch den Laserscanner: Zum Einen reagiert das Tracking deutlich schneller auf Änderungen der Geschwindigkeit (siehe Szenario 4). Zum Anderen ist die Messung sehr genau und unabhängig von Modellannahmen der Positionsbestimmung, was auch in Szenen mit wenig Dynamik zu einer besseren Schätzung der Geschwindigkeit führt (siehe Szenario 2). Durch eine Fusion aus Laserscanner und Radar kann die sehr gute Qualität der Geschwindigkeitsschätzung der Konfiguration  $K_2$  nur noch unwesentlich verbessert werden. Im Vergleich zur  $K_1$  ist hingegen ein deutlicher Mehrwert zu erkennen.

In einem Vergleich der Szenen untereinander werden folgende Aspekte deutlich: Die Entfernungs- und Geschwindigkeitsschätzung wird direkt durch den Anteil an Längsdynamik beeinflusst. So nimmt die Qualität mit steigender Dynamik deutlich ab. Dies ist gut an einem Vergleich der Szenen 3, 1 und 4 zu erkennen, in denen der Anteil an Längsdynamik von kaum vorhanden bis zu einer Vollbremsung ansteigt. Die Schätzung der lateralen Ablage ist hingegen im Wesentlichen von der in der Szene enthaltenen Querdynamik abhängig (vergleiche Szene 1 und 2).

### Evaluierung für Fußgänger

Die Szenarien, die in Abschnitt 7.2.3 eingeführt wurden, werden in diesem Abschnitt bewertet. Dabei wird der Einfluss von unterschiedlichen Sensorkonfigurationen auf die Systemleistung untersucht. In dieser Evaluierung werden zwei Sensorkonfigurationen betrachtet, um den Mehrwert des Fusionssystems in Bezug auf das Tracking von Fußgängern zu bewerten. Die erste Konfiguration beinhaltet nur

den Laserscanner. In der zweiten Konfiguration wird eine Fusion aus Wärmebildkamera und Laserscanner untersucht. In der Tabelle 7.6 sind die verwendeten Konfigurationen nochmals dargestellt.

Konfiguration	Laserscanner	Ferninfrarotkamera
$K_1$	x	
$K_2$	x	x

Tabelle 7.6: Übersicht der verwendeten Sensorkonfigurationen.

In Tabelle 7.7 sind die ermittelten Ergebnisse zusammengefasst. Die Tabellenspalten enthalten die untersuchten Szenen. Jeweils zwei Tabellenzeilen fassen die Werte einer Metrik für alle Sensorkonfigurationen zusammen. Als Qualitätsmaß der Evaluierung wurde der RMSE für die Schätzgröße der Entfernung  $r$  sowie der lateralen Ablage  $\phi$  gewählt.

	Einheit	Szenario 5	Szenario 6	Szenario 7	Szenario 8	Szenario 9	Szenario 10
$RMSE_r^P(K_1)$	[m]	0.12	0.11	0.10	0.12	0.10	0.12
$RMSE_r^P(K_2)$	[m]	0.13	0.11	0.10	0.12	0.10	0.12
$RMSE_\phi^P(K_1)$	[°]	1.33	0.72	0.81	0.77	0.94	0.60
$RMSE_\phi^P(K_2)$	[°]	0.78	0.72	0.76	0.77	0.96	0.60

Tabelle 7.7: Auswertung der Szenen 5 bis 10 (siehe Tabelle 7.3) mithilfe der Metriken aus Abschnitt 7.2.1. Es werden zwei Sensorkonfigurationen betrachtet: Die erste Konfiguration verwendet nur den Laserscanner. In der zweiten Konfiguration wird eine Fusion aus Wärmebildkamera und Laserscanner evaluiert.

Aus dem Vergleich der beiden Sensorkonfigurationen können folgende Ergebnisse abgeleitet werden:

**Schätzung der Entfernung:** Die Qualität der Entfernungsschätzung ist für beide Sensorkonfigurationen unabhängig von der Dynamik der Fußgänger (vergleiche die ersten beiden Zeilen der Tabelle 7.7). Dies war zu erwarten, da zum Einen durch den Laserscanner Entfernungen sehr genau geschätzt werden können und der Einfluss der hierzu vergleichsweise geringen Dynamik eines Fußgängers vernachlässigt werden kann. Zum Anderen kann durch die niedrige Verbauposition der Wärmebildkamera die Entfernung des Fußgängers nur sehr ungenau abgeleitet werden. Aus diesem Grund konnte die Wärmebildkamera in der zweiten Konfiguration im Vergleich zum Laserscanner keinen signifikanten Beitrag leisten.

**Schätzung der lateralen Ablage:** Für die Schätzung der lateralen Ablage  $\phi$  ergibt sich ein deutlich anderes Bild: Hier ist zu beobachten, dass für Sensorkonfiguration 1 in den Szenarien, die eine Bewegung des Fußgängers quer zur Sensorik beinhalten (Szenario 5, 7 und 9), die Qualität der Schätzung im Vergleich zu den restlichen Szenen deutlich abnimmt. Dies ist dadurch zu erklären, dass dynamische laterale Bewegungen durch die geringe Winkelauflösung des Laserscanners schlecht erfasst werden können. An dieser Stelle birgt die Wärmebildkamera zwei entscheidende Vorteile: Zum Einen können dynamische Bewegungen durch die höhere Abtastrate

schneller erfasst werden. Zum Anderen lässt sich Aufgrund der höheren Winkelauflösung die laterale Ablage von Fußgängern sehr präzise messen. Durch diese Effekte kann für die Konfiguration 2 in den Szenarien 5 und 7 eine verbesserte Schätzung der lateralen Ablage erreicht werden.

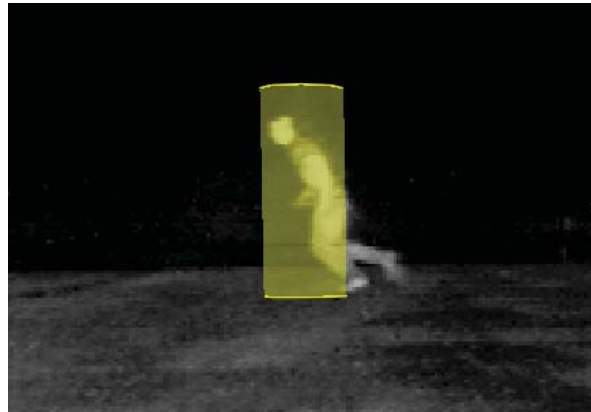


Abbildung 7.7: Rennender Fußgänger. Hier befindet sich der Kopf nicht mehr über dem Schwerpunkt, sondern ist in Bewegungsrichtung nach vorne verlagert.

Für Szenario 9, in dem der Fußgänger durch die Szene rennt, ist kein Vorteil durch eine Fusion mit der Wärmebildkamera zu erkennen. Dies ist darauf zurückzuführen, dass in diesem Szenario die Annahme des Beobachtungsmodells für die Kopfkontur nicht mehr zutrifft: Bei einem rennenden Fußgänger befindet sich der Kopf nicht mehr über den Beinen, sondern ist in Bewegungsrichtung nach vorne verlagert (siehe auch Abbildung 7.7). Somit ergibt sich ein systematischer Widerspruch zwischen der Messung in der Kamera und dem Laserscanner bzw. den Referenzdaten. Für die Szenarien, in denen sich der Fußgänger längs zum Eigenfahrzeug bewegt (Szenario 6 und 8), ist ebenfalls kein Mehrwert der Sensorfusion zu erkennen.

### 7.2.3 Bewertung des Gesamtsystems

Nachdem im vorherigen Abschnitt die Trackingleistung unter idealen bzw. „sterilen“ Randbedingungen untersucht wurde, soll im Folgenden das Gesamtsystem zur Fahrumfelderfassung, bestehend aus Klassifikation und Tracking, hinsichtlich Genauigkeiten und Erkennungsraten unter realistischen Umweltbedingungen bewertet werden. Nach einem Überblick über die verwendeten Szenarien und einer Beschreibung der Testsystemparametrisierung erfolgt eine Diskussion der Ergebnisse.

#### Beschreibung der Testszenarien

Zur Bewertung des Gesamtsystems wurde eine repräsentative Stichprobe aus allen aufgezeichneten Testszenarien ausgewählt. Tabelle 7.8 listet die verwendeten Testszenarien mit einer kurzen Beschreibung sowie einzelnen Szenenausschnitten auf. Die dargestellten Ausschnitte zeigen das Ergebnis der



maschinellen Erkennung zu einem ausgewählten Zeitpunkt.

Die Szenen können hinsichtlich des Aufzeichnungsortes in Stadt-, Landstraßen- und Autobahnsszenarien untergliedert werden. Zudem kann eine Unterteilung bezüglich der Witterungs- und Umgebungseinflüsse erfolgen. Es werden Schönwetter-, Schnee- und Regen- sowie Tag- und Nachtszenarien abgedeckt. In sämtlichen Szenen wurden mittels einer semiautomatischen algorithmusgestützten Referenzdatenerzeugung (siehe auch Abschnitt 7.1.2) alle relevanten Verkehrsteilnehmer annotiert. Dadurch wird für die nachfolgende Auswertung eine Bewertungsgrundlage geschaffen.

Tabelle 7.8: In dieser Tabelle wird ein Überblick über die Szenen gegeben, die zur Bewertung herangezogen wurden. Neben einer kurzen Szenenbeschreibung sind in der rechten Tabellenhälfte Szenenausschnitte mit maschinell erkannten Objekten zu sehen.




Szenenbeschreibung	Szenenausschnitte
<b>Szenarien bei trockenen Witterungsbedingungen</b>	
<p>Szene 1: Stadtszene in München mit 17 Fahrzeugen und 10 Fußgängern. Starkes Eigenfahrzeugnicken wegen der Überquerung einer Brücke. (Dauer 0:30 Minuten)</p>	
<p>Szene 2: Szene in einem Wohngebiet in Garching mit 11 fahrenden und 24 parkenden Fahrzeugen sowie 13 Fußgängern. (Dauer 1:43 Minuten)</p>	
<p>Szene 3: Landstraßenszene mit 20 entgegenkommenden Fahrzeugen und vier Fahrzeugen in gleicher Fahrtrichtung. (Dauer 1:39 Minuten)</p>	
<p>Szene 4: Autobahnsszene mit 86 PKW und 12 LKW bei trockener Fahrbahn. (Dauer 3:55 Minuten)</p>	



Tabelle 7.8: In dieser Tabelle wird ein Überblick über die Szenen gegeben, die zur Bewertung herangezogen wurden. Neben einer kurzen Szenenbeschreibung sind in der rechten Tabellenhälfte Szenenausschnitte mit maschinell erkannten Objekten zu sehen.





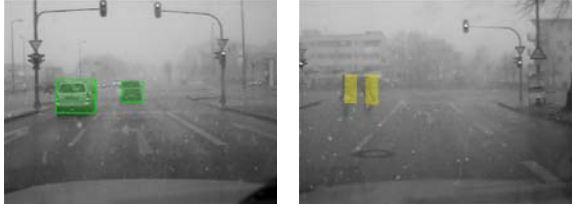
Szenenbeschreibung	Szenenausschnitte
Nachtszenarien	
<p>Szene 5: Stadtszene mit vier stationären, 15 entgegenkommenden, vier in Fahrtrichtung fahrenden Fahrzeugen sowie 11 Fußgängern bei Nacht und regennasser Fahrbahn. (Dauer 1:40 Minuten)</p>	
<p>Szene 6: Autobahnszene mit 38 Fahrzeugen bei Nacht und regennasser Fahrbahn. (Dauer 4:04 Minuten)</p>	
Szenarien bei Niederschlag	
<p>Szene 7: Stadtszene in Dachau auf einer Ausfallstraße mit 12 entgegenkommenden Fahrzeugen, 13 Fahrzeugen in gleicher Fahrtrichtung und 7 Fußgängern. Mittelstarker Regen. (Dauer 2:36 Minuten)</p>	
<p>Szene 8: Autobahnszene mit 9 Fahrzeugen bei starkem Regen und Gischtschleppen hinter den Fahrzeugen. Sichtweiten von unter 80 m. (Dauer 0:54 Minuten)</p>	

Tabelle 7.8: In dieser Tabelle wird ein Überblick über die Szenen gegeben, die zur Bewertung herangezogen wurden. Neben einer kurzen Szenenbeschreibung sind in der rechten Tabellenhälfte Szenenausschnitte mit maschinell erkannten Objekten zu sehen.

Szenenbeschreibung	Szenenausschnitte
<p>Szene 9: Stadtszene mit 17 Fahrzeugen und 2 Fußgängern bei starkem Schneefall. Sichtweiten von unter 100 m. Der Schnee verursacht in der zweiten Szenenhälfte einen Totalausfall des Laserscanners. (Dauer 1:30 Minuten)</p>	

## Testsystem

Für die Evaluierung der Gesamtsystemleistung wird das im Grundlagenkapitel 2 in Abbildung 2.9 illustrierte multisensorielle Fahrumfelderfassungssystem verwendet. Die einzelnen Systemkomponenten sind wie folgt parametrisiert:

**Datenvorverarbeitung** Für die Datenvorverarbeitung der Laserscannerrohdaten wird die inkrementelle Linienaggregation verwendet (siehe auch Abschnitt 4.2.1). Um die Robustheit der Vorverarbeitung gegenüber Messfehlern zu erhöhen, wird zusätzlich ein Verfahren nach RANSAC verwendet (siehe Abschnitt 4.2.2).

**Hypothesenerzeugung** Fahrzeughypothesen werden auf Basis von Radar- und Laserscannerrohdaten erzeugt. Fußgängerhypothesen werden ausschließlich auf Laserscannermesswerten aufgesetzt. Für eine detaillierte Beschreibung sei auf die Arbeit von Leonhard Walchshäusl [Wal08] verwiesen.

**Klassifikation** Der Klassifikator verwendet Gradientenhistogramme im Video- und Wärmebild sowie Dynamik- und Reflektionsmerkmale. Das arithmetische Mittel als funktionaler Kombinator kommt als Klassifikatorfusionsverfahren zum Einsatz. Für weitere Details zu diesen Verfahren sei ebenso auf die Arbeit von Leonhard Walchshäusl [Wal08] verwiesen.

**Tracking** Für das Verfolgen der Objekte wird ein Extended-Kalman-Filter verwendet (siehe auch Abschnitt 5.2.2), da die Evaluierung in Bezug auf zukünftige Fahrerassistenzsysteme durchgeführt wurde und somit Echtzeitfähigkeit gefordert war. Für Fahrzeuge erfolgt eine Fusion der Messwerte des Laserscanners (siehe Abschnitt 4.2.1) und der Radarsensoren (siehe Abschnitt 4.3) nach dem Prinzip Messwertfusion (siehe Abschnitt 5.2.4). Fußgänger werden durch Laserscannermesswerte verfolgt.

Für die Auswertungen wurden zwei Systemkonfigurationen untersucht. In der ersten Konfiguration (K1) wurde ausschließlich der Laserscanner für das Tracking und die Klassifikation der Objekte verwendet. Alle restlichen Sensoren wurden deaktiviert. In dieser Konfiguration werden Fahrzeuge an-

hand der Heckreflektanz und der Objektdynamik klassifiziert. Bei Fußgängern wird die Pendelbewegung als Merkmal herangezogen. Diese Konfiguration wurde gewählt, da durch den winkelauflösenden entfernungsgebenden Sensor, im Vergleich zu den anderen untersuchten Einzelsensorsystemen, die höchste Gesamtleistung erzielt werden konnte.

Um den Mehrwert des Fusionssystems zu evaluieren, wurde eine zweite Konfiguration K2 erstellt, die alle Sensoren umfasst. Der Laserscanner und das Radar werden für das Objekttracking verwendet. Für die Klassifikation wird der Laserscanner, die Weitwinkel- und die Normalwinkelvideokamera, sowie die Wärmebildkamera herangezogen. Für Testszenarien, in denen eine Teilmenge aller Sensoren eine bessere Leistung als K2 liefert, wird diese gesondert diskutiert. Die sich ergebenden Teilkonfigurationen von K2 sind  $K2^\dagger$ , das kein Wärmebild verwendet, sowie Konfiguration  $K2^\ddagger$ , bei der alle bildgebenden Sensoren deaktiviert sind. In der Tabelle 7.9 sind alle verwendeten Konfigurationen in einer Übersicht dargestellt.

Konfiguration	Laserscanner	Radar	Ferninfrarotkamera	Weitwinkelvideokamera	Normalwinkelvideokamera
$K1$	x				
$K2$	x	x	x	x	x
$K2^\dagger$	x	x	x		
$K2^\ddagger$	x	x			

Tabelle 7.9: Übersicht der verwendeten Sensorkonfigurationen. Die Konfiguration K1 verwendet lediglich einen Laserscanner. Die Konfiguration K2 verwendet alle verfügbaren Sensoren. Konfiguration  $K2^\dagger$  ist eine Teilkonfiguration von K2, die kein Wärmebild verwendet. Analog dazu sind bei Konfiguration  $K2^\ddagger$  alle bildgebenden Sensoren deaktiviert.

## Ergebnisse

Die Szenarien, die in Abschnitt 7.2.3 eingeführt wurden, werden in diesem Abschnitt hinsichtlich der Metriken aus Abschnitt 7.2.1 bewertet. Dabei wird der Einfluss von unterschiedlichen Sensorkonfigurationen auf die Systemleistung untersucht.

Die Evaluierung der Szenen wurde durch den maschinellen Abgleich mit einer semiautomatisch generierten Referenz durchgeführt. Da mit dieser Art der Referenzdatengewinnung die Geschwindigkeit nicht attribuiert werden kann, wird diese in der folgenden Evaluierung nicht berücksichtigt. In Tabelle 7.10 sind die ermittelten Ergebnisse der Szenenauswertung zusammengefasst. Die Tabellenspalten enthalten die untersuchten Szenen. Jeweils zwei Tabellenzeilen beinhalten die Werte einer Metrik für die unterschiedlichen Sensorkonfigurationen K1 und K2. Für die Szenen 5, 6 und 7, 8 wurde Konfiguration  $K2^\dagger$  und  $K2^\ddagger$  verwendet (siehe auch Abschnitt 7.2.3 für eine Erläuterung zu den Konfigurationen). Im Folgenden wird ein Leistungsvergleich zwischen den beiden Sensorkonfigurationen sowie zwischen den einzelnen Szenen vorgenommen.

Tabelle 7.10: Auswertung der Fahrscenen 1 bis 9 (siehe Tabelle 7.8) mithilfe der Metriken aus Abschnitt 7.2.1. Die Konfiguration K1 verwendet lediglich einen Laserscanner. Die Konfiguration K2 verwendet alle verfügbaren Sensoren. Konfiguration K2<sup>†</sup> ist eine Teilkonfiguration von K2, die kein Wärmebild verwendet. Analog dazu sind bei Konfiguration K2<sup>‡</sup> alle bildgebenden Sensoren deaktiviert.

		Trockene Witterungsbedingungen				Nacht		Schlechtes Wetter		
		Stadt Szene 1	Stadt Szene 2	Land Szene 3	Autobahn Szene 4	Stadt Szene 5 <sup>†</sup>	Autobahn Szene 6 <sup>†</sup>	Stadt Szene 7	Stadt Szene 8 <sup>‡</sup>	Autobahn Szene 9 <sup>‡</sup>
<i>RPR</i> K1	[%]	0.24	0.32	0.78	0.91	0.24	0.91	0.46	0.57	0.53
<i>RPR</i> K2	[%]	0.76	0.80	0.96	0.87	0.62	0.82	0.79	0.71	0.67
<i>RPR<sub>V</sub></i> K1	[%]	0.40	0.26	0.78	0.91	0.35	0.91	0.62	0.57	0.62
<i>RPR<sub>V</sub></i> K2	[%]	0.93	0.81	0.96	0.87	0.61	0.82	0.90	0.71	0.69
<i>RPR<sub>P</sub></i> K1	[%]	0.00	0.46	–	–	0.00	–	0.00	–	0.00
<i>RPR<sub>P</sub></i> K2	[%]	0.50	0.77	–	–	0.64	–	0.43	–	0.50
<i>FER</i> K1	[1/h]	240.55	380.71	51.02	12.97	45.24	26.65	183.31	0.00	0.00
<i>FER</i> K2	[1/h]	721.66	103.83	153.07	0.00	67.86	53.29	297.87	7.48	18.71
<i>FER<sub>V</sub></i> K1	[1/h]	120.28	34.61	51.02	12.97	45.24	26.65	183.31	0.00	0.00
<i>FER<sub>V</sub></i> K2	[1/h]	481.11	34.61	102.05	0.00	45.24	53.29	297.87	7.48	18.71
<i>FER<sub>P</sub></i> K1	[1/h]	120.28	346.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<i>FER<sub>P</sub></i> K2	[1/h]	240.55	69.22	51.02	0.00	22.62	0.00	0.00	0.00	0.00
<i>MCC</i> K1		0.42	0.42	0.84	0.94	0.43	0.93	0.54	0.76	0.73
<i>MCC</i> K2		0.76	0.86	0.87	0.93	0.73	0.85	0.70	0.77	0.78
<i>MCC<sub>V</sub></i> K1		0.58	0.48	0.84	0.94	0.53	0.93	0.62	0.76	0.78
<i>MCC<sub>V</sub></i> K2		0.85	0.88	0.90	0.93	0.73	0.85	0.73	0.77	0.79
<i>MCC<sub>P</sub></i> K1		–	0.42	–	–	0.00	0.00	–	–	0.00
<i>MCC<sub>P</sub></i> K2		0.60	0.80	–	–	0.75	0.00	0.65	–	0.71
<i>ZA</i> K1	[%]	0.49	0.36	0.75	0.95	0.64	0.95	0.81	0.40	0.42
<i>ZA</i> K2	[%]	0.69	0.74	0.84	0.92	0.75	0.71	0.95	0.82	0.76
<i>ZA<sub>V</sub></i> K1	[%]	0.61	0.39	0.75	0.95	0.78	0.95	0.85	0.40	0.42
<i>ZA<sub>V</sub></i> K2	[%]	0.82	0.74	0.84	0.92	0.81	0.71	0.97	0.82	0.77
<i>ZA<sub>P</sub></i> K1	[%]	–	0.30	–	–	0.00	–	–	–	0.00
<i>ZA<sub>P</sub></i> K2	[%]	0.16	0.74	–	–	0.49	–	0.45	–	0.42
<i>RMSE<sub>r</sub></i> K1	[m]	0.18	0.18	0.18	0.18	0.40	0.20	0.33	0.24	0.23
<i>RMSE<sub>r</sub></i> K2	[m]	0.29	0.18	0.21	0.19	0.38	0.22	0.31	0.47	0.62
<i>RMSE<sub>r</sub><sup>V</sup></i> K1	[m]	0.18	0.17	0.18	0.18	0.40	0.20	0.33	0.24	0.23
<i>RMSE<sub>r</sub><sup>V</sup></i> K2	[m]	0.28	0.19	0.21	0.19	0.41	0.22	0.32	0.47	0.62
<i>RMSE<sub>r</sub><sup>P</sup></i> K1	[m]	0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<i>RMSE<sub>r</sub><sup>P</sup></i> K2	[m]	0.40	0.12	0.00	0.00	0.13	0.00	0.18	0.00	0.97

Aus dem Leistungsvergleich der beiden Sensorkonfigurationen können folgende Ergebnisse abgeleitet werden:

**Stadt und Autobahn** Erwartungsgemäß ist die Systemperformanz sowohl bei Autobahn- als auch bei Landstraßenszenen besser, als bei den wesentlich komplexeren Stadtszenarien (vergleiche Metriken der Szenen 1 und 2 mit den Metriken der Szenen 3 und 4). Sowohl die Erkennungsrate, die zeitliche Abdeckung und die MCC-Metrik sind entsprechend höher. Insbesondere Fehlerkennungen ( $FER_p$ ) von Fußgängern treten im urbanen Umfeld häufiger auf, da der Hintergrund im Vergleich zu den Außerstadtszenarien stärker strukturiert ist. Zudem wird die Fahrumfelderfassung häufig durch Verdeckungen von Verkehrsteilnehmern erschwert, wodurch sowohl die Qualität des Trackings (ZA) als auch die Klassifikation (RPR, MCC) negativ beeinflusst wird.

**Unterschied Tag und Nacht** Die Auswertungen der Autobahnszene 4, die bei Tag aufgenommen wurde und der Autobahnszene 6, die bei Nacht aufgenommen wurde, sind für die Konfiguration K1 weitgehend identisch (vergleiche MCC und ZA), da der Laserscanner beleuchtungsunabhängig arbeitet. Für den fusionierten Ansatz (K2) bricht hingegen die Klassifikationsleistung für die Nachtszene stark ein, da aufgrund der eingeschränkten Videobildqualität lediglich veräuschte Merkmale extrahiert werden können. Unter der Annahme dass dieser „Sensorausfall“ detektiert werden kann, wurde statt Konfiguration K2, die um das Videobild reduzierte, Konfiguration  $K2^\dagger$  für die Evaluierung gewählt. Mit dieser Konfiguration werden für die Nachtszene 6 etwas schlechtere Ergebnisse als mit K2 für die Tagszene 4 erzielt. Dies ist dadurch zu erklären, dass der verwendete Wärmebildsensor eine geringere Entfernungreichweite als der Videosensor aufweist. Dies führt zu einer verspäteten Klassifikation und somit zu einer geringeren zeitlichen Abdeckung. Im Vergleich der Stadtszene bei Nacht (Szene 5) mit den Stadtszenen bei Tag (Szenen 1 und 2) ist eine reduzierte Erkennungsleistung für Fahrzeuge zu beobachten. Dies ist darauf zurückzuführen, dass die Wärmebildsignatur von entgegenkommenden Fahrzeugen stark von der Heckansicht abweicht. Bei den Tagszenen kann dies durch den Videosensor kompensiert werden.

**Einfluss von schlechtem Wetter** Bei der Stadtszene 6 mit mittelstarkem Regen sind im Vergleich zu den Stadtszenen 1 und 2 keine signifikanten Leistungsabfälle zu beobachten, da die Sensorik nur geringfügig beeinflusst wird. Autobahnszenen mit starkem Regen und/oder Gischtschleppen sowie Szenen mit starkem Schneefall (vergleiche Szene 7 bzw. Szene 8) reduzieren insbesondere die Sichtweiten der optischen Sensorsysteme. Dadurch nimmt für K1 sowohl die Klassifikations- als auch die Trackingleistung erheblich ab (vergleiche ZA und MMC). Für K2 sinkt insbesondere die Klassifikationsleistung, da durch die witterungsbedingten Effekte sowohl die Video- als auch die Wärmebildmerkmale stark veräuscht sind. Unter der Annahme dass diese „Sensorausfälle“ detektiert werden können, wurde statt Konfiguration K2, Konfiguration  $K2^\ddagger$  verwendet, die ausschließlich die entfernungsgebende Sensorik beinhaltet. Da die Objektdetektion und -verfolgung im Nahbereich durch den Laserscanner und im Fernbereich durch den Radarsensor übernommen wird, ist ein erheblicher Anstieg der zeitlichen Abdeckung (ZA) im Vergleich zu K1 zu beobachten.

**Alle Szenarien** Im Mittel über die Szenen werden Verkehrsteilnehmer durch die Konfiguration K2 besser (vergleiche MCC) erkannt. Dies ist im Wesentlichen dadurch zu erklären, dass Fußgänger und parkende Fahrzeuge im städtischen Umfeld, allein durch den Laserscanner, lediglich sehr eingeschränkt klassifiziert werden können. Die Merkmale, die aus der bildgebenden Sensorik gewonnen werden, sind hierfür besser geeignet. Aufgrund der höheren Erkennungsraten

und der besseren Trackingleistung in den Schlechtwetzerszenarien ergibt sich zudem eine verbesserte zeitliche Abdeckung (ZA). Ausnahmen sind die Autobahnszenarien 4 und 6. Dort sind die Ergebnisse (MCC und ZA) der beiden Konfigurationen nahezu identisch. Der leichte Abfall kann durch den negativen Einfluss der Videobildmerkmale weiter entfernter Objekte auf die Klassifikation erklärt werden.

Die Trackinggenauigkeit  $r$  ist in den Szenarien 1 bis 7 unabhängig von der jeweiligen Sensor-konfiguration. Da die evaluierten Szenen kaum hochdynamisches Objektverhalten beinhalten, schafft die zusätzliche Geschwindigkeitsbestimmung der Radarsensorik gegenüber dem Laserscanner, keinen signifikanten Mehrwert in Bezug auf die Entfernungsschätzung. In den Schlechtwetzerszenarien 8 und 9, in denen die optische Sichtweite stark eingeschränkt ist, können durch den Laserscanner Objekte lediglich im Nahbereich detektiert und verfolgt werden. Dieses Defizit wird durch die Radarsensorik zwar ausgeglichen (vergleiche ZA), gleichzeitig sinkt jedoch die Genauigkeit der Entfernungsbestimmung. Dies ist durch die multiplen Reflexionsquellen an den detektierten Fahrzeugen zu erklären. Diese können nicht, wie in Abschnitt 4.3 beschrieben, kompensiert werden, da ein zweiter Sensor, aufgrund der eingeschränkten Sichtweite, nicht zur Verfügung steht.

#### 7.2.4 Charakteristische Fehlerfälle

In Bezug auf Fahrumfelderfassungssysteme werden Fehler als Abweichungen von den Referenzdaten definiert. Die Fehlerursachen sind vielfältig und liegen zum Einen in der Algorithmik der einzelnen Systemkomponenten und zum Anderen in äußeren Einflüssen, wie ungünstigen Witterungsbedingungen. Tabelle 7.11 fasst charakteristische Fehler für das Tracking sowie die Klassifikation und deren Ursachen anhand von Beispielen zusammen.

In der ersten Tabellenzeile sind Fehler zu sehen, die durch die Datenvorverarbeitung verursacht werden. Sinkt der Abstand zwischen Objekten unter die Sensorauflösung, kann keine Separierung mehr erfolgen. Fehlsegmentierungen treten insbesondere bei Objekten auf, die einen geringen Abstand zueinander aufweisen. Zudem können Fehlsegmentierungen bei Verdeckungen auftreten, da Objekte in diesem Fall meist nur partiell beobachtet werden können.

Bei der Klassifikation von Objekten treten falsch-positiv und falsch-negativ Erkennungen auf. Bei der ersten Fehlerart werden Hintergrundobjekte fälschlicherweise als relevante Verkehrsteilnehmer erkannt. Dies ist insbesondere der Fall, wenn Hintergrundobjekte eine ähnliche Merkmalssignatur wie Vordergrundobjekte aufweisen (vergleiche Straßenschilder und -baken in der zweiten Tabellenzeile). Bei der zweiten Fehlerart werden Vordergrundobjekte nicht als solche erkannt (vergleiche entgegenkommendes Fahrzeug in der zweiten Tabellenzeile).

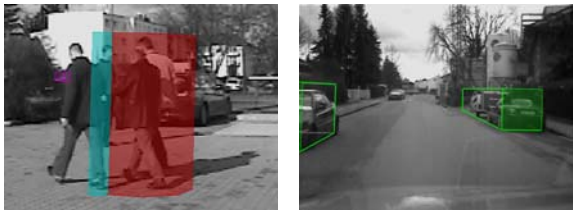



Verlorene Objekte sind ein Hauptfehler beim Verfolgen von Objekten. Dies ist insbesondere bei hochdynamischen Objekten zu beobachten, da dort die verwendeten Bewegungsmodelle und Trackingverfahren die Zustandsänderungen zu langsam abbilden (siehe linkes Bild in Tabellenzeile 3). Weitere Gründe sind Zuordnungsfehler, wie sie bei Objekten mit geringem Abstand zueinander, Verdeckungen oder multiplen Reflexionsquellen auftreten können (siehe Ampelpfosten im rechten Bild in Tabellenzeile 3). Eine weitere Fehlerquelle ist eine zu einfache Modellierung der zu verfolgenden Objekte oder falsche Annahmen in der Vorverarbeitung der Sensordaten. So können insbesondere bei LKWs oder entgegenkommenden Fahrzeugen verfälschte Messwerte entstehen, da ihr Erscheinungsbild im



Laserscanner von der Annahme eines Quadermodells abweicht. Diese systematischen Fehler können nur sehr begrenzt durch den verwendeten Kalman Filter kompensiert werden.

Grundsätzlich können aber auch widrige Umweltbedingungen eine Ursache für die oben angegebenen Fehler sein. Durch verrauschte Sensordaten, partielle Sensorausfälle und zusätzliche Störquellen wie beispielsweise Wassertropfen und Schneeflocken können Geisterziele, Fehlklassifikationen und fehlerbehaftete Datenvorverarbeitungen verursacht werden (vergleiche auch Szenen 8 und 9 sowie Tabellenzeile 4).

Tabelle 7.11: Charakteristische Fehler des multisensoriellen Fahrumfelderfassungssystems mit kurzer Beschreibung und Beispielen.

Problembeschreibung	Beispiele
Fehler von Komponenten des Fahrumfelderfassungssystems	
<p>Datenvorverarbeitung: Objekte mit einem zu geringen Abstand können nicht getrennt werden. Beispiele dafür sind Fußgängergruppen (linkes Bild) und parkende Fahrzeuge (rechtes Bild).</p>	
<p>Klassifizierung: Das Fahrzeug in der entgegenkommenden Spur im linken Bild ist ein Beispiel für eine falsch-positiv Erkennung. Das Schild im linken Bild und die Bake im rechten Bild sind Beispiele für falsch-negativ Erkennungen.</p>	
<p>Tracking: Abfliegender Track bei einem entgegenkommenden Fahrzeug (linkes Bild). An einem Ampelpfosten „hängen gebliebener“ Fußgänger-track durch einen Zuordnungsfehler beim nahen Passieren eines Fußgängers (rechtes Bild).</p>	
Umweltverursachte Fehler	
<p>Eingeschränkte Sichtbereiche der Sensorik durch widrige Witterungsbedingungen wie starkem Regen (linkes Bild), Nebel oder Schneefall (rechtes Bild).</p>	



### 7.2.5 Laufzeitanalyse

Da die Fahrumfelderfassung als automotiv Anwendung konzipiert ist, sind Echtzeitanprüche von großer Bedeutung. Die Komplexität der einzelnen Szenen ist entscheidend für die durchschnittliche Gesamtlaufzeit  $t_{Zyk}$  der Fahrumfelderfassung. In Abbildung 7.8 sind für ein typisches Autobahn- und Stadtszenario (siehe auch Szenarien 4 und 1 in Tabelle 7.8) die zeitlichen Verläufe der durchschnittlichen Laufzeiten für das Tracking, die Klassifikation sowie die Vorverarbeitung dargestellt. Darüber hinaus ist die Anzahl an Objekten aufgetragen, die in jedem Zeitschritt verarbeitet werden müssen. Sie beinhaltet sowohl bereits klassifizierte Fahrzeuge und Fußgänger, als auch noch nicht klassifizierte Objekthypothesen und dient als Indikator für die Komplexität der Szene. Die folgende Laufzeitanalyse basiert auf einem Core2Duo Prozessor mit 2 GHz, wobei nur ein Prozessorkern verwendet wurde. Die Skalierbarkeit der vorgestellten Fahrumfelderfassung auf aktuelle Multiprozessorarchitekturen wird im zweiten Teil dieses Abschnitts thematisiert.

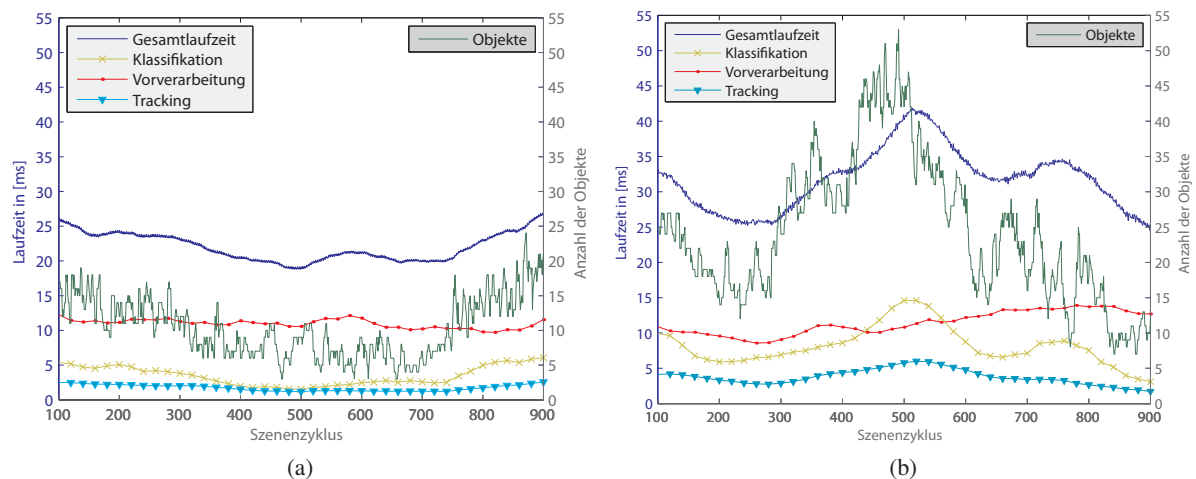


Abbildung 7.8: Vergleich der System- und Komponentenlaufzeiten für das Autobahnszenario 4 (a) und das Stadtszenario 1 (b) abhängig vom Szenenzeitpunkt. Es ist ein deutlicher Zusammenhang zwischen Zykluslaufzeiten und Objektaufkommen erkennbar.

Die Vorverarbeitung für den Laserscanner und die Wärmebildkamera ist weitgehend unabhängig von der Anzahl der Objekte und benötigt für die beiden Szenarien im Schnitt 13 ms. Die durchschnittlichen Laufzeiten für die Klassifikation und das Tracking sind hingegen direkt proportional zur Komplexität der Szene. So benötigt das Tracking 2.5 ms in der Autobahnszene mit bis zu 25 Objekten. Bei bis zu 50 gleichzeitig zu verarbeitenden Objekten, wie es für eine Stadtszene typisch ist, steigt die durchschnittliche Laufzeit für das Tracking hingegen auf 5 ms. Für die Klassifikation der Objekthypothesen werden 5 bzw. 12 ms benötigt. Zusammengefasst lässt sich eine echtzeitfähige Fahrumfelderfassung für Autobahn- und Landstraßenszenarien mit einer Systemrate von  $f_{Zyk} = 30\text{Hz}$  realisieren. Für komplexe Stadtszenarien kann unter Verwendung eines Einprozessorsystems ebenfalls Echtzeitfähigkeit bei  $f_{Zyk} = 25\text{Hz}$  erreicht werden. An dieser Stelle bieten aktuelle Multiprozessorarchitekturen weiteres Optimierungspotenzial.

Durch eine parallele Abarbeitung von ProgrammROUTINEN kann auf Multiprozessorssystemen Laufzeit eingespart werden. In der Architekturentscheidung für das Fahrumfelderfassungssystem wurde eine

parallele Ausführung speziell berücksichtigt. So werden die Objekte des Fahrumfeldes (Hypothesen, Verkehrsteilnehmer, Sensordaten) in einer hierarchischen Baumstruktur verwaltet. Dieser kreisfreie Graph kann durch ein Visitor-Designmuster<sup>5</sup> [GHJV95] effizient durchlaufen werden, um Operationen auf den einzelnen Knoten auszuführen. Für die Fahrumfelderfassung werden alle Operationen, die auf dieser Datenstruktur operieren, als Visitor angelegt. Eine parallele Ausführung kann demnach erzielt werden, indem die Baumstruktur durch mehrere Prozesse zeitgleich durchlaufen wird. Dabei müssen Serialisierungsanforderungen berücksichtigt werden.

Alle Rechenoperationen, die per Objekt oder per Sensor angewendet werden sind parallelisierbar. Dazu zählen die Module Datenvorverarbeitung, Messwertaktualisierung, Zustandsprädiktion, Merkmalsgewinnung und Klassifikation. Die Datenvorverarbeitung wird für alle beteiligten Sensoren durchgeführt und kann demnach pro Sensor parallelisiert werden. Die restlichen Module werden objektwise ausgeführt und können dadurch parallelisiert werden. Die Laufzeitmessungen mit und ohne Parallelisierung sind in Tabelle 7.12 aufgeführt. Dabei wurden die durchschnittlichen Laufzeiten für Szenario 4 und Szenario 1 (siehe auch Abschnitt 7.2.3) mit einem Multiprozessorsystem ausgewertet.

	Datenvorverarbeitung	Tracking	Merkmalsgewinnung	Klassifikation	Gesamtlaufzeit
Autobahnscene (Szenario 4 aus Tabelle 7.8)					
1 CPU	12.89 ms	0.71 ms	0.98 ms	3.19 ms	22.96 ms
2 CPUs	10.16 ms	0.48 ms	0.57 ms	1.75 ms	18.43 ms
Stadtscene (Szenario 1 aus Tabelle 7.8)					
1 CPU	15.34 ms	0.98 ms	3.13 ms	6.46 ms	38.04 ms
2 CPUs	12.29 ms	0.5 ms	1.62 ms	3.75 ms	29.46 ms

Tabelle 7.12: Vergleich der durchschnittlichen Zykluslaufzeiten der einzelnen Systemkomponenten bei einer parallelen Abarbeitung. Die Gesamtlaufzeit beinhaltet nicht aufgeführte Komponenten wie das Objektmanagement und die Visualisierung. Als Testsystem kam ein Zweiprozessorrechner mit 2GHz zum Einsatz.

Die Komponenten Prädiktion und Messwertaktualisierung, Merkmalsgewinnung und Klassifizierung, die für jedes Objekt durchgeführt werden, erreichen einen Speedup<sup>6</sup> von nahezu 0.5. Die Datenvorverarbeitung erzielt lediglich einen geringen Speedup, da die Vorverarbeitung für den Laserscanner deutlich mehr Zeit beansprucht, als die Vorverarbeitung für die bildgebende Sensorik. Insgesamt kann durch die parallele Ausführung eine Laufzeitsteigerung um 20% im Autobahnscenario und um 25% im Stadtscenario erzielt werden.

<sup>5</sup> Das Visitor-Designmuster ist ein Entwurfsmuster aus der Softwareentwicklung wodurch Operationen auf Datenstrukturen gekapselt werden können.

<sup>6</sup> Speedup (Beschleunigung) definiert das Laufzeitverhältnis eines parallelisierten Programms auf einem Prozessor, verglichen mit der Laufzeit desselben Programms auf  $p$  Prozessoren.

## 7.3 Zusammenfassung

Im ersten Teil des Kapitels wurden mehrere Möglichkeiten aufgezeigt, wie für ein Fahrumfelderfassungssystem Referenzdaten erzeugt werden können. Die vorgestellten Verfahren adressieren im Wesentlichen eine effiziente und präzise Gewinnung dieser Daten. Das semiautomatische Verfahren bzw. die algorithmusgetriebene Vorgehensweise optimiert den gesamten Annotierungsprozess in Bezug auf den Arbeitsaufwand. Durch den Einsatz eines dedizierten Sensors können für die Evaluierung von Trackingverfahren sehr genaue und von der verwendeten Sensorik unabhängige Referenzdaten generiert werden.

Im zweiten Teil des Kapitels wurde eine Bewertung der Tracking- sowie Gesamtsystemleistung durchgeführt, die in realistischen Umgebungen, wie Stadt, Land und Autobahn, sowie unter verschiedenen Witterungsbedingungen, den Mehrwert einer multisensoriellen Fahrumfelderfassung untersucht. Dabei ergab sich für Fußgänger und Fahrzeuge ein ähnliches Bild. Unter normalen Witterungsbedingungen konnte bei Autobahn- und Landstraßenszenen mit einem Laserscanner eine sehr gute Trackingleistung erreicht werden. Durch die Hinzunahme von weiteren Sensoren war in diesen Domänen keine weitere Leistungssteigerung möglich. In komplexeren städtischen Szenarien oder in sehr dynamischen Szenen ist ein leichter Abfall der Trackingqualität zu beobachten. An dieser Stelle kann durch die Fusion mit weiterer Sensorik die Qualität der Schätzung einzelner Zustandsgrößen optimiert werden. So profitiert bei Fahrzeugen insbesondere die Schätzung der Geschwindigkeit durch die Hinzunahme der Radarsensorik. Für Fußgänger kann durch die Fusion mit der Wärmebildkamera die Qualität der Schätzung der lateralen Ablage verbessert werden. Darüber hinaus zeigt sich unter widrigen Umweltbedingungen, wie starkem Regen oder Schneefall, dass mit einer multisensoriellen Fahrumfelderfassung eine Steigerung der Zuverlässigkeit erzielt werden kann.

Wird die Gesamtsystemleistung betrachtet, ergibt sich ein ähnliches Bild: Bei Autobahn- und Landstraßenszenen kann mit einem Laserscanner unter normalen Witterungsbedingungen eine gute Gesamtleistung erzielt werden. Durch die Fusion mit weiteren Sensoren konnte in diesen Fällen kein Leistungszuwachs erreicht werden. Erst in komplexen Stadtszenarien sowie unter widrigen Umweltbedingungen zeigt sich ein deutlicher Nutzen einer multisensoriellen Fahrumfelderfassung. Die durchgeführte Laufzeitanalyse belegt, dass das damit verbundene erhöhte Datenaufkommen mit aktuellen Rechensystemen in Echtzeit bewältigt werden kann.



# FAZIT

---

## 8.1 Zusammenfassung

Ziel dieser Dissertation war es, durch die Kombination (Fusion) von Radar, Laserscanner und bildgebender Sensorik eine präzise und zuverlässige Erfassung der Positionen, Geschwindigkeiten und Raummaße von Fußgängern und Fahrzeugen zu realisieren und den Mehrwert dieser multisensoriellen Fahrumfelderfassung im Vergleich zu konventionellen Ein-Sensor-Systemen zu untersuchen.

Ein wesentlicher Aspekt, der in einem Multi-Sensor System betrachtet werden muss, ist die gemeinsame Interpretation von Daten. Um Bildausschnitte der Kameras mit den Daten des Laserscanners oder der Radarsensorik zu verknüpfen, ist deshalb eine exakte Kalibrierung der einzelnen Sensoren erforderlich. In dieser Arbeit wurden hierzu Möglichkeiten dargestellt und diskutiert, wie ein multispektrales Kamerasystem, bestehend aus Ferninfrarot- und Videokameras, kalibriert werden kann. Die erste Variante, bestehend aus einem Schachbrett aus Peltierelementen, liefert deutlich genauere Ergebnisse im Vergleich zu der zweiten Vorgehensweise, die eine multispektrale Punkt-Lichtquelle verwendet. Allerdings ist bei ersterer die Durchführung komplexer und nur bedingt für einen mobilen Einsatz geeignet.

Insbesondere in zeitkritischen Fahrerassistenzsystemen müssen Latenzzeiten berücksichtigt werden. Diese können durch einen asynchronen Systemaufbau, sensorinterne Vorverarbeitung oder durch die Kommunikation zwischen den einzelnen Komponenten entstehen. Die genaue Zusammensetzung lässt sich in vielen Fällen nicht bestimmen. Aus diesem Grund wurden verschiedene Möglichkeiten vorgestellt, wie sich zumindest der Gesamtbeitrag für bild- und entfernungsgebende Sensorik experimentell bestimmen lässt. In der Praxis hat sich gezeigt, dass bei den verwendeten Sensoren bzw. Systemaufbau von einer nicht unerheblichen zeitlichen Verzögerung von 20 ms bis zu 100 ms auszugehen ist und dass die bildgebenden Sensoren eine deutlich geringere Latenzzeit besitzen als der Laserscanner oder die Radarsensoren.

Neben den Aspekten der räumlichen und zeitlichen Kalibrierung wurde in dieser Arbeit beschrieben, wie Fußgänger und Fahrzeuge durch die einzelnen Sensoren beobachtet werden können. Die Grundlegende Annahme für das Beobachtungsmodell des Laserscanners besteht darin, dass sich Fahrzeuge als quaderförmige Objekte und Fußgänger durch einen Zylinder approximieren lassen. Zunächst müssen in einem ersten Vorverarbeitungsschritt die Rohdaten des Laserscanners in zusammengehörige Bereiche unterteilt werden. Hierzu wurde eine Erweiterung des Region-Growing Verfahrens vorgestellt, die im Rahmen einer Fahrumfelderfassung eine zuverlässigere und vollständigere Segmentierung von

Fahrzeugen und Fußgängern gewährleistet. Basierend auf diesen Segmenten werden schließlich Merkmale extrahiert, die die zu erwartenden Messwerte der Beobachtungsmodelle beschreiben. Der wesentliche Bestandteil dieser Merkmalsextraktion wurde durch ein Verfahren nach RANSAC erweitert. Es konnte gezeigt werden, dass sich dadurch der gesamte Vorverarbeitungsprozess deutlich robuster gegenüber Messfehlern verhält.

Für die Radarsensoren wurde ein Messmodell vorgestellt, das die charakteristischen Rückstrahlungseigenschaften von Fahrzeugen explizit berücksichtigt. So kann ein Messwert zum Beispiel vom Fahrzeugheck, den Rückleuchten, der Hinterachse oder dem Motorblock erzeugt worden sein. In der Praxis hat sich gezeigt, dass mit dem eingeführten Modell der multiplen Reflexionsquellen im Vergleich zu einem einfachen Punktziel eine höhere laterale Genauigkeit erreicht werden kann. Allerdings ist das vorgestellte Verfahren auf einen zweiten Sensor (Laserscanner) angewiesen, um Mehrdeutigkeiten zu vermeiden bzw. aufzulösen. Fußgänger werden hingegen als einfache Punktziele modelliert. Hierbei ist zu beachten, dass Fußgänger mit der in dieser Arbeit verwendeten Radarsensorik nur sehr eingeschränkt erfasst werden können, da insbesondere die Fernbereichsradare auf Fahrzeuge und hohe Reichweiten optimiert sind und zusätzlich einen sehr eingeschränkten Öffnungswinkel besitzen.

Neben den entfernungsgebenden Sensoren (Laserscanner und Radar) wurde in dieser Arbeit ein Beobachtungsmodell vorgestellt, das dem charakteristischen Erscheinungsbild von Fußgängern in einer Ferninfrarotkamera nachempfunden ist. Hierbei hat sich die Kontur des Kopfes als sehr gutes Merkmal erwiesen, da dieser im Allgemeinen nicht mit Kleidung bedeckt und seine Kontur weitgehend invariant gegenüber dem Blickwinkel auf den Fußgänger ist. Ab einer Umgebungstemperatur  $\geq 25^\circ$  nimmt jedoch der Kontrast zwischen Fußgänger und Hintergrund deutlich ab und die Position des Kopfes kann nur noch sehr bedingt beobachtet werden.

Fahrzeuge besitzen sehr viele vertikale oder horizontale Strukturen und heben sich durch ihre klaren Flächen und Farben deutlich von Hintergrund ab. Aus diesem Grund wurde ein Konturmodell beschrieben, das den Umriss eines Fahrzeuges durch einen Quader approximiert.

Für Fahrerassistenzsysteme der aktiven Sicherheit, wie zum Beispiel die aktive Gefahrenbremsung, sind die Präzision und Robustheit der Fahrumfelderfassung sowie die Erfüllung von Echtzeitanforderungen von wesentlicher Bedeutung. Im Hinblick auf diese Kriterien wurden in dieser Arbeit verschiedene Varianten des Kalman-Filters betrachtet und Erweiterungen beschrieben, die zum Einen die Verwendung nichtlinearer Objekt- und Messmodelle gestatten und zum Anderen Messungen mehrerer Sensoren fusionieren. Zudem wurden Möglichkeiten aufgezeigt, wie die experimentell ermittelten Latenzzeiten in den Vorgang des Trackings integriert und somit kompensiert werden können. Es konnte gezeigt werden, dass damit eine signifikante Steigerung der Schätzgenauigkeiten erreicht werden kann. Bei ausreichender Rechenkapazität ist der Unscented-Kalman-Filter gegenüber dem Extended-Kalman-Filter zu bevorzugen, da er in der Umsetzung einfacher zu handhaben ist und für nichtlineare Systeme die besten Ergebnisse hinsichtlich der Genauigkeit der Zustandsschätzung erreicht werden können.

Neben dem klassischen Kalman-Filter wurden in dieser Arbeit auch nichtdeterministische Verfahren der Objektverfolgung betrachtet. Als Hauptvertreter dieser Gruppe wurden verschiedene Erweiterungen der Partikel-Filter evaluiert, die im Hinblick auf eine Fahrumfelderfassung eine gesteigerte Genauigkeit versprechen bzw. den Rechenaufwand im Vergleich zum Standard-Partikel-Filter verringern. Der Annealed-Partikel-Filter erwies sich hierbei als bestes Verfahren aus dieser Gruppe, da er schon bei kleiner Partikelanzahl konvergiert und somit bei geringerem Rechenaufwand genauere Ergebnisse liefert. Zudem war ein Zusammenhang zwischen der Komplexität der Objektmodelle und der benötigten Partikelanzahl zu beobachten. So ließen sich Fußgänger deutlich effizienter verfolgen als Fahrzeuge. In einem Vergleich mit dem Kalman-Filter konnten in stark dynamischen Szenarien

bessere Schätzgenauigkeiten erreicht werden. Allerdings ist der Annealed-Partikel-Filter noch nicht echtzeitfähig und somit mit Hinblick auf zeitkritische Fahrerassistenzsysteme nicht geeignet.

Aufgrund der großen Datenmenge, die in einer multisensoriellen Fahrumfelderfassung verarbeitet werden müssen, kann die Entwicklung und Diagnose des Gesamtsystems schnell unübersichtlich und ineffektiv werden. Deshalb wurden in dieser Arbeit Visualisierungsmöglichkeiten vorgestellt, die den Systementwickler, unter anderem durch den Einsatz von erweiterter Realität, im Umsetzungsprozess unterstützen. Dabei wird ihm ein Instrument an die Hand gegeben, das interne Abläufe und Zusammenhänge transparenter macht, die Fehlererkennung- und -suche erleichtert und das Gespür für die Daten und Datenbeziehungen fördert. Ein quantitativer Performanzvergleich der einzelnen Visualisierungsmöglichkeiten für automotiv Perzeptionssysteme ist nicht trivial. Selbst wenn sinnvolle Vergleichsmetriken zur Verfügung stünden, könnten erst aufwendige Nutzerstudien eine abgesicherte Aussage treffen. Die aufgestellten Thesen stützen sich deshalb vorwiegend auf wissenschaftlich fundierten Visualisierungsrichtlinien und die im Rahmen dieser Arbeit gewonnene Erfahrung bei der Umsetzung der Fahrumfelderfassung. In diesem Kontext haben sie sich als äußerst hilfreich erwiesen und wesentlich zum Gelingen dieser Arbeit beigetragen.

Für die Evaluierung einer Fahrumfelderfassung werden im Allgemeinen Referenzdaten von typischen Stadt- und Autobahnscenarien verwendet. In dieser Arbeit wurden Möglichkeiten aufgezeigt und diskutiert, wie multisensorielle Referenzdaten erzeugt werden können. Die vorgestellten Verfahren adressieren im Wesentlichen eine effiziente und präzise Gewinnung dieser Daten. So optimiert das vorgestellte semiautomatische Verfahren den gesamten Annotierungsprozess in Bezug auf den Arbeitsaufwand um den Faktor 5 bis 10. Für die Evaluierung von Trackingverfahren können allerdings erst durch den Einsatz eines dedizierten Sensors sehr genaue und von der verwendeten Sensorik unabhängige Referenzdaten generiert werden.

Diese Daten dienen als Grundlage einer abschließenden Bewertung der Tracking- sowie Gesamtsystemleistung, die in realistischen Umgebungen, wie Stadt, Land und Autobahn, sowie unter verschiedenen Witterungsbedingungen, den Mehrwert einer multisensoriellen Fahrumfelderfassung untersucht. Dabei ergab sich für Fußgänger und Fahrzeuge ein ähnliches Bild. Unter normalen Witterungsbedingungen konnte bei Autobahn- und Landstraßenszenen mit einem Laserscanner eine sehr gute Trackingleistung erreicht werden. Durch die Hinzunahme von weiteren Sensoren war in diesen Domänen keine weitere Leistungssteigerung möglich. In komplexeren städtischen Szenarien oder in sehr dynamischen Szenen ist ein leichter Abfall der Trackingqualität zu beobachten. An dieser Stelle kann durch die Fusion mit weiterer Sensorik die Qualität der Schätzung einzelner Zustandsgrößen optimiert werden. So profitiert bei Fahrzeugen insbesondere die Schätzung der Geschwindigkeit durch die Hinzunahme der Radarsensorik. Für Fußgänger kann durch die Fusion mit der Wärmebildkamera die Qualität der Schätzung der lateralen Ablage verbessert werden. Darüber hinaus zeigt sich unter widrigen Umweltbedingungen, wie starker Regen oder Schneefall, dass mit einer multisensoriellen Fahrumfelderfassung eine Steigerung der Zuverlässigkeit erzielt werden kann.

Zusammenfassend ist ein Mehrwert der in dieser Arbeit betrachteten Sensorfusion in zwei Bereichen zu beobachten. Zum Einen kann in hochdynamischen und komplexen Szenarien die Präzision der Schätzung einzelner Zustandsgrößen erhöht werden. So profitiert zum Beispiel die Geschwindigkeitsschätzung von Fahrzeugen durch die Radarsensorik und die Schätzung der lateralen Ablage von Fußgängern durch die zusätzliche Ferninfrarotkamera. Zum Anderen lässt sich insbesondere unter widrigen Umweltbedingungen die Robustheit des Gesamtsystems steigern, da die verschiedenen physikalischen Messprinzipien davon unterschiedlich stark beeinflusst werden.



## 8.2 Wissenschaftlicher Beitrag

Ein wesentlicher wissenschaftlicher Beitrag dieser Arbeit ist die konsequente Betrachtung der Objektverfolgung von Fahrzeugen und Fußgängern in Bezug auf eine Fusion von vier Sensoren unterschiedlicher Messphysik. Dabei wurden alle relevanten Bereiche in der Umsetzung eines Fahrumfelderfassungssystems behandelt und gegebenenfalls durch neuartige Methoden ergänzt und deren Mehrwert diskutiert.

Für das Beobachtungsmodell des Laserscanners wurden in Kapitel 4 Optimierungen der Vorverarbeitung vorgestellt. Im Vergleich zum einfachen „Region-Growing“-Vorgehen, wird durch die eingeführte inkrementellen Linienextraktion eine vollständigere und robustere Segmentierung von Fußgängern und Fahrzeugen erreicht. Wird zusätzlich RANSAC in den Prozess der Merkmalsgewinnung integriert, so lässt sich der Einfluss von Messfehlern reduzieren.

In Kapitel 6 wurde der Visualisierungsprozesses von heterogenen Sensordaten behandelt und Konzepte, die zum Teil auf pragmatischen Visualisierungsansätzen aus der Sensordomäne beruhen, einer systematischen Ordnung zugeführt. Darüber hinaus wurden in diesem Zusammenhang die Möglichkeiten und Grenzen der erweiterten Realität an einem realen Fallbeispiel diskutiert.

Die effiziente Gewinnung von präzisen Referenzdaten ist eines der Probleme in der Evaluierung einer Fahrumfelderfassung. Hierzu wurden in Kapitel 7 Lösungsansätze dargestellt, die den Arbeitsaufwand um den Faktor 5-10 reduzieren oder durch zusätzliche Sensorik hochgenaue Referenzdaten zur Verfügung stellen.

## 8.3 Ausblick

Abschließend werden einige Vorschläge zur Erweiterung und Verbesserung des in dieser Arbeit vorgestellten Fahrumfelderfassungssystems aufgeführt.

**Filterverfahren** Im Rahmen dieser Arbeit konnte nur eine kleine Auswahl an Filtervarianten betrachtet werden. Für sehr dynamische Szenen verspricht der IMM-Kalman-Filter [BBS88] eine Steigerung der Trackingqualität, da er gleichzeitig mehrere dynamische Modelle berücksichtigt. Ein ähnliches Verfahren könnte auch den Partikel-Filter weiter optimieren. Allerdings ist bei den Partikel-Filtern im Augenblick die ungenügende Echtzeitfähigkeit von größerer Bedeutung. An dieser Stelle könnte zum Beispiel der Rao-Blackwellization Partikel-Filter [DdFMR00] eine viel versprechende Alternative darstellen, da er den Kalman-Filter und den Partikel-Filter kombiniert und nur für die kritische Zustandsgrößen den zeitaufwendigen Partikel-Filter verwendet. Darüber hinaus gewinnen Occupancy Grids [Thr01] zunehmend an Bedeutung und sollten deshalb näher betrachtet werden.

Das im Rahmen dieser Dissertation verwendete Verfahren zur Kompensation von Latenzzeiten hat sich als ausreichend erwiesen. Dennoch sollte in weiterführenden Arbeiten das Potential komplexerer Algorithmen, wie sie in [vdMWJ04] beschrieben sind, evaluiert werden.

**Beobachtungsmodelle** Die in Kapitel 4 betrachteten Modelle stellen lediglich eine Untermenge der vielfältigen Möglichkeiten dar, wie sich Fahrzeuge und Fußgänger durch Laserscanner, Radarsensoren bzw. bildgebende Sensorik beobachten lassen. Durch eine zusätzliche Auswertung der

Rückstrahlintensitäten des Laserscanners könnte zum Beispiel eine Verbesserung in der Orientierungsschätzung erreicht werden. Die Messmodelle für die bildgebende Sensorik lassen sich ebenfalls weiter optimieren. So könnte für Fußgänger durch eine konturbasierte Fusion zwischen Video und Ferninfrarot [BHWB08] die Entfernungsschätzung und Robustheit verbessert werden. Zudem sollte in weiteren Arbeiten geklärt werden, in welchem Maß sich dieses Verfahren auf Fahrzeuge übertragen lässt. In sehr dynamischen Fußgängerszenen hat sich gezeigt, dass die Modellannahme für die Kopfkontur zu rigide ist. An dieser Stelle könnte ein erweitertes Beobachtungsmodell die Genauigkeit des Trackings erhöhen.

Im Rahmen dieser Arbeit wurden Verdeckungen zwischen den einzelnen Verkehrsteilnehmern nicht behandelt. Für Autobahn- und Landstraßenszenarien kann deren Einfluss im Allgemeinen vernachlässigt werden. Allerdings ist insbesondere in Stadtszenarien und für Fußgängergruppen ein Mehrwert zu erwarten, da in diesen Szenarien Verdeckungen besonders häufig auftreten.

**Visualisierung** Die in Kapitel 6 aufgestellten Thesen stützen sich vorwiegend auf wissenschaftlich fundierten Visualisierungsrichtlinien und die im Rahmen dieser Arbeit gewonnene Erfahrung bei der Umsetzung der Fahrumfelderfassung. Dennoch sollte eine Nutzerstudie durchgeführt werden, um das Potential unterschiedlicher Visualisierungsmethoden im Entwicklungsprozess von Fahrumfelderfassungssystem weiter zu ergründen. Darüber hinaus bieten andere Visualisierungsmedien wie die Virtual-Reality oder das Head-Up-Display Raum für weitere Forschungen.

**Fahrumfelderfassung** Um der Vision einer allumfassenden Fahrumfelderfassung näher zu kommen, sollten in weiterführenden Arbeiten zusätzlich Fahrradfahrer, Motorräder, LKWs oder Fußgängergruppen adressiert werden. Noch einen Schritt weiter geht die Erkennung von Tieren auf nächtlichen Landstraßen, was mit der in dieser Arbeit verwendeten Sensorik prinzipiell möglich ist. Neben der Erweiterung der Objektklassen gewinnt für zukünftige Fahrerassistenzsysteme zunehmend der maximale Erfassungsbereich an Bedeutung. So ist zum Beispiel der Spurwechselassistent auf einen Sichtbereich von nahezu 360° angewiesen. Obwohl die in dieser Arbeit verwendete Sensorik auf den Bereich vor dem Fahrzeug ausgerichtet ist, sind die vorgestellten Ansätze hiervon weitgehend unabhängig und können somit auch für größere Erfassungsbereiche verwendet werden.

Durch zukünftige Kommunikationsnetzwerke wie Car2Car lässt sich der Begriff der Sensorfusion über die eigene Fahrzeuggrenze hinweg erweitern. An dieser Stelle bieten zusätzliche Problemstellungen wie eingeschränkte Verfügbarkeit der Sensoren, eine schmale Bandbreite der Kommunikation, eine dynamische Kalibrierung der Sensoren oder große und schwankende Latenzzeiten Raum für weitere Forschungsaktivitäten.

Im Rahmen des europäischen Forschungsprojektes COMPOSE demonstrierte BMW als Beispiel eines visionären Fahrerassistenzsystems die „aktiven Gefahrenbremsung“ für Fußgänger. Grundlage des vorgestellten Systems war die in dieser Arbeit präsentierte multisensorielle Fahrumfelderfassung. Der Mehrwert dieser Sensorfusion in Bezug auf zukünftige Fahrerassistenzsysteme konnte live im Fahrzeug erlebt werden. Darüber hinaus wurde den Insassen vor Augen geführt, in welchem Maße die allgemeine Verkehrssicherheit durch den Einsatz von mehreren Sensoren und intelligenter Assistenzfunktionen gesteigert werden kann.



# LITERATURVERZEICHNIS

---

- [ABB<sup>+</sup>01] AZUMA, R.T., Y. BAILLOT, R. BEHRINGER, S. FEINER, S. JULIER und B. MACINTYRE: *Recent advances in augmented reality*. Computer Graphics and Applications, IEEE, 21(6):34–47, 2001.
- [ABW01] ALLEN, B.D., G. BISHOP und G. WELCH: *Tracking: Beyond 15 Minutes of Thought*. Course Notes, Ann. Conf. Computer Graphics and Interactive Techniques, 2001.
- [Azu97] AZUMA, R.T.: *A survey of augmented reality*. Presence: Teleoperators and Virtual Environments(1054-7460), 6(4):355–385, 1997.
- [Bau07] BAUER, MARTIN A.: *Tracking Errors in Augmented Reality*. Doktorarbeit, Technische Universität München, 2007.
- [BBC<sup>+</sup>00] BALDI, P., S. BRUNAK, Y. CHAUVIN, C.A.F. ANDERSEN und H. NIELSEN: *Assessing the accuracy of prediction algorithms for classification: an overview*. Bioinformatics, 16(5):412–424, 2000.
- [BBG<sup>+</sup>04] BERTOZZI, M., A. BROGGI, P. GRISLERI, A. TIBALDI und M.D. ROSE: *A tool for vision based pedestrian detection performance evaluation*. Intelligent Vehicles Symposium, 2004 IEEE, Seiten 784–789, 2004.
- [BBK04] BROADHURST, A., S. BAKER und T. KANADE: *A Prediction and Planning Framework for Road Safety Analysis, Obstacle Avoidance and Driver Information*. In: *Proceedings of the 11th World Congress on Intelligent Transportation Systems*, October 2004.
- [BBS88] BLOM, HAP und Y. BAR-SHALOM: *The interacting multiple model algorithm for systems with Markovian switching coefficients*. Automatic Control, IEEE Transactions on, 33(8):780–783, 1988.
- [Bed92] BEDWORTH, M.: *Probability Moderation for Multilevel Information Processing*. Personal Communication, 1992.
- [Ben81] BENEŠ, VE: *Exact finite-dimensional filters for certain diffusions with nonlinear drift*. Stochastics An International Journal of Probability and Stochastic Processes, 5(1):65–92, 1981.
- [BER03] BLACK, J., T. ELLIS und P. ROSIN: *A novel method for video tracking performance evaluation*. Joint IEEE Int. Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS), Seiten 125–132, 2003.

- [BGZ02] BUNGARTZ, H.J., M. GRIEBEL und C. ZENGER: *Einführung in die Computergraphik: Grundlagen, geometrische Modellierung, Algorithmen*. Vieweg, 2002.
- [BH94] BAUMBERG, AM und DC HOGG: *An efficient method for contour tracking using active shape models*. Motion of Non-Rigid and Articulated Objects, 1994., Proceedings of the 1994 IEEE Workshop on, Seiten 194–199, 1994.
- [BHWB08] BAUER, G., F. HOMM, L. WALCHSHÄUSL und D. BURSCHKA: *Multi Spectral Pedestrian Detection and Localization*. Advanced Microsystems for Automotive Applications 2008, 2008.
- [BLN<sup>+</sup>98] BAK, M., T.D. LARSEN, M. NORGAARD, N.A. ANDERSEN, N.K. POULSEN und O. RAVN: *Location estimation using delayed measurements*. Advanced Motion Control, 1998. AMC '98-Coimbra., 1998 5th International Workshop on, Seiten 180–185, 1998.
- [BM06] BOCK, T. und M. MAURER: *Vehicle in the Loop Simulation*. In: *4. Workshop Fahrerassistenzsysteme*, Seiten 125–134, Löwenstein / Hößlinsülz, October 2006.
- [BMF07] BOCK, T., M. MAURER und G. FÄRBER: *Validation of the Vehicle in the Loop (VIL) – A milestone for the simulation of driver assistance systems*. In: *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Juni 2007. IEEE.
- [BO00] BEDWORTH, M. und J. O'BRIEN: *The Omnibus model: a new model of data fusion?* IEEE Aerospace and Electronic Systems Magazine, 15(4):30–36, 2000.
- [Bor06] BORNER, A.: *Optische Sensoren für Fahrzeuge*. VDI BERICHTE, 1944:37, 2006.
- [Bou07] BOUGUET, JEAN-YVES: *Camera Calibration Toolbox for Matlab*. [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/), 2007.
- [Boy87] BOYD, J.R.: *A Discourse on Winning and Losing*. Maxwell AFB lecture, 1987.
- [BSC86] BAR-SHALOM, Y. und L. CAMPO: *The effect of the common process noise on the two-sensor fused-track covariance*. IEEE Transactions on Aerospace and Electronic Systems, 22:803–05, 1986.
- [BSL93] BAR-SHALOM, Y. und X.R. LI: *Estimation and tracking- Principles, techniques, and software*. Norwood, MA: Artech House, Inc, 1993., 1993.
- [Bun07] BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG: *INVENT*. <http://www.invent-online.de/index.html>, 2007.
- [CA03] CHANG, C. und R. ANSARI: *Kernel particle filter: iterative sampling for efficient visual tracking*. Image Processing, 2003. IICIP 2003. Proceedings. 2003 International Conference on, 3, 2003.
- [Che03] CHEN, Z.: *Bayesian filtering: From Kalman filters to particle filters, and beyond*. Adaptive Syst. Lab., McMaster Univ., Hamilton, ON, Canada.[Online]. Available: [http://soma.crl.mcmaster.ca/zhechen/download/ieee\\_bayesian.ps](http://soma.crl.mcmaster.ca/zhechen/download/ieee_bayesian.ps), 2003.
- [Che04] CHEN, C.: *Information Visualization: Beyond the Horizon*. Springer, 2004.
- [CRM03] COMANICIU, D., V. RAMESH und P. MEER: *Kernel-based object tracking*. Transactions on Pattern Analysis and Machine Intelligence, 25(5):564–577, 2003.

- [CS86] COMISO, JC und CW SULLIVAN: *Satellite Microwave and in situ observations of the Weddell Sea ice cover and its marginal ice zone*. Journal of Geophysical Research, 91(C8):9663–9682, 1986.
- [DAR07] DARPA: *Grand Challenge*. <http://www.darpa.mil/grandchallenge/>, 2007.
- [Das97] DASARATHY, B.V.: *Sensor fusion potential exploitation-innovative architectures and illustrative applications*. Proceedings of the IEEE, 85(1):24–38, 1997.
- [Das00] DASARATHY, B.V.: *More the merrier...or is it? Sensor suite augmentation benefits assessment*. Information Fusion, 2000. FUSION 2000. Proceedings of the Third International Conference on, 2:WEC3/20–WEC3/25 vol.2, 2000.
- [Dau86] DAUM, F.: *Exact finite-dimensional nonlinear filters*. Automatic Control, IEEE Transactions on, 31(7):616–622, 1986.
- [DBR00] DEUTSCHER, J., A. BLAKE und I. REID: *Articulated body motion capture by annealed particle filtering*. Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, 2, 2000.
- [DdFMR00] DOUCET, A., N. DE FREITAS, K. MURPHY und S. RUSSELL: *Rao-Blackwellised particle filtering for dynamic Bayesian networks*. Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence, Seiten 176–183, 2000.
- [DGA00] DOUCET, A., S. GODSILL und C. ANDRIEU: *On sequential Monte Carlo sampling methods for Bayesian filtering*. Statistics and Computing, 10(3):197–208, 2000.
- [Dic02] DICKMANN, E.D.: *The development of machine vision for road vehicles in the last decade*. Intelligent Vehicle Symposium, 2002. IEEE, 1:268–281 vol.1, 2002.
- [DM00] DOERMANN, D. und D. MIHALCIK: *Tools and techniques for video performance evaluation*. Pattern Recognition, 2000. Proceedings. 15th International Conference on, 4, 2000.
- [DW88] DURRANT-WHYTE, HF: *Sensor models and multisensor integration*. International Journal of Robotics Research, 7(6):97–113, 1988.
- [Eur07a] EUROPEAN COMMISSION: *PreVENT*. <http://www.prevent-ip.org/>, 2007.
- [Eur07b] EUROPEAN COMMISSION: *Radar Net*. <http://www.radarnet.org/>, 2007.
- [FB81] FISCHLER, M.A. und R.C. BOLLES: *Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography*. Communications of the ACM, 24(6):381–395, 1981.
- [FD04] FUERSTENBERG, KC und K. DIETMAYER: *Object tracking and classification for multiple active safety and comfort applications using a multilayer laser scanner*. Intelligent Vehicles Symposium, 2004 IEEE, Seiten 802–807, 2004.
- [FD05] FUERSTENBERG, K.C. und K. DIETMAYER: *Fußgängerschutz unter Nutzung mehrzeiliger Laserscanner*. FAS Workshop Fahrerassistenzsysteme, Seiten 111–120, 2005.
- [FH75] FUKUNAGA, K. und L. HOSTETLER: *The estimation of the gradient of a density function, with applications in pattern recognition*. Information Theory, IEEE Transactions on, 21(1):32–40, 1975.



- [Foe06] FOELSTER, FLORIAN: *Erfassung ausgedehnter Objekte durch ein Automobil-Radar*. Doktorarbeit, Technischen Universität Hamburg-Harburg, 2006.
- [Fol95] FOLEY, J.D.: *Computer Graphics: Principles and Practice*. Addison-Wesley Professional, 1995.
- [FP03] FORSYTH, D.A. und J. PONCE: *Computer Vision: A Modern Approach*. Prentice Hall, 2003.
- [Fre04] FREYMAN, RAYMOND: *Möglichkeiten und Grenzen von Fahrerassistenz- und Aktiven Sicherheitssystemen*. In: *Tagung Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, 03 2004.
- [FS05] FÜRSTENBERG, KAY CH. und JOCHEN SCHOLZ: *Reliable Pedestrian Protection using Laserscanners*. In: *The IEEE Intelligent Vehicles Symposium (IV'05)*, Las Vegas, Nevada, U.S.A., June 2005.
- [FTSW01] FLEMING, WJ, S. TECHNOL, T.R.W.O.S. SYST und MI WASHINGTON: *Overview of automotive sensors*. *Sensors Journal, IEEE*, 1(4):296–308, 2001.
- [GA93] GREWAL, M.S. und A.P. ANDREWS: *Kalman filtering: theory and practice*. Prentice-Hall Information And System Sciences Series, Seite 381, 1993.
- [GH01] GAN, Q. und C.J. HARRIS: *Comparison of two measurement fusion methods for Kalman-filter-based multisensor data fusion*. *Aerospace and Electronic Systems, IEEE Transactions on*, 37(1):273–279, 2001.
- [GH02] GAO, JB und CJ HARRIS: *Some remarks on Kalman filters for the multisensor fusion*. *Information Fusion*, 3(3):191–201, 2002.
- [GHJV95] GAMMA, E., R. HELM, R. JOHNSON und J. VLISSIDES: *Design patterns: elements of reusable object-oriented software*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1995.
- [GJE<sup>+</sup>04] GRESHAM, I., A. JENKINS, R. EGRI, C. ESWARAPPA, N. KINAYMAN, N. JAIN, R. ANDERSON, F. KOLAK, R. WOHLERT, SP BAWELL et al.: *Ultra-wideband radar sensors for short-range vehicular applications*. *Microwave Theory and Techniques, IEEE Transactions on*, 52(9):2105–2122, 2004.
- [GM07] GAVRILA, D.M. und S. MUNDER: *Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle*. *International Journal of Computer Vision*, 73(1):41–59, 2007.
- [GPS<sup>+</sup>07] GALL, JÜRGEN, JÜRGEN POTTHOFF, CHRISTOPH SCHNÖRR, BODO ROSENHAHN und HANS-PETER SEIDEL: *Interacting and Annealing Particle Filters: Mathematics and a Recipe for Applications*. *J. Math. Imaging Vis.*, 28(1):1–18, 2007.
- [HGS07] HARATI, A., S. GACHTER und R. SIEGWART: *Fast range image segmentation for indoor 3D-SLAM*. 6th IFAC Symposium on Intelligent Autonomous Vehicles, Toulouse, France, 2007.
- [HM54] HAMMERSLEY, JM und KW MORTON: *Poor Man's Monte Carlo*. *Journal of the Royal Statistical Society. Series B (Methodological)*, 16(1):23–38, 1954.



- [HZ03] HARTLEY, R. und A. ZISSERMAN: *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
- [IB98] ISARD, M. und A. BLAKE: *CONDENSATION—Conditional Density Propagation for Visual Tracking*. International Journal of Computer Vision, 29(1):5–28, 1998.
- [IB05] IYENGAR, SS und R.R. BROOKS: *Distributed Sensor Networks*. Chapman & Hall/CRC, 2005.
- [INR02] INRIA: *CARSENSE*. <http://www.carsense.org/en/>, 2002.
- [JH97] J. HEIKKILA, O. SILVEN: *A Four-Step Camera Calibration Procedure with Implicit Image Correction*. In Proc. of IEEE Computer Vision and Pattern Recognition, Seiten 1106–1112, 1997.
- [JJ02] JANSSON, J. und J. JOHANSSON: *Decision Making for Collision Avoidance Systems*. Technischer Bericht 2002-01-0403, Society of Automotive Engineers, Detroit, 2002.
- [JU04] JULIER, S.J. und J.K. UHLMANN: *Unscented filtering and nonlinear estimation*. Proceedings of the IEEE, 92(3):401–422, 2004.
- [JUDW95] JULIER, SJ, JK UHLMANN und HF DURRANT-WHYTE: *A new approach for filtering nonlinear systems*. American Control Conference, 1995. Proceedings of the, 3, 1995.
- [JUDW00] JULIER, S., J. UHLMANN und H.F. DURRANT-WHYTE: *A new method for the nonlinear transformation of means and covariances in filters and estimators*. Automatic Control, IEEE Transactions on, 45(3):477–482, 03 2000.
- [JWSX02] JAYNES, C., S. WEBB, R.M. STEELE und Q. XIONG: *An Open Development Environment for Evaluation of Video Surveillance Systems*. Proceedings of the Third International Workshop on Performance Evaluation of Tracking and Surveillance (PETS'2002), Copenhagen, June, 2002.
- [Kal60] KALMAN, R.E.: *A new approach to linear filtering and prediction problems*. Journal of Basic Engineering, 82(1):35–45, 1960.
- [KBD05] KÄMPCHEN, N., M. BUEHLER und K. DIETMAYER: *Feature-level fusion for free-form object tracking using laserscanner and video*. Intelligent Vehicles Symposium, 2005. Proceedings. IEEE, Seiten 453–458, 2005.
- [KGJV83] KIRKPATRICK, S., CD GELATI JR und MP VECCHI: *Optimization by Simulated Annealing*. Science, 220(4598):671680, 1983.
- [KJG04] KARLSSON, R., J JANSSON und F. GUSTAFSSON: *Model-based statistical tracking and decision making for collision avoidance application*. American Control Conference, 2004. Proceedings of the 2004, 4:3435–3440 vol.4, 2004.
- [KKNK03] KLAPPER, S., R. KYLE, R. NICKLIN und A. KORMOS: *Night Vision—Changing the Way We Drive [White Paper] Texas*. USA: Raytheon Electronic Systems, 2003.
- [Käm07] KÄMPCHEN, N.: *Feature-level fusion of laser scanner and video data for advanced driver assistance systems*. Doktorarbeit, Universität Ulm. Fakultät für Ingenieurwissenschaften und Informatik, 06 2007.

- [KTÅ00] KAHL, F., B. TRIGGS und K. ÅSTRÖM: *Critical Motions for Auto-Calibration When Some Intrinsic Parameters Can Vary*. Journal of Mathematical Imaging and Vision, 13(2):131–146, 2000.
- [LJ03] LAVIOLA JR, JJ: *A comparison of unscented and extended Kalman filtering for estimating quaternion motion*. American Control Conference, 2003. Proceedings of the 2003, 3, 2003.
- [LWP07] LINDL, R., L. WALCHSHÄUSL und V. PARASCHOUDIS: *Autonomous Braking for Collision Mitigation Purposes by Means of Multi Sensor Perception*. In: *6th European Congress and Exhibition on Intelligent Transport Systems and Services*, Aalborg, Denmark, Juni 2007. IEEE.
- [Mac86] MACKINLAY, J.: *Automating the design of graphical presentations of relational information*. ACM Transactions on Graphics, 5(2):110–141, 1986.
- [MBN04] MENDES, A., LC BENTO und U. NUNES: *Multi-target detection and tracking with a laser scanner*. Intelligent Vehicles Symposium, 2004 IEEE, Seiten 796–801, 2004.
- [MDB87] MCCORMICK, B.H., T.A. DEFANTI und M.D. BROWN: *Visualization in scientific computing-a synopsis*. IEEE Computer Graphics & Applications, 7(7):61–70, 1987.
- [MDG<sup>+</sup>05] MERTZ, C., D. DUGGINS, J. GOWDY, J. KOZAR, R. MACLACHLAN, A. STEINFELD, A. SUPPE, T. CHUCK, THORPE und C. WANG: *Collision Warning and Sensor Data Processing in Urban Areas*. In: *Proceedings of the 5th international conference on ITS telecommunications*, Seiten 73–78, June 2005.
- [MDW94] MANYIKA, J. und H. DURRANT-WHYTE: *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*. Ellis Horwood, New York, 1994.
- [Mol95] MOLNÁR, PÉTER: *Modellierung und Simulation der Dynamik von Fußgängerströmen*. Doktorarbeit, Institut für Theoretische Physik der Universität Stuttgart, 1995.
- [MS05] MAURER, M. und C. STILLER: *Fahrerassistenzsysteme mit maschineller Wahrnehmung*. Springer, 2005.
- [Naa04] NAAB, K.: *Sensorik-und Signalverarbeitungsarchitekturen für Fahrerassistenz und Aktive Sicherheit*. In: *Tagung Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, 03 2004.
- [Nev05] NEVE, ANTJE: *Fußgängererkennung in FIR und Video Sequenzen mithilfe entfernungsgebender Sensorik*. Diplomarbeit, Technische Universität München, 2005.
- [NMTS05] NGUYEN, V., A. MARTINELLI, N. TOMATIS und R. SIEGWART: *A comparison of line extraction algorithms using 2D laser rangefinder for indoor mobile robotics*. Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on, Seiten 1929–1934, 2005.
- [NO87] NIELSON, G. und D. OLSEN: *Direct manipulation techniques for 3D objects using 2D locator devices*. Proceedings of the 1986 workshop on Interactive 3D graphics, Seiten 175–182, 1987.
- [Nor94] NORM, D.: *Din 70 000 Fahrzeugdynamik und Fahrverhalten (Vehicle dynamics and handling)*, 1994.

- [NWB04] NEWMAN, J., M. WAGNER und M. BAUER: *Ubiquitous Tracking for Augmented Reality*. In: *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR'04)*, Arlington, VA, USA, Nov. 2004.
- [Oxf08] OXFORD TECHNICAL SOLUTIONS: *RT3003*. <http://www.oxts.co.uk/>, 2008.
- [Par62] PARZEN, E.: *On Estimation of a Probability Density Function and Mode*. *The Annals of Mathematical Statistics*, 33(3):1065–1076, 1962.
- [PB88] PHILLIPS, C. und N. BADLER: *JACK: a toolkit for manipulating articulated figures*. *Proceedings of the 1st annual ACM SIGGRAPH symposium on User Interface Software*, Seiten 221–229, 1988.
- [PHBK06] PUSTKA, D., M. HUBER, M. BAUER und G. KLINKER: *Spatial Relationship Patterns: Elements of Reusable Tracking and Calibration Systems*. In: *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR'06)*, October 2006.
- [PHK08] PUSTKA, D., M. HUBER und G. KLINKER: *Integrating Gyroscopes into Ubiquitous Tracking Environments*. *Virtual Reality Conference, 2008. VR'08. IEEE*, Seiten 283–284, 2008.
- [RA04] RISTIC, B. und S. ARULAMPALAM: *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, 2004.
- [RBW05] REGENBRECHT, H., G. BARATOFF und W. WILKE: *Augmented reality projects in the automotive and aerospace industries*. *IEEE Computer Graphics and Applications*, 25(6):48–56, 2005.
- [RC02] ROGALSKI, A. und K. CHRZANOWSKI: *Infrared devices and techniques*. *Optoelectronics Review*, 10(2):111–36, 2002.
- [RC04] ROBERT, C.P. und G. CASELLA: *Monte Carlo Statistical Methods*. Springer, 2004.
- [RDB01] ROLLAND, J.P., L.D. DAVIS und Y. BAILLOT: *A Survey of Tracking Technology for Virtual Environments*. *Fundamentals of Wearable Computers and Augmented Reality*, 2001.
- [RG05] RASSHOFER, R.H. und K. GRESSER: *Automotive Radar and Lidar Systems for Next Generation Driver Assistance Functions*. *Advances in Radio Science*, 3(10):205–209, 2005.
- [RH95] ROBINETT, W. und R. HOLLOWAY: *The visual display transformation for virtual reality*. *Presence: Teleoperators and Virtual Environments*, 4(1):1–23, 1995.
- [RM88] ROECKER, J.A. und C.D. MCGILLEM: *Comparison of two-sensor tracking methods based on state vector fusion and measurement fusion*. *Aerospace and Electronic Systems, IEEE Transactions on*, 24(4):447–449, 1988.
- [RPL07] RUSER, H. und F. PUENTE LEÓN: *Informationsfusion - Eine Übersicht*. *Technisches Messen*, 74(3):93–102, 2007.
- [RR91] ROBINETT, W. und J.P. ROLLAND: *Computational model for the stereoscopic optics of a head-mounted display*. *Proceedings of SPIE*, 1457:140, 1991.

- [SBM06] SUN, Z., G. BEBIS und R. MILLER: *On-Road Vehicle Detection: A Review*. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, Seiten 694–711, 2006.
- [SBW99] STEINBERG, A.N., C. BOWMAN und F.E. WHITE: *Revisions to the JDL data fusion process model*. Sensor Fusion: Architectures, Algorithms, and Applications III, SPIE Proceedings, 3719:430–441, 1999.
- [Sch04] SCHNEBERGER, M.: *Infrared Optical Tracking Systems Mathematical and Operation Principles*. Technischer Bericht, A.R.T. GmbH, 2004.
- [SDS01] SPARBERT, J., K. DIETMAYER und D. STRELLER: *Lane detection and street type classification using laser range images*. Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE, Seiten 454–459, 2001.
- [SG06] SEECK, A. und T.M. GASSER: *Klassifizierung und Würdigung der rechtlichen Rahmenbedingungen im Zusammenhang mit der Einführung moderner FAS*. In: *Tagung Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, 04 2006.
- [Shn96] SHNEIDERMAN, B.: *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations*. In: *IEEE Visual Languages*, Seiten 336–343, College Park, Maryland 20742, U.S.A., 1996.
- [Shn03] SHNEIDERMAN, B.: *Why not make interfaces better than 3D reality?* Computer Graphics and Applications, IEEE, 23(6):12–15, 2003.
- [Sho92] SHOEMAKE, K.: *ARCBALL: A User Interface for Specifying Three-Dimensional Orientation Using a Mouse*. Graphics Interface, 92:151–156, 1992.
- [Sil86] SILVERMAN, BW: *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, 1986.
- [SKK<sup>+</sup>97] SIADAT, A., A. KASKE, S. KLAUSMANN, M. DUFAUT und R. HUSSON: *An Optimized Segmentation Method for a 2D Laser-Scanner Applied to Mobile Robot Navigation*. Proceedings of the 3rd IFAC Symposium on Intelligent Components and Instruments for Control Applications, 1997.
- [SM00] SCHUMAN, H. und W. MÜLLER: *Visualisierung*. Springer Verlag, 2000.
- [SMP05] SVOBODA, T., D. MARTINEC und T. PAJDLA: *A Convenient Multi-Camera Self-Calibration for Virtual Environments*. Teleoperators and Virtual Environments, 14:4, 2005.
- [Spe00] SPENCE, R.: *Information Visualization*. Addison Wesley, Dezember 2000.
- [SS06] STOSSEL, D. und G. SAGERER: *Kernel Particle Filter for Visual Quality Inspection from Monocular Intensity Images*. LECTURE NOTES IN COMPUTER SCIENCE, 4174:597, 2006.
- [ST88] SAMET, H. und M. TAMMINEN: *Efficient component labeling of images of arbitrary dimension represented by linear bintrees*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 10(4):579–586, 1988.

- [Sta06] STATISTISCHES BUNDESAMT: *Verkehrsunfälle - Fachserie 8 Reihe 7 - 2006*. Technischer Bericht 8, Statistisches Bundesamt, 2006. Reihe 7.
- [Svo03] SVOBODA, T.: *Quick guide to multi-camera self-calibration*. Technischer Bericht, BiWi-TR-263, Computer Vision Lab, Swiss Federal Institute of Technology, August 2003.
- [SWND05] SHREINER, D., M. WOO, J. NEIDER und T. DAVIS: *OpenGL(R) Programming Guide: The Official Guide to Learning OpenGL(R), Version 2 (5th Edition) (OpenGL)*. Addison-Wesley Professional, 2005.
- [TFF<sup>+</sup>07] TATSCHKE, T., F. FÄRBER, E. FUCHS, L. WALCHSHÄUSL und R. LINDL: *Semi-autonomous reference data generation for perception performance evaluation*. In: *Proceedings of the 10th International Conference on Information Fusion (FUSION)*, Seiten 1–7, Quebec, Kanada, Juli 2007.
- [Thr01] THRUN, S.: *Learning occupancy grids with forward models*. Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on, 3, 2001.
- [TLWK07] TÖNNIS, M., R. LINDL, L. WALCHSHÄUSL und G. KLINKER: *Visualization of Spatial Sensor Data in the Context of Automotive Environment Perception Systems*. In: *Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR)*, Nara, Japan, November 2007.
- [TN00] TUCERYAN, M. und N. NAVAB: *Single point active alignment method (SPAAM) for optical see-through HMD calibration for AR*. Augmented Reality, 2000.(ISAR 2000). Proceedings. IEEE and ACM International Symposium on Augmented Reality, Seiten 149–158, 2000.
- [TRC01] TAN, D.S., G.G. ROBERTSON und M. CZERWINSKI: *Exploring 3D navigation: combining speed-coupled flying with orbiting*. Proceedings of the SIGCHI conference on Human factors in computing systems, Seiten 418–425, 2001.
- [Var97] VARSHNEY, P.K.: *Multisensor data fusion*. Electronics & Communication Engineering Journal, 9(6):245–253, 12 1997.
- [vdHvW01] HEIJDEN, R. VAN DER und K. VAN WEES: *Introducing Advanced Driver Assistance Systems: Some Legal Issues*. European Journal for Transport and Infrastructure Research, 1(3):309–326, 2001.
- [vdMWJ04] MERWE, R. VAN DER, E. WAN und S. JULIER: *Sigma-Point Kalman Filters for Nonlinear Estimation and Sensor-Fusion: Applications to Integrated Navigation*. Proceedings of the AIAA Guidance, Navigation & Control Conference, Seiten 16–19, 2004.
- [Vic01] VICCI, L.: *Quaternions and Rotations in 3-Space: The Algebra and its Geometric Interpretation*. TR01-014, Department of Computer Science, UNC, Chapel Hill, NC, 2001.
- [Vog07] VOGEL, K.: *High-Accuracy Reference Data Acquisition for Evaluation of Active Safety Systems by means of a RTK-GNSS-Surveying System*. 6th European Congress and Exhibition on Intelligent Transport Systems and Services (ITS), June 2007.
- [Wal08] WALCHSHÄUSL, LEONHARD: *Klassifikation von Verkehrsteilnehmern mit heterogener Sensorik*. Doktorarbeit, TU München, 2008.



- [WB01] WELCH, G. und G. BISHOP: *An Introduction to the Kalman Filter*. ACM SIGGRAPH 2001 Course Notes, 2001.
- [WCD<sup>+</sup>06] WENGERT, C., P.C. CATTIN, J.M. DUFF, C. BAUR und G. SZEKELY: *Markerless Endoscopic Registration and Referencing*. LECTURE NOTES IN COMPUTER SCIENCE, 4190:816, 2006.
- [WCH92] WENG, J., P. COHEN und M. HERNIOU: *Camera Calibration with Distortion Models and Accuracy Evaluation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(10):965–980, 1992.
- [Wei03] WEISSTEIN, E.W.: *Least Squares Fitting–Perpendicular Offsets*. Online, 18(06), 2003.
- [WGS<sup>+</sup>04] WISSELMANN, D., K. GRESSER, H. SPANNHEIMER, K. BENGLER und A. HUESMANN: *ConnectedDrive–ein methodischer Ansatz für die Entwicklung zukünftiger Fahrerassistenzsysteme*. Garching, Tagung Aktive Sicherheit durch Fahrerassistenz, 2004.
- [Wik07] WIKIPEDIA: *Head-Mounted Display* — *Wikipedia, Die freie Enzyklopädie*, 2007. [Online; Stand 9. November 2007].
- [WO90] WARE, C. und S. OSBORNE: *Exploration and virtual camera control in virtual three dimensional environments*. Proceedings of the 1990 symposium on Interactive 3D graphics, Seiten 175–183, 1990.
- [XLF05] XU, F., X. LIU und K. FUJIMURA: *Pedestrian Detection and Tracking With Night Vision*. IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, 6(1):63, 2005.
- [YRYA05] YASUNO, M., S. RYOUSUKE, N. YASUDA und M. AOKI: *Pedestrian detection and tracking in far infrared images*. Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE, Seiten 182–187, Sept. 2005.
- [Zha99] ZHANG, Z.: *Flexible Camera Calibration By Viewing a Plane From Unknown Orientations*. International Conference on Computer Vision (ICCV’99), Seiten 666–673, Sep. 1999.