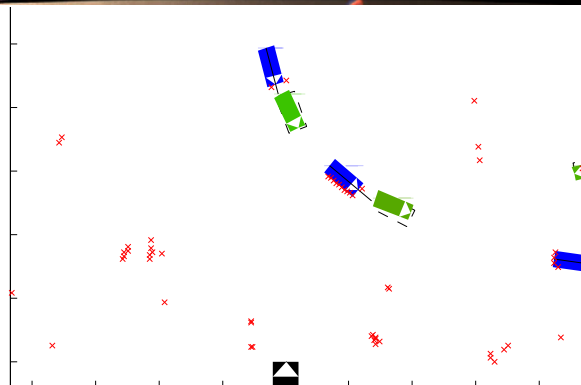




Masterarbeit

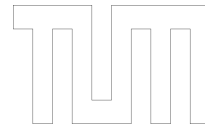
Erkennung und Prädiktion anderer Verkehrsteilnehmer



Verfasser: David Wittmann <Mat.-Nr.: 3600510>

ausgegeben am: 15. April 2013
eingereicht am: 15. Oktober 2013

Betreuer: Dipl.-Ing. Frederic Chucholowski



Masterarbeit für
Herrn David Wittmann
Erkennung und Prädiktion anderer Verkehrsteilnehmer

Um die bei der Teleoperation von Straßenfahrzeugen auftretenden Zeitverzögerungen in der Übertragung zu kompensieren wird die Position des eigenen Fahrzeugs prädiziert und in das übertragene Videobild eingezeichnet. Für eine korrekte Darstellung der Umgebung ist es zudem wichtig, das Verkehrsgeschehen um das Fahrzeug zu prädizieren. Je nach Geschwindigkeit der prädizierten Verkehrsteilnehmer werden sich diese von wenigen Zentimetern bis zu einigen Metern weiter bewegt haben. Ziel der Arbeit ist es, die am Lehrstuhl für Fahrzeugtechnik bestehenden Methoden zur Objekterkennung anhand von Laserscannern zu verbessern und mit den erkannten Objekten des Radarscanners zu fusionieren. Auf Basis der so gewonnenen Informationen sollen anschließend die Bewegungen der erkannten Objekte vorausberechnet werden. Nach Möglichkeit sollen dafür neben Eigenentwicklungen auch bewährte Methoden nach aktuellem Stand der Technik angewandt werden. Die Implementierung der Methoden erfolgt in Matlab/Simulink zunächst anhand von aufgezeichneten Messfahrten, in welchen die Videobilddaten zusammen mit den Laser- und Radarscannerdaten vorhanden sind. Bereits bei der Entwicklung soll die Echtzeitfähigkeit der Algorithmen berücksichtigt werden.

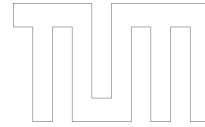
Die Ausarbeitung soll die einzelnen Arbeitsschritte in übersichtlicher Form dokumentieren. Der Kandidat verpflichtet sich, die Masterarbeit selbständig durchzuführen und die von ihm verwendeten wissenschaftlichen Hilfsmittel anzugeben. Die eingereichte Arbeit verbleibt als Prüfungsunterlage im Eigentum des Lehrstuhls und darf Dritten nur unter Zustimmung des Lehrstuhlinhabers zugänglich gemacht werden.

Prof. Dr.-Ing. M. Lienkamp

Betreuer: Dipl.-Ing. Frederic Chucholowski

Ausgabe: _____

Abgabe: _____



Geheimhaltungsverpflichtung

Herr/Frau: **Wittmann, David**

Im Rahmen der Angebotserstellung und der Bearbeitung von Forschungs- und Entwicklungsverträgen erhält der Lehrstuhl für Fahrzeugtechnik der Technischen Universität München regelmäßig Zugang zu vertraulichen oder geheimen Unterlagen oder Sachverhalten industrieller Kunden, wie z.B. Technologien, heutige oder zukünftige Produkte, insbesondere Prototypen, Methoden und Verfahren, technische Spezifikationen oder auch organisatorische Sachverhalte.

Der Unterzeichner verpflichtet sich, alle derartigen Informationen und Unterlagen, die ihm während seiner Tätigkeit am Lehrstuhl für Fahrzeugtechnik zugänglich werden, strikt vertraulich zu behandeln.

Er verpflichtet sich insbesondere

- derartige Informationen betriebsintern zum Zwecke der Diskussion nur dann zu verwenden, wenn ein ihm erteilter Auftrag dies erfordert,
- keine derartigen Informationen ohne die vorherige schriftliche Zustimmung des betreffenden Kunden an Dritte weiterzuleiten,
- keine Fotografien, Zeichnungen oder sonstige Darstellungen von Prototypen oder technischen Unterlagen hierzu anzufertigen,
- auf Anforderung des Lehrstuhls für Fahrzeugtechnik oder unaufgefordert spätestens bei seinem Ausscheiden aus dem Lehrstuhl für Fahrzeugtechnik alle Dokumente und Datenträger, die derartige Informationen enthalten, an Lehrstuhl für Fahrzeugtechnik zurückzugeben.

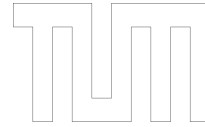
Die Verpflichtung zur Geheimhaltung endet nicht mit dem Ausscheiden aus dem Lehrstuhl für Fahrzeugtechnik, sondern bleibt 5 Jahre nach dem Zeitpunkt des Ausscheidens in vollem Umfang bestehen

Datum:

Unterschrift: _____



TECHNISCHE UNIVERSITÄT MÜNCHEN
LEHRSTUHL FAHRZEUGTECHNIK
PROF. DR.-ING. M. LIENKAMP



Eidesstattliche Erklärung

Hiermit versichere ich, die vorliegende Arbeit selbstständig und unter ausschließlicher Verwendung der angegebenen Literatur und Hilfsmittel erstellt zu haben. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

David Wittmann
München, 15. Oktober 2013

Formelzeichen und Abkürzungen

Formelzeichen

α	[°]	Winkelauflösung des LIDAR-Sensors
β_{detect}	[°]	minimaler Winkel zwischen Objekt und LIDAR-Strahl bei der Segmentierung
δ	[m]	maximaler Abstand zwischen Objekt und Referenzobjekt
σ	[-]	Standardabweichung
ς	[-]	Distanz im Zustandsraum
τ	[s]	Laufzeit des Laserstrahls
$\tau_{segment}$	[-]	maximale prozentuale Distanzänderung innerhalb eines Segments
τ_{I-Form}	[m]	maximaler mittlerer Abstand der Regressionsgerade für I-Form
τ_{L-Form}	[m]	maximaler mittlerer Abstand der Regressionsgerade für L-Form
$\tau_{L-Form,\psi}$	[°]	Grenzwert des Winkels zwischen den Seiten der L-Form
τ_{C-Form}	[m]	Maximaler Durchmesser einer C-Form
$\tau_{E-Form,1}$	[m]	Maximale große Halbachse der E-Form
$\tau_{E-Form,2}$	[m]	Maximale kleine Halbachse der E-Form
$\tau_{association}$	[m]	Grenzwert für die Assoziation von Segmenten und Objekten
$\tau_{seg,\psi}$	[°]	Orientierungs-Grenzwert für die Hypothesenbestätigung
$\tau_{seg,vel}$	[m/s]	Geschwindigkeits-Grenzwert für die Berücksichtigung der Bewegungsrichtung bei der Hypothesenbestätigung
$\tau_{seg,vel,\psi}$	[°]	maximale Richtungsdivergenz zwischen Orientierung und Geschwindigkeit
$\tau_{dist,\perp}$	[m]	maximaler senkrechter Abstand zu einer Form
$\tau_{dist,tot}$	[m]	maximaler totaler Abstand zu einer Form
$\tau_{max,length}$	[m]	maximale Länge von Fahrzeugobjekten
$\tau_{max,length,lkw}$	[m]	maximale Länge von LKW-Objekten
$\tau_{\Delta Form}$	[m]	maximaler Abstand zwischen Messpunkt und Form zur Auslösung eines Formwechsels
$\tau_{max,r,C}$	[m]	maximaler Radius zur Zuordnung bei einer C-Form
$\tau_{min,v,b}$	[m/s]	Mindestgeschwindigkeit für Fahrradobjekte
$\tau_{min,v,radar}$	[m/s]	Mindestgeschwindigkeit für Radarobjekte
$\hat{\gamma}$	[%]	Klassenübereinstimmung
γ	[%]	Klassenzugehörigkeit
φ	[rad]	Wankwinkel
ϑ	[rad]	Nickwinkel
ψ	[rad]	Gierwinkel
ω	[rad/s]	Gierrate

Formelzeichen und Abkürzungen

a	[m/s ²]	Beschleunigung
a_r	[m]	Achsenabschnitt der Regressionsgeraden
b_r	[-]	Steigung der Regressionsgeraden
c	[m/s]	Lichtgeschwindigkeit
D	[m]	mittlerer Abstand der Punkte von der Regressionsgeraden
Δd_{max}	[m]	maximal zulässige Differenz zwischen zwei konsekutiven Werten bei der Segmentierung
f_C	[Hz]	Frequenz des Radarsenders
f_D	[Hz]	Dopplerfrequenz
f_n	[-]	Anzahl an Fehlerkennungen
l_i	[s]	Lebensdauer des Objekts i
ΔT	[s]	Lidar-Zeitschritt
t_{Latenz}	[s]	Latenzzeit zwischen Lidar- und Radarmessungen
$t_{kl,i}$	[s]	Zeitdauer, die das Objekt i korrekt klassifiziert ist
t_p	[-]	Anzahl richtig erkannter Objekte
v	[m/s]	Geschwindigkeit
v_{ego}	[m/s]	Geschwindigkeit des Versuchsfahrzeugs
v_{rel}	[m/s]	Relativgeschwindigkeit zwischen Versuchsfahrzeug und Radarobjekt
v_x	[m/s]	Geschwindigkeit in x-Richtung
v_y	[m/s]	Geschwindigkeit in y-Richtung
x	[m]	x-Position
y	[m]	y-Position
\hat{x}	[m]	prädizierter x-Radarmesswert
\hat{y}	[m]	prädizierter y-Radarmesswert
\bar{x}	[m]	Mittelwert der x-Werte
\bar{y}	[m]	Mittelwert der y-Werte
$\mathbf{x}(t)$		Zustandsvektor
$\mathbf{u}(t)$		Stelleingriff
$\mathbf{w}(t)$		Prozessrauschen
$\mathbf{Q}(t)$		Kovarianzmatrix des Prozessrauschens
$\mathbf{F}(t)$		kontinuierliche Systemmatrix
$\mathbf{G}(t)$		kontinuierliche Eingangsmatrix
$\mathbf{z}(t)$		Messgrößen
$\mathbf{v}(t)$		Messrauschen
$\mathbf{R}(t)$		Kovarianzmatrix des Messrauschens
$\mathbf{H}(t)$		Messmatrix
Φ_k		diskrete Systemmatrix
Γ_k		diskrete Eingangsmatrix
\mathbf{M}_k		Gewichtungsmatrix des quadratischen Fehlers
\mathbf{P}_k		Fehlerkovarianzmatrix
$\bar{\mathbf{K}}_k$		Kalmanfaktor
$\mathbf{T}_{t,t-1}$		Drehmatrix zur Kompensation der Eigenbewegung des Fahrzeugs

Abkürzungen

ABS	Antiblockiersystem
ACC	Adaptive Cruise Control, Abstandsregeltempomat
ADAS	Advanced Driver Assistance Systemes, fortschrittliche Fahrerassistenzsysteme
ADTF	Automotive Data and Time Triggered Framework (Entwicklungstool)
CAN	Controller Area Network
CEP	Kreisfehlerwahrscheinlichkeit, engl. Circular Error Probable
COG	Schwerpunkt (engl. Center of Gravity)
ESP	Elektronisches Stabilitätsprogramm
FAS	Fahrerassistenzsystem
FER	Fehlererkennungsrate
FKR	Falschklassifizierungsrate
FSRA	Full Speed Range Adaptive Cruise Control
GPS	Global Positioning System
IMM	Interacting Multiple Model (-Filter)
IMU	inertiale Messeinheit, engl. Internal Measurement Unit
ISO	International Organization for Standardization
LIDAR	Light Detection and Ranging
LKW	Lastkraftwagen
LTE	Long Term Evolution (Mobilfunkstandard)
PKW	Personenkraftwagen
RADAR	Radio Detection and Ranging
RMS	Quadratischer Mittelwert, engl. Root Mean Square
RMSE	Quadratischer mittlerer Fehler, engl. Root Mean Square Error
RPR	Richtig-Positiv-Rate
SG	Steuergerät
USB	Universal Serial Bus
ZA	Zeitliche Abdeckung

Inhaltsverzeichnis

Formelzeichen und Abkürzungen	I
1 Einleitung	1
1.1 Motivation und Zielsetzung	1
1.2 Stand der Technik	2
1.2.1 Teleoperiertes Fahren	2
1.2.2 Umfelderkennung	2
1.3 Aufbau der Arbeit	4
2 Versuchsträger	5
2.1 Systemarchitektur	5
2.2 Koordinatensystem	6
2.3 Sensoren	8
2.3.1 LIDAR	8
2.3.2 RADAR	10
2.3.3 Kameras	11
2.3.4 Inertiale Messeinheit	11
3 Grundlagen der Objektverfolgung	13
3.1 Kalman-Filter	13
3.1.1 Lineares Schätzproblem	13
3.1.2 Klassischer Kalman-Filter	15
3.1.3 Nichtlineares Schätzproblem	17
3.1.4 Extended Kalman-Filter	18
3.1.5 Unscented Kalman-Filter	19
3.2 Assoziation	22
4 Objekterkennung und Tracking LIDAR	23
4.1 Übersicht über Programmablauf	24
4.2 Segmentierung	25
4.3 Objekterzeugung	28
4.3.1 Merkmalsextraktion	28
4.3.2 Hypothesenerstellung	33
4.3.3 Koordinatentransformation in das aktuelle Sensorkoordinatensystem	34
4.3.4 Assoziation	35
4.3.5 Erzeugung neuer Objekte	36
4.4 Objektverfolgung	37
4.4.1 Rohdatenauswertung	37
4.4.2 Aktualisierung von Objekten	42
4.4.3 Klassifizierung	43
4.4.4 Filterung	46
4.4.5 Löschen von Objekten	49
5 Objekterkennung Radarsensor	51

6	Sensordatenfusion	54
6.1	Grundlagen der Datenfusion	54
6.1.1	Fusionsarten	54
6.1.2	Fusionsebenen	55
6.2	Sensorsynchronisation und Transformation	56
6.2.1	Räumliche Kalibrierung	56
6.2.2	Temporale Kalibrierung	57
6.3	Umsetzung der Fusion	64
6.3.1	Sensorbereiche	66
6.3.2	Fusion der Objekte	67
6.3.3	Integration in die Lidar-Objekterkennung	68
6.3.4	Fazit	69
7	Prädiktion	70
7.1	Grundlagen der Prädiktion	70
7.2	Systembetrachtung	71
7.3	Modellierung	71
8	Systembewertung	73
8.1	Grundlagen der Systembewertung	73
8.1.1	Erzeugung von Referenzdaten	75
8.1.2	Definition geeigneter Gütemaße	75
8.2	Bewertung der entwickelten Umfelderkennung	77
8.2.1	Optische Beurteilung	77
8.2.2	Verwendete Referenzdaten	78
8.2.3	Testszenarien	80
8.2.4	Systemkonfigurationen	80
8.2.5	Referenzbereich	82
8.2.6	Ergebnisse	82
8.2.7	Bewertung durch externe Referenzdaten	86
9	Zusammenfassung und Ausblick	88
9.1	Zusammenfassung	88
9.2	Ausblick	88
	Abbildungsverzeichnis	i
	Tabellenverzeichnis	ii
	Literaturverzeichnis	ii

1 Einleitung

1.1 Motivation und Zielsetzung

Diese Arbeit wurde im Rahmen des Projekts „Teleoperiertes Fahren“ des Lehrstuhls für Fahrzeugtechnik der TU München erstellt. Das Ziel dieses Projekts ist es, ein Versuchsfahrzeug durch einen Controllerarbeitsplatz fernsteuern zu können. Bei autonomen Fahrzeugen, an deren Entwicklung seit Jahren intensiv geforscht wird, treten vor allem auch juristische Probleme auf, die wohl auf absehbare Zeit nicht gelöst werden können. Deshalb zielt das Projekt darauf ab, zunächst einmal ferngesteuertes Fahren zu ermöglichen. Die größte Herausforderung liegt dabei in der Übertragung zwischen Fahrer und Fahrzeug. Das Fahrzeug sendet die von zahlreichen Sensoren aufgezeichneten Informationen zum Controller und empfängt die Steuerungsbefehle. Problematisch ist jedoch die Zeitverzögerung, die durch die Übertragung verursacht wird. Neben der sicherheitsrelevanten Problematik der erhöhten Reaktionszeit wird das Steuern dadurch erschwert, dass die Zeitverzögerung je nach Übertragungsstandard und Netzauslastung schwankt. Durch dieses Stocken im Video ist es für einen Menschen fast unmöglich das Fahrzeug kontrolliert zu steuern. Deshalb wird in diesem Projekt der Lösungsansatz verfolgt, eine möglichst konstante Zeitverzögerung von $500ms$ zu halten ([Fre+13]). Das Ziel dieser Arbeit ist nun, die Position der anderen Verkehrsteilnehmer zu präzisieren, um die Verzögerung zu kompensieren und somit dem Fahrer die Einschätzung der Situation zu erleichtern. Die Grundvoraussetzung für eine sinnvolle Vorhersage ist eine möglichst genaue Erfassung des aktuellen Zustands der anderen Verkehrsteilnehmer.

Die dafür nötige Umfelderkennung erlangt auch in einem anderen Bereich immer größere Bedeutung. Der Trend bei der Entwicklung von Fahrerassistenzsystemen (FAS) geht ebenfalls immer mehr zur Berücksichtigung der Umgebung. Während die ersten Systeme, wie beispielsweise der Tempomat oder das Antiblockiersystem (ABS), nur Informationen des eigenen Fahrzeugs verwenden, wurden in den letzten Jahren Systeme entwickelt, die auch das Umfeld des Fahrzeugs berücksichtigen. Um diese Systemklassen unterscheiden zu können, werden erstere als „konventionelle FAS“ und diejenigen mit Umfelderkennung als „FAS mit maschineller Wahrnehmung“ bezeichnet ([Win12]). Diese werden häufig auch als Advanced Driver Assistance Systems (ADAS, deutsch: fortschrittliche Fahrerassistenzsysteme) bezeichnet. Eines der ersten, serienreifen Assistenzsysteme dieser Art ist die Adaptive Fahrgeschwindigkeitsregelung (engl. Adaptive Cruise Control (ACC)), die mit Hilfe eines Radarsensors den Abstand zum voraus fahrenden Fahrzeug misst und dadurch sowohl Geschwindigkeit als auch Sicherheitsabstand einhalten kann. Ein anderes frühes Beispiel sind die Einparkhilfen, die mit Hilfe von Ultraschallsensoren die unmittelbare Umgebung des Fahrzeugs erfassen.

In den letzten Jahren wurden mehrere, Umfelderkennende FAS zur Serienreife gebracht, wie beispielsweise Verkehrszeichenerkennung, Spurassistenten, Kollisionswarnsysteme und automatische Bremsassistenten. Neben den hier erwähnten Systemen gibt es noch eine Vielzahl weiterer FAS die durch die Detektion der Umwelt zum Fahrkomfort beitragen oder gefährliche Situationen vermeiden. Je nach Anwendungsfall werden dafür unterschiedliche Sensoren verwendet und bestimmte Anteile der Umgebung erfasst und interpretiert.

Bei dieser Arbeit geht es nun darum, die am Versuchsträger vorhandenen Sensoren für eine Umfelderkennung zu nutzen. Der Versuchsträger stammt noch von dem Vorgängerprojekt im Rahmen des Sonderforschungsbereichs/TR 28 „Kognitive Automobile“ und ist neben dem serienmäßigen Radar des ACC-Systems noch mit weiteren Sensoren ausgestattet worden. An der Fahrzeugfront, dem Heck und am Dach sind drei LIDAR-Sensoren (light detection and ranging) angebracht, die ge-

naue Entfernungsdaten der Umgebung liefern. Zur präzisen Erfassung der Bewegung des Versuchsfahrzeugs dient eine inertielle Messeinheit (Internal Measurement Unit, IMU) bestehend aus Gyro-, Beschleunigungs- und GPS-Sensoren. Schließlich sind im Bereich des Rückspiegels noch fünf Videokameras angebracht, die sowohl zur Bildübertragung an den Controller als auch zur Erkennung und Verifizierung von Objekten verwendet werden können. Mit Hilfe dieser Sensoren sollen nun die Verkehrsteilnehmer in der Umgebung erkannt und deren Bewegung präzisiert werden. Diese Informationen sollen dann dem Controller in einem Predictive Display angezeigt werden und ihn somit bei der zeitverzögerten Steuerung des teleoperierten Fahrzeugs unterstützen.

1.2 Stand der Technik

1.2.1 Teleoperiertes Fahren

Obwohl es unzählige Forschungsprojekte zu autonomen Fahrzeugen gibt, ist das Teleoperierte Fahren bisher kein weit verbreiteter Forschungsansatz. Dabei ist dieser Ansatz aufgrund der Integration des Menschen und dessen Fähigkeiten auch in komplexen Situationen vielversprechend. Ein guter Überblick über die Teleoperation im Allgemeinen sowie die Herausforderungen im Hinblick auf das teleoperierte Fahren ist in [GSL] zu finden. Darin werden drei wichtige Felder für teleoperiertes Fahren identifiziert: die Automatisierung des Fahrzeugs, die Übertragung zwischen Fahrzeug und Controller, sowie die Schnittstelle zwischen Mensch und Maschine.

Die bei der Übertragung auftretende Verzögerung stellt die größte Herausforderung dar und muss durch geeignete Konzepte zur Schnittstelle zum Controller oder durch Automatisierung des Fahrzeugs kompensiert werden. Dafür wird hier ein „Prädiktives Display“ (engl. predictive display) verwendet. Dieses visualisiert dem Controller den voraussichtlichen Zustand zu dem Zeitpunkt, in dem seine Steuersignale beim Fahrzeug ankommen. Diese Idee wurde bereits 1963 von Arnold und Braisted [AB63] bei der Fernsteuerung von Mondfahrzeugen untersucht. Seither wurde dieses Konzept in unterschiedlichen Anwendungen erprobt und weiterentwickelt. Für einen ausführlicheren Überblick sei auf die Einleitung von [Fre+13] verwiesen. Die darin entwickelten Prädiktionsmodelle sind speziell für Straßenfahrzeuge ausgelegt und dienen der Prädiktion des teleoperierten Fahrzeugs. Um das Prädiktive Display auf die gesamte Umgebung ausweiten zu können, soll hier eine entsprechende Umfelderkennung und -prädiktion entwickelt werden.

1.2.2 Umfelderkennung

Die Umfelderkennung im automobilen Bereich ist ein sehr aktives Forschungsfeld. Vor allem im Bereich der Fahrerassistenzsysteme wurden in den letzten Jahren zahlreiche Systeme eingeführt, die die Umgebung des Fahrzeugs detektieren. Die ersten umfelderfassenden ADAS waren für sehr beschränkte Einsatzgebiete konzipiert. Beispielsweise die Einparkhilfe, die per Ultraschallsensoren lediglich den Abstand bei langsamer Fahrt und im engsten Nahbereich erkennt, oder die ersten ACC-Systeme, die nur in Autobahnsituationen und beschränkten Geschwindigkeitsbereichen anwendbar sind. Um den steigenden Anforderungen an die FAS gerecht zu werden, muss auch die Qualität der Umfelderkennung steigen. Für die Erweiterung des konventionellen ACC auf den niedrigen Geschwindigkeitsbereich, damit auch im stockenden Verkehr eine automatische Geschwindigkeitsregelung ermöglicht wird, reicht der geringe Öffnungswinkel des Longrange-Radars nicht mehr aus. Deshalb werden bei dem sogenannten full speed range ACC (FSRA) von mehreren Herstellern zusätzlich zwei Nahbereichs-Radarsensoren fusioniert ([Win12]). Durch den großen Öffnungswinkel und die gute Winkelauflösung können damit alle relevanten Objekte detektiert werden.

Da beim FSRA nur Fahrzeuge im Fließverkehr berücksichtigt werden müssen, ist die Komplexität dieser Situation noch relativ überschaubar. Der Trend bei den ADAS geht aber zur Interpretation von

immer komplexeren Situationen, wie sie beispielsweise im Stadtverkehr und im Kreuzungsbereich auftreten. Da sich ein Großteil der Unfälle an Kreuzungen und Abbiegungen ereignen ([Sta12]), wird seit einigen Jahren an Systemen geforscht, die den Fahrer in diesen Situationen unterstützen oder warnen sollen. Dabei ist die hohe Informationsdichte und die möglicherweise daraus folgende Überforderung des Fahrers ein Grund für das hohe Unfallrisiko [Win12]. Durch die große Komplexität der Situation sind die Anforderungen an die Umfelderkennung auch dementsprechend hoch.

Als Grundlage für diese Systeme gibt es zahlreiche Forschungsarbeiten im Bereich der Umfelderkennung. Dabei werden unterschiedliche Ansätze und Sensorkonfigurationen zur Erkennung und Verfolgung von Verkehrsteilnehmern untersucht. Vor allem die Fusion von Sensoren spielt eine immer wichtigere Rolle. Dabei werden vor allem Kamerasysteme (auch Infrarot), Radarsensoren (Nah- und Fernbereich) und Lidarscanner eingesetzt und kombiniert. So fusionieren Tehrani et al [TN+] beispielsweise einen Lidar-Scanner mit einer kamerabasierten Objekterkennung um durch die Einschränkung des Suchbereichs den Rechenaufwand zu reduzieren. Ein ähnlicher Ansatz findet sich in [Wu+], bei dem zunächst durch einen Lidar-Scanner mögliche Fußgänger gesucht und diese durch eine Kamera anschließend validiert werden. Kaempchen et al. [Kae+ne] gehen noch einen Schritt weiter. Sie verwenden ebenfalls einen Lidar-Scanner um den Suchbereich im Kamerabild einzuschränken, jedoch werden anschließend die Ergebnisse beider Sensoren fusioniert. Außerdem wird darin ein Ansatz vorgestellt, um undefinierte Formen in den Lidardaten verfolgen zu können, wie sie beispielsweise bei LKWs auftreten. In [Lab+05] wird eine Stereokamera mit einem Lidar-Scanner fusioniert, um eine möglichst umfangreiche und robuste Hinderniserkennung zu implementieren. Ein allgemeiner Ansatz zur Sensorfusion für Fahrerassistenzsysteme ist in [KKne] mit der Kombination von Radarsensor und Kamera vorgestellt.

Die Fusion von Lidar-Scanner und Radarsensor wird beispielsweise in [LW] untersucht. Darin wird auch ein Multilevel Fusionsalgorithmus vorgestellt, der die Sensordaten je nach Objekt auf unterschiedlichen Ebenen fusioniert. In [BAB] wird die selbe Sensorkombination verwendet. Burlet und Aycard kombinieren dabei auch die Kartierung der Umgebung mit einer Objekterkennung. Für die Objektverfolgung wird dann ein Interacting Multiple Models Filter (IMM) angewandt, der parallel verschiedene Modelle verwendet und dadurch unterschiedliche Manöver gut abbilden kann. Darüber hinaus gibt es auch Ansätze, die noch mehr Sensoren fusionieren. Amditis et al [Amd+13] kombinieren beispielsweise Radarsensoren, Lidarscanner und Kameras und realisieren damit ein Assistenzsystem, das sowohl die laterale als auch longitudinale Steuerung in Kolonnensituationen übernehmen kann. In der Dissertation von Lindl [Lin08] werden neben den bisher verwendeten Radar-, Lidar- und Videosensoren auch noch Infrarotkameras behandelt. Dabei wird auf alle wichtigen Bereiche der Umfelderkennung mit mehreren Sensoren ausführlich eingegangen.

Am Lehrstuhl für Fahrzeugtechnik gibt es bereits Ansätze, die die Umfelderkennung mit den vorhandenen Sensoren untersuchen. Im Vorgängerprojekt wurde der Fokus dabei auf die Erfassung von Autobahnsituationen gelegt. Aus diesem Grund wird dabei der Radarsensor als wichtigste Informationsquelle verwendet und mit den Daten der Lidarsensoren fusioniert ([GSL]). Durch die beschränkte Komplexität von Autobahnsituationen wurde die Umfelderkennung sehr einfach konzipiert. Darauf aufbauend entwickelte Wolter [Wol12] eine Objekterkennung auf Lidar-Basis, die auch für komplexere Verkehrssituationen geeignet ist. Dabei wird der in der Literatur weit verbreitete Ansatz verfolgt, aus den Lidardaten Merkmale zu extrahieren und diese für die Objektverfolgung zu verwenden. Die dabei entwickelte Umfelderkennung funktioniert grundsätzlich, jedoch hat sich gezeigt, dass die erzielten Ergebnisse für die hier gesuchte Anwendung nicht ausreichend sind. Das liegt größtenteils daran, dass die von dem Lidarsensor gelieferten Daten häufig keine sinnvolle Merkmalsextraktion erlauben. Deshalb wird in dieser Arbeit ein eigenständiger und innovativer Lösungsansatz entwickelt, um eine bessere Umfelderkennung von komplexen Situationen zu ermöglichen.

1.3 Aufbau der Arbeit

In den vorherigen Abschnitten wurde auf die Motivation für diese Arbeit und den Stand der Technik im Bereich des Teleoperierten Fahrens und der Umfelderkennung im Automotive Bereich eingegangen. Der Aufbau der restlichen Arbeit gliedert sich folgendermaßen. Im zweiten Kapitel wird ausführlich auf den verwendeten Versuchsträger eingegangen. Dabei wird auch ein besonderes Augenmerk auf die einzelnen Sensoren und deren Eigenschaften gelegt. Die theoretischen Grundlagen für die Objektverfolgung mit Hilfe des Kalman-Filters werden in Kapitel 3 behandelt. Im darauffolgenden Kapitel wird die Objekterkennung und -verfolgung mit Hilfe des Lidarscanners beschrieben. Dabei wird neben dem verwendeten Trackingalgorithmus auch die Klassifizierung vorgestellt. Im Kapitel 5 wird die radarbasierte Objekterkennung kurz beschrieben, die größtenteils bereits in der internen Sensorverarbeitung abläuft. Die Fusion von Lidar- und Radarsensoren ist in Kapitel 6 behandelt. Darin werden zunächst auch die Grundlagen zur Sensorfusion erläutert. Die Prädiktion der Positionen der umgebenden Verkehrsteilnehmer ist Inhalt des 7. Kapitels. Im achten Kapitel wird der implementierte Algorithmus mit Hilfe von aufgezeichneten Messfahrten evaluiert. Abschließend findet sich im Kapitel 9 noch eine kurze Zusammenfassung sowie der Ausblick auf mögliche Erweiterungen und zukünftige Aufgaben.

2 Versuchsträger

Wie in der Einleitung bereits erwähnt, wird der Versuchsträger des Sonderforschungsbereichs/TR 28 „Kognitive Automobile“ in dem Projekt „Teleoperiertes Fahren“ verwendet. Es handelt sich dabei um einen AUDI Q7, der mit zusätzlichen Sensoren und Computerhardware aufgerüstet wurde. In Abbildung 2.1 ist der Versuchsträger mit den entsprechenden Aufbauten dargestellt. Im Gegensatz zu dem serienmäßigen ACC-Radar-Sensor, der hinter der Frontschürze verborgen ist, sind die zusätzlich installierten Sensoren gut erkennbar. An dem Fahrzeug sind drei Lidar-Sensoren installiert. Einer an der Fahrzeugfront, der parallel zur Fahrzeugachse ausgerichtet ist und somit bei verschwindendem Nickwinkel des Fahrzeugs parallel zur Fahrbahnoberfläche scannt. Der Zweite ist an dem auffälligen Dachaufbau installiert, wobei dieser jedoch nach unten gerichtet ist und im vorhergehenden Projekt der Straßenerkennung diente. Der dritte Scanner ist am Fahrzeugheck angebracht und scannt wie der Front-LIDAR parallel zur Fahrbahnoberfläche. An dem Dachaufbau sind außerdem die zwei zur IMU gehörenden GPS-Antennen angebracht. Schließlich ist das Fahrzeug noch mit acht Kameras ausgestattet. Fünf dieser Kameras sind im Bereich des Rückspiegels angebracht. Davon sind drei Kameras nach vorne gerichtet, wobei zwei davon gerade in einem anderen Projekt zu einer Stereokamera umkonfiguriert werden und die dritte mit Hilfe eines Weitwinkelobjektivs den Großteil des vor dem Fahrzeug befindlichen Bereichs aufnimmt und zur Visualisierung dient. Zwei weitere Kameras im Bereich des Rückspiegels sind auf die Fenster der Fahrer- bzw. Beifahrertür gerichtet und nehmen somit den Bereich neben dem Versuchsfahrzeug auf. Darüber hinaus gibt es noch zwei Kameras, die jeweils an den Außenspiegeln befestigt sind, um diesen Sichtbereich zu erfassen. Die letzte Kamera befindet sich an der Heckscheibe und zeichnet den Bereich hinter dem Fahrzeug auf.

Neben den von außen gut sichtbaren Sensoren zur Erfassung der Umwelt ist das Versuchsfahrzeug noch mit einigen Einbauten versehen, um Teleoperiertes Fahren zu ermöglichen. Im Kofferraum ist neben der bereits erwähnten IMU auch die benötigte Computerhardware untergebracht. Für die Sicherheitsrelevanten und damit zwingend in Echtzeit ausgeführten Programme steht eine Dspace Autobox zur Verfügung. Für weitere Berechnungen sind noch zwei weitere Rechner untergebracht, auf denen mit Hilfe der Simulationsumgebung ADTF weniger echtzeitkritische Anwendungen ausgeführt werden. Um das Fahrzeug teleoperiert steuern zu können, sind die Lenkung, Schaltung und Bremsen mit der Autobox verbunden. Die Betätigung der Bremse übernimmt ein Pneumatikzylinder. Die anderen Steuerfunktionen sind über die serienmäßigen Aktuatoren verwirklicht. Lediglich die Schaltung wurde angepasst, damit teleoperiert zwischen den unterschiedlichen Modi des Automatiktriebes gewechselt werden kann.

2.1 Systemarchitektur

In Abbildung 2.2 sind die für diese Arbeit entscheidenden Systemkomponenten und deren Verschaltung dargestellt. Neben der Autobox sind noch die beiden anderen Rechner Car-PC und Opteron installiert. Die Autobox ist über unterschiedliche Schnittstellen mit allen wichtigen Steuergeräten und Sensoren des Fahrzeugs verbunden. Dadurch können sowohl die zum fahren notwendigen Aktuatoren angesteuert, als auch wichtige Informationen, wie beispielsweise die vom ESP-SG ermittelte Fahrzeuggierrate, ausgelesen werden. Die drei Rechner sind über den AutoboxCAN miteinander verbunden. Der CAN des serienmäßigen Radar-Geräts ist sowohl an der Autobox als auch am Opteron angeschlossen. Für die Kommunikation mit der IMU wurde eine CAN-Verbindung zu allen drei Rechnersystemen hergestellt. Die drei LIDAR-Scanner sind dagegen mit einer seriellen Schnittstelle

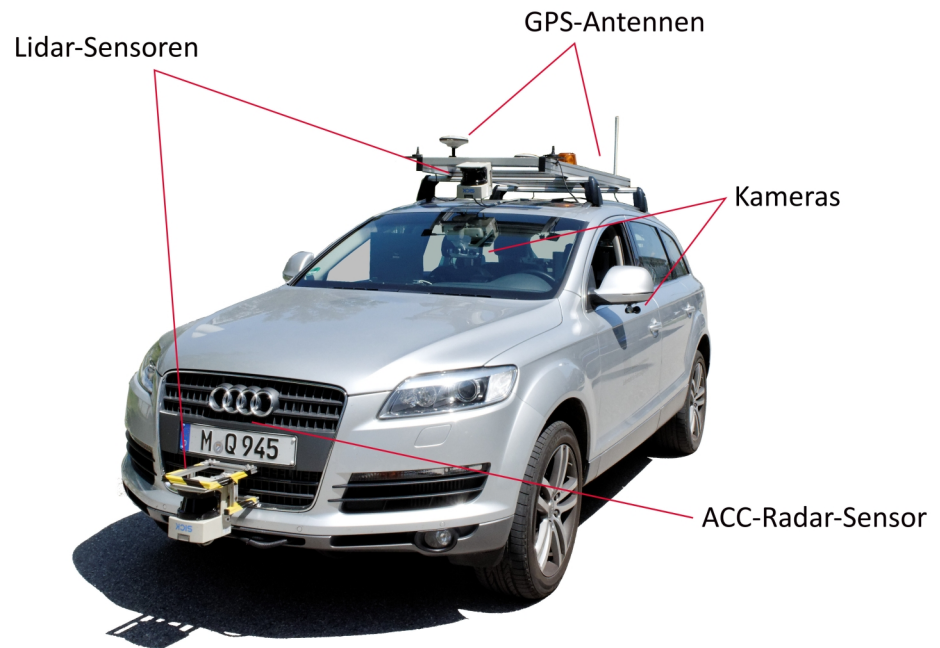


Abbildung 2.1: Versuchsfahrzeug mit Sensoren

an den Opteron angeschlossen. Um eine möglichst hohe Auflösung und farbige Bilder zu ermöglichen, sind die acht Kameras sowohl auf den Opteron und den Car-PC, als auch auf verschiedene Schnittstellen (USB, Ethernet) verteilt. Die Kommunikation mit dem Controller wird über den LTE-Stick ermöglicht.

2.2 Koordinatensystem

Für die eindeutige Angabe der Positionen anderer Verkehrsteilnehmer ist ein klar definiertes Koordinatensystem notwendig. Vor allem bei Systemen mit mehreren Sensoren ist eine Betrachtung der verschiedenen Koordinatensysteme unumgänglich. Die Messwerte der unterschiedlichen Sensoren werden immer im jeweiligen Sensorkoordinatensystem ausgegeben. Da bei dieser Arbeit sowohl ein LIDAR-Scanner als auch ein Radar-Sensor zum Einsatz kommen, gibt es demnach zwei Sensorkoordinatensysteme. Darüber hinaus wird ein sensorunabhängiges Fahrzeugkoordinatensystem nach ISO 8855 definiert, dessen Ursprung im auf die Fahrbahn projizierten Schwerpunkt (engl. Center of Gravity (COG)) des Fahrzeugs liegt. Die drei Koordinatensysteme sind in Abbildung 2.3 dargestellt.

Wie aus der Abbildung ersichtlich wird, sind alle drei Systeme gleich orientiert. Die x-Achse zeigt dabei jeweils in Fahrtrichtung, die y-Achse nach links und die z-Achse nach oben. Bei Sensorkoordinatensystemen ist es üblich, dass die x-Achse in Sensorrichtung zeigt. Da beide Sensoren parallel zur Fahrzeugachse ausgerichtet sind, ergibt sich die identische Orientierung. Der LIDAR-Scanner ist mittig an der Fahrzeugfront angebracht. Der Radar befindet sich hinter der Blende neben dem Nummernschild. Aufgrund der gleichen Orientierung sind die für die Datenfusion notwendigen Trans-

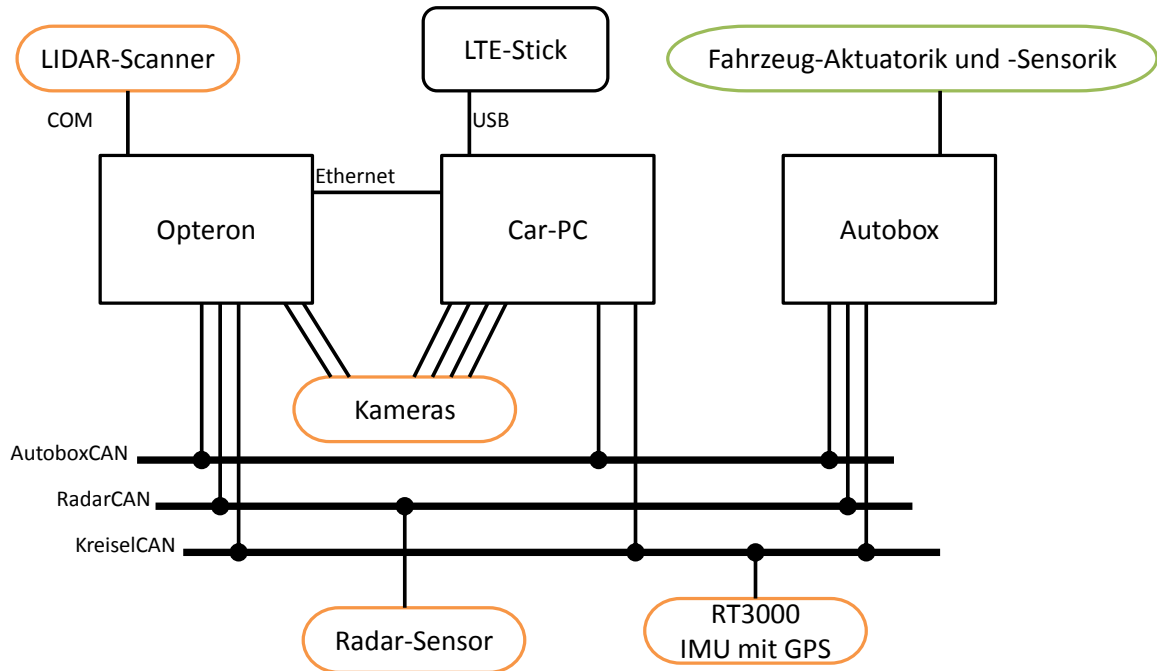


Abbildung 2.2: Systemarchitektur des Versuchsträgers

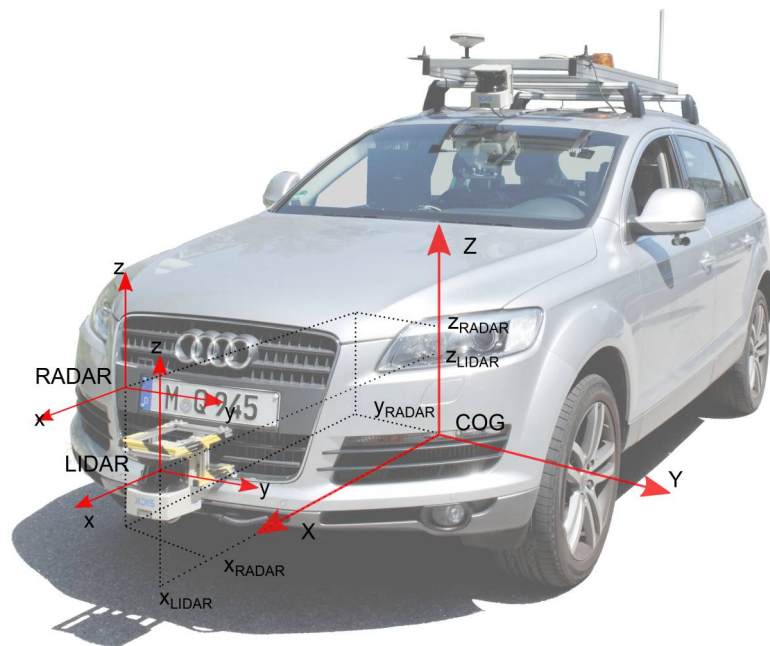


Abbildung 2.3: Verwendete Koordinatensysteme

Tabelle 2.1: Translationsvektoren zwischen den Koordinatensystemen

Transformation	x	y	z
LIDAR → COG	2,50m	0m	0,53m
RADAR → COG	2,20m	0m	0,70m
RADAR → LIDAR	-0,30m	0m	0,17m

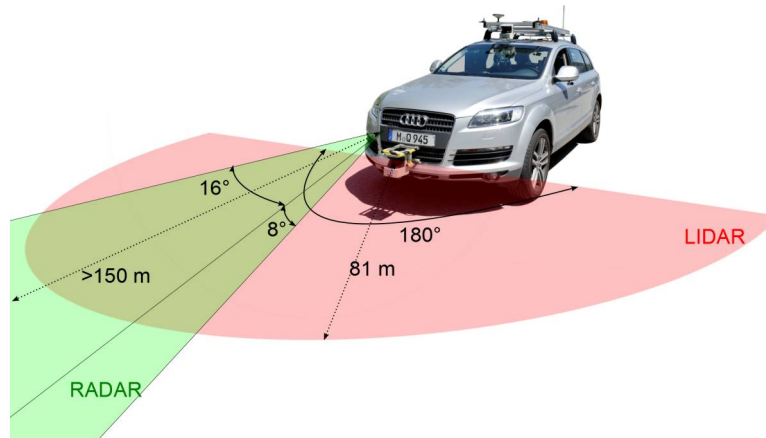


Abbildung 2.4: Abdeckungsbereiche der Sensoren

formationen zwischen den Koordinatensystemen nur Verschiebungen:

$${}_{COG}\mathbf{x} = {}_S\mathbf{x} + \mathbf{x}_{COG,S}, \text{ mit } \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \mathbf{x}_{COG,S} = \begin{bmatrix} x_S \\ y_S \\ z_S \end{bmatrix} \quad (2.1)$$

Dabei gibt der vorangestellte Index jeweils an in welchem Koordinatensystem der Vektor angegeben ist. Der Vektor $\mathbf{x}_{COG,S}$ beschreibt die nötige Verschiebung, um vom Sensorkoordinatensystem S in das COG System zu transferieren. Die entsprechenden Werte wurden gemessen und sind in Tabelle 2.1 zusammengefasst. Der Versatz des Radarsensors zur Fahrzeugachse ist bei der Ausgabe der Radardaten bereits berücksichtigt, wodurch das Koordinatensystem keine y -Verschiebung aufweist.

2.3 Sensoren

In diesem Abschnitt werden die verwendeten Sensoren genauer beschrieben. In Abbildung 2.4 sind die Messbereiche der hauptsächlich verwendeten Sensoren dargestellt. Die Bereiche sind nicht maßstäblich sondern dienen nur der Veranschaulichung. Die vom LIDAR-Scanner abgetastete Ebene ist rot dargestellt. Der vom Radar erfasste Bereich ist grün eingefärbt.

2.3.1 LIDAR

LIDAR steht wie bereits erwähnt für „Light Detection And Ranging“. Im deutschsprachigen Raum ist auch die Bezeichnung Laserradar geläufig. Es handelt sich dabei um Sensoren, die mit Hilfe eines Lichtimpulses die Entfernung messen. Da der Impuls vom Sensor ausgesendet wird, gehört der LIDAR zu den aktiven Sensoren. Die Messung basiert auf dem Pulslaufzeit-Prinzip (im Englischen als time-of-flight bezeichnet). Dabei wird ein kurzer Lichtimpuls ausgesendet und die Zeit gemessen, bis der reflektierte Strahl wieder bei dem Sensor ankommt [SIC03]. Aus der Laufzeit und der



Abbildung 2.5: SICK Lidar-Sensor [SIC03]

Lichtgeschwindigkeit kann dann die Entfernung des reflektierenden Objekts d durch

$$d = \frac{1}{2} c \tau \quad (2.2)$$

berechnet werden. Dabei ist c die Lichtgeschwindigkeit und τ die Laufzeit, bis der reflektierte Strahl wieder beim Sensor eintrifft und dementsprechend die Distanz zwei mal zurück gelegt hat. Für den Lichtimpuls wird bei den meisten Sensoren ein Laser im nahen Infrarotbereich verwendet, wobei die Intensität so gering ist, dass er der Laserklasse 1 (augensicher) angehört.

Auch wenn das Messprinzip dieser Sensoren immer das Selbe ist, gibt es doch deutliche Unterschiede in deren Bauformen. Zum einen kann zwischen Mehrkanalsensoren und Lidar-Scannern unterschieden werden. Erstere bestehen aus mehreren Sendern und Empfängern, die in feste Richtungen aussenden und messen. Die sogenannten Lidar-Scanner besitzen meist nur einen Laser und Sensor. Die Lichtstrahlen werden über einen beweglichen Spiegel abgelenkt. Damit hängt der vermessene Bereich lediglich von der Bewegung des Spiegels ab und kann somit auch variabel eingestellt werden. Mit diesem Prinzip ist auch ein 3D-Scan möglich. Jedoch wird im Automotive-Bereich meistens ein rotierender Spiegel verwendet und somit von einem Sensor nur eine Ebene gescannt.

Zum anderen wird zwischen einschichtigen und mehrschichtigen Sensoren unterschieden. Einschichtige Sensoren tasten nur eine Ebene ab und können somit keine Informationen zur Ausdehnung senkrecht zur Ebene (in diesem Anwendungsfall die Höhe) liefern. Darüber hinaus sind sie sehr anfällig gegenüber Nickbewegungen des eigenen Fahrzeugs. Deshalb werden bei den meisten Forschungsarbeiten auf diesem Gebiet Sensoren verwendet, die gleichzeitig vier Ebenen abtasten (z.B. in [TN+] oder [Wal08]).

Ein großer Vorteil der LIDAR-Sensoren ist, dass das zu messende Objekt keine besonderen Anforderungen erfüllen muss. Es sind beispielsweise keine Reflexionsmarken oder Ähnliches nötig. Allerdings hängt die Reichweite der Sensoren stark von der Remission der Objekte sowie den Wetterverhältnissen ab. Des Weiteren ist für die Detektion eine direkte Sichtverbindung zwischen Sensor und Objekt nötig, wodurch Objekte leicht verdeckt werden können. Diese Voraussetzung limitiert auch die möglichen Einbaupositionen, was bei einer künftigen Serienreife Restriktionen für das Fahrzeugdesign mit sich bringt. Dennoch eignen sich LIDAR-Sensoren gut für die maschinelle Erfassung der näheren Umgebung, da sie sehr genaue Entfernungsmesswerte liefern. Darüber hinaus können große Winkelbereiche bis 180° mit hohen Taktraten abgetastet werden.

Der hier verbaute LIDAR-Sensor ist ein SICK LMS-291 ([SIC03]) wie in Abbildung 2.5 dargestellt. Es handelt sich dabei um einen einschichtigen Sensor mit rotierendem Spiegel. Im Vergleich zu den mehrschichtigen Modellen sind vor allem die höhere Abtastrate (bei gleicher Übertragungsgeschwindigkeit) und die günstigere Anschaffung von Vorteil. Die maximale Abtastrate beträgt 75 Hz, wobei ein Bereich von 0° - 180° mit einer Auflösung von 1° abgetastet werden kann. Der Sensor bietet auch

2 Versuchsträger

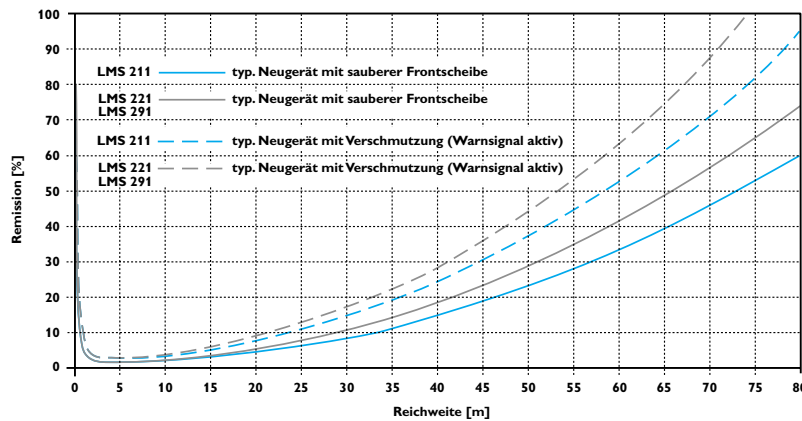


Abbildung 2.6: Reichweite des LIDARs in Abhängigkeit des Remissionswerts des Objekts [SIC03, S.14]

Betriebsmodi mit Winkelauflösungen von $0,25^\circ$ und $0,5^\circ$ an. Allerdings benötigt er dafür dann 2 respektive 4 Umdrehungen des Spiegels. Dadurch halbiert bzw. viertelt sich dann auch die Taktrate und die unterschiedlichen Aufnahmezeiten der einzelnen Messpunkte erschweren die Objekterkennung. Die maximale Reichweite des Sensors beträgt 81,91 m, allerdings ist diese stark vom der Remission des Objekts abhängig, wie aus Abbildung 2.6 ersichtlich. So können beispielsweise Objekte mit 10% Remission von dem verwendeten Sensor LMS 291 nur bis knapp 30 m detektiert werden. Ist die Remission eines Objekts zu gering, gibt der Sensor die maximale Reichweite als Messwert zurück. Neben der Abnahme der Detektionswahrscheinlichkeit verringert sich die Güte der LIDAR-Messung mit steigender Entfernung noch aufgrund des Abtastprinzips. Durch die Öffnungswinkel zwischen den Messrichtungen steigen mit zunehmender Sensorentfernung die Abstände zwischen den Strahlen. Aus diesen beiden Gründen liefert der LIDAR-Scanner vor allem im Nahbereich gute Ergebnisse, ist für Entfernungen über ca. 30m aber nur noch bedingt geeignet. Unabhängig von der geringer werdenden Detektionsrate sind die Entfernungsangaben von tatsächlich erkannten Objekten sehr genau. So liefert der hier verwendete Sensor Genauigkeiten von $10\text{mm} \pm 35\text{mm}$ ([SIC03]).

2.3.2 RADAR

Der sehr geläufige Begriff Radar ist eine Abkürzung von „Radio Detection and Ranging“, womit das Funktionsprinzip des Sensors beschrieben wird. Ähnlich wie beim LIDAR werden Impulse ausgesandt und deren Reflexionen empfangen. Allerdings sendet der Radar elektromagnetische Impulse deutlich größerer Wellenlänge aus. Im Automotive-Bereich wird für größere Reichweiten der Frequenzbereich 76-77 GHz und für den Nahbereich 24 GHz verwendet. Die ausgesendeten Wellen werden dabei durch elektrisch leitende Objekte reflektiert. Damit eignen sich Radar-Sensoren sehr gut zur Detektion von Verkehrsteilnehmern. Durch die Laufzeit des ausgesandten und reflektierten Signals kann dann der Abstand des Objekts wieder durch die Gleichung (2.2) bestimmt werden. Aufgrund der deutlich größeren Wellenlängen ist es bei RADAR-Geräten praktikabel auch die Frequenz des reflektierten Signals zu messen. Bewegen sich Sensor und Objekt relativ zueinander, ergibt sich aufgrund des Dopplereffekts eine Frequenzverschiebung. Dementsprechend kann durch die Frequenzmessung des reflektierten Signals direkt die Relativgeschwindigkeit des Objekts mit der Formel

$$v_{rel} = -\frac{1}{2}c \frac{f_D}{f_C} \quad (2.3)$$

bestimmt werden [200]. Dabei ist f_D die Dopplerfrequenz, f_C die gesendete Frequenz des Radars und c die Lichtgeschwindigkeit. Aus diesem Grund eignen sich Radarsensoren hervorragend für die Verwendung in ACC-Systemen, da sie die für die Anwendung wichtigsten Zustandsgrößen der an-

Tabelle 2.2: Kenndaten der Radar-Einheit [200, S.160]

Reichweite	2...150m
Messbare Relativgeschwindigkeit	-50... + 50 $\frac{m}{s}$
Winkelbereich	$\pm 8^\circ$
Trennfähigkeit	0,85m; 1 $\frac{m}{s}$
Messrate	10Hz
Frequenzbereich	76...77GHz
Mittlere Sendeleistung	ca. 5MW
Bandbreite	ca. 200MHz

deren Verkehrsteilnehmer direkt messen können. Neben dem Abstand des Objekts wird auch der Winkel bzw. die laterale Ablage benötigt. Dafür verwenden die meisten Radarsensoren sogenannte Radar-“Keulen“. Das heißt, dass mehrere unabhängige Sensoren in leicht versetzte Richtungen strahlen. Aus den Amplituden der reflektierten Wellen kann dann die laterale Ablage des Objekts bestimmt werden [200].

Ein weiterer Vorteil von RADAR-Systemen ist die geringe Wetterabhängigkeit. Während optische Systeme stark durch Niederschlag oder Nebel beeinflusst werden, bleibt die Leistungsfähigkeit des Radars nahezu unbeeinträchtigt. Lediglich eine Schnee- oder Eisschicht die sich direkt auf der Linse des Sensor festsetzt, reduziert die Reichweite und kann vor allem auch die Winkelbestimmung negativ beeinflussen. Deshalb werden die entsprechenden Bauteile häufig durch eine Heizung vor dem Festsetzen von Schnee und Eis bewahrt.

Das Versuchsfahrzeug ist mit dem bereits erwähnten ACC-System ausgestattet. Zur Erfassung des vorausfahrenden Verkehrs wird dabei ein 77 GHz-Radar verwendet, der hinter der Frontschürze verbaut ist. Der Radar-Sensor hat eine maximale Reichweite von ca. 150m und einen horizontalen Öffnungswinkel von $\pm 8^\circ$, der mit drei Radar-“Keulen“ abgedeckt wird. Alle relevanten Leistungsdaten sind in Tabelle 2.2 zusammengefasst. Dieser Radar besitzt eine CAN-Schnittstelle, über die er die erkannten Radarobjekte ausgeben kann. Allerdings wurde der Radar explizit für ein ACC-System konzipiert und die CAN-Schnittstelle nur für den Fall eines externen ACC-Steuergeräts integriert. Das bedeutet, dass die gesamte Vorverarbeitung der Daten bereits im Gerät stattfindet und nur die fertig getrackten Objekte ausgegeben werden.

2.3.3 Kameras

Der Versuchsträger ist mit insgesamt 8 Kameras ausgerüstet. Diese erfassen alle relevanten Bereiche der Umgebung die gewöhnlich von dem Fahrersitz einsehbar sind. So erfassen drei Kameras den Sichtbereich aus der Windschutzscheibe und den beiden Seitenfenstern, zwei Kameras decken jeweils die Rückspiegelbereiche ab und eine ist nach hinten gerichtet. Mit den letzten beiden Kameras wird momentan eine Stereokamera entwickelt. In dieser Arbeit werden nur die Aufnahmen der vorderen, frontalen Kamera zur Visualisierung verwendet.

2.3.4 Inertiale Messeinheit

Im Kofferraum des Versuchsfahrzeugs ist die Inertiale Messeinheit RT3000 von Oxford Technical Solutions installiert. Diese ermittelt aus den zwei auf dem Dach befindlichen GPS-Antennen und jeweils drei Beschleunigungs- und Drehratensensoren die Bewegung des Versuchsträgers. Die GPS- und Beschleunigungsdaten werden durch mathematische Algorithmen gefiltert und fusioniert, sodass die IMU alle relevanten Größen zum abbilden der Bewegung liefert. Durch die Kombination der Daten können Sprünge oder Abrisse des GPS-Signals kompensiert bzw. überbrückt werden. Außerdem

Tabelle 2.3: Genauigkeiten der inertialen Messeinheit [Oxf11]

Messgröße	Genauigkeit
Position	0,4m CEP (DGPS)
Geschwindigkeit	0,05 $\frac{km}{h}$ RMS
Beschleunigung	10 $\frac{mm}{s^2}$ 1σ
Roll-/Nickwinkel	0,03° 1σ
Orientierung	0,03° 1σ
Drehrate	0,01 $\frac{^\circ}{s}$ 1σ

zeichnet sich die Messeinheit durch eine hohe Abtastrate von 100Hz sowie eine sehr geringe Latenzzeit von 3,5ms aus ([Oxf11]).

Die IMU benötigt nach dem Einschalten eine Aufwärm- und Initialisierungsphase von ca. 15 Minuten um die maximale Genauigkeit der Messwerte zu erreichen. Dabei muss das Versuchsfahrzeug mindestens 5 Minuten lang freien Kontakt zu den GPS-Satelliten haben und Beschleunigungen in allen Richtungen und um alle Achsen erfahren. Während dieser Phase ermittelt der Kalman-Filter die Sensorfehler und dadurch werden Genauigkeiten wie in Tabelle 2.3 erreicht. Die erzielbare Positionsgenauigkeit hängt vom verfügbaren GPS-Signal ab. Die verwendete Plattform unterstützt mehrere Arten der GPS-Korrektur, jedoch sind die dafür benötigten Signale teilweise militärisch reserviert bzw. nicht frei verfügbar. Bei dem hier verwendete Differential-GPS werden mit Hilfe von festen Referenzstationen die atmosphärischen Störungen der Laufzeitmessung kompensiert. Der CEP-Wert (Kreisfehlerwahrscheinlichkeit, engl. Circular Error Probable) beschreibt in welchem Radius 50% der Messwerte liegen. Bei der Geschwindigkeit ist der quadratische mittlere Fehler angegeben. Bei der Beschleunigung, den Winkeln und den Drehraten steht aufgrund fehlender Referenzdaten nur die 1σ Standardabweichung zur Verfügung. Wie an den sehr geringen Abweichungen zu sehen ist, liefert die Inertialplattform in Echtzeit hochpräzise Zustandswerte des Versuchsfahrzeugs, wodurch die Eigenbewegungen bei der Auswertung sehr genau berücksichtigt werden kann.

3 Grundlagen der Objektverfolgung

Bei der Objektverfolgung geht es darum, den Zustand von Objekten durch Messungen über einen längeren Zeitraum zu ermitteln und zu verfolgen. Häufig wird dafür auch im Deutschen der englische Begriff „Tracking“ verwendet. In den meisten Fällen wird versucht, die Position und Bewegung der Objekte zu ermitteln. Als Mitte des letzten Jahrhunderts die Entwicklung von Trackingverfahren deutliche Fortschritte gemacht hat, war das treibende Problem die Verfolgung von Radarobjekten. Es wurde nach Möglichkeiten gesucht, trotz des Messrauschens die Position und die Bewegung der Objekte möglichst genau abbilden zu können. Prinzipiell sind Trackingverfahren jedoch auf eine Vielzahl von Problemstellungen anwendbar.

Der grundsätzliche Ablauf und die relevanten Komponenten einer solchen Verfolgung sind in Abbildung 3.1 dargestellt. Der Sensor erfasst ein Signal, das Rückschlüsse auf die Position des Objekts zulässt. Allerdings wird dieses Signal von einem Messrauschen überlagert bevor es im Signalprozessor ausgewertet werden kann. Der Trackingalgorithmus soll aus diesen Messungen eine möglichst genaue Schätzung des Objektzustands ermitteln. Der Durchbruch gelang 1960, als Rudolf Emil Kalman die wahrscheinlich größte Erkenntnis der statistischen Schätztheorie des 20. Jahrhunderts veröffentlichte ([GA08]). Der von ihm veröffentlichte und auf den Vorarbeiten von Schmidt und Kolmogorov basierende Kalmanfilter ist der wohl wichtigste Trackingalgorithmus. Dieser wird in den folgenden Kapiteln beschrieben.

Aufgrund von stetig steigender Rechenleistung werden für die Verfolgung auch immer häufiger Partikelfilter eingesetzt. Bei diesen Filtern werden die Wahrscheinlichkeitsverteilungen, im Gegensatz zum Kalmanfilter, nicht durch statistische Momente, sondern durch eine Vielzahl von Partikeln modelliert. Diese Partikel werden dann einzeln verfolgt und aus den Zuständen aller Partikel können dann die Zustände des Objekts abgeleitet werden. Aufgrund des hohen Rechenaufwands und der guten Anwendbarkeit des Kalman-Filters auf die vorliegenden Probleme werden hier jedoch keine Partikelfilter verwendet.

3.1 Kalman-Filter

Seit der ursprünglichen Veröffentlichung durch Kalman [Kal] ist dieser Filter Gegenstand unzähliger Veröffentlichungen und Forschungsarbeiten. Die hier dargestellten Ausführungen und Bezeichnungen stammen größtenteils von Grewal und Andrews [GA08]. Der ursprüngliche Kalman-Filter ist ein Schätzer für ein lineares, dynamisches System, wobei sowohl das System als auch die Messwerte durch weißes Rauschen überlagert werden. Zusätzlich zu dem klassischen Kalman-Filter wurden mehrere Varianten entwickelt, die auch für nichtlineare Systeme verwendet werden können. Bevor die wichtigsten Typen des Filters vorgestellt werden, wird zunächst das zu lösende Problem erläutert.

3.1.1 Lineares Schätzproblem

Die gute Anwendbarkeit und mathematische Einfachheit des Kalman-Filters beruhen auf der Verwendung des Zustandsraummodells. Dabei wird ein dynamisches System durch eine endliche Zahl von Zustandsvariablen und deren Dynamik beschrieben. Ein System in kontinuierlicher Zeit kann

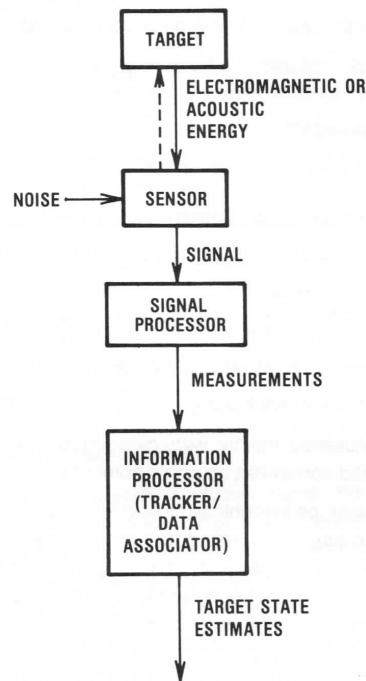


Abbildung 3.1: Komponenten eines Trackingsystems [BSF88, S.3]

demnach durch die Bewegungsgleichung

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{w}(t), \text{ mit } \mathbf{x}(t_0) = \mathbf{x}_0 \quad (3.1)$$

beschrieben werden. Dabei ist \mathbf{x} der Zustandsvektor, \mathbf{u} ein möglicher Eingriff in das System und \mathbf{w} eine statistische Zufallsgröße die das Prozessrauschen abbildet. Die Matrizen \mathbf{F} und \mathbf{G} definieren die dynamischen Zusammenhänge zwischen den Zuständen und Stellgrößen. Zur vollständigen Beschreibung der Dynamik gehört neben der Bewegungsgleichung auch der definierte Anfangszustand \mathbf{x}_0 . Um die Bewegungen des Systems nun ermitteln zu können, muss zumindest ein Teil der Zustandsvariablen gemessen werden. Das wird durch die Messgleichung

$$\mathbf{z}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{v}(t) \quad (3.2)$$

abgebildet. Die Messgrößen \mathbf{z} hängen dabei von den Zuständen \mathbf{x} ab und werden wiederum durch ein Rauschen \mathbf{v} überlagert. Die Matrix \mathbf{H} beschreibt, welche Zustände gemessen werden. In vielen Anwendungen wird nur ein geringer Teil der Zustände von Sensoren erfasst, da für die Messung der restlichen Zustände entweder der Aufwand zu groß ist, oder die Größen in der Situation schlicht nicht messbar sind.

Die beiden Rauschgrößen \mathbf{w} und \mathbf{v} sollten die tatsächlichen Störgrößen möglichst gut abbilden. Allerdings sind die Störungen in vielen Fällen nicht hinreichend bekannt, wodurch sie oft als weißes Rauschen angenommen werden. Bei der Herleitung des Kalman-Filters wird nun vorausgesetzt, dass die beiden Störgrößen normalverteilt sind, einen Erwartungswert von Null haben und unkorreliert sind. Für das Modellrauschen \mathbf{w} folgt demnach

$$E \langle \mathbf{w}(t) \rangle = \mathbf{0} \quad (3.3)$$

und

$$E \langle \mathbf{w}(t)\mathbf{w}^T(s) \rangle = \delta(t-s)\mathbf{Q}(t) \quad (3.4)$$

3 Grundlagen der Objektverfolgung

Dabei ist $E\langle \cdot \rangle$ der Erwartungswert, dieser wird bei Vektoren oder Matrizen elementweise angewendet. $\delta(\cdot)$ steht für die Dirac-Funktion und \mathbf{Q} für die Kovarianzmatrix des Prozessrauschens. Da die Dirac-Funktion in dieser Gleichung (3.4) nur für $t = s$ einen Wert ungleich Null liefert, ist das Rauschen zeitlich unkorreliert, es hängt somit nicht mit dem Rauschen anderer Zeitschritte zusammen. Die linke Seite dieser Gleichung entspricht im allgemeinen Fall der Korrelation und nicht der Kovarianz. Durch den Erwartungswert von Null sind diese beiden Größen jedoch identisch. Eine gaußverteilte (oder normalverteilte) Zufallsgröße hat einige spezielle Eigenschaften und wird bereits durch die ersten beiden zentralen Momente einer Zufallsgröße, des Mittelwerts und der Kovarianz, vollständig beschrieben. Somit ist \mathbf{w} durch die Gleichungen (3.3) mit (3.4) und \mathbf{v} durch

$$E\langle \mathbf{v}(t) \rangle = \mathbf{0}; \quad E\langle \mathbf{v}(t)\mathbf{v}^T(s) \rangle = \delta(t-s)\mathbf{R}(t) \quad (3.5)$$

bereits definiert.

Die Darstellung im kontinuierlichen Zeitbereich entspricht dem realen Verhalten von vielen physikalischen Systemen. Da die meisten Mess- und Trackingsysteme auf digitalen Rechnern ablaufen und dementsprechend zu diskreten Zeitpunkten bearbeitet werden, muss das Modell in den zeitdiskreten Bereich transformiert werden. Dabei wird die Differentialgleichung (3.1) zu der Differenzgleichung

$$\mathbf{x}_k = \Phi_{k-1}\mathbf{x}_{k-1} + \Gamma_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1} \quad (3.6)$$

Darin bezeichnet k den diskreten Zeitschritt k und alle anderen Größen sind analog zur kontinuierlichen Gleichung zu verstehen. Für die diskrete Messgleichung ergibt sich

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \quad (3.7)$$

und für die Zufallsgrößen

$$\begin{aligned} E\langle \mathbf{w}_k \rangle &= \mathbf{0}; & E\langle \mathbf{w}_k\mathbf{w}_i^T \rangle &= \Delta(k-i)\mathbf{Q}_k \\ E\langle \mathbf{v}_k \rangle &= \mathbf{0}; & E\langle \mathbf{v}_k\mathbf{v}_i^T \rangle &= \Delta(k-i)\mathbf{R}_k \end{aligned} \quad (3.8)$$

$\Delta(\cdot)$ steht dabei für das Kronecker-Delta für das gilt

$$\Delta(k-i) = \begin{cases} 1, & \text{falls } k = i \\ 0, & \text{falls } k \neq i \end{cases} \quad (3.9)$$

Mit den hier dargestellten Gleichungen lässt sich ein lineares, dynamisches System mit Mess- und Prozessrauschen beschreiben. Das Ziel eines Schätzers besteht darin, einen möglichst genauen Schätzwert $\hat{\mathbf{x}}_k$ für den Zustandsvektor des Systems zu ermitteln. Dafür muss der gewichtete, mittlere, quadratische Fehler

$$E\langle [\mathbf{x}_k - \hat{\mathbf{x}}_k]^T \mathbf{M} [\mathbf{x}_k - \hat{\mathbf{x}}_k] \rangle \quad (3.10)$$

minimiert werden. Durch die Matrix \mathbf{M} können die Zustände unterschiedlich gewichtet werden. Dieses Problem wird als lineares, quadratisches, Gauß-Schätz-Problem bezeichnet ([GA08]).

3.1.2 Klassischer Kalman-Filter

Der klassische Kalman-Filter liefert für das im vorherigen Absatz erläuterte Schätzproblem eine optimale Lösung, in dem Sinne, dass der mittlere, quadratische Fehler (3.10) minimiert wird. Diese optimale Lösung kann über das sogenannte Orthogonalitätsprinzip hergeleitet werden. Die Grundidee ist dabei, dass der Messfehler der besten Schätzung unkorreliert zu den Messwerten ist und

3 Grundlagen der Objektverfolgung

diese sozusagen orthogonal zueinander sind ([BSF88]. Für eine ausführliche Herleitung sei auf das hier hauptsächlich zitierte Werk von Grewal und Andrews [GA08] oder beispielsweise auf die Ausführungen von Bar-Shalom und Fortmann [BSF88] verwiesen.

Der daraus herleitbare Kalman-Filter besteht aus zwei Schritten. Zunächst werden aus dem geschätzten Zustandsvektor des letzten Zeitschritts und dessen Fehlerkovarianz die aktuellen Größen prädiert. Anschließend werden die prädierten Größen mit Hilfe der Messwerte aktualisiert. Die für diese beiden Schritte relevanten Gleichungen werden hier entsprechend der Problemstellung mit diskreter Zeit und für Systeme ohne Stelleingriff vorgestellt. Die diskrete Zeitdarstellung wird aufgrund der zeitdiskreten Verarbeitung im Rechner benötigt. Der fehlende Stelleingriff bedeutet, dass die zu verfolgenden Objekte nicht direkt beeinflusst werden können, wie es für andere Verkehrsteilnehmer der Fall ist. Über einen Stelleingriff könnte zwar auch die Bewegung des Ego-Fahrzeugs modelliert werden, aber da in dieser Arbeit die Bewegung auch für nicht verfolgte Größen kompensiert werden muss, wird diese Berechnung außerhalb des Filters durchgeführt.

Das dynamische System lässt sich damit analog zu Gleichung (3.6) durch

$$\begin{aligned} \mathbf{x}_k &= \Phi_{k-1} \mathbf{x}_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{z}_k &= \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \end{aligned} \quad (3.11)$$

beschreiben. Der Kalman-Filter soll für dieses System eine Schätzung des Zustandsvektors und dessen Fehlerkovarianz liefern. Der Schätzwert wird hier mit $\hat{\mathbf{x}}_k$ und dessen Kovarianzmatrix mit \mathbf{P}_k bezeichnet. Damit gilt für den Schätzfehler $\tilde{\mathbf{x}}_k = \hat{\mathbf{x}}_k - \mathbf{x}_k$. Wie bereits erwähnt, besteht der Filter aus zwei konsekutiven Schritten. Zur klaren Unterscheidbarkeit werden prädierte Werte mit $(-)$, und aktualisierte Werte mit $(+)$ gekennzeichnet.

Für die Initialisierung des Kalman-Filters müssen Anfangswerte

$$\begin{aligned} E \langle \mathbf{x}_0 \rangle &= \hat{\mathbf{x}}_0 \\ E \langle \tilde{\mathbf{x}}_0 \tilde{\mathbf{x}}_0^T \rangle &= \mathbf{P}_0 \end{aligned} \quad (3.12)$$

angegeben werden. Dabei ist vor allem darauf zu achten, dass \mathbf{P}_0 nicht zu klein gewählt wird, da der Filter den Anfangsfehler dann nur sehr langsam reduziert. Aus diesen Anfangswerten können anschließend durch rekursives Vorgehen die Werte für alle weiteren Zeitschritte berechnet werden.

Die prädierten Größen werden durch

$$\begin{aligned} \hat{\mathbf{x}}_k(-) &= \Phi_{k-1} \hat{\mathbf{x}}_{k-1}(+) \\ \mathbf{P}_k(-) &= \Phi_{k-1} \mathbf{P}_{k-1}(+) \Phi_{k-1}^T + \mathbf{Q}_{k-1} \end{aligned} \quad (3.13)$$

berechnet. Wie zu sehen ist, entspricht die Vorhersage des Zustandsvektors der Bewegungsgleichung des Modells ohne dem Rauschterm \mathbf{w} . Dieser entfällt aufgrund des Mittelwerts von Null, wie in Gleichung (3.8) definiert. Die Vorhersage der Fehlerkovarianz ergibt sich aus der Propagation des letzten Wertes und der Addition der Kovarianz des Prozessrauschens \mathbf{w} (siehe Gleichung (3.8)).

Der prädierte Wert für den Zustand wird anschließend mit den Messwerten \mathbf{z}_k aktualisiert:

$$\hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) + \bar{\mathbf{K}}_k [\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k(-)] \quad (3.14)$$

Darin ist $\bar{\mathbf{K}}_k$ der sogenannte Kalmanfaktor (engl. Kalman gain) der die Gewichtung zwischen Prädiktion und Messwert steuert. Dieser Faktor wird über

$$\bar{\mathbf{K}}_k = \mathbf{P}_k(-) \mathbf{H}_k^T [\mathbf{H}_k \mathbf{P}_k(-) \mathbf{H}_k^T + \mathbf{R}_k]^{-1} \quad (3.15)$$

3 Grundlagen der Objektverfolgung

berechnet. Die definierten Messungenauigkeiten werden durch \mathbf{R}_k berücksichtigt. Abschließend wird noch die Fehlerkovarianzmatrix mit Hilfe des Kalmanfaktors angepasst

$$\mathbf{P}_k(+)=\left[\mathbf{I}-\bar{\mathbf{K}}_k\mathbf{H}_k\right]\mathbf{P}_k(-) \quad (3.16)$$

Diese Gleichungen liefern trotz der einfachen, rekursiven Struktur die optimale Lösung des in Abschnitt 3.1.1 definierten Problems. Gerade diese einfache und effiziente Handhabung hat dem Kalman-Filter zu seiner herausragenden Bedeutung verholfen.

Eine weitere interessante Größe des Kalman-Filters ist die sogenannte Innovation. Diese beschreibt, wie gut die gemessenen Werte zu den prädizierten Werten passen. Die sogenannte informationsgewichtete, quadratische Innovation lässt sich nach [GA08] durch

$$\mathbf{I}_k^2=\left(\mathbf{z}_k-\mathbf{H}_k\hat{\mathbf{x}}_k(-)\right)^T\left(\mathbf{H}_k\mathbf{P}_k(-)\mathbf{H}_k^T+\mathbf{R}_k\right)^{-1}\left(\mathbf{z}_k-\mathbf{H}_k\hat{\mathbf{x}}_k(-)\right) \quad (3.17)$$

berechnen. Diese Größe ermöglicht, während der Verfolgung deren aktuelle Leistung zu beurteilen.

3.1.3 Nichtlineares Schätzproblem

Die größte Einschränkung beim klassischen Kalman-Filter besteht in der Voraussetzung linearer Systeme. Die meisten technisch relevanten Problemstellungen beinhalten jedoch nichtlineares Verhalten. Diese Nichtlinearitäten können sowohl im dynamischen Verhalten als auch bei der Messung auftreten. Im Vergleich zum linearen Modell werden die linearen Multiplikationen durch allgemeine Funktionsterme ersetzt, die auch nichtlineare Funktionen beschreiben können. Bei dieser Verallgemeinerung wird aus Gleichung (3.1)

$$\dot{\mathbf{x}}(t)=\mathbf{f}(\mathbf{x},t)+\mathbf{w}(t), \quad (3.18)$$

wobei der Stelleingriff $\mathbf{u}(t)$ vernachlässigt wurde. Da die Nichtlinearitäten auch bei der Messung auftreten können, ergibt sich für die Messgleichung

$$\mathbf{z}(t)=\mathbf{h}(\mathbf{x},t)+\mathbf{v}(t). \quad (3.19)$$

Im weiteren Verlauf wird wieder nur die zeitdiskrete Beschreibung betrachtet, die dementsprechend durch

$$\mathbf{x}_k=\phi_{k-1}(\mathbf{x}_{k-1})+\mathbf{w}_{k-1} \quad (3.20)$$

und

$$\mathbf{z}_k=\mathbf{h}_k(\mathbf{x}_k)+\mathbf{v}_k. \quad (3.21)$$

beschrieben werden kann. Handelt es sich bei $\phi_{k-1}(\cdot)$ und $\mathbf{h}_k(\cdot)$ um lineare Funktionen, liefert der klassische Kalman-Filter aus Abschnitt 3.1.2 eine optimale Lösung. Der Kalman-Filter verwendet lediglich die Mittelwerte und Kovarianzen der statistischen Größen. Die Transformationen dieser Größen in den Gleichungen des Filters können aus zwei Gründen angewendet werden. Zum einen werden normalverteilte Größen durch ihren Mittelwert und ihre Kovarianz bereits eindeutig definiert. Und andererseits werden normalverteilte Größen durch eine lineare Abbildung wieder auf normalverteilte Größen abgebildet ([GA08]).

Bei nichtlinearen Funktionen ergeben sich dadurch einige Probleme, da die Abbildung des Mittelwerts und der Kovarianz nicht mehr genau definiert ist. Zur Lösung dieses Problems gibt es nach Grewal [GA08] sowohl praktikable als auch unpraktikable Ansätze. Mit praktikabel bezeichnet er dabei Lösungen, die aufgrund des überschaubaren Rechenaufwands auch für Echtzeit-Anwendungen

in Frage kommen, auch wenn diese eine gewisse Näherung beinhalten. Die unpraktikablen Lösungen transformieren dagegen ganze Wahrscheinlichkeitsverteilungen und werden aufgrund des hohen Aufwands nur zur theoretischen Beurteilung der Fehler der praktischen Verfahren eingesetzt.

Die für Anwendungen interessanten Lösungen lassen sich in zwei Klassen einteilen. Die erste und intuitive Lösungskategorie arbeitet mit einer Linearisierung des Problems. Zu dieser Klasse gehört auch der für nichtlineare Probleme am häufigsten verwendete Extended Kalman-Filter, der im nächsten Abschnitt vorgestellt wird. Die zweite Klasse verwendet Abtastverfahren, bei denen einzelne Punkte mit bekanntem Mittelwert und Kovarianz mit der nichtlinearen Funktion abgebildet werden und aus dieser Abbildung dann die transformierten Größen ermittelt werden. Im Gegensatz zur Linearisierung können damit auch Effekte von höheren statistischen Momenten als dem Mittelwert und der Kovarianz berücksichtigt werden. Der Unscented Kalman-Filter ist eine wichtige Variante des Kalman-Filters, die dieses Verfahren nutzt und wird im Abschnitt 3.1.5 behandelt.

3.1.4 Extended Kalman-Filter

Der sogenannte Extended Kalman-Filter wurde von Stanley F. Schmidt bereits kurz nach der Veröffentlichung des Kalman-Filters entwickelt. Bei dieser Variante werden die Schätzwerte direkt durch die nichtlinearen Funktionen berechnet. Anschließend werden die benötigten partiellen Ableitungen für die Linearisierung an dieser Stelle gebildet ([GA08]).

Demnach wird der Schätzwert des Zustands mit Hilfe der nichtlinearen Bewegungsgleichung berechnet:

$$\hat{\mathbf{x}}_k(-) = \phi_{k-1}(\hat{\mathbf{x}}_{k-1}(+)). \quad (3.22)$$

Über die Messgleichung ergibt sich aus dem geschätzten Zustand der erwartete Messvektor

$$\hat{\mathbf{z}}_k = \mathbf{h}_k(\hat{\mathbf{x}}_k(-)). \quad (3.23)$$

Ähnlich wie im linearen Fall (Gleichung (3.13)) wird der aktuelle Wert der Fehlerkovarianz durch

$$\mathbf{P}_k(-) = \Phi_{k-1}^{[1]} \mathbf{P}_{k-1}(+) \Phi_{k-1}^{[1]T} + \mathbf{Q}_{k-1} \quad (3.24)$$

berechnet. Allerdings wird im nichtlinearen Fall die Zustandsänderung nicht einfach durch eine Matrixmultiplikation mit Φ_{k-1} beschrieben, sondern durch eine nicht weiter spezifizierte Funktion $\phi_k(\cdot)$. Um die Fehlerkovarianzmatrix $\mathbf{P}_k(-)$ dennoch ohne größeren Aufwand berechnen zu können, wird $\phi_k(\cdot)$ über eine Taylorreihenentwicklung um $\hat{\mathbf{x}}_k$ angenähert. Da für die Fehlerkovarianz nur die Abweichungen relevant sind und die Terme höherer Ordnung vernachlässigt werden, ergibt sich für die Näherung

$$\Phi_{k-1}^{[1]} \approx \left. \frac{\partial \phi_{k-1}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}(-)} \quad (3.25)$$

Dabei bezeichnet $^{[1]}$ die Ordnung der Taylorreihenentwicklung. Für diese Näherung muss die Funktion $\phi_k(\cdot)$ an der Stelle der Schätzung differenzierbar sein. Die Güte der Näherung hängt zum einen von der Funktion ab. Stark nichtlineare Funktionen können durch so eine einfache Näherung selbst auf kleinen Bereichen nur ungenügend dargestellt werden. Deshalb sollte die Funktion im Bereich um die Schätzung möglichst quasilinear sein. Das heißt, dass die Linearisierungsfehler im Verhältnis zu den Modellfehlern klein sind. Zum anderen bedingt eine gute Schätzung nur geringe Abweichungen vom Linearisierungspunkt und erhöht demnach die Chancen auf quasilineares Verhalten.

3 Grundlagen der Objektverfolgung

Wie im linearen Fall werden anschließend die Messwerte zur Aktualisierung des geschätzten Zustands verwendet:

$$\hat{\mathbf{x}}_k(+)=\hat{\mathbf{x}}_k(-)+\bar{\mathbf{K}}_k(\mathbf{z}_k-\hat{\mathbf{z}}_k) \quad (3.26)$$

Allerdings stammt der erwartete Messvektor $\hat{\mathbf{z}}_k$ jetzt von der nichtlinearen Messfunktion und kann nicht wie in Gleichung (3.14) durch eine einfache Multiplikation gewonnen werden. Außerdem tritt bei der Berechnung des Kalmanfaktors nach Gleichung (3.15) dasselbe Problem wie bei der Fehlerkovarianzmatrix auf. Dementsprechend wird der Kalmanfaktor durch

$$\bar{\mathbf{K}}_k=\mathbf{P}_k(-)\mathbf{H}_k^{[1]T}\left[\mathbf{H}_k^{[1]}\mathbf{P}_k(-)\mathbf{H}_k^{[1]T}+\mathbf{R}_k\right]^{-1} \quad (3.27)$$

berechnet, wobei $\mathbf{h}_k(\cdot)$ durch

$$\mathbf{H}_k^{[1]}\approx\left.\frac{\partial\mathbf{h}_k}{\partial\mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}_k(-)} \quad (3.28)$$

genähert wird. Als letzter Schritt wird dann noch die aktualisierte Fehlerkovarianzmatrix berechnet

$$\mathbf{P}_k(+)=\left[\mathbf{I}-\bar{\mathbf{K}}_k\mathbf{H}_k^{[1]}\right]\mathbf{P}_k(-) \quad (3.29)$$

Der hier vorgestellte Ansatz ermöglicht eine relativ einfache Filterung von nichtlinearen Problemen. Im Gegensatz zum klassischen Kalman-Filter liefert der Extended Kalman-Filter aufgrund der Näherungen von Nichtlinearitäten keine optimale Lösung des Problems. Da sehr viele Anwendungen jedoch mit quasilinearem Verhalten approximiert werden können, hat diese Variante ein breites Anwendungsspektrum. Um Mögliche Näherungsprobleme zu reduzieren wurden auch leichte Abwandlungen mit höheren Ordnungen der Taylorreihenentwicklung (siehe z.B. [BSF88] sowie eine iterative Form des Extended Kalman-Filter entwickelt ([GA08]).

3.1.5 Unscented Kalman-Filter

Wie bereits erwähnt, liefert der Extended Kalman-Filter nur bei quasilinearen Problemen brauchbare Ergebnisse. Außerdem müssen die nichtlinearen Funktionen an den entsprechenden Stellen differenzierbar sein um die Linearisierung durchführen zu können. Die Motivation für die Entwicklung des Unscented Kalman-Filter war, diese Einschränkungen zu umgehen. Um das zu erreichen, wird bei dieser Variante nicht das nichtlineare Problem durch ein lineares Problem angenähert, sondern versucht, die nichtlinearen Abbildungen der benötigten Mittelwerte und Kovarianzen nachzubilden. Ein guter Überblick über die Vorteile und die Herleitung des Unscented Kalman-Filter findet sich beispielsweise in [JU04]. Die folgenden Ausführungen stammen wieder größtenteils aus [GA08].

Die nichtlineare Transformation der entscheidenden ersten beiden statistischen Momente (Mittelwert und Kovarianz) wird beim Unscented Kalman-Filter mit Hilfe eines Abtastverfahrens (engl: sampling method) abgebildet. Grundsätzlich werden bei diesen Verfahren mehrere Punkte transformiert und aus den transformierten Werten die entsprechenden statistischen Momente ermittelt. Bei der ursprünglichen, bereits 1946 von Stanislaw Ulam veröffentlichten Methode werden die Abtastpunkte rein zufällig gewählt, was als Monte Carlo Analyse bezeichnet wird ([GA08]). Beim Unscented Kalman-Filter wird ein deutlich effektiveres Verfahren verwendet, die sogenannte Unscented Transformation nach der der Filter auch benannt wurde. Bei dieser Transformation werden die Abtastpunkte durch den Mittelwert und die Kovarianz der ursprünglichen Verteilung ermittelt, also genau den Werten, die der Kalman-Filter benötigt. Das grundsätzliche Vorgehen ist in Abbildung 3.2 (aus [JU04]) dargestellt. Die gewählten Abtastpunkte auf der linken Seite, die den selben Mittelwert und

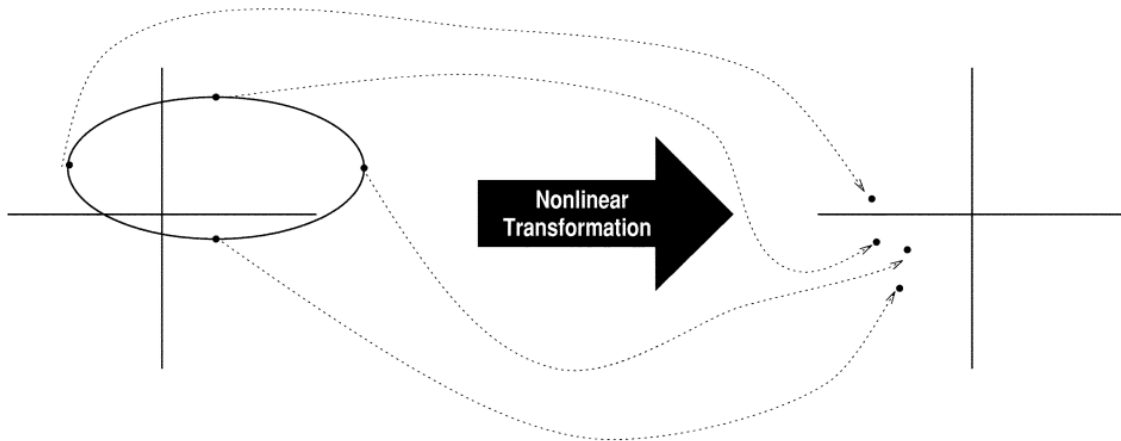


Abbildung 3.2: Grundprinzip eines Abtastverfahrens [JU04, S. 5]

die selbe Kovarianz der Verteilung aufweisen, werden durch die nichtlineare Transformation abgebildet. Aus diesen Bildpunkten werden anschließend die transformierten Momente bestimmt.

Die Unscented Transformation für eine Funktion $y = g(x)$ besteht somit aus drei Teilen. Zunächst werden die Abtastpunkte χ abhängig vom Mittelwert $\hat{\mathbf{x}}$ und der Kovarianz $\mathbf{P}_{\mathbf{xx}}$ so gewählt, dass die Gleichungen

$$\hat{\mathbf{x}} = \sum_i W_i \chi_i \quad (3.30)$$

und

$$\mathbf{P}_{\mathbf{xx}} = \sum_i W_i (\chi_i - \hat{\mathbf{x}})(\chi_i - \hat{\mathbf{x}})^T \quad (3.31)$$

erfüllt sind. Dabei hat jeder Abtastpunkt auch einen Gewichtungsfaktor W_i , wodurch eine größere Vielfalt an Varianten der Unscented Transformation ermöglicht wird. Im zweiten Schritt werden die gewählten Punkte χ transformiert und somit auf die Bildpunkte ζ abgebildet:

$$\zeta_i \stackrel{\text{def}}{=} g(\chi_i). \quad (3.32)$$

Aus diesen Bildpunkten können schließlich durch

$$\hat{\mathbf{y}} \approx \sum_i W_i \zeta_i \quad (3.33)$$

und

$$\mathbf{P}_{\mathbf{yy}} \approx \sum_i W_i (\zeta_i - \hat{\mathbf{y}})(\zeta_i - \hat{\mathbf{y}})^T \quad (3.34)$$

der transformierte Mittelwert und die transformierte Kovarianz näherungsweise bestimmt werden.

Da die Abtastpunkte lediglich die beiden Gleichungen 3.30 und 3.31 erfüllen müssen und die Gewichtungsfaktoren für zusätzliche Freiheitsgrade sorgen, sind unzählige Varianten der Abtastpunkte denkbar. Für eine ausführlichere Beschreibung der drei wichtigsten Abtaststrategien Simplex, Symmetric und Scaled sei auf [GA08] verwiesen. Beispielhaft wird hier lediglich die Berechnungsvorschrift der Symmetric Strategie vorgestellt. Mit der Dimension n des Zustandsvektors werden für diese Stra-

3 Grundlagen der Objektverfolgung

tegie $2n + 1$ Abtastpunkte folgendermaßen bestimmt:

$$\begin{aligned} \chi_0 &= \hat{\mathbf{x}}, & W_0 &= \frac{1}{n+1} \\ \chi_i &= \hat{\mathbf{x}} + \sqrt{n+1} \mathbf{c}_i, & W_i &= \frac{1}{2(n+1)}, \quad 1 \leq i \leq n \\ \chi_{n+i} &= \hat{\mathbf{x}} + \sqrt{n+1} \mathbf{c}_i, & W_{n+i} &= \frac{1}{2(n+1)}, \quad 1 \leq i \leq n \end{aligned} \quad (3.35)$$

Dabei ist \mathbf{c}_i die i -te Spalte eines Choleskyfaktors von $\mathbf{P}_{\mathbf{xx}}$. Wie aus den Gleichungen ersichtlich, wird mit χ_0 der Mittelwert direkt übernommen und die restlichen Abtastpunkte symmetrisch dazu ausgerichtet.

Mit Hilfe der Unscented Transformation können nun die benötigten Mittelwerte und Kovarianzen transformiert werden und damit ergibt sich für den Unscented Kalman-Filter folgender Algorithmus. Zunächst wird ein Cholesky-Faktor $\mathbf{C}_{k-1}(+)$ von $\mathbf{P}_{k-1}(+)$ berechnet. Dabei kann \mathbf{C} je nach Cholesky-Zerlegung eine untere oder obere Dreiecksmatrix, eine symmetrische Matrix aus der Singulärwertzerlegung, oder der sogenannte Eigenvektor-Cholesky-Faktor sein. Je nach Wahl der Faktorisierungsmethode können sich die erzielbaren Ergebnisse unterscheiden. Für genauere Ausführungen zu den Faktorisierungsmethoden und deren Vor- und Nachteile sei auf [GA08] verwiesen.

Damit lassen sich die prädizierten Werte durch die Unscented Transformation UT bestimmen:

$$[\hat{\mathbf{x}}_k(-), \mathbf{P}_{\mathbf{xx}}] = UT(\hat{\mathbf{x}}_{k-1}(+), \mathbf{C}_{k-1}(+), \phi(\cdot)) \quad (3.36)$$

Um auf die prädizierte Kovarianz des Zustandes zu kommen, muss noch das Prozessrauschen durch

$$\mathbf{P}_k(-) = \mathbf{P}_{\mathbf{xx}} + \mathbf{Q} \quad (3.37)$$

berücksichtigt werden. Nach dem erfolgreichen Prädiktionsschritt müssen nun die geschätzten Größen mit Hilfe der Messwerte angepasst werden. Zur Ermittlung der erwarteten Messwertmomente wird ein Cholesky-Faktor $\mathbf{C}_k(-)$ von $\mathbf{P}_k(-)$ benötigt. Damit lassen sich

$$[\hat{\mathbf{z}}_k, \mathbf{P}_{\mathbf{zz}}] = UT(\hat{\mathbf{x}}_k(-), \mathbf{C}_k(-), h(\cdot)) \quad (3.38)$$

berechnen. Zur Berechnung des Kalmanfaktors $\bar{\mathbf{K}}_k$ wird neben der Kovarianzmatrix der Messung auch die Kreuzkovarianzmatrix zwischen Zustand und Messung benötigt. Diese lässt sich aus den Werten der UT mittels

$$\mathbf{P}_{\mathbf{zx}} = \left[\sum_i W_i (\chi_i - \hat{\mathbf{x}}_k(-)) (\zeta_i - \hat{\mathbf{z}}_k)^T \right] \quad (3.39)$$

bestimmen. Der Kalmanfaktor ergibt sich dann aus

$$\bar{\mathbf{K}}_k = \mathbf{P}_{\mathbf{zx}} [\mathbf{P}_{\mathbf{zz}} + \mathbf{R}_k]^{-1}. \quad (3.40)$$

Die Aktualisierung der prädizierten Werte ist wieder analog zu den oben vorgestellten Varianten des Kalman-Filters

$$\begin{aligned} \hat{\mathbf{x}}_k(+) &= \hat{\mathbf{x}}_k(-) + \bar{\mathbf{K}}_k (\mathbf{z}_k - \hat{\mathbf{z}}_k) \\ \mathbf{P}_k(+) &= \mathbf{P}_k(-) + \bar{\mathbf{K}}_k \mathbf{P}_{\mathbf{zx}}^T \end{aligned} \quad (3.41)$$

Obwohl der hier vorgestellte Unscented Kalman-Filter für nichtlineare Probleme anwendbar ist, ist der benötigte Rechenaufwand relativ gering. Ein weiterer Vorteil gegenüber dem EKF ist die bessere

Anwendbarkeit, da das System nicht mehr differenzierbar sein muss. Darüber hinaus können mit dem UKF im Vergleich zum EKF auch stark nichtlineare Systeme abgebildet werden. Abschließend lässt sich festhalten, dass die optimale Filterwahl vom Grad der Nichtlinearitäten abhängt. Je nichtlinearer sich ein System verhält, desto sinnvoller ist eine Verwendung des UKF.

3.2 Assoziation

Die hier vorgestellten Filter benötigen jeweils den aktuellen Messwert des zu verfolgenden Objekts. Im einfachsten Fall, wenn ein Sensor nur ein Objekt detektiert, ist diese Zuordnung direkt gegeben. Im Allgemeinen werden jedoch meistens mehrere Objekte gemessen, die jeweils einzeln verfolgt werden sollen. Deshalb muss in jedem Zeitschritt vor der Verwendung des Filters zu jedem Objekt der korrespondierende Messwert assoziiert werden. Diese Zuordnung kann in manchen Fällen sehr aufwendig und teilweise nicht eindeutig bestimmbar sein. Vor allem wenn sich die Trajektorien der Objekte kreuzen wird eine eindeutige Zuordnung schwierig.

Ein einfaches und weit verbreitetes Assoziationsverfahren ist das Nearest-Neighbour-Verfahren, bei dem jeweils die Objekte und Messungen mit dem geringsten Abstand assoziiert werden. Da jedoch gerade bei dem ursprünglichen Trackingproblem von Radarobjekten dieses Verfahren häufig falsche Assoziationen erzeugt, wurden in den letzten Jahrzehnten viele unterschiedliche Assoziationsverfahren entwickelt. Die Zuordnung wird umso schwieriger, je kleiner der Objektabstand im Vergleich zu den Sensorgenauigkeiten wird. Das hier verwendete Assoziationsverfahren beruht wie das Nearest-Neighbour-Verfahren auf dem Abstand zwischen Objekt und Messwert und wird im Rahmen der Objektverfolgung im nächsten Kapitel genauer behandelt. Für die Beschreibung von komplexeren Assoziationsmethoden sei auf [BSF88] verwiesen.

4 Objekterkennung und Tracking LIDAR

In diesem Kapitel wird die Objekterkennung und Verfolgung mit Hilfe des Lidar-Scanners an der Fahrzeugfront beschrieben. Die Funktionsweise dieses Sensors ist in Abschnitt 2.3.1 ausführlich beschrieben. Der Sensor wird mit 1° Winkelauflösung und dem maximalen Scanbereich von 0° bis 180° betrieben. Somit liefert er in jedem Zeitschritt 181 Entfernungsmesswerte, die den Konturverlauf der Umgebung in der Sichte Ebene des Sensors darstellen. Detektiert der Sensor keine Reflexion, entweder aufgrund zu geringer Remission oder zu großer Entfernung des Objekts, gibt er den Maximalwert von 81,91m aus.

Zur Veranschaulichung der Messdaten ist in Abbildung 4.1 das Bild einer beispielhaften Verkehrssituation (a) mit den korrespondierenden Messwerten des Sensors aus der Vogelperspektive (b) dargestellt. Dabei liegt der an der Fahrzeugfront angebrachte Sensor im Koordinatenursprung, da die Daten im Sensorkoordinatensystem dargestellt sind. Zur Erinnerung sei noch einmal darauf hingewiesen, dass die x-Achse in Fahrtrichtung zeigt, das Ego-Fahrzeug somit gerade nicht mehr abgebildet ist. Die genauen Zusammenhänge zwischen den verwendeten Koordinatensystemen sind in Abschnitt 2.2 beschrieben.

Diese Szene wurde ausgewählt, da bereits alle wichtigen Formen von Verkehrsteilnehmern vorkommen. Um das Verständnis der folgenden Abschnitte zu erleichtern, werden hier kurz die Begriffe I-, L-, IS-, C- und E-Form eingeführt. Für Fahrzeuge wird in den meisten Arbeiten ein annähernd rechteckiger Grundriss angenommen. Ein Rechteck wird in den LIDAR-Daten als I- oder L-Form abgebildet, da der Scanner entweder nur eine oder zwei Seiten erfassen kann. Das Heck des vor dem Versuchsfahrzeugs parkenden, schwarzen PKWs wird von dem Sensor mit fünf, auf einer Linie liegenden Messpunkten abgebildet und ist ein Beispiel für eine solche I-Form. Wie in Abbildung 4.1(b) zu sehen, wird der weiße Van ebenfalls durch eine Linie aus mehreren Punkten repräsentiert. Offensichtlich handelt es sich dabei aber nicht um das Heck bzw. die Front des Fahrzeugs, sondern um eine Seite. Da die in den folgenden Kapiteln vorgestellte Herangehensweise eine solche Unterscheidung ermöglicht, wird diese Form als IS-Form (I Seite) bezeichnet. In der Abbildung der LIDAR-Daten ist auf der linken Seite (bei $y = 5m$) eine L-Form eines parkenden Fahrzeugs zu erkennen, das allerdings aufgrund des deutlich kleineren Öffnungswinkels der Kamera nicht mehr im Bild zu sehen ist. Außerdem wird noch die C-Form eingeführt, wie sie beispielsweise in [Wal08] verwendet wird. Streng genommen handelt es sich dabei um keine Form, sondern um eine zylinderförmige Repräsentation von kleinen Objekten mit weniger als 90 cm Durchmesser. Da vor allem Fußgänger aufgrund unterschiedlicher Beinstellungen keine definierte Form in den Lidar-Daten ergeben, werden die einzelnen Messpunkte einfach in einem Kreis zusammengefasst. Da die Objekte in der Realität auch eine Höhe besitzen wird von einem Zylinder gesprochen (engl. cylinder). Neben den Fahrzeugen und Fußgängern, die durch diese Formen gut beschrieben werden können, sollen in dieser Arbeit auch Radfahrer berücksichtigt werden, da diese im städtischen Verkehr eine wichtige Rolle spielen. Das große Problem bei der Detektion von Radfahrern mit Hilfe eines LIDAR-Scanners ist, ähnlich wie bei Fußgängern die undefinierte Oberfläche. Außerdem kommt noch erschwerend hinzu, dass die Lichtimpulse des Scanners teilweise durch die Lücken im Objekt, wie beispielsweise in den Felgen oder im Rahmen, dahinter liegende Flächen messen. Da diese Problemstellung in der Literatur nach meinen Erkenntnissen noch nicht veröffentlicht wurde, wurde hier ein eigener Ansatz entwickelt. Ähnlich zur C-Form, wird für Repräsentation von Radfahrern eine Ellipse verwendet. Damit kann die längliche, undefinierte Form gut beschrieben werden. Allerdings ist die Zuordnung und Identifizierung solcher E-Formen aufgrund von Überschneidungen mit andern Formen (z.B. I-Form) nicht trivial. Auf alle diese Formen und ihre Verwendung wird in Abschnitt 4.3.1 noch genauer eingegangen.

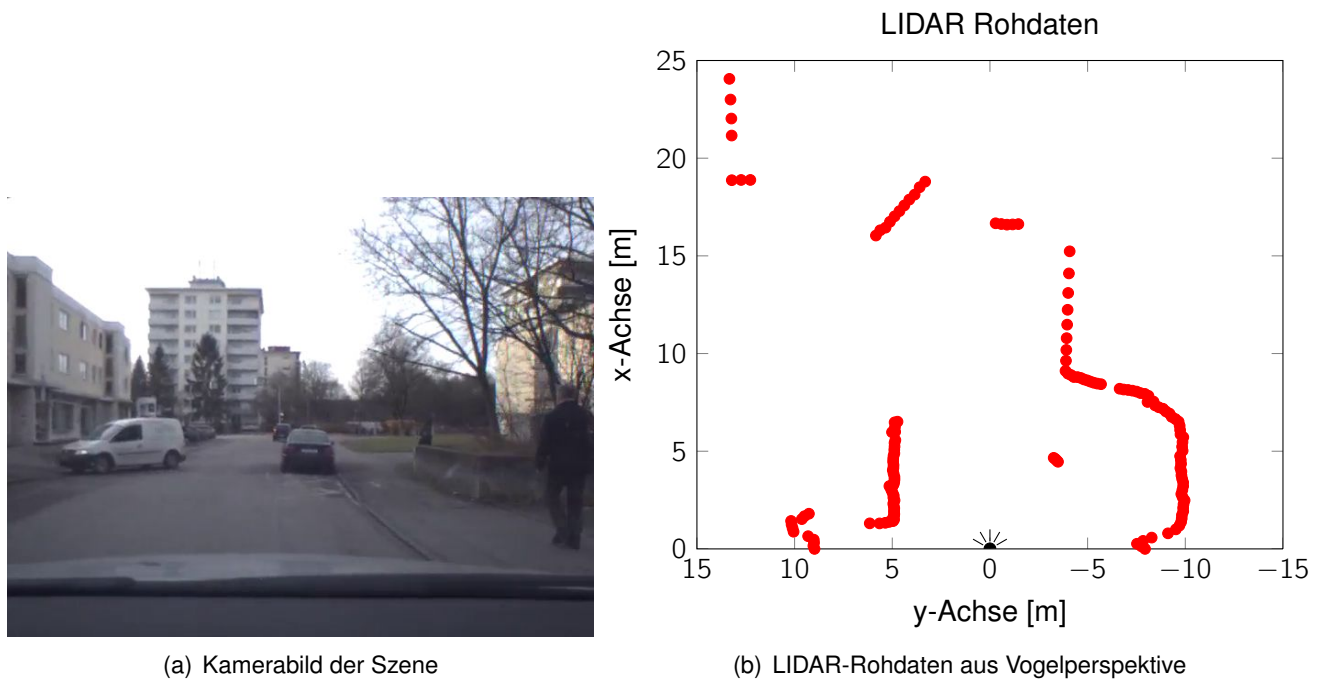


Abbildung 4.1: Lidar Rohdaten

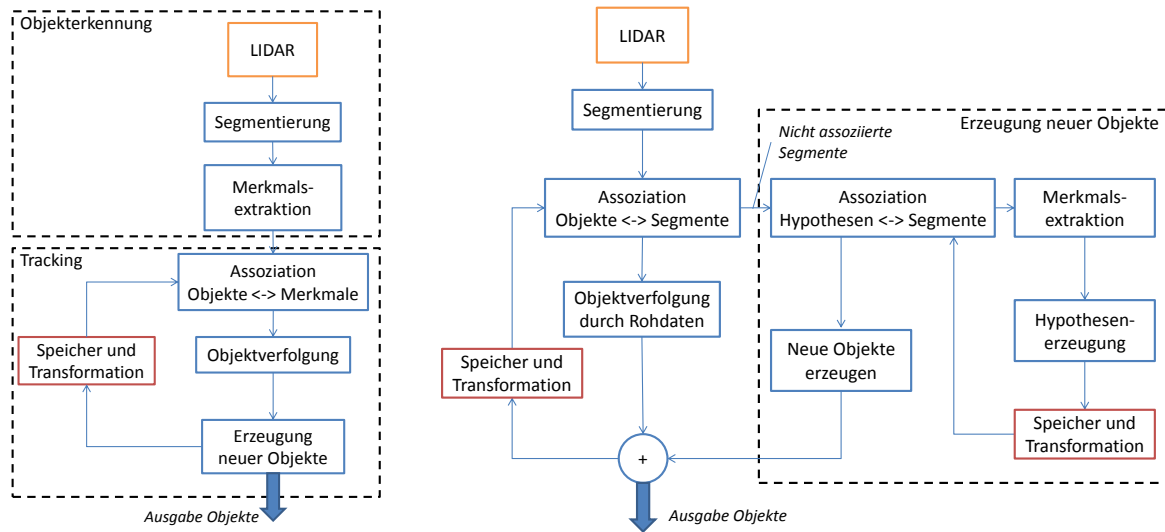
Das Ziel ist es nun, aus den Messpunkten Objekte zu gewinnen und diese zu verfolgen. Mit dieser Problematik beschäftigen sich zahlreiche Veröffentlichungen. Dabei gibt es viele unterschiedliche Ansätze, wie aus den Daten die Objekte extrahiert und verfolgt werden können. Da sich der hier verwendete Algorithmus deutlich von den meisten Veröffentlichungen unterscheidet, wird zunächst im Abschnitt 4.1 der Ablauf übersichtlich dargestellt und im Vergleich mit dem weit verbreiteten Ansatz der Unterschied verdeutlicht. Anschließend wird dann in den folgenden Kapiteln auf die nötigen Einzelschritte eingegangen.

4.1 Übersicht über Programmablauf

In fast allen Arbeiten, die sich mit diesem Thema beschäftigen, existiert eine strikte Trennung zwischen der Objekterkennung und der Verfolgung (beispielsweise [Wal08], [KBD], [MBNne] oder [SDS 2]). Das bedeutet, dass zunächst aus der Lidar-Messung eines Zeitschritts Objekte extrahiert werden, die die oben vorgestellten Formen aufweisen, und diese dann mit den Spuren des letzten Zeitschritts verglichen werden. Dieser Ablauf ist in Abbildung 4.2 (a) dargestellt. Prinzipiell gibt es in den meisten dieser Arbeiten zwei entscheidende Schritte der Objekterkennung. Zunächst werden die Rohdaten in zusammengehörende Segmente eingeteilt (Segmentierung). Anschließend wird dann versucht, die Form dieser Segmente zu ermitteln bzw. deren Merkmale zu extrahieren. Das Tracking funktioniert dann wie in Kapitel 3 beschrieben. Zunächst werden die Merkmale mit den alten Objekten assoziiert und anschließend die aktuellen Objektzustände mittels Kalmanfilter ermittelt. Aus nicht assoziierten Merkmalen werden dann neue Objekte erzeugt. Um die aktuellen Objekte im nächsten Zeitschritt wieder verwenden zu können, werden diese gespeichert und in das neue Fahrzeugkoordinatensystem transformiert.

Diese Trennung vereinfacht zwar den Programmaufbau, führt aber zu einigen Problemen bei der Objekterkennung. Ein bekanntes Problem ist beispielsweise die Veränderung der Form eines entgegenkommenden Fahrzeugs. Dabei wechselt die erkannte Form von einem I, zu einem L. Dadurch springt auch der aus den Daten ermittelte Referenzpunkt der für das Verfolgen verwendet wird. Die-

4 Objekterkennung und Tracking LIDAR



(a) gewöhnlicher Trackingalgorithmus

(b) verwendeter Trackingalgorithmus

Abbildung 4.2: Übersicht über Trackingalgorithmen

ses Problem und auch mögliche Lösungsansätze werden in [FLD] übersichtlich dargestellt und diskutiert. Ein deutlich größeres Problem ergibt sich aber bei schlechten Messwerten. Wird ein Objekt in einem Zeitschritt nur ungenügend oder Bruchstückhaft erkannt, wird es nahezu unmöglich darin eine bekannte Form zu erkennen. Da bei dem hier verwendeten Sensor nur eine Ebene mit 1° Winkelauflösung abgetastet wird, stehen oft nur einzelne Messpunkte für ein Objekt zur Verfügung, aus denen ohne Vorwissen keine vernünftige Form abgeleitet werden kann. Aus diesem Grund wurde hier ein Ansatz entwickelt, die Objekterkennung und die Objektverfolgung zu kombinieren. Der Ablauf ist in Abbildung 4.2 (b) abgebildet. Zunächst findet wie im konventionellen Ansatz eine Vorverarbeitung der Messdaten statt. Auf die Besonderheiten und unterschiedlichen Umsetzungsmöglichkeiten dieser Segmentierung wird in Abschnitt 4.2 eingegangen. Die Besonderheit des verwendeten Ansatzes ist nun, dass diese Segmente direkt mit den alten Objekten assoziiert werden und die neuen Messwerte im Bezug auf das alte Objekt gewonnen werden. Die Funktionsweise ist in Abschnitt 4.4 ausgeführt. Die Erzeugung neuer Objekte läuft ähnlich wie in der konventionellen Methode. Dafür werden ebenfalls die Segmente verwendet, die nicht mit alten Objekten assoziiert werden konnten. Allerdings werden aus den extrahierten Merkmalen nicht direkt neue Objekte sondern zunächst Hypothesen erzeugt, die zurückgeführt werden und im nächsten Zeitschritt bestätigt werden müssen (siehe Abschnitt 4.3).

Die einzige vom Autor gefundene Arbeit in diesem Bereich, in der diese beiden Schritte ebenfalls kombiniert werden, stammt von Vu und Aycard [TA]. Darin werden immer die Messwerte von mehreren Zeitschritten über ein gleitendes Zeitfenster betrachtet und die wahrscheinlichste Trackzuordnung gesucht. Es handelt sich dabei um einen relativ aufwendigen Ansatz, da sowohl die Daten mehrere Zeitschritte verarbeitet werden müssen, als auch der Raum möglicher Trackzuordnungen sehr groß ist.

4.2 Segmentierung

Um aus den Lidardaten Objekte zu extrahieren, müssen zunächst zusammenhängende Bereiche identifiziert werden. Das Ziel dabei ist, dass möglichst jedem Objekt genau ein Segment zugeordnet werden kann. Selbst wenn dieser Schritt für die nachfolgende Bearbeitung nicht nötig wäre, kann

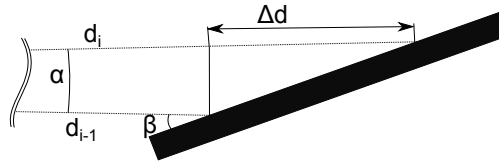


Abbildung 4.3: Veranschaulichung des Segmentierungsgrenzwerts

durch eine erste Einteilung die Anzahl der zu untersuchenden Punkte reduziert und somit Rechenaufwand gespart werden. Deshalb findet sich in fast allen Veröffentlichungen eine Form dieser Vorverarbeitung der Rohdaten. Für die Segmentierung werden in der Literatur einige unterschiedliche Ansätze verwendet. Ein intuitiver Ansatz besteht einfach darin, den euklidischen Abstand zwischen zwei aufeinander folgenden Messpunkten zur Segmentierung heranzuziehen. Das heißt: Sobald der Abstand zwischen zwei Punkten einen gewissen Schwellwert überschreitet, gehört der zweite Punkt einem neuen Segment an. Ein weit verbreiteter Ansatz ist die inkrementelle Linienextraktion, die in den zusammenhängenden Arbeiten von Lindl [Lin08] und Walchshäusl [Wal08] anschaulich erklärt ist und zusätzlich zum Abstand auch Linienverläufe berücksichtigt. Bei diesem Verfahren werden Messpunkte so lange zu einem Segment hinzugefügt, bis der durchschnittliche Abstand zu der Regressionsgeraden durch alle Punkte des Segments einen Schwellwert überschreitet. Falls der euklidische Abstand zwischen den letzten beiden Punkten ebenfalls einen Grenzwert überschreitet, wird ein neues Segment begonnen. Andernfalls wird zu dem aktuellen Segment ein neues Linienelement hinzugefügt. Dadurch werden die Messdaten durch dieses Verfahren nicht nur in Segmente unterteilt sondern auch gleich durch zusammengehörige Geradenabschnitte repräsentiert.

Darüber hinaus gibt es noch zahlreiche andere Verfahren wie beispielsweise in [Lab+05], wo Kovarianzen berücksichtigt werden, oder in [FLN12] bei dem geometrische Invarianten in Polarkoordinaten ausgenutzt werden .

Wie bereits erwähnt, werden in dieser Arbeit direkt die Rohdaten für das Verfolgen von Objekten verwendet. Deshalb müssen aus den Daten keinerlei zusätzliche Informationen wie Linienelemente oder ähnliches gewonnen werden, bevor sie mit den alten Objekten assoziiert werden. Es könnte theoretisch für jeden einzelnen Messpunkt überprüft werden, ob er zu dem aktuellen Objekt passt und könnte dann dementsprechend verarbeitet werden. Aber auch hier gilt, dass durch eine grobe Vorverarbeitung, der Rechenaufwand massiv reduziert werden kann.

In dieser Arbeit wird ein sehr einfaches und damit rechenarmes Verfahren für die Segmentierung verwendet. Ähnlich wie bei Sparbert et al. [SDS 2] und Mendes et al. [MBNne] wird lediglich ein Entfernungskriterium herangezogen. Im Gegensatz zum häufig verwendeten euklidische Abstand, der zunächst berechnet werden muss, wird direkt die Differenz der gemessenen Entfernung zum Sensor zweier aufeinander folgender Messpunkte verwendet. Es wird angenommen, dass der Lidar-Sensor Objekte nur ab einem bestimmten Winkel zwischen Oberfläche und LIDAR-Strahl β_{detect} erkennen kann. Wird der Winkel zu klein, wird höchstwahrscheinlich zu wenig Licht zum Sensor zurück reflektiert. Für diesen Grenzwinkel kann der resultierende Distanzunterschied zwischen zwei benachbarten (um 1° versetzten) Strahlen ermittelt werden. Aus den geometrischen Beziehungen ergibt sich für die maximale Differenz zwischen zwei gemessenen Entfernungen d_i und d_{i-1} :

$$\Delta d_{max} = d_{min} \frac{2 \sin(\frac{\alpha}{2}) \sin(90^\circ - \beta_{detect} + \frac{\alpha}{2})}{\sin(\beta_{detect} - \alpha)} \quad (4.1)$$

Dabei ist Δd_{max} die maximal zulässige Differenz zwischen zwei konsekutiven Werten, d_{min} der kleinere Wert der beiden und α die Winkelauflösung des Sensors. Die Situation ist zur Verdeutlichung in Abbildung 4.3 dargestellt. Aus dem flachen Auftreffwinkel β und der Winkelauflösung des Sensors α

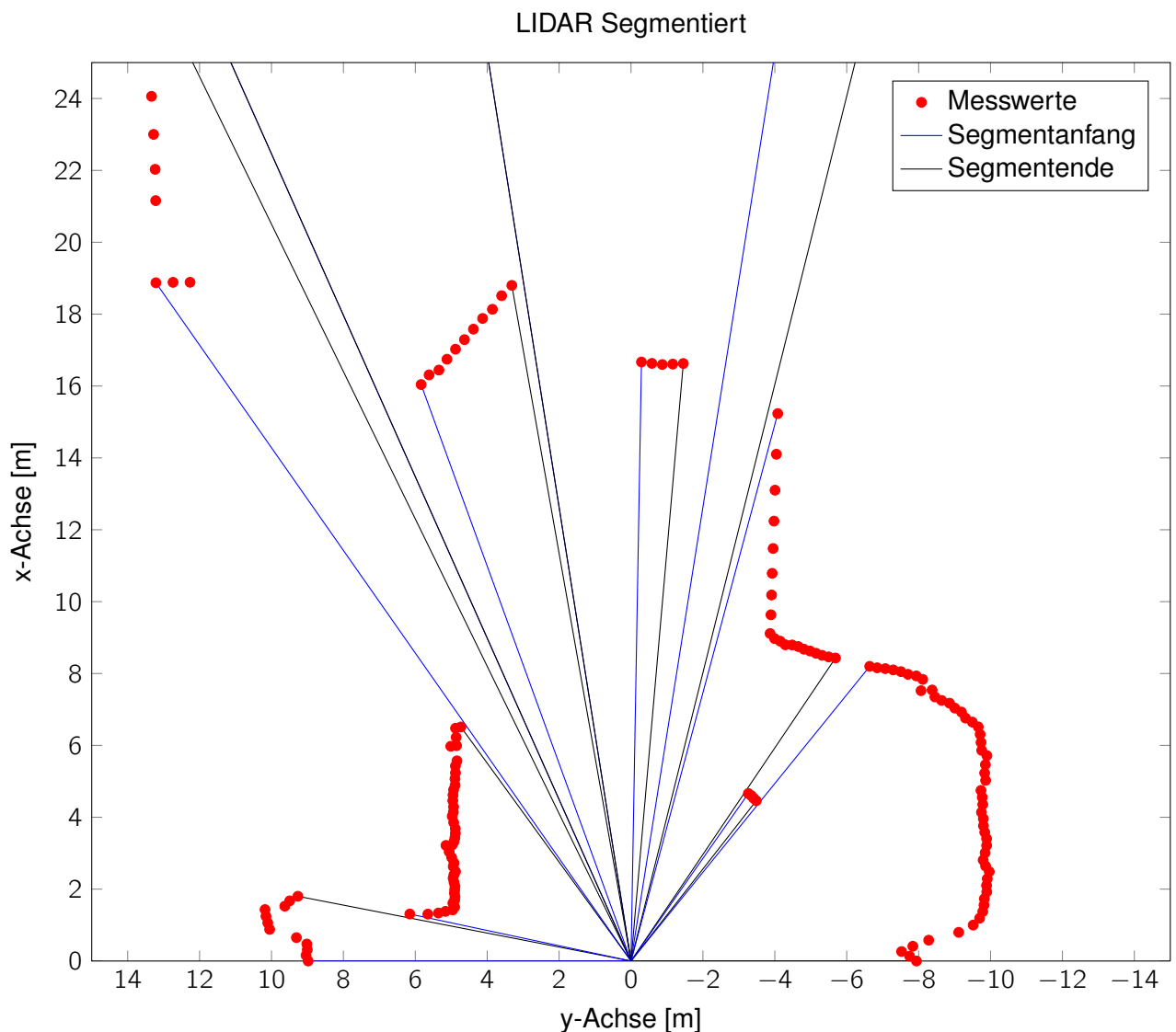


Abbildung 4.4: Segmentierte Lidar-Messung

ergibt sich die Differenz, die bei der Vermessung eines geraden Objekts zu erwarten ist. Da sowohl die Winkelauflösung des Sensors als auch der Wert von β_{detect} konstant sind, kann der Bruch zu einem festen Schwellwert $\tau_{segment}$ zusammengefasst werden. Zur Segmentierung muss dann lediglich für jeden Messpunkt überprüft werden, ob die Gleichung

$$|d_i - d_{i-1}| < d_{min} \tau_{segment} \quad (4.2)$$

erfüllt ist. Ist der Abstand größer, gehört d_i zu einem neuen Segment. Das heißt, dass Objekte die unter einem kleineren Winkel als β_{detect} erfasst werden, fälschlicherweise in mehrere Segmente zerfallen. Deshalb ergibt sich für die Wahl von β_{detect} ein Zielkonflikt. Wird der Winkel zu groß gewählt, werden die Objekte durch mehrere Segmente dargestellt, ist er zu klein, können unterschiedliche Objekte nicht mehr unterschieden werden. Da für den hier verwendeten Algorithmus der zweite Fall problematischer ist, wurde $\beta_{detect} = 10^\circ$ empirisch festgelegt.

In Abbildung 4.4 ist das Ergebnis der Segmentierung für die bereits am Anfang des Kapitels dargestellten Situation zu sehen. Darin markieren die blauen Linien jeweils den Beginn und die schwarzen Linien das Ende eines Segments. Die über den Rand der Grafik hinausgehenden Linien zeigen,

dass der Sensor auch noch weiter entfernte Punkte aufgezeichnet hat die hier aus Gründen der Übersichtlichkeit nicht mehr im abgebildeten Bereich liegen. Wie zu erkennen ist, werden in dieser übersichtlichen Situation alle Objekte mit einem einzelnen Segment erfasst. Somit kann bereits mit geringstem Rechenaufwand eine erste Einteilung der Rohdaten vorgenommen werden.

4.3 Objekterzeugung

Neue Objekte werden bei Trackingalgorithmen für gewöhnlich aus Messungen erzeugt, die zu keinem der bisher bekannten Objekte zugeordnet werden können. Dementsprechend ist die Objekterzeugung im Algorithmusablauf einer der späteren Schritte. Aus zwei Gründen wird dieser Abschnitt dennoch nach vorne gezogen. Zum einen gibt es am Anfang des Trackings noch gar keine bekannten Objekte, dementsprechend ist die Objekterzeugung bei der Initialisierung der wichtigste Schritt. Und andererseits ist die Merkmalsextraktion, einer der entscheidenden Schritte der Erzeugung, bei der konventionellen Herangehensweise noch vor dem Tracking angesiedelt, wie aus Abbildung 4.2(a) ersichtlich. In dem entwickelten Algorithmus sind noch weitere Schritte der Objekterzeugung zugeordnet (siehe Abbildung 4.2(b)), die in den folgenden Abschnitten ausführlich erläutert werden.

4.3.1 Merkmalsextraktion

Bei der Merkmalsextraktion wird versucht, in den einzelnen Messpunkten bestimmte Objekte oder Formen zu erkennen und deren Merkmale wie Position, Orientierung und Ausdehnung zu ermitteln. Meistens werden die Rohdaten davor bereits in zusammengehörige Bereiche unterteilt (Abschnitt 4.2). Bevor die Segmente dann weiterverarbeitet werden, wird häufig noch nach Verbindungen zwischen benachbarten Segmenten gesucht. Zur Merkmalsextraktion gibt es unterschiedlich komplexe Ansätze. Fuerstenberg et al. [FLD] beschreiben und vergleichen die weit verbreiteten Ansätze zur Bestimmung des Referenzpunkts mittels des Schwerpunkts (engl. Center of Gravity Approach) und mittels des Objektrechtecks (engl. Object-Box Approach). Die Verwendung des Schwerpunkts eines Segments ist denkbar einfach, jedoch für Objekte die wie Fahrzeuge eine gewisse Ausdehnung haben, äußerst ungenau, da die Lage des Schwerpunkts stark von den erkannten Seiten und der Messpunktverteilung auf diesen Seiten abhängt. Für Fahrzeuge wird deshalb oft der Object-Box Ansatz verwendet, bei dem um die erkannte Form ein Rechteck gezogen wird und dessen Mittelpunkt verwendet wird. Damit die Ausdehnung des Rechtecks möglichst der des Objekts entspricht, werden die maximalen detektierten Objektdimensionen gespeichert (z.B. [FLD], [FC]).

In dieser Arbeit wird ähnlich vorgegangen, wie es von Walchshäusl [Wal08, S.36 ff.] ausführlich beschrieben wird. Allerdings wird hier keine inkrementelle Linienextraktion durchgeführt, da sie nicht zwingend notwendig erscheint. Zunächst wird überprüft, ob benachbarte Segmente zu dem selben realen Objekt gehören könnten und dann anschließend die mögliche Form extrahiert.

4.3.1.1 Verbinden von Segmenten

Das Ziel der Segmentierung ist bekanntlich, dass möglichst alle Messpunkte die zu einem realen Objekt gehören, zu einem Segment gehören. Es gibt allerdings drei Fälle, in denen Segmente fälschlicherweise getrennt sind. Zwei davon sind in Abbildung 4.5 dargestellt.

In Abbildung (a) sind die Messpunkte eines auf der Gegenseite geparkten, schwarzen PKWs und die entsprechenden Linien der Segmentierung, die bereits aus der Abbildung 4.4 bekannt sind, zu sehen. Der Lidar-Sensor detektiert mit sieben Messwerten einen Großteil der Fahrzeugfront, die linke vordere Ecke und in etwa den linken Kotflügel. Bei den nächsten vier Messwerten, symbolisiert durch die gestrichelten Linien, empfängt der Sensor jedoch nicht genügend Remission. Erst der fünfte

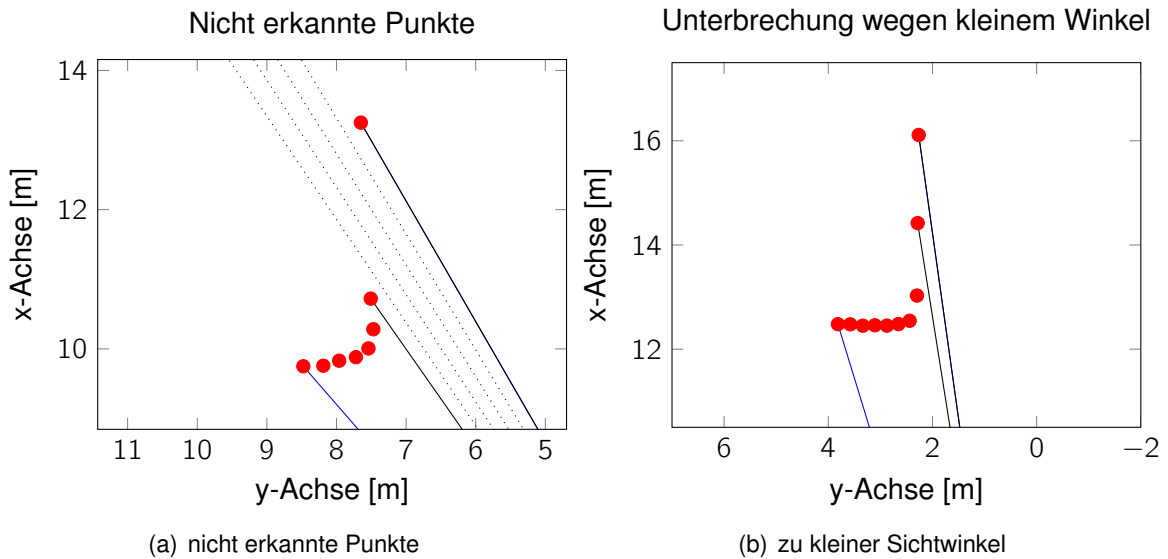


Abbildung 4.5: Verbinden von Segmenten

Messwert detektiert wieder das Objekt. Solche partiellen Nichtdetektionen treten besonders häufig an den Seiten dunkler Fahrzeuge auf. Werden diese nur unter einem relativ kleinen Winkel vom Lidar bestrahlt, wird der Großteil der Energie an der glatten Oberfläche reflektiert und die remittierte Energie reicht nicht mehr aus um das Objekt zu detektieren.

Das selbe Problem von fälschlicherweise getrennten Segmenten tritt bei Verdeckung auf. Die Konstellation der in Abbildung 4.5 (a) dargestellten Messpunkte kann auch dadurch entstehen, dass ein näheres Objekt diese vier Messstrahlen blockiert. Auch dann sollten die beiden Segmente als Einheit betrachtet werden, um das reale Objekt möglichst gut abzubilden.

Der dritte Fall entsteht durch die Segmentierung und ist in Abbildung 4.5 (b) gezeigt. Es handelt sich dabei um ein überholendes Fahrzeug, dessen L-Form sehr gut detektiert wird. Allerdings betragen die Winkel jeweils zwischen den beiden letzten abgebildeten Strahlen und der detektierten Fahrzeugseite $9,66^\circ$ bzw. $8,66^\circ$ und sind demnach kleiner wie der Grenzwert $\beta_{detect} = 10^\circ$. Aus diesem Grund wurde der letzte Punkt auch einem eigenen Segment zu gewiesen. Wie im Abschnitt 4.2 bereits erwähnt, unterliegt die Wahl dieses Grenzwerts einem Zielkonflikt. Da zu stark unterteilte Segmente nachträglich verbunden werden können, fälschlicherweise verbundene Objekte aber schlecht getrennt werden können, ist es wichtig den Grenzwert trotz solcher Beispiele nicht zu niedrig anzusetzen. Das resultierende Problem ähnelt wieder sehr stark den beiden vorangegangenen Fällen.

Entscheidend für die Verbindung ist die richtige Auswahl der Segmente. Diese sollten nur dann verbunden werden, wenn sie wirklich zu dem selben realen Objekt gehören. Walchshäusl [Wal08, S.38] fordert, dass die beiden benachbarten Linien an den Segmentenden kollinear sind und die Distanz zwischen den Endpunkten unter einem Grenzwert liegt. Die darin ebenfalls erwähnte Segmentverlängerung ist aufgrund des anderen Segmentierungsverfahrens hier nicht notwendig. Wie in Abbildung 4.5 veranschaulicht, müssen teilweise auch Einzelpunkte auf mögliche Segmentzugehörigkeit überprüft werden. Deshalb wird hier zunächst der senkrechte Abstand des Punkts zur verlängerten Gerade des vorherigen Segments als Kriterium verwendet. Bestehen beide Segmente aus mindestens zwei Punkten, wird ebenfalls die Kollinearität vorausgesetzt.

Neben diesen Kriterien, die rein geometrisch motiviert sind und auch in vielen Arbeiten verwendet werden, wird hier noch ein Kriterium aus der Funktionsweise des Lidarscanners abgeleitet. Wenn der Scanner während eines Messvorgangs von links nach rechts schwenkt, gibt es drei Möglichkeiten für den Übergang von einem Objekt auf das nächste. Im ersten Fall liegt das nächste Objekt näher

am Scanner, wodurch das erste Objekt möglicherweise verdeckt wird. Im zweiten Fall verhält es sich genau umgekehrt, wodurch das zweite Objekt möglicherweise verdeckt wird, das erste aber sicher an diesem Messpunkt endet. Der dritte Fall beschreibt einen Übergang zu einem nicht detektierten Punkt. Dadurch lässt sich keine eindeutige Aussage treffen, da das Objekt zu Ende sein kann oder aber einfach nicht detektiert wurde. Mit Hilfe dieser Unterscheidung kann nun noch überprüft werden, ob die zu verbindenden Segmente überhaupt zueinander passen können. Wird in der Lücke zwischen zwei Segmenten im Hintergrund noch ein Objekt erkannt, existiert diese Lücke mit hoher Wahrscheinlichkeit auch in der realen Welt und die Segmente sollten demnach nicht verbunden werden. Da der hier verwendete Lidar auf einer Höhe von 53 cm angebracht ist, sind die Momente in denen der Lidarstrahl tatsächlich durch ein Fahrzeug dringt, wie beispielsweise durch eine Scheibe, zu vernachlässigen.

4.3.1.2 Form extrahieren

Nach dem Verbinden der Segmente wird jetzt davon ausgegangen, dass jedes Segment ein reales Objekt repräsentiert. Das bedeutet, dass nun aus einer relativ kleinen Menge von Messpunkten charakteristische Merkmale gewonnen werden müssen. Wie bereits am Anfang dieses Kapitels erwähnt, werden dafür spezielle Formen in den Messdaten gesucht. Zunächst werden die I-, IS-, L-, C- und E-Form und ihre spezifischen Größen erläutert, bevor auf die Berechnung dieser Formen und Größen eingegangen wird.

Diese fünf Formen sind in Abbildung 4.6 dargestellt. Dabei ist der erkannte Umriss als schwarze Linie, die Orientierung als blaue Linie und der Referenzpunkt als schwarzes X eingezeichnet. Die I-Form repräsentiert in dieser Arbeit jeweils Fahrzeugfront oder Fahrzeugheck. Deshalb zeigt die Orientierung auch senkrecht zum detektierten Fahrzeugumriss und somit parallel zur erwarteten Bewegungsrichtung. Dazu ist noch hinzuzufügen, dass die Orientierung immer in Richtung der positiven x-Achse zeigt. In den meisten Fällen entspricht das der Richtung, in der der Rest des Objekts erwartet wird, da es ins sensorabgewandte Gebiet zeigt. Als Referenzpunkt wird der Mittelpunkt der Regressionsgeraden verwendet. Die Länge der Strecke wird entsprechend als Objektbreite aufgefasst. In der Literatur wird häufig noch zusätzlich zur I-Form eine U-Form eingeführt, um stärker abgerundete Fahrzeugfronten oder Hecks gut nachbilden zu können. Aufgrund der relativ geringen Winkelauflösung des Scanners kann dieser Fall aber vernachlässigt werden.

Da die IS-Form zur Beschreibung von Fahrzeugseiten verwendet wird, sind die Merkmale etwas anders angeordnet. Die Orientierung zeigt in Richtung der Fahrzeugseite um wie bei der I-Form parallel zur erwarteten Bewegungsrichtung zu liegen. Als Referenzpunkt wird nicht der Mittelpunkt sondern das Ende mit dem größeren x-Wert verwendet. Der Grund dafür ist die bessere Beobachtbarkeit dieses Punktes. Wird beispielsweise das Ego-Fahrzeug überholt, schiebt sich eine IS-Form langsam in das Sichtfeld des Scanners. Da das eine Ende des Objekts außerhalb des Sichtfelds liegt, bewegt sich der Mittelpunkt der Form nur halb so schnell wie das eigentliche Objekt. Der vorderste Punkt ist dagegen die ganze Zeit gut zu beobachten. Die Länge der Strecke entspricht in diesem Fall der Objektlänge.

Wird im Gegensatz zu den letzten beiden Formen nicht nur eine sondern zwei Objektseiten vom Lidarscanner erfasst, handelt es sich um eine L-Form. Als repräsentativer Punkt wird dabei der Eckpunkt verwendet. Die Orientierung zeigt in Richtung der längeren Seite, die auch als Objektlänge verwendet wird. Die kürzere Seite ist dementsprechend die Breite.

Bei der C-Form wird der Schwerpunkt als Referenzpunkt verwendet und der größte Abstand zwischen Schwerpunkt und einem Einzelpunkt definiert den Radius. Diese Form wird allerdings nur für Objekte mit einem maximalem Durchmesser von $r_{C-Form} = 0,9m$ verwendet.

Die speziell für Zweiradfahrer eingeführte Ellipsen-Form stellt hier eine Ausnahme dar. Im Gegensatz zu den anderen vier Formen wird diese nicht zur Objekterzeugung verwendet, sondern lediglich zur Objektverfolgung von als Radfahrer klassifizierten Objekten verwendet. Daran wird ein großer

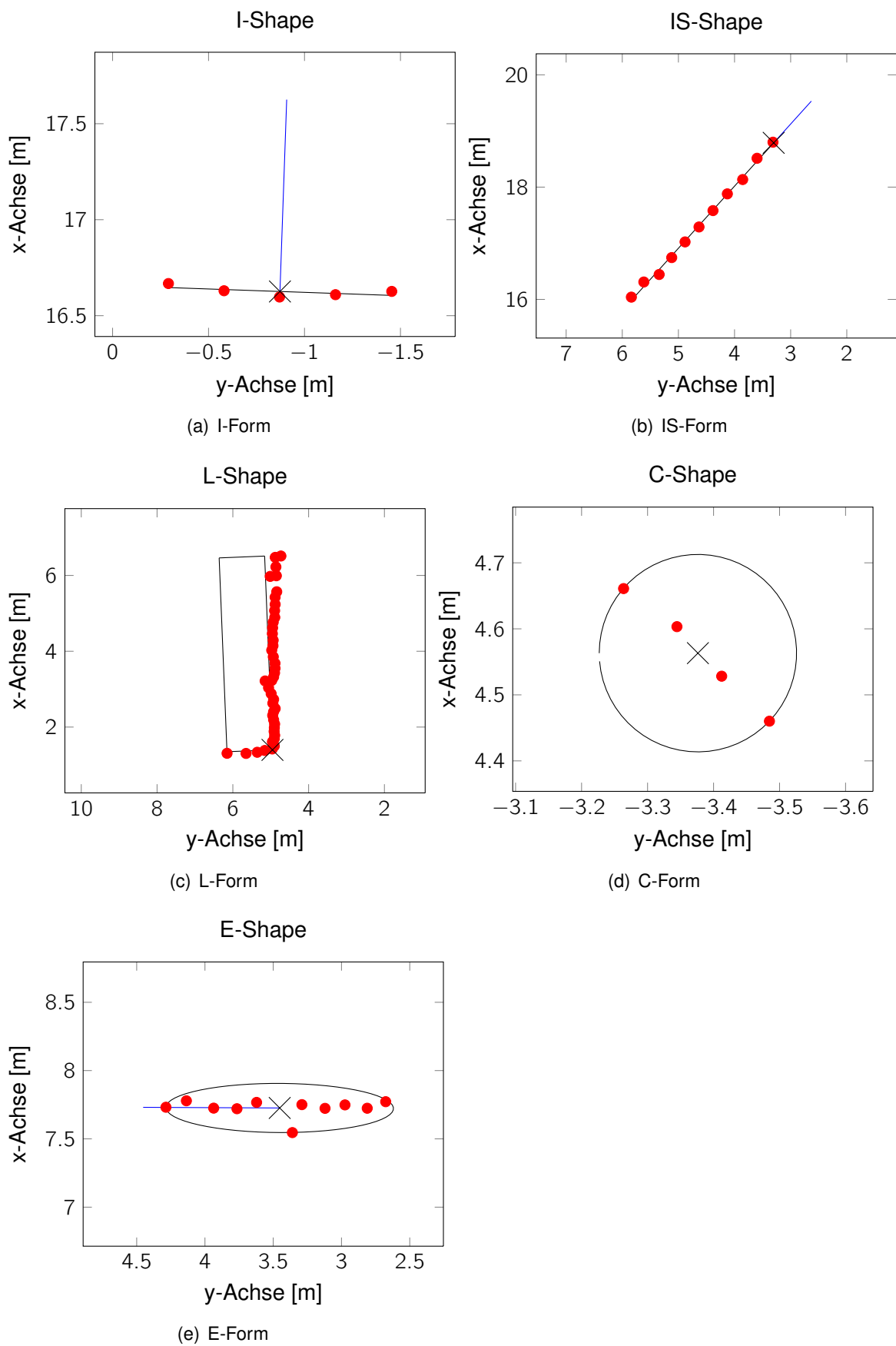


Abbildung 4.6: charakteristische Segmentformen

Vorteil des hier verwendeten und in Abschnitt 4.1 vorgestellten Verfahrens sichtbar. Aufgrund des objektorientierten Ansatzes kann das Vorwissen über das Objekt verwendet werden, um die Rohdaten adäquat zu interpretieren. Das genaue Vorgehen wird wie bei den anderen Formen in Abschnitt 4.4.1 beschrieben. Die E-Form ist für den konventionellen Ansatz und dementsprechend auch für die Erzeugung neuer Objekte ungeeignet, da sie gerade deshalb verwendet wird, um die undefinierte Erscheinung eines Fahrrads zu beschreiben. Dementsprechend gibt es auch keine definierten Formen nach denen in den Messdaten gesucht werden könnte. Die hier verwendeten Einschränkungen von $\tau_{E-Form,1} = 1m$ bzw. $\tau_{E-Form,2} = 0,25m$ für die beiden Halbachsen der Ellipse in der die Messwerte liegen dürfen, würden zu unzähligen Ellipsenobjekten führen. Auch wenn diese Form erst bei der Objektverfolgung benötigt wird, ist sie der Vollständigkeit halber hier bereits erwähnt. Als Referenzpunkt wird wie bei der C-Form der Schwerpunkt der Messwerte verwendet. Die Halbachsen werden über den maximalen Abstand der Messpunkte zu der jeweils anderen Achse bestimmt.

Zur Bestimmung der richtigen Form aus den Messpunkten wird zunächst die gesamte Ausdehnung des Segments betrachtet. Ist der maximale Abstand zweier Punkte kleiner als $0,9m$ wird das Segment als C-Form behandelt und dementsprechend Schwerpunkt und Radius berechnet (vgl. [KBD]). Ist das betrachtete Segment größer, wird wie in [Wal08] vorgegangen. Zunächst wird eine Regressionsgerade durch alle Punkte des Segments gelegt und der mittlere Abstand D aller Punkte zu dieser Gerade berechnet. Bei der Regression wird der quadratische mittlere Abstand minimiert. Bei n Messwerten berechnet sich der mittlere Abstand zu:

$$D = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - (a_r + b_r x_i)|}{\sqrt{1 + b_r^2}} \quad (4.3)$$

Aus den partiellen Ableitungen nach den Geradenparametern a_r und b_r der Geradengleichung $y = a_r + b_r x$ folgt:

$$B = \frac{1}{2} \frac{(\sum_{i=1}^n y_i^2 - n\bar{y}^2) - (\sum_{i=1}^n x_i^2 - n\bar{x}^2)}{n\bar{x}\bar{y} - \sum_{i=1}^n x_i y_i} \quad (4.4)$$

Darin sind \bar{x} und \bar{y} jeweils die Mittelwerte aus allen x bzw. y -Werten dieses Segments. Aus diesem B kann mit $b_r = -B \pm \sqrt{B^2 + 1}$ die Steigung der Geraden und dann mit $a_r = \bar{y} - b_r \bar{x}$ der Achsenabschnitt bestimmt werden (aus [Wal08]). Da es sich um eine quadratische Gleichung handelt, ergeben sich zwei Lösungen deren Geraden senkrecht aufeinander stehen. Dementsprechend muss dann noch die richtige Lösung beispielsweise über eine grobe Abschätzung mit den beiden äußersten Punkten ausgewählt werden. Ist der durchschnittliche Abstand kleiner als ein empirisch festgelegter Grenzwert $\tau_{I-Form} = 0.09m$ wird das Segment als I-Form behandelt. Die Orientierung ergibt sich dann aus der um 90° gedrehten Geradensteigung. Für die Berechnung des Referenzpunktes werden die beiden Endpunkte auf die Regressionsgerade projiziert und davon der Mittelpunkt berechnet.

Kann dieses Segment nicht mit einer Gerade angenähert werden, wird überprüft ob es als L-Form darstellbar ist. Dafür wird zunächst der Eckpunkt ermittelt. Prinzipiell könnte dafür der Punkt mit dem geringsten Abstand zum Sensor verwendet werden, laut Walchshäusl [Wal08] ist das Verfahren aber robuster, wenn stattdessen der Punkt mit dem maximalen Abstand zu beiden Endpunkten des Segments verwendet wird. Anschließend werden zwei Regressionsgeraden jeweils inklusive des Eckpunkts berechnet. Liegen die beiden mittleren Abstände unter $\tau_{L-Form} = 0.18m$ wird der Winkel zwischen diesen beiden Geraden überprüft. Wenn der Winkel um weniger als $\tau_{L-Form,\psi} = 15^\circ$ von 90° abweicht, wird die L-Form verwendet. Der Referenzpunkt ergibt sich durch die konsekutive Projektion des Eckpunkts auf beide Regressionsgeraden. Als Orientierung könnte einfach die Steigung der längeren Seite verwendet werden. Aber auch hier empfiehlt Walchshäusl [Wal08] ein robusteres Verfahren um den Einfluss von einzelnen Ausreißern zu reduzieren. Deshalb wird aus beiden Geraden die Winkelhalbierende ermittelt und um 45° in Richtung der längeren Seite gedreht. Die Längere der beiden Seiten wird als Objektlänge und die kürzere als Objektbreite gespeichert.

Tabelle 4.1: erwartete Abmessungen der Verkehrsteilnehmer

Verkehrsteilnehmer	Länge	Breite	Formen
PKW, Van	2 – 6m	0,9 – 2m	I,IS,L
LKW, Bus	6 – 19m	0,9 – 3m	I,IS,L
Fußgänger	Durchmesser	0,25 – 0,9m	C

Mit diesem Vorgehen können die in Abbildung 4.6 dargestellten Formen in den Messwerten identifiziert und berechnet werden. Erfüllen die Messpunkte eines Segments nicht die Anforderungen für diese Formen, handelt es sich höchstwahrscheinlich um Objekte am Fahrbahnrand wie zum Beispiel Hecken oder Häuserfronten die auch nicht weiter verfolgt werden müssen. Zur Identifizierung und Verfolgung von Verkehrsteilnehmern sind diese vier Formen ausreichend.

4.3.2 Hypothesenerstellung

Damit ein Objekt als möglicher Verkehrsteilnehmer in Frage kommt, muss nicht nur der erkannte Umriss einer der vier Formen entsprechen, sondern auch die Abmessungen der Form mit den erwarteten Werten übereinstimmen (vgl. [Wal08]). Wird beispielsweise eine L-Form mit 4m Breite um 5m Länge erkannt, handelt es sich eher um ein Hauseck als um einen Verkehrsteilnehmer. Deshalb wird anhand der Dimensionen entschieden, ob aus der Form eine Hypothese erzeugt wird. Dabei wird mit Hypothese ein möglicher Verkehrsteilnehmer bezeichnet, der im nächsten Zeitschritt bestätigt werden muss bevor er zu einem Objekt wird.

Damit aus einer Form eine Hypothese erzeugt wird, müssen die gemessenen Abmessungen innerhalb den erwarteten Abmessungen von Verkehrsteilnehmern liegen, die in Tabelle 4.1 zusammengefasst sind. Die darin definierten oberen Grenzen entsprechen gängigen Fahrzeugabmessungen. Da nur selten die komplette Länge und Breite der Objekte erfasst werden kann, sind die unteren Grenzen relativ niedrig angesetzt. Wird in einem Segment eine L-Form erkannt, müssen sowohl Objektlänge als auch Objektbreite in den entsprechenden Intervallen liegen. Dem aufmerksamen Leser wird aufgefallen sein, dass bei der Ermittlung der Segmentform lediglich die I-Form, nicht jedoch die IS-Form behandelt wurde. Die Unterscheidung ob es sich bei einer geraden Linie um ein Ende oder eine Seite handelt, kann durch die gemessene Länge bereits abgeschätzt werden. Ist diese kürzer als 3m wird sie als I-Form und somit als Heck oder Front behandelt. Bei einer Länge zwischen 3 und 6m handelt es sich um eine Fahrzeugseite.

Dabei handelt es sich nur um eine erste Abschätzung, da im Bereich zwischen 2 und 3 Metern beides Möglich ist. Die endgültige Entscheidung kann dann erst bei der Objektverfolgung getroffen werden, wenn auch die Geschwindigkeitsrichtung bekannt ist. Die untere Schwelle für Fußgänger sorgt für den Ausschluss von Leitpfosten und anderen Pfählen.

In dem konventionellen Algorithmus, wie in Abbildung 4.2(a) dargestellt, werden die Merkmale wie Position, Orientierung und Abmessungen zum Verfolgen der Objekte verwendet. Dafür werden sie zunächst mit den alten Objekten assoziiert und diese dann entsprechend auf den neuen Stand gebracht. Hier wird das in den letzten Kapiteln beschriebene Vorgehen nur zur Erzeugung neuer Objekte genutzt. Bevor aus diesen Hypothesen neue Objekte erzeugt werden, müssen diese im nächsten Zeitschritt noch einmal auftauchen. Dieses Vorgehen wurde gewählt, da vor allem in den Bereichen neben der Straße, also beispielsweise in Grünstreifen alle mögliche Formen vorkommen und somit häufig falsche Hypothesen erstellt werden. Da sich die Messwerte in diesen Bereichen ohne definierte Form relativ schnell ändern, je nachdem auf was die Lidarstrahlen treffen, wird die Erzeugung falscher Objekte durch die zweifache Überprüfung stark reduziert. Das ist bei diesem Algorithmus besonders wichtig, da er darauf ausgelegt ist, Objekte auch mit wenigen passenden Messpunkten verfolgen zu können und nicht jedes mal eine der bekannten Formen verlangt.

4.3.3 Koordinatentransformation in das aktuelle Sensorkoordinatensystem

Um die erzeugten Hypothesen mit den Messwerten im nächsten Zeitschritt vergleichen zu können, müssen die Werte gespeichert und in das neue Koordinatensystem transformiert werden. Da die Messwerte ohnehin im Koordinatensystem des Lidarscanners ausgegeben werden, bietet es sich an, die Objektverfolgung ebenfalls in diesem Koordinatensystem durchzuführen. Allerdings befindet sich der Sensor an dem fahrenden Versuchsfahrzeug. Das bedeutet, dass das Koordinatensystem der nächsten Messwerte mit dem Fahrzeug mitgefahren ist. Um nun die Koordinatentransformation durchführen zu können, sind möglichst genaue Informationen über die Bewegungen des Ego-Fahrzeugs nötig. Die wichtigste Information ist dabei die Geschwindigkeit, da sie in den meisten Situationen den größten Einfluss hat. Bei Kurvenfahrten ist aber auch die Gierrate und die damit verbundenen Orientierungsänderung von entscheidender Bedeutung. Aufgrund der hohen Abtastrate des Lidarsensors von 75 Hz sind die Zeitunterschiede zwischen den Aufnahmepunkten so klein, dass die Fahrzeugbewegung stark vereinfacht betrachtet werden kann. Die wohl einfachste Transformation, die sowohl Geschwindigkeit und Orientierungsänderung berücksichtigt, besteht aus einer Drehung und einer Verschiebung entlang der neuen Fahrtrichtung des Versuchsfahrzeugs. Die Drehung lässt sich mit Hilfe der Drehmatrix

$$\mathbf{T}_{t,t-1} = \begin{bmatrix} \cos(\Delta\psi) & \sin(\Delta\psi) \\ -\sin(\Delta\psi) & \cos(\Delta\psi) \end{bmatrix} \quad (4.5)$$

darstellen. Dabei ist $\Delta\psi$ die Orientierungsänderung zwischen den Zeitpunkten $t-1$ und t . Zur Kompensation der Eigengeschwindigkeit wird das System anschließend noch entlang der neuen X-Achse mit

$$\Delta x = \Delta t v_{ego} \quad (4.6)$$

verschoben. Dabei ist v_{ego} die Eigengeschwindigkeit des Fahrzeugs zum Zeitpunkt t und Δt der Zeitunterschied zwischen zwei konsekutiven Scans und dementsprechend etwa 13,3ms. Für die Transformation von Positionswerten ergibt sich dann die Gleichung

$$\mathbf{x}_t = \mathbf{T}_{t,t-1} \mathbf{x}_{t-1} + \begin{bmatrix} \Delta x \\ 0 \end{bmatrix}, \text{ mit } \mathbf{x}_t = \begin{bmatrix} x \\ y \end{bmatrix}_t \quad (4.7)$$

Zusätzlich zur Position der Hypothesen muss auch noch die Orientierung über

$$\psi_t = \psi_{t-1} - \Delta\psi \quad (4.8)$$

in das neue Koordinatensystem transformiert werden. Zur Überprüfung der Transformation wurden die Lidarmesspunkte bei einer Kurvenfahrt über mehrere Zeitschritte gespeichert und jeweils transformiert und dann mit den aktuellen Messwerten verglichen. Mit einer perfekten Transformation und exakten Messwerten der Fahrzeugbewegung und des Lidarscanners müssten statische Objekte dann aufeinander abgebildet werden. In Abbildung 4.7(a) sind die Lidarmesspunkte eines Hausecks am Ende einer Kurve dargestellt. Neben der aktuellen Messung sind auch die Messungen von vor 20 bzw. 40 Messschritten abgebildet. Diese wurden in jedem Zeitschritt jeweils in das nächste Koordinatensystem transformiert. Wie aus der Abbildung ersichtlich, liegen die Werte der Ecke nicht genau übereinander sondern im Laufe der 20 bzw. 40 Transformationen hat sich ein Fehler aufgebaut. Da so wohl die Messungen der Kreiselpattform als auch die Messwerte des Lidar nur geringe Fehler aufweisen, wird diese Abweichung durch die Transformationsvorschrift verursacht. Die Ursache davon ist der Verbauort des Lidarscanners an der Fahrzeugfront. Durch den Abstand zwischen Drehpunkt und Scanner d_{LIDAR} bewirkt eine Drehung des Fahrzeugs auch eine Verschiebung des Scanners, die zusätzlich berücksichtigt werden muss. Da als Drehpunkt die Mitte der Hinterachse angenommen wird ergibt sich bei dem Versuchsfahrzeug ein Abstand von ca. 4m. Damit ergibt sich

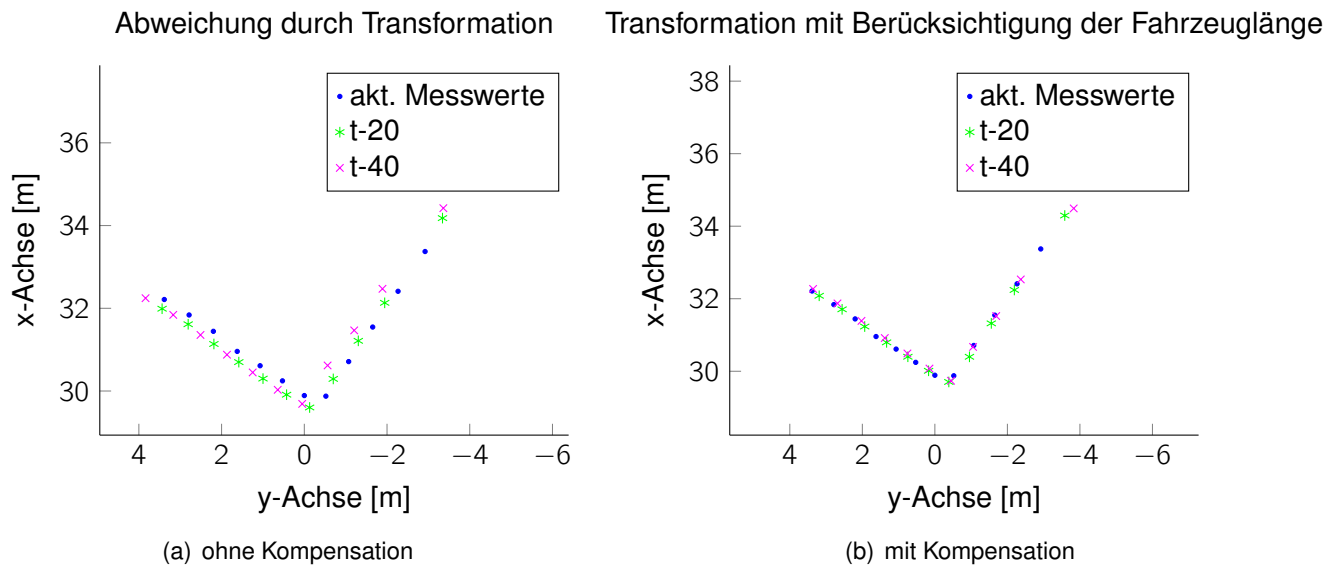


Abbildung 4.7: Überprüfung der Koordinatentransformation

für die Transformation

$$\mathbf{x}_t = T_{t,t-1} \mathbf{x}_{t-1} + \begin{bmatrix} \Delta x \\ 0 \end{bmatrix} + \begin{bmatrix} 1 - \cos(\Delta\psi) \\ \sin(\Delta\psi) \end{bmatrix} d_{LIDAR} \quad (4.9)$$

Werden die selben Messwerte mit der verbesserten Transformation bearbeitet, tritt selbst bei der 40-fachen Transformation ein vernachlässigbarer Fehler auf, wie in Abbildung 4.7(b) zu sehen.

Allerdings liefert diese Transformation nur solange gute Ergebnisse, wie sich das Versuchsfahrzeug hauptsächlich in Fahrtrichtung bewegt. Für Grenzsituationen in denen das Fahrzeug stark seitlich rutscht, müsste die laterale Geschwindigkeit zusätzlich berücksichtigt werden. Da hier zunächst der normale Betriebsfall betrachtet wird und für Grenzsituationen auch keine Messfahrten zur Evaluierung vorliegen, wird auf eine aufwendigere Transformation verzichtet.

Wie in Abbildung 4.2 dargestellt, muss diese Transformation nicht nur auf die Hypothesen sondern auch auf die Objekte angewendet werden um diese in das aktuelle Koordinatensystem zu transformieren. Der Vollständigkeit halber wird deshalb hier auch kurz darauf eingegangen. Da bei Objekten neben der Position und Orientierung auch die Geschwindigkeiten zur Verfügung stehen, müssen diese Werte ebenfalls transformiert werden. Die Geschwindigkeiten müssen lediglich in das neue Koordinatensystem gedreht werden:

$$\mathbf{v}_t = T_{t,t-1} \mathbf{v}_{t-1}, \text{ mit } \mathbf{v}_t = \begin{bmatrix} v_x \\ v_y \end{bmatrix} \quad (4.10)$$

4.3.4 Assoziation

Ein wichtiger Schritt für die kontinuierliche Verfolgung von Objekten ist die Assoziation zwischen Objekten des letzten Zeitschritts und aktuellen Messwerten. In dieser Arbeit wird dieser Schritt sowohl beim Verfolgen der Objekte benötigt als auch bei der Erzeugung neuer Objekte, da die Hypothesen erst mit aktuellen Messwerten assoziiert werden müssen bevor ein neues Objekt erstellt werden kann. Bei den meisten Trackingverfahren und auch in den meisten Arbeiten über die Verfolgung in Lidarscannern werden zunächst Merkmale extrahiert und diese mit den alten Objekten assoziiert. Das wohl bekannteste Assoziationsverfahren verwendet den Abstand zwischen den alten Objekten und den Merkmalen, meistens dem Schwerpunkt, und assoziiert jeweils die nächsten Paare. Dieses

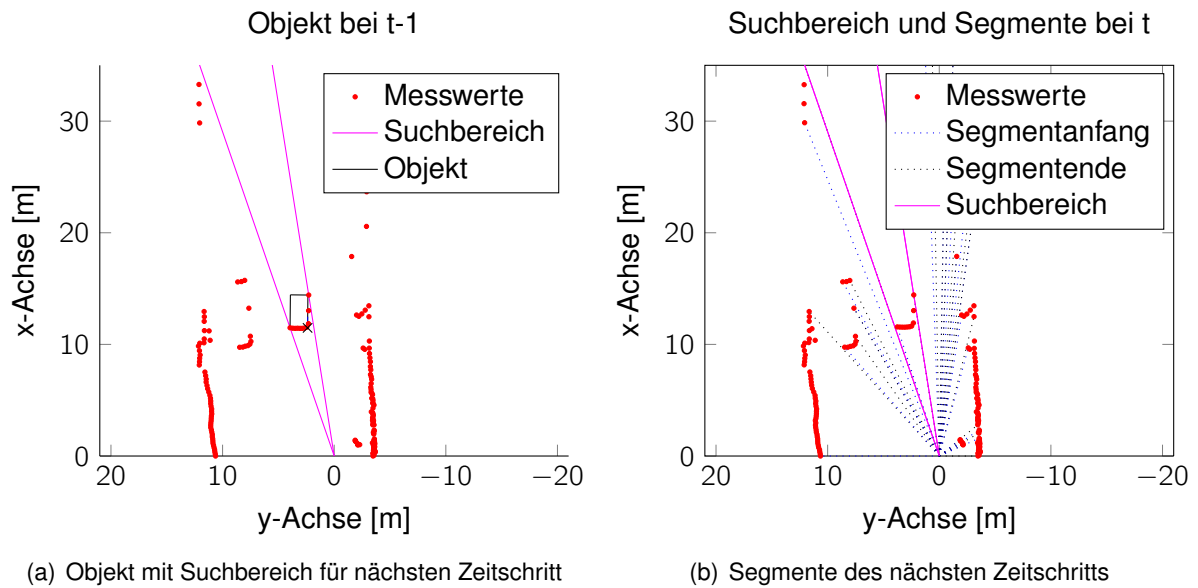


Abbildung 4.8: Messwertassoziation

vorgehen ist z.B. in [FLD] gut beschrieben. Darin wird zunächst auch mit Hilfe einer Objektumgebung eine erste Voreinteilung vorgenommen und auf die Problematik von wechselnden Referenzpunkten eingegangen.

In dieser Arbeit werden für die Objektverfolgung die Rohpunkte herangezogen, wie bereits in Abschnitt 4.1 erläutert. Deshalb stehen für die Assoziation auch keine gesonderten Merkmale sondern nur die in Segmente unterteilten Rohdaten zur Verfügung. Die erste Vorauswahl wird deshalb aus der Funktionsweise des Lidarscanners abgeleitet. Dafür werden bei jedem Objekt die Lidarwinkel bzw. die Indizes der korrespondierenden Messwerte gespeichert. Im nächsten Zeitschritt werden dann zunächst die Segmente gesucht, die diesen Bereich abdecken. Da sich sowohl das Objekt als auch der Sensor etwas bewegen kann, wird dieser Suchbereich je nach Entfernung des Objekts noch etwas vergrößert. Zur Veranschaulichung ist das Vorgehen in Abbildung 4.8 dargestellt. Im linken Bild ist ein Objekt und der daraus resultierende Suchbereich abgebildet. In der rechten Abbildung sind die Messwerte und Segmente des nächsten Zeitschritts mit dem Suchbereich des alten Objekts abgebildet. Nun werden alle Segmente ausgewählt, die diesen Bereich zumindest teilweise abdecken. In diesem Fall liegen in dem Suchbereich zwei Segmente. Eines davon repräsentiert das Objekt, das andere liegt deutlich im Hintergrund.

Durch diese Vorauswahl wird die Anzahl möglicher Segmente schon stark reduziert. In diesem Fall kommen von ursprünglich 19 erkannten Segmenten nur noch zwei in Frage. Als zweiter Schritt wird nun noch der Abstand des Segmentes zur erwarteten Objektkante herangezogen. Dazu wird der Abstand des Segmentendes, das im Suchbereich liegt, zu der erkannten Form berechnet. Ist der Abstand kleiner als $\tau_{association} = 0.5m$ wird das Segment zu der Hypothese oder dem Objekt assoziiert. Dieser Grenzwert berücksichtigt sowohl eine mögliche Geschwindigkeit des Objekts als auch mögliche Messungenauigkeiten. Aufgrund der geringen Zeitspanne zwischen zwei Messungen sind somit neue Objekte mit Geschwindigkeiten bis mindestens $135 \frac{km}{h}$ erfassbar und dennoch können falsche Segmentzuordnungen weitgehend ausgeschlossen werden.

4.3.5 Erzeugung neuer Objekte

Damit aus den Hypothesen neue Objekte erstellt werden, müssen die Informationen des aktuellen Zeitschritts zu denen der Hypothese aus dem letzten Zeitschritt passen. Dieser Schritt dient da-

zu, möglichst wenige falsche Objekte zu erzeugen. Dafür wird zunächst gefordert, dass jeder alten Hypothese genau ein Segment assoziiert wurde. Erfüllen mehrere Segmente die Assoziationsbedingungen aus dem vorherigen Abschnitt, handelt es sich wahrscheinlich nicht um ein freistehendes Objekt sondern eher um eine zufällige Objektform in einer Umgebung ohne definierte Formen, wie sie beispielsweise in Grünstreifen auftreten.

Besitzt eine Hypothese genau ein assoziiertes Segment, werden dafür anschließend die selben Bearbeitungsschritte wie für die Hypothesenerzeugung durchgeführt. Zunächst wird die Form ermittelt (siehe Abschnitt 4.3.1 und anschließend die ermittelten Abmessungen auf Plausibilität überprüft (siehe Abschnitt 4.3.2). Dementsprechend muss auch im aktuellen Zeitschritt eine bekannte Form mit verkehrsteilnehmertypischen Abmessungen detektiert werden. Abschließend wird noch überprüft, ob die Informationen der beiden Hypothesen plausibel sind, d.h. ein reales Objekt darstellen können. Dafür müssen eine Reihe von Bedingungen erfüllt werden. Zunächst müssen beide Hypothesen die selbe Form aufweisen und die ermittelten Orientierungen dürfen maximal um $\tau_{seg,\psi} = 10^\circ$ voneinander abweichen. Außerdem dürfen die Enden der Hypothesen nicht aufgrund von Verdeckung entstanden sein, da solche Hintergrundobjekte vom Lidarscanner nicht vollständig erfasst werden können und somit keine ausreichenden Aussagen über die Größe und Form der Objekte getroffen werden können. Abschließend wird noch untersucht, ob die Richtung der Bewegung zwischen den letzten beiden Zeitschritten zur Orientierung der Form passt. Dies wird allerdings nur bei Hypothesen durchgeführt, deren Position sich in diesem Zeitschritt um mehr als $0.04m$ ändert. Das entspricht einer Geschwindigkeit von $\tau_{seg,vel} = 3\frac{m}{s}$. Die Winkeldifferenz zwischen Bewegungsrichtung und Orientierung darf dann $\tau_{seg,vel,\psi} = 15^\circ$ nicht überschreiten.

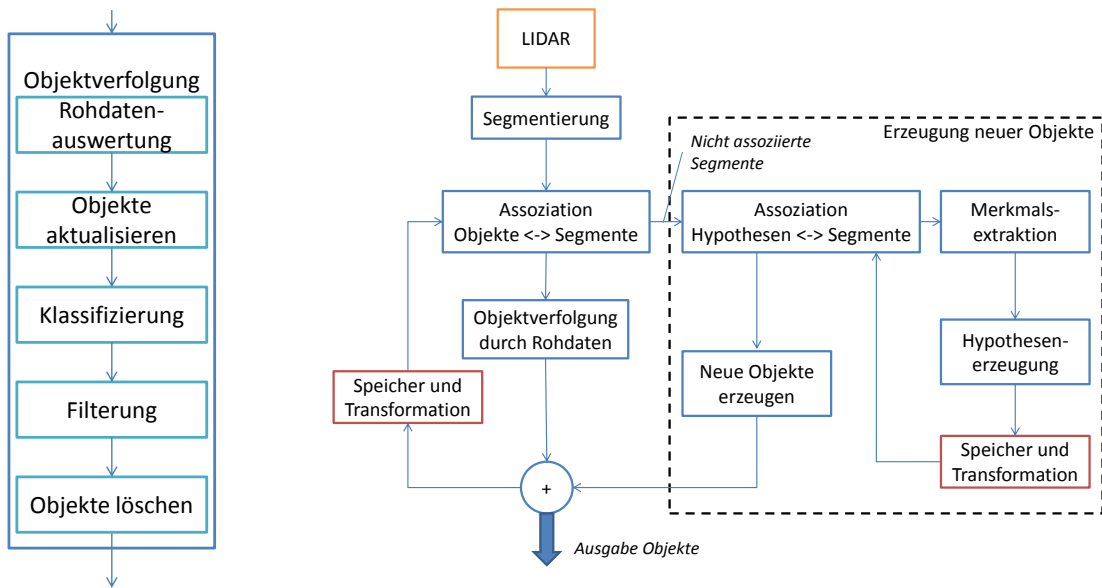
Durch diese engen Bestimmungen werden nur solche Objekte erzeugt, die in zwei aufeinander folgenden Messungen gute und zueinander passende Detektionen liefern. Die Wahl der Grenzwerte unterliegt auch hier wieder einem Zielkonflikt. Je enger die Werte gewählt werden, desto weniger falsche Objekte werden detektiert. Jedoch werden dadurch auch reale Objekte schlechter bzw. erst in geringerer Entfernung erkannt. Die benutzten Werte wurden empirisch anhand aufgezeichneter Messfahrten ermittelt.

4.4 Objektverfolgung

Wie in Abschnitt 4.1 erläutert, werden in der Literatur Objekte gewöhnlich durch eine Merkmalsextraktion in jedem Zeitschritt verfolgt. Um Objekte auch bei schlechter oder teilweiser Erkennung durch den Lidarscanner verfolgen zu können, werden in dieser Arbeit die alten Objekte direkt mit den Rohdaten assoziiert. Dadurch können die Messwerte mit Hilfe des Vorwissens über die Objektform und dessen möglicher Bewegung verarbeitet werden. Wie aus den Messwerten die aktuelle Position und Orientierung des Objekts gewonnen wird, wird in Abschnitt 4.4.1 beschrieben. Zur Verfolgung von Objekten sind neben dieser Rohdatenauswertung noch weitere Schritte notwendig, die in Abbildung 4.9 dargestellt sind und in den folgenden Abschnitten beschrieben werden.

4.4.1 Rohdatenauswertung

Zu Beginn dieser Auswertung stehen die Objekte des letzten Zeitschritts und die dazu assoziierten Segmente (siehe Abschnitt 4.3.4) zur Verfügung. Dabei werden auch mögliche Verbindungen zu den benachbarten Segmenten (siehe Abschnitt 4.3.1.1) geprüft. Die Lage eines Objekts wird durch die erkannte Form (I, IS, U, C oder E) und den dazugehörigen Werten des Referenzpunkts, der Orientierung und den Dimensionen eindeutig beschrieben (siehe Abschnitt 4.3.1). Zusätzlich zu den erkannten Abmessungen des letzten Zeitschrittes können aus früheren Messungen auch die tatsächlichen Abmessungen bekannt sein. Wurde ein Fahrzeug beispielsweise zunächst als L-Form erkannt, steht die Objektlänge auch noch zur Verfügung wenn das Objekt aktuell nur noch als I-Form erkannt wird.



(a) Schritte der Objektverfolgung

(b) gesamter Algorithmus

Abbildung 4.9: Übersicht über Trackingalgorithmus

Wie in Kapitel 3 beschrieben, werden durch die Verfolgung Zustandsgrößen ermittelt, die die Bewegung der Objekte beschreiben. Die verwendeten Bewegungsmodelle werden in Abschnitt 4.4.4 noch genau beschrieben. An dieser Stelle ist es ausreichend, dass die Bewegung des Objekts beschrieben wird und somit auch vorhergesagt werden kann. Dieser Prädiktionsschritt ist fester Bestandteil des Kalmanfilters. Bei gewöhnlichen Trackingverfahren wird nun eine Messung der Zustandsgrößen verwendet, um diese Prädiktion anzupassen und zu korrigieren. Die hier verwendete Rohdatenauswertung ermittelt dagegen den Unterschied zwischen dem prädizierten Objekt und den gemessenen Rohdaten um den Kalmanfilter zu korrigieren.

Das Vorgehen bei der Auswertung wird anhand von Abbildung 4.10 erläutert. Zur Veranschaulichung ist in Abbildung (a) das entsprechende Kamerabild der Situation dargestellt. In der Darstellung (b) ist der Frontbereich des entgegenkommenden, weißen Lieferwagens aus der Vogelperspektive abgebildet. Die erfassten Messpunkte des Lidarscanners im aktuellen Zeitschritt sind wieder als rote Punkte dargestellt. Zusätzlich ist der prädizierte Objektumriss eingezeichnet. Zur Verdeutlichung wurde dieses Objekt erst im letzten Zeitschritt initialisiert, wodurch noch keine Geschwindigkeit erfasst wurde und damit die prädizierte Position der Position des letzten Zeitschritts entspricht. Wie zu erkennen ist, hat sich das Fahrzeug inzwischen einige Zentimeter weiter bewegt. Um die aktuelle Position zu ermitteln, wird nun die Positionsdifferenz zwischen prädiziertem Objekt und Messwerten ermittelt.

4.4.1.1 Positionsdifferenz

Das jeweilige Vorgehen variiert etwas mit der erkannten Form des Objekts. Prinzipiell wird jedoch immer der Abstandsvektor jedes einzelnen Messpunkts zur erwarteten Form berechnet. Diese Vektoren sind ebenfalls in Abbildung 4.10(b) dargestellt.

Handelt es sich, wie in diesem Fall, um eine L-Form, so wird zunächst der Eckpunkt der Rohpunkte ermittelt. Dabei wird analog zu dem Vorgehen in Abschnitt 4.3.1 der Punkt mit dem maximalen Abstand zu beiden Endpunkten des Segments gewählt. Damit ist jeder Punkt einer der beiden erkannten Seiten zugeordnet. Für jede dieser Seiten wird dann der durchschnittliche Differenzvektor bestimmt, der die Bewegung senkrecht zu der Seite beschreibt. Da bei einer L-Form zwei Seiten

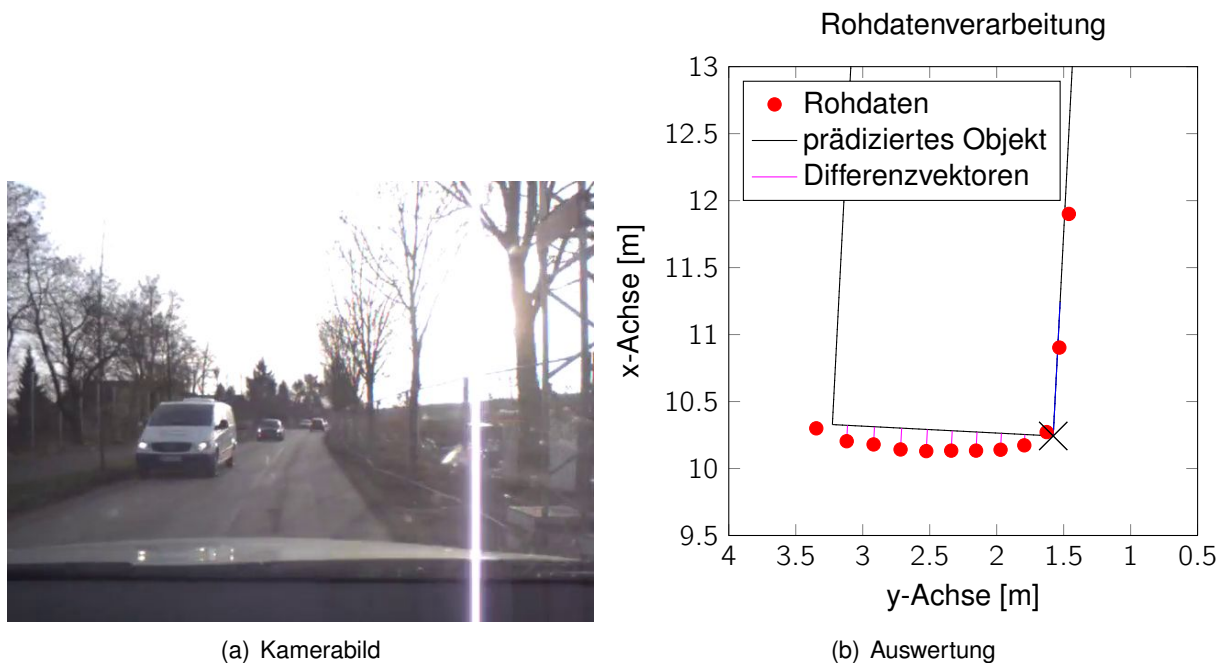


Abbildung 4.10: Rohdatenauswertung

erfasst werden, kann die Positionsänderung damit bereits beschrieben werden. Für die Berechnung der aktuellen Orientierung werden wie in Abschnitt 4.3.1 Regressionsgeraden durch die Messpunkte beider Seiten verwendet. Da neben der Position und der Orientierung auch die Abmessungen der Objekte erfasst werden sollen, wird bei der Berechnung der Differenzvektoren auch die Lage des Lotfußpunktes relativ zur Seite ausgegeben. Da die Differenzvektoren dem jeweiligen Lot von dem Messpunkt auf die verlängerte Seite entsprechen, wird der Fußpunkt ohnehin berechnet. Aus dem Abstand zwischen den beiden äußersten Lotfußpunkten wird die Ausdehnung des Objekts in Richtung dieser Seite bestimmt.

Wird ein Objekt dagegen nur als I-Form oder IS-Form erkannt, muss die Auswertung etwas angepasst werden. Die Berechnung der Positionsänderung senkrecht zur Seite, der Objektbreite bzw. Länge und Orientierung ist analog zur L-Form. Bewegungen in Seitenrichtung werden über die äußersten Lotfußpunkte detektiert. Verschieben sich beide in dieselbe Richtung, kann daraus die Längsbewegung bestimmt werden. Wie eingangs erwähnt, wird diese Form der Bewegung eigentlich nur von IS-Formen also Fahrzeugseiten erwartet. Deshalb wird in diesem Fall bei I-Formen (Fahrzeugfront/Heck) überprüft, ob es sich nicht doch eher um eine Fahrzeugseite handelt.

Für kleine Objekte in Zylinderrepräsentation ist die Auswertung denkbar einfach, da lediglich die Ausdehnung des assoziierten Segments überprüft werden muss. Ist die Ausdehnung der aktuelle Messung wieder kleiner als $0,9m$ (Tabelle 4.1) wird der neue Schwerpunkt und der neue Radius berechnet.

Bei den für Radfahrer eingesetzten E-Formen wird ebenfalls zunächst der neue Schwerpunkt und die Halbachsen der Ellipse bestimmt. Die Schwierigkeit ergibt sich bei der Berechnung der richtigen Orientierung. Im Idealfall wird die Längsachse des Zweirads durch zahlreiche Messpunkte dargestellt, wie es beispielsweise in Abbildung 4.6(e) der Fall ist. Dann wird zur Berechnung der Orientierung wie bei den I-, IS- und L-Formen eine Regressionsgerade verwendet. Voraussetzung dafür ist jedoch, dass die äußersten beiden Punkte mindestens $1m$ voneinander entfernt sind. Andernfalls kann die Orientierung aufgrund der versetzten Messwerte der Beine stark verzerrt sein. In diesen Fällen kann die Orientierung nicht sinnvoll aus den Messwerten abgeleitet werden.

4.4.1.2 Formänderung

Allerdings können Objekte im Laufe der Verfolgung ihr repräsentative Form ändern. Wird beispielsweise ein vorausfahrendes Fahrzeug überholt, ändert sich dessen Form von einem I zu einem L und schließlich zu einem IS. Um solche Formänderungen zu detektieren und somit die gelieferten Messpunkte optimal interpretieren zu können, wird die Lage der einzelnen Messpunkte relativ zu den Formen anhand von Bereichen beurteilt. Dabei wird zwischen drei unterschiedlichen Bereichen unterschieden. Der erste Bereich umfasst die nahe Umgebung der Form und beinhaltet zur Form passende Messpunkte. Messpunkte im zweiten Bereich zeigen einen möglichen Formwechsel an. Der letzte Bereich umfasst den ganzen restlichen Messbereich und die darin enthaltenen Messpunkte passen nicht zum betrachteten Objekt. Die Definition dieser Bereiche stellt wieder einen Kompromiss zwischen dem Ausschluss falscher Formwechsel und der Erkennung sinnvoller Formwechsel da. Die hier verwendeten Bereiche sind für die jeweiligen Formen in Abbildung 4.11 dargestellt.

Darin sind jeweils die jeweilige Form, deren Bereiche und passende Messpunkte dargestellt. Der passende Bereich der I-, IS- und L-Form wird über einen maximalen senkrechten Abstand zur verlängerten Seite von $\tau_{dist,\perp} = 0,5m$ und einen maximalen, absoluten Abstand von den Endpunkten der Seiten $\tau_{dist,tot} = 1,5m$ begrenzt. Allerdings ist der Grenzwert für die Längsseite der L-Form von der Objektlänge abhängig. Die maximale Länge ist hier für Fahrzeuge auf $\tau_{max,length} = 6m$ festgelegt. Damit werden nur Segmente verwendet, die maximal $6m$ von dem Eckpunkt entfernt liegen. Für größere Objekte wie LKWs und Busse wird dieser Grenzwert auf $\tau_{max,length,lkw} = 19m$ erhöht. Außerdem wird der senkrechte Abstand dynamisch mit der normalisierten, quadratischen Innovation (vgl. Gleichung (3.17) des Kalmanfilters) angepasst. Sobald der Filter die Bewegung des Objekts adaptiert hat, sind die prädizierten Werte näher an den gemessenen und der Bereich kann eingeschränkt werden. Dadurch ergeben sich für den Abstand $\tau_{dist,\perp}$ durch Interpolation dann Werte zwischen $0,25m - 0,5m$. Allerdings macht diese Reduzierung des Gebiets nur in der näheren Umgebung des Sensors Sinn. Mit wachsender Entfernungen steigt auch der Abstand zwischen den einzelnen Messpunkten so dass diese deutlich versetzte Objektbereiche erfassen können und somit eine größere Toleranz nötig ist.

Wie in den Abbildungen 4.11 (a) und (b) dargestellt, gibt es bei den I- und IS-Formen Bereiche, die auf einen Formwechsel hindeuten. Dabei wird zunächst in beiden Fällen auf einen Wechsel zu einer erkannten L-Form ausgegangen. Im Idealfall würde die bis dahin unerkannte Seite der L-Form senkrecht am Objektende anschließen. Dabei wäre die Ecke der nächste Punkt zum Sensor. Da aufgrund von Messungenauigkeiten oder einer abgerundeten Fahrzeugecke, weder das alte Objekt, noch die neuen Messpunkte diesem Idealfall entsprechen müssen, sind die möglichen Bereiche für einen solchen Formwechsel durch eine Reihe von Bedingungen definiert. Zunächst wird überprüft, an welchem Objektende die neue Seite anschließen kann. Dafür kommt prinzipiell nur das Ende mit der geringeren Distanz zum Scanner in Frage. Dabei wird jedoch ein Viertel Objektausdehnung als Toleranz verwendet, wodurch auch beide Enden in Betracht gezogen werden können (siehe Abbildung (a)). Dann muss der Lotfußpunkt des Messwerts im Randbereich der Form liegen. Die hier verwendeten Grenzwerte schränken den Bereich auf die letzten 30% der Form bis zur doppelten Formlänge ein. Außerdem muss der Winkel zwischen dem alten Objekt und dem Messwert mehr als 45° betragen und der maximale, absolute Abstand darf $\tau_{\Delta Form} = 4m$ nicht überschreiten. Dieser Grenzwert wurde im Hinblick auf die I-Form gewählt, bei der möglicherweise die Fahrzeuglänge zusätzlich erkannt werden muss. Da bei einem solchen Übergang für gewöhnlich nicht sofort die komplette Länge erkannt wird, ist der Wert kleiner als die sonst verwendeten $6m$. Bei der IS-Form wird dagegen nur die zusätzliche Erkennung der Fahrzeugbreite erwartet, weshalb nur der halbe Grenzwert ($2m$) verwendet wird.

Bei der L-Form verhält sich der Formwechsel etwas anders. Da hier zunächst beide Seiten erwartet werden, wird ein solcher Wechsel durch fehlende Messwerte angezeigt. Wird bei der Auswertung erkannt, dass für eine Seite keine oder zu wenige Messwerte zur Verfügung stehen, wird nur die andere Seite verwendet und dementsprechend zu einer I- oder IS-Form gewechselt. Für eine gültige Seite müssen mindestens zwei Messpunkte erkannt werden. Außerdem müssen diese Messpunkte

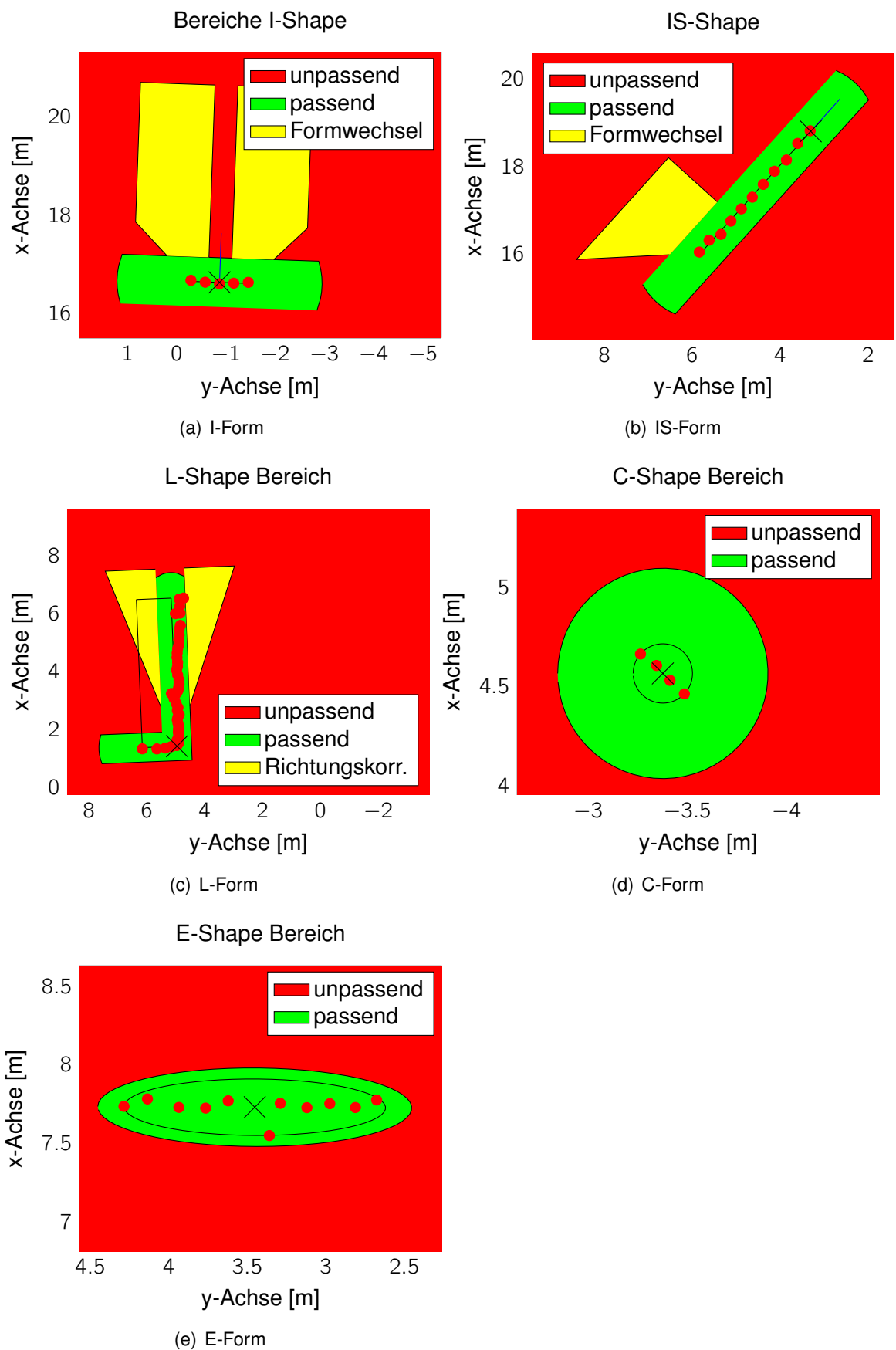


Abbildung 4.11: Gültigkeitsbereiche für Messpunkte

mindestens 40% der Objektbreite bzw. mindestens $1m$ der Objektlänge abdecken. Für die Objektlänge ist die prozentuale Abdeckung nicht sonderlich aussagekräftig, da die Länge oft erst bei dem Formwechsel zum ersten mal gemessen wird und somit zu 100% abgedeckt ist. Deshalb wird hier die abgedeckte Länge berücksichtigt. Wie in Abbildung 4.11 (c) zu sehen, werden bei der L-Form auch noch Bereiche zur sogenannten Richtungskorrektur definiert. Die Motivation dafür liegt in der schwierigen Orientierungsbestimmung bei Objekten mit abgerundeten Ecken, vor allem wenn diese nur mit wenigen Punkten erfasst werden. Wird von einem entgegenkommenden Fahrzeug eine solche Rundung detektiert, weißt die Orientierung der erzeugten I-Form nicht in die Fahrtrichtung des Objekts sondern ist deutlich verdreht. Wird nun im nächsten Zeitschritt ein Punkt des hinteren Teils der Längsseite erkannt, liegt diese mit großer Wahrscheinlichkeit nicht mehr in dem passenden Bereich der L-Form. Aufgrund des großen Abstands zur erwarteten Form sollte der Differenzvektor des Punktes auch nicht zur Ermittlung der Positionsveränderung verwendet werden. da dieser Vektor nicht der Bewegung des Objektes entspricht. Deshalb werden Messpunkte in diesem zweiten Bereich nur zur Berechnung und somit zur Korrektur der Orientierung verwendet.

Der passende Bereich für eine C-Form ergibt sich einfach aus einem Kreis mit dem Radius $\tau_{max,r,C} = 0.53m$ um den Referenzpunkt. Bei diesem Grenzwert sind neben der maximalen Ausdehnung ($0.45m$) auch eine mögliche Bewegung oder eine verschobene Lage des Schwerpunkts berücksichtigt. Allerdings wird die C-Form etwas anders behandelt wie die vorangehenden Formen. Während bei den anderen Formen ein Punkt des Segments zum Ausschluss oder zum Formwechsel führen kann, werden bei der C-Form immer Anfangs- und Endpunkt des Segments überprüft. Liegen beide Punkte innerhalb, wird das Segment verwendet, liegen beide außerhalb, passt das Segment nicht zu dem betrachteten Objekt. Im gemischten Fall mit einem Ende innerhalb und einem Ende außerhalb des Zylinders, wird ein Formwechsel erwartet. Die neue Form wird dabei wie bei der Objekterzeugung aus den Messwerten des Segments berechnet. Dieser Fall tritt beispielsweise bei einem aus dem Sichtschatten fahrenden Fahrzeugs oder Radfahrer auf. Zunächst werden nur wenige Punkte mit geringer Ausdehnung erkannt, die dementsprechend als C-Form verarbeitet werden. Sobald das Fahrzeug mit einer größeren Ausdehnung als den $0.9m$ erkannt wird, wird die neue Form untersucht. Falls sich das Objekt mit mehr als $\tau_{min,v,b} = 2\frac{m}{s}$ bewegt, wird zusätzlich auch auf die E-Form getestet. Durch diese Einschränkung wird die Anzahl der fälschlicherweise in der Umgebung detektierten Ellipsenformen stark reduziert.

Die E-Form wird wieder ähnlich wie die C-Form gehandhabt. Es werden ebenfalls die beiden Endpunkte der Segmente auf die Zugehörigkeit zur maximalen Ellipsenausdehnung mit den Halbachsen $1m$ und $0.25m$ untersucht. Aufgrund den eng umrissenen Initialisierungsbedingungen für die E-Form wird dabei nur ein Formwechsel berücksichtigt. Wie bereits erwähnt, kann die E-Form auch fälschlicherweise bei neu auftauchenden Fahrzeugen initialisiert werden. Wird das Segment länger als die maximalen $2m$ für ein Zweirad, so handelt es sich höchstwahrscheinlich um eine Fahrzeugseite und dementsprechend wird auf die IS-Form gewechselt, falls das Segment die Bedingungen für eine IS-Form erfüllt.

4.4.2 Aktualisierung von Objekten

Durch die Auswertung der Messwerte können für erkannte Objekte die aktuelle Position, Orientierung und die gemessene Ausdehnung ermittelt werden. Um die Bewegung des Objekts zu verfolgen, wird ein Kalmanfilter verwendet. Bevor die für die Bewegung relevanten Größen Position und Orientierung dem Kalmanfilter als aktuelle Messwerte übergeben werden, muss noch die Objektausdehnung und die Lage des Referenzpunktes berücksichtigt werden. Da sich die gemessene Position stets auf den Referenzpunkt der erkannten Form bezieht, muss bei auftretenden Formänderungen der verursachte Sprung vom Referenzpunkt auch bei den prädizierten Werten kompensiert werden.

Im Gegensatz zu den dynamischen Größen sind die Abmessungen eines Objektes normalerweise konstante Werte, die jedoch von dem Scanner nicht immer vollständig erfasst werden. Um die Objekte dennoch möglichst realitätsnah darstellen zu können, sollte zur Ermittlung der Objektdimensionen

nicht nur die aktuelle Messung, sondern auch vergangene Messungen in Betracht gezogen werden. Fayad und Cherfaoui [FC] verwenden für die gemessene Länge und Breite jeweils einen „Fixed Gain Filter“. Durch den Filter wird der Einfluss von Messrauschen reduziert, allerdings schrumpft der gefilterte Wert falls über längere Zeit eine Abmessung nur teilweise erkannt wird. Deshalb kommt hier ein einfacher wie effektiver Ansatz zum Einsatz, bei dem immer das Maximum aus festgehaltener und gemessener Dimension gespeichert wird (vgl. z.B. [FLD]). Dadurch bleiben einmal erkannte Objektdimensionen für die gesamte Erkennungsdauer des Objekts erhalten. Allerdings werden zu große Objektdimensionen, die beispielsweise durch falsch zugeordnete Messpunkte entstehen können, nicht korrigiert. Es hat sich gezeigt, dass dieses Problem sehr selten auftritt und dieser Ansatz ausreichende Ergebnisse liefert. Schließlich werden noch die äußersten, erkannten Punkte des Objektes für die Segmentierung im nächsten Zeitschritt gespeichert.

4.4.3 Klassifizierung

Für die Verfolgung von Objekten in der Umgebung ist eine Klassifizierung zwar nicht notwendig, bringt jedoch in diesem Anwendungsfall einige Vorteile mit sich. Neben der Darstellung kann auch das zugrunde liegende Bewegungsmodell an die Klasse der Verkehrsteilnehmer angepasst werden und damit das Tracking verbessert werden. Ein Fahrzeug bewegt sich beispielsweise anders als ein Fußgänger und sollte auch anders dargestellt werden.

Die Klassifizierung beruht auf den durch die Erkennung und die Verfolgung der Objekte ermittelten Größen. Dementsprechend stehen Abmessungen, Geschwindigkeit und Orientierung zur Verfügung. Aufgrund der beschränkten Informationen wird hier nur zwischen vier Klassen (Fahrzeuge, Zweiräder, Fußgänger und Unbekannt) unterschieden. Die Hauptmotivation ist die Klassifizierung von Fahrzeugen und Zweirädern, da diese mit einem eigenen Bewegungsmodell verfolgt werden, wie im nächsten Abschnitt beschrieben. Die Unterscheidung zwischen Fußgängern und unbekanntem Objekten hat auf das Bewegungsmodell keinen Einfluss, allerdings sollen Fußgänger möglichst eindeutig identifiziert werden, um sich von Hintergrundobjekten zu unterscheiden. Der verwendete Klassifizierungsalgorithmus ist aber nicht auf diese vier Klassen beschränkt und könnte je nach Bedarf und verfügbaren Unterscheidungsmerkmalen noch verfeinert werden.

Zur Klassifizierung wird ein ähnlicher Ansatz wie der von Mendes [MBNne] verwendet. Für jede Klasse und für jede entscheidende Größe wird eine Gewichtungsfunktion definiert. Diese Funktionen beschreiben jeweils wie gut der Wert einer Größe zu dieser Klasse passt. Die hier verwendeten Gewichtungsfunktionen für die Klassen Fahrzeug, Zweirad und Fußgänger sind in Abbildung 4.12 dargestellt. Wie zu erkennen ist, werden fünf Größen zur Klassifizierung verwendet. Die Funktionen wurden intuitiv definiert. So wurden die charakteristischen Punkte der Gewichtungsfunktionen von den jeweils erwarteten Dimensionen bzw. Geschwindigkeiten abgeleitet. Beispielhaft werden kurz die Wahl der Größen für die Klasse Fahrzeug erklärt. In den ersten beiden Funktionen werden die Objektausdehnungen bewertet. Da die meisten Fahrzeuge eine Länge zwischen 3 und 5m, sowie eine Breite zwischen 1.5 und 2m haben, erzielen diese Werte hohe Gewichtungswerte für diese Klasse. Die breitere Ausdehnung der Funktionen über die erwarteten Maße hinaus ist der meist unvollständigen Erkennung durch den LIDAR-Scanner geschuldet, da je nach Blickwinkel und durch abgerundete Kanten die Dimensionen selten komplett erfasst werden. Zusätzlich fließt auch die erkannte Form in die Gewichtung ein. Da für ein Fahrzeug annähernd eine Rechtecksform angenommen wird, werden die entsprechenden Formen I-, L- und IS-Form positiv gewichtet. Abschließend wird noch die Bewegung des Objekts berücksichtigt. Da in dieser Arbeit vor allem bewegte Objekte interessant sind, werden erst Geschwindigkeitswerte ab $2\frac{m}{s}$ voll gewichtet. Eine Klassifizierung von statischen Objekten wie zum Beispiel parkenden Fahrzeugen ist nicht notwendig. Das wichtigste Kriterium zur Identifikation von Fahrzeugen ist jedoch nicht der Geschwindigkeitsbetrag, sondern deren Richtung. Wie bei der Einführung der Formen in Kapitel 4.3.1.2 bereits erwähnt, ist deren Orientierung immer parallel zur erwarteten Bewegungsrichtung. Dementsprechend kann die Winkeldifferenz zwischen der Geschwindigkeitsrichtung und der Orientierung der Form zur Klassifizierung verwendet werden

4 Objekterkennung und Tracking LIDAR

und wird hier im Vergleich zu den anderen Größen doppelt gewichtet. Da sich Fahrzeuge entweder in Orientierungsrichtung oder genau entgegengesetzt bewegen können, sind Werte um 0° und um 180° positiv gewertet. Natürlich ist diese Parallelität vor allem bei dynamischen Fahrmanövern oder Rangiermanövern nicht immer gegeben, jedoch sind die erzielten Klassifizierungen in normalen Verkehrssituationen vollkommen ausreichend. Das liegt zum einen an dem linearen Toleranzbereich für kleine Winkelabweichungen und zum anderen daran, dass dieses Kriterium nur für unbekannte Objekte verwendet wird. Ist ein Objekt bereits als Fahrzeug oder Zweirad klassifiziert, entfällt diese Unterscheidung aufgrund des verwendeten Bewegungsmodells.

Bei Fußgängern entfällt dieses Kriterium, da die C-Form keine definierte Orientierung hat. Dafür wird bei Fußgängern der Geschwindigkeitsbetrag doppelt gewichtet, um eine bessere Abgrenzung zu statischen Objekten wie Laternenpfählen o.ä. zu gewährleisten.

Die aus diesen Funktionen stammenden Werte $w_1 - w_5$ werden anschließend gemittelt, wobei die doppelt gewichteten Werte die Nenner entsprechend erhöhen. Damit wird einem Objekt für jede Klasse ein Wert $\hat{\gamma}_{Klasse}$ zwischen 0 und 1 zugewiesen, der die aktuelle Übereinstimmung mit dieser Klasse anzeigt:

$$\hat{\gamma}_c = \frac{w_{1,c} + w_{2,c} + w_{3,c} + w_{4,c} + w_{5,c}}{6} \quad (4.11)$$

$$\hat{\gamma}_b = \frac{w_{1,b} + w_{2,b} + w_{3,b} + w_{4,b} + w_{5,b}}{6} \quad (4.12)$$

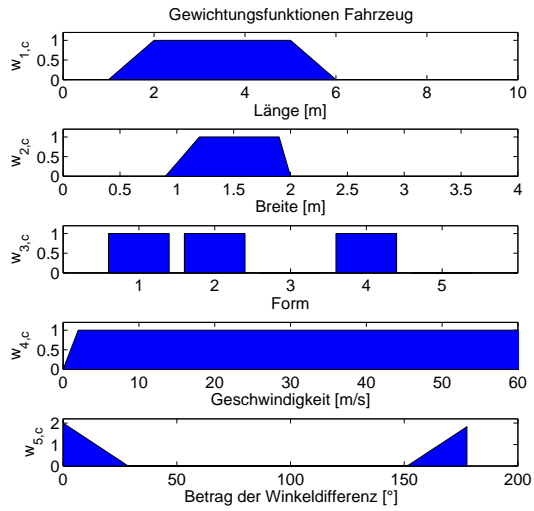
$$\hat{\gamma}_p = \frac{w_{1,p} + w_{2,p} + w_{3,p} + w_{4,p}}{5} \quad (4.13)$$

Dabei steht der Index c für die Klasse Fahrzeug (engl. car), b für Zweirad (engl. bike) und p für Fußgänger (engl. pedestrians).

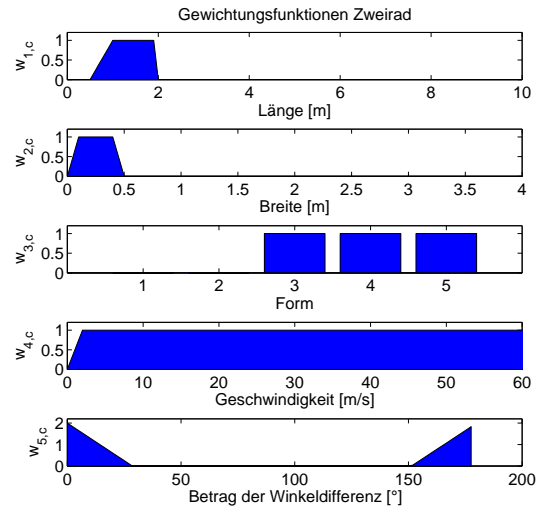
Da eine solche Klassifizierung nicht ständig wechseln, sondern möglichst eindeutige Zuordnungen erzeugen sollte, werden diese Werte nicht direkt als Klassifizierung sondern nur zur Anpassung der Klasse verwendet. Deshalb werden zusätzlich zu den Werten für die aktuelle Klassenübereinstimmung $\hat{\gamma}_{Klasse}$ auch Werte für die Zugehörigkeitswahrscheinlichkeit γ_{Klasse} definiert. Für diese gilt

$$\sum_{Klassen} \gamma_i = 1. \quad (4.14)$$

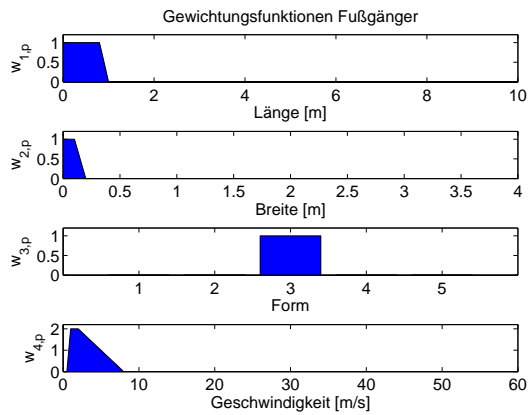
Grundsätzlich werden alle Objekte als unbekannte Objekte initialisiert, d.h. $\gamma_u = 1$ mit u als Index für die Klasse Unbekannt. Als Grenzwert für die Klassenübereinstimmung wird hier zunächst 0,7 verwendet. Erfüllt einer der aktuellen Werte die Ungleichung $\hat{\gamma}_{Klasse} > 0,7$ handelt es sich sehr wahrscheinlich um ein Objekt dieser Klasse und die dazugehörige Wahrscheinlichkeit γ_{Klasse} wird um den doppelten Differenzbetrag $2(\hat{\gamma}_{Klasse} - 0,7)$ erhöht. Liegt der Wert $\hat{\gamma}_{Klasse}$ darunter, wird die Wahrscheinlichkeit dementsprechend reduziert. Anschließend werden die Wahrscheinlichkeitswerte wieder normiert, so dass 4.14 erfüllt ist. Durch dieses Vorgehen steigt nur bei Objekten die sehr gut zu einer Klasse passen die entsprechende Wahrscheinlichkeit an. Schließlich wird ein Objekt erst einer Klasse zugeordnet, falls die Zugehörigkeitswahrscheinlichkeit $\hat{\gamma}_{Klasse}$ größer als 60% ist. Durch dieses Vorgehen lässt sich eine zügige Unterscheidung zwischen den verschiedenen Klassen verwirklichen, ohne dass die Klassenzuordnung ständig wechselt.



(a) Fahrzeug



(b) Zweirad



(c) Fußgänger

Abbildung 4.12: Gewichtungsfunktionen

4.4.4 Filterung

Zur Objektverfolgung werden in diesem Ansatz die in Kapitel 3 vorgestellten Kalman-Filter verwendet. Entscheidend für eine gute Verfolgung ist dabei das zugrunde liegende Bewegungsmodell. Zur Verfolgung von Verkehrsteilnehmer werden in der Literatur unterschiedlich komplexe Modelle aufgeführt. Das Einfachste beschreibt schlicht die Bewegung einer Punktmasse in zwei Dimensionen mit den Positionen und Geschwindigkeiten als Zustandsgrößen. Für die zeitdiskrete Darstellung der Bewegungsgleichung, wie in Gleichung (3.6), ergibt sich demnach

$$\begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_k = \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_{k-1} \quad (4.15)$$

Bei diesem Bewegungsmodell, das beispielsweise in [FC], [LH] oder [SRS] zu finden ist, wird eine über den Zeitschritt konstante Geschwindigkeit des Objekts angenommen. Geschwindigkeitsänderungen werden dann über das Prozessrauschen \mathbf{w} modelliert.

Auch die Erweiterung dieses Modells um die entsprechenden Beschleunigungsgrößen auf den Zustandsvektor $\mathbf{x} = [xyv_xv_ya_xa_y]^T$ wird beispielsweise in [Amd+13] verwendet. Dabei werden die Beschleunigungen zunächst konstant angenommen und Änderungen wieder über das Prozessrauschen abgebildet.

Fuerstenberg et al [FLD] erweitert das beschleunigungsfreie Punktmassenmodell (4.15) um eine Orientierung und die zugehörige Drehrate

$$\begin{bmatrix} x \\ y \\ \psi \\ v_x \\ v_y \\ \omega \end{bmatrix}_k = \begin{bmatrix} 1 & 0 & 0 & \Delta T & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta T & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta T \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ \psi \\ v_x \\ v_y \\ \omega \end{bmatrix}_{k-1} \quad (4.16)$$

Wie zu erkennen ist, sind die Bewegungen in den beiden Koordinatenrichtungen und die Drehung des Objekts komplett unabhängig voneinander.

Der Vorteil der bisher vorgestellten Bewegungsmodelle ist ihre Linearität. Damit können sie direkt mit dem klassischen Kalman-Filter (siehe Abschnitt 3.1.2) verfolgt werden. Da sich Fahrzeuge im Normalfall aber nicht unabhängig von ihrer Orientierung bewegen können, wird häufig auf aufwendigere Bewegungsmodelle zurückgegriffen. Im einfachsten Fall wird dabei angenommen, dass sich das Fahrzeug nur in Orientierungsrichtung bewegen kann. Damit ergibt sich für die beschreibende Differentialgleichung

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\psi} \\ \dot{v} \\ \dot{\omega} \\ \dot{a} \end{bmatrix} = \begin{bmatrix} v \cos(\psi) \\ v \sin(\psi) \\ \omega \\ a \\ 0 \\ 0 \end{bmatrix} \quad (4.17)$$

Dieses Modell wird beispielsweise in [Lin08] und [Kae+ne] verwendet. Um von dieser Differentialgleichung auf die benötigte Differenzengleichung zu kommen, muss über den Zeitschritt integriert

werden. Nach [Lin08] ergibt sich damit

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \begin{bmatrix} \frac{v+a\Delta T}{\omega} \sin(\psi + \omega\Delta T) + \frac{a}{\omega^2} \cos(\psi + \omega\Delta T) - \frac{v}{\omega} \sin(\psi) - \frac{a}{\omega^2} \cos(\psi) \\ -\frac{v+a\Delta T}{\omega} \cos(\psi + \omega\Delta T) + \frac{a}{\omega^2} \sin(\psi + \omega\Delta T) + \frac{v}{\omega} \cos(\psi) - \frac{a}{\omega^2} \sin(\psi) \\ \omega\Delta T \\ a\Delta T \\ 0 \\ 0 \end{bmatrix} \quad (4.18)$$

Die bei kleinen Gierraten $\omega \approx 0$ auftretenden, numerischen Probleme umgeht Lindl mit der Näherung

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \begin{bmatrix} \cos(\psi)v\Delta T + \frac{1}{2}\cos(\psi)a\Delta T^2 \\ \sin(\psi)v\Delta T + \frac{1}{2}\sin(\psi)a\Delta T^2 \\ \omega\Delta T \\ a\Delta T \\ 0 \\ 0 \end{bmatrix} \quad (4.19)$$

Wie aus den Gleichungen (4.18) bzw. (4.19) ersichtlich, handelt es sich dabei um nichtlineare Bewegungsmodelle, für die entsprechende Erweiterungen des Kalman-Filters benötigt werden.

Wie bereits an einigen Stellen erwähnt, werden in dieser Arbeit, ähnlich wie in [Lin08], zwei unterschiedliche Bewegungsmodelle verwendet. Eines für Fußgänger und noch nicht klassifizierte, unbekannte Objekte, das Andere zur Beschreibung von Fahrzeugen und Zweirädern. Durch diese Unterscheidung können die Bewegungen genauer modelliert werden und damit ein deutlich bessere Verfolgung erzielt werden.

4.4.4.1 Bewegungsmodell für Fußgänger und unbekannte Objekte

Das Bewegungsmodell für Fußgänger und unbekannte Objekte ist sehr ähnlich zu dem Modell von Fuerstenberg aus Gleichung (4.16), jedoch wird auf die Drehrate verzichtet. Durch diese Vereinfachung wird von einer konstanten Orientierung dieser Objekte ausgegangen. Für Fußgänger bzw. C-Formen trifft diese Annahme immer zu, da bei diesen die Orientierung nicht berücksichtigt wird und somit konstant Null ist. Es hat sich gezeigt, dass diese Vereinfachung auch für alle anderen unbekanntem Objekte ausreicht. Bei statischen Hintergrundobjekten sollte ohnehin keine Drehung auftreten und die Verkehrsteilnehmer, bei denen die Drehraten relevant sind, werden nur bis zur Klassifizierung mit diesem einfachen Modell gefiltert. Die dabei auftretenden Drehungen können über das Prozessrauschen abgefangen werden. Das verwendete Bewegungsmodell lässt sich somit durch

$$\begin{bmatrix} x \\ y \\ \psi \\ v_x \\ v_y \end{bmatrix}_k = \begin{bmatrix} 1 & 0 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & 0 & \Delta T \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ \psi \\ v_x \\ v_y \end{bmatrix}_{k-1} \quad (4.20)$$

beschreiben. Die zusätzliche Filterung der Orientierung, im Vergleich zu dem Modell (4.15), ist notwendig, da sich sonst fälschlicherweise erkannte Formen im Gebüsch ständig neu ausrichten können, ohne dass diese Drehung erkannt wird. Solche falschen Objekte können über die gefilterte Orientierung erkannt und gelöscht werden. Durch die Vernachlässigung der Drehrate im Vergleich zu dem Modell von Fuerstenberg werden zum einen die erwarteten Bewegungen besser abgebildet und zum anderen wird durch die geringere Zustandsdimension der Berechnungsaufwand reduziert.

4.4.4.2 Bewegungsmodell für Fahrzeuge und Zweiräder

Zur Beschreibung der Bewegungen von Fahrzeugen und Zweiräder wird grundsätzlich angenommen, dass diese sich nur in Richtung ihrer Orientierung bewegen können. Damit lässt sich die Bewegung durch die Differentialgleichung (4.17) beschreiben. Die von Lindl daraus abgeleitete Differenzgleichung (4.18) ist zwar mathematisch korrekt, jedoch auch relativ aufwendig in der Berechnung und darüber hinaus ergeben sich auch numerische Instabilitäten bei $\omega \approx 0$. Da die Abtastfrequenz des LIDAR-Scanners mit 75 Hz sehr hoch und damit die Zeitschrittweite eines Filterschritts sehr klein ist, bietet sich eine einfache Linearisierung an, um diese Probleme zu umgehen. Das verwendete Bewegungsmodell ist demnach

$$\mathbf{x}_k = \begin{bmatrix} x \\ y \\ \psi \\ v \\ \omega \\ a \end{bmatrix}_k = \mathbf{x}_{k-1} + \begin{bmatrix} \cos(\psi)v\Delta T \\ \sin(\psi)v\Delta T \\ \omega\Delta T \\ a\Delta T \\ 0 \\ 0 \end{bmatrix} \quad (4.21)$$

und entspricht fast der Näherung von Lindl (4.19). Die darin enthaltenen Terme mit ΔT^2 können aufgrund des kleinen ΔT vernachlässigt werden. Trotz der sehr einfachen Berechnungsvorschrift ergeben sich für die relevanten Wertebereiche keine entscheidenden Abweichungen von der richtigen Differenzgleichung.

Da bei diesem Bewegungsmodell die Drehung nicht mehr unabhängig von der Bewegung ist, ist die Wahl eines sinnvollen Drehpunkts entscheidend für eine gute Verfolgbarkeit von Fahrzeugen. Bei den unbekanntem Objekten wird einfach der Referenzpunkt der entsprechenden Form als Drehpunkt angenommen. Da dieser Punkt bei vielen Formen jedoch auf der Außenkante der Objekte liegt, ist die Entfernung zum tatsächlichen Drehpunkt eines Fahrzeugs teilweise so groß, dass die Bewegung nicht mehr mit dem verwendeten Modell dargestellt werden kann. Deshalb wird der Referenzpunkt bei Fahrzeugobjekten zunächst transformiert, bevor der Kalmanfilter verwendet wird. Als Drehpunkt eines Fahrzeugs wird hier näherungsweise die Mitte der Hinterachse angenommen. Da jedoch nicht immer die gesamte Ausdehnung eines Fahrzeugs erfasst werden kann, und auch die genaue Lage der Hinterachse vom LIDAR-Scanner nicht gemessen wird, wird die Lage dieses Drehpunktes nur grob angenähert. Deshalb wird der Punkt mittig mit einem Meter Abstand zur Hinterkante des Objekts gewählt. Auch wenn diese Näherung nicht für alle Fahrzeuge passt, sind die damit erzielten Ergebnisse dennoch ausreichend. Solche kleineren Ungenauigkeiten können durch das modellierte Prozessrauschen ausgeglichen werden. Für Zweiradfahrer wird auf eine Anpassung des Referenzpunkts verzichtet, da die Entfernung des Mittelpunkts zu dem erwarteten Drehpunkt ohnehin relativ klein ist.

4.4.4.3 Parametrisierung des Kalman-Filters

Neben dem Bewegungsmodell müssen auch das erwartete Mess- und Prozessrauschen vorgegeben werden. Wie in Kapitel 3 beschrieben, werden diese Größen durch die Matrizen \mathbf{R} und \mathbf{Q} definiert. Je genauer die gewählten Werte mit den tatsächlichen Störeinflüssen übereinstimmen, desto exakter sind die ermittelte Fehlerkovarianz und somit die Messwertgewichtung des Kalman-Filters.

Für das Fußgänger-Bewegungsmodell finden sich beispielsweise in [LH] und [Amd+13] mögliche Parameter. Dabei werden jeweils die Standardabweichung der Beschleunigung in einer Koordinatenrichtung approximiert. Die hier zusätzlich auftretende Orientierung, müsste ebenfalls mit einer Störgröße berücksichtigt werden. Aus der Bewegungsgleichung könnte dann die Matrix \mathbf{Q} ermittelt werden. Während die Fehlerfortpflanzung in diesem Modell aufgrund der Linearität einfach zu berechnen ist, ist die Situation beim Fahrzeugmodell ungleich schwerer. Allerdings finden sich dazu in den Arbeiten die ähnliche Modelle verwenden, keine näheren Angaben. Deshalb wurde hier auf

eine genaue Betrachtung der Fehlerzusammenhänge verzichtet und die Ungenauigkeiten für jede Zustandsvariable einzeln abgeschätzt, sowie anhand von Messfahrten evaluiert.

Anschließend muss noch das auftretende Messrauschen modelliert werden. Die Herausforderung liegt darin, dass nicht nur die Störungen in den Messwerten, sondern auch die Ungenauigkeiten der Auswertung berücksichtigt werden müssen, da der Kalman-Filter erst die verarbeiteten Daten filtert. Aufgrund des Messprinzips des Lidar-Scanners treten die Messfehler nur in Strahlrichtung und senkrecht dazu auf. Die Modellierung der Sensorfehler müsste demnach auch in einem Polarkoordinatensystem durchgeführt werden. Allerdings sind die Messfehler des Sensors im Vergleich zu den Ungenauigkeiten bei der Auswertung vernachlässigbar. Der angegebenen Entfernungsmessfehler von $10\text{mm} \pm 35\text{mm}$ ist viel kleiner als die Fehler, die bei der Interpretation von unvollständig erkannten Objekten auftreten. Deshalb wurden die hier verwendeten Werte, genauso wie das Prozessrauschen, zunächst geschätzt und anschließend durch die optische Evaluation von Messfahrten angepasst.

4.4.5 Löschen von Objekten

Ein wichtiger Schritt bei der Verfolgung von mehreren, wechselnden Objekten ist das Löschen von nicht mehr erkannten Objekten. Aufgrund des beschränkten Sichtfelds und der Bewegung, sowohl der anderen Verkehrsteilnehmer als auch des Versuchsträgers, verlassen Objekte regelmäßig den relevanten Bereich. Dabei können Objekte entweder über den maximalen Öffnungswinkel des Sensors wandern, sich zu weit vom Sensor entfernen oder von anderen Objekten verdeckt werden. In allen drei Fällen stehen für das Objekt keine passenden Messdaten mehr zu Verfügung und dementsprechend sollte das Objekt entfernt werden. Neben den geschilderten Fällen sollten auch falsche Objekte identifiziert und gelöscht werden. Aufgrund von zufälligen Formen in den Messwerten oder durch Fehlinterpretationen kann es immer wieder vorkommen, dass Objekte erzeugt werden die in der Realität nicht existieren.

Um diesen beiden Anwendungen gerecht zu werden, wird das Löschen von Objekten durch mehrere Funktionen gesteuert. Als wichtigstes Kriterium wird die aktuelle Messbarkeit herangezogen. Wird ein Objekt nicht mehr vom Sensor erfasst, könnte es den Messbereich bereits verlassen haben. Allerdings ist es auch möglich, dass das Objekt nur sehr kurz verdeckt wurde bzw. die Messdaten falsch interpretiert wurden. Deshalb führt eine fehlende Detektion nicht zum sofortigen Löschen. Stattdessen werden zwei Variablen eingeführt. Die Variable μ_{life} beschreibt die Lebensdauer des Objekts, also wie oft dieses Objekt bereits vom Sensor detektiert wurde. $\mu_{nondetect}$ steht für die Anzahl der fehlenden Detektionen. Kann einem Objekt kein Messwert zugeordnet werden, wird $\mu_{nondetect}$ um eins erhöht und μ_{life} um eins reduziert. Bei einer erfolgreichen Detektion werden die Variablen genau gegenteilig verändert. Ein Objekt wird dann gelöscht, sobald $\mu_{nondetect}$ größer als μ_{life} oder größer als $\tau_{nondetect} = 12$ ist. Durch dieses Vorgehen wird sichergestellt, dass Objekte die seit längerer Zeit erfolgreich verfolgt werden, nicht gleich bei einer kurzen Unterbrechung gelöscht werden. Stattdessen wird das Objekt weiter prädiziert und die gemessenen Dimensionen und die ermittelte Klasse bleiben erhalten. Wurde ein Objekt dagegen gerade erst erkannt, führt bereits eine fehlende Messung zum Löschen des Objekts. Damit können viele falsch erzeugte Objekte schnell wieder entfernt werden.

Neben dieser Routine, die so oder so ähnlich in vielen Trackingalgorithmen enthalten ist, werden hier noch weitere Überprüfungen durchgeführt, um noch mehr falsche Objekte zu erkennen und zu löschen. Das ist vor allem der Tatsache geschuldet, dass der hier verwendete Algorithmus anfällig für falsche Detektionen von nicht existierenden Objekten ist. Wie am Anfang dieses Kapitels beschrieben, sollen die I-, IS- und L-Formen Fahrzeuge repräsentieren. Dementsprechend sollte der LIDAR-Scanner keine Punkte innerhalb dieser Formen detektieren. Seltene Ausnahmefälle in denen der Scanner beispielsweise durch ein Fenster einen Punkt innerhalb des Fahrzeugs erkennt, werden hier vernachlässigt, da sie bei einer Installationshöhe des Scanners von ca. 40 cm so gut wie nicht auftreten. Wird nun ein Messpunkt innerhalb einer Form erkannt, deutet das auf ein falsches Objekt hin und wird dementsprechend gelöscht. Der zum Löschen eines Objekts führende Bereich

entspricht dem unpassenden Gebiet aus Abbildung 4.11, dass innerhalb der Objektfläche liegt. In dieser Abbildung sind jedoch nur bei der L-Form beide Dimensionen eingezeichnet. Wie zu sehen ist, können dementsprechend nur Objekte gelöscht werden, deren gemessene Ausdehnung größer als die erlaubte Toleranz $\tau_{dist,\perp} = 0,5m$ ist. Mit dieser Funktion können auch zu groß gewordene Objekte entfernt werden. Da bei diesem Algorithmus immer die maximal erkannten Dimensionen gespeichert werden, kann es in manchen Fällen dazu kommen, dass Objekte durch falsch interpretierte Messungen zu groß werden. Beispielsweise können zwei hintereinander stehende Fahrzeuge als ein langes Fahrzeug interpretiert werden. Sobald der Scanner aufgrund des veränderten Blickwinkels einen Punkt zwischen den Fahrzeugen detektiert, kann dieser Fehler korrigiert werden.

Darüber hinaus werden auch die gemessenen Dimensionen explizit mit den erwarteten Dimensionen von Verkehrsteilnehmern verglichen. Sind die Maße größer als die in Tabelle 4.1, ist das Objekt entweder aufgrund von Fehlinterpretationen zu groß gewachsen oder es handelt sich ein reales, großes Objekt wie beispielsweise ein Haus. In beiden Fällen sollte das Objekt entfernt werden. Als letztes Kriterium wird noch die quadratische Innovation aus Gleichung (3.17) herangezogen. Je größer dieser Wert ist, desto schlechter passen die vom Filter erwarteten Werte zu den gemessenen Werten. Demnach ist die Messwertinterpretation fehlerhaft oder aber das gemessene Erscheinungsbild hat sich stark geändert, was bei falschen Objekten im Randbereich häufig auftritt. Tritt bei einer Messung eine sehr hohe Innovation auf, wird diese Messung nicht verwendet, da es sich höchst wahrscheinlich um eine Fehlinterpretation des Algorithmus handelt. Stattdessen wird das Objekt mit den präzisierten Werten weiterverfolgt. Weißt die darauf folgende Messungen ebenfalls eine hohe Innovationen auf, wird das Objekt gelöscht.

5 Objekterkennung Radarsensor

Wie in Abschnitt 2.3.2 erwähnt, ist das Versuchsfahrzeug auch mit dem RADAR-Sensor des serienmäßigen ACC-Systems ausgestattet. Allerdings eignet sich der Sensor nur bedingt für die hier betrachtete Anforderung, da der Radar und die interne Datenverarbeitung für das ACC entwickelt und ausgelegt sind. Erstens hat der Sensor einen relativ schmalen Öffnungswinkel von ca. $\pm 8^\circ$, womit nur ein geringer Bereich vor dem Fahrzeug abgedeckt werden kann, und zweitens werden von dem Sensorsystem nur verfolgte Objekte ausgegeben und keine Rohdaten. Die interne Datenverarbeitung ist aber auf die ACC-Funktionalität und somit auf parallelen Verkehr spezialisiert. Querverkehr oder abbiegende Fahrzeuge werden deshalb schlecht bzw. fehlerhaft ausgegeben. Da es aber leider keine Möglichkeit gibt, auf die Rohdaten des Radarsensors zuzugreifen, müssen die gelieferten Ergebnisse verwendet werden.

Der Sensor liefert Daten von bis zu 32 verfolgten Objekten. Die Position eines Objekts wird dabei über den Abstand zum Sensor und die laterale Ablage ausgegeben. Allerdings ist vor allem bei Fahrzeugen nicht klar, von welcher Stelle des Fahrzeugs das empfangene Radarecho stammt. So kann entweder eine beliebige Position der zugewandten Seite oder beispielsweise auch die Position des Motorblocks erfasst werden. Neben der direkt gemessenen Relativgeschwindigkeit zwischen Objekt und Sensor wird auch eine Relativbeschleunigung ausgegeben, die über den internen Filter berechnet wird. Außerdem wird noch ausgegeben, ob das Objekt bei der letzten Messung erfasst wurde und ob es bereits aus dem vorherigen Zeitschritt bekannt ist. Der Sensor hat eine Zykluszeit von 80 bis 120 ms und ist demnach deutlich langsamer als der LIDAR-Scanner. Grundsätzlich erfasst der Sensor elektrisch leitfähige Objekte, wodurch neben Fahrzeugen auch viele Hintergrundobjekte wie Schilder oder Ampeln detektiert werden. Diese Problematik wird bei dem ACC-System dadurch umgangen, dass nur bewegte Objekte berücksichtigt werden.

Aufgrund der geschilderten Sensoreigenschaften und der Spezialisierung der internen Auswertung werden die Radarobjekte noch einer Nachbearbeitung unterzogen, um die Messdaten für die Anwendung aufzubereiten. Da bei dieser Arbeit in erster Linie bewegte Objekte von Interesse sind, werden wie beim ACC alle Radarobjekte vernachlässigt, deren gemessene Relativgeschwindigkeit der negativen Eigengeschwindigkeit des Versuchsfahrzeugs entspricht. Deshalb wird für Radarobjekte gefordert, dass

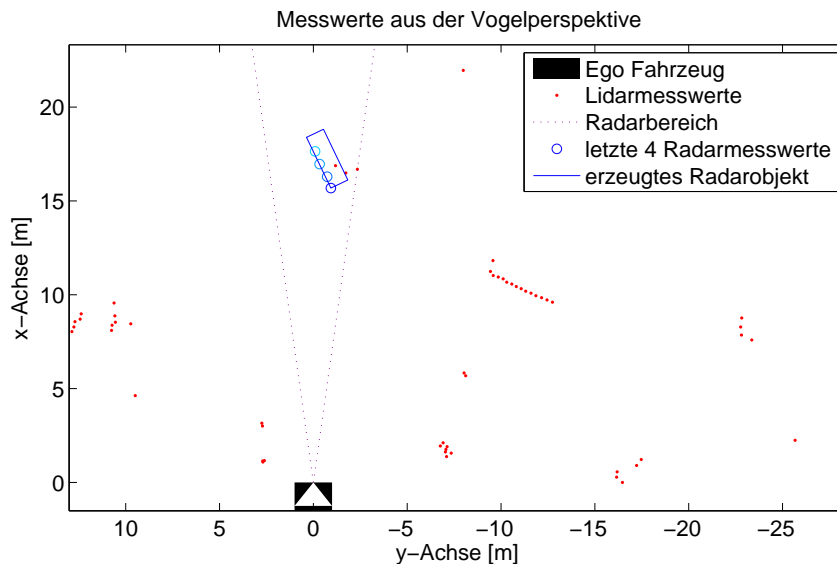
$$v_{rel} + v_{ego} < \tau_{min,v,radar} = 4 \frac{m}{s} \quad (5.1)$$

gilt. Da der Radar nur die Relativgeschwindigkeit v_{rel} misst, muss die Eigengeschwindigkeit des Versuchsfahrzeugs kompensiert werden. Aufgrund des geringen Öffnungswinkels von maximal 8° kann die Winkeldifferenz zwischen Eigengeschwindigkeit und Relativgeschwindigkeit vernachlässigt werden. Da der Radar jedoch auch teilweise falsche Objekte erkennt bzw. falsche Geschwindigkeiten misst, wird zusätzlich gefordert, dass die Bedingung (5.1) in drei aufeinander folgenden Messungen erfüllt sein muss, bevor das Radarobjekt verwendet wird. Die damit verbundene Verzögerung von ca. 0.3 Sekunden bis ein Radarobjekt verarbeitet wird stellt aus zweierlei Gründen kein Problem da. Einerseits sind die Messungen von Objekten, die gerade erst durch eine Drehung des Eigenfahrzeugs oder durch die Bewegung des Objekts seitlich in den Radarbereich eingetreten sind, noch sehr ungenau. Deshalb ist es sogar sinnvoll, die Objekte erst nach einem kurzen Einschwingvorgang des internen Filters zu verwenden. Und andererseits eignet sich der Radar vor allem für die Erfassung von Fließverkehr und Gegenverkehr. Bei der großen Reichweite von ca. 150m fällt diese kurzzeiti-

5 Objekterkennung Radarsensor



(a) Kamerabild der Szene



(b) Messwerte und Radarobjekt aus Vogelperspektive

Abbildung 5.1: Radarauswertung

ge Verzögerung nicht ins Gewicht, da die Objekte selbst bei sehr großen Relativgeschwindigkeiten deutlich länger zum Erreichen des für die Prädiktion interessanten Bereichs benötigen.

Neben der gemessenen Position und der Relativgeschwindigkeit ist für eine sinnvolle Objektdarstellung vor allem noch die Orientierung des Objektes interessant. Da bei erkannten Radarobjekten davon ausgegangen wird, dass es sich entweder um Fahrzeuge oder Zweiräder handelt, entspricht die Orientierung auch gleichzeitig der Bewegungsrichtung des Objekts (siehe Abschnitt 4.4.4.2). Da leider nicht auf die internen Werte des Radarfilters zugegriffen werden kann, muss diese Größe nachträglich berechnet werden. Da die ausgegebenen Werte bereits gefiltert wurden, ist dafür aber keine erneute Filterung nötig. Die Orientierung wird aus den letzten drei bis vier gemessenen Positionswerten ermittelt. Dafür wird eine Regressionsgerade durch die Werte gelegt, welche eine Abschätzung der groben Orientierung ermöglicht. Wegen der relativ ungenauen Winkelauflösung des Radarsensors sowie der undefinierten Position der Reflexion am Objekt ist eine genaue Bestimmung der Orientierung ohnehin nicht möglich.

5 Objekterkennung Radarsensor

Zur Veranschaulichung ist das beschriebene Vorgehen in Abbildung 5.1 dargestellt. In Abbildung (a) ist das Kamerabild der entsprechenden Situation mit projizierten Lidar-Messwerten zu sehen. Die Lidar-Messwerte sind als weiße Kreise in das Bild projiziert. Wie im Vergleich mit Abbildung 4.1 zu sehen ist, wurde während dieser Arbeit eine neue Kamera mit einem Weitwinkelobjektiv und dementsprechend einem deutlich größeren Öffnungswinkel installiert. Allerdings ist die Kameratransformation noch nicht optimal, wie beispielsweise am Fuß des Umleitungsschildes zu erkennen ist. In Darstellung (b) ist die Szene aus der Vogelperspektive dargestellt. Dabei befindet sich das Versuchsfahrzeug im Ursprung und die Lidar-Messwerte sind wie in Kapitel 4 durch rote Punkte dargestellt. Zusätzlich ist der vom Radar erfasste Bereich durch gestrichelte Linien gekennzeichnet. Die letzten vier vom Radar gemessenen Positionen des abbiegenden PKWs sind mit blauer werdenden, blauen Kreisen eingezeichnet. Aus diesen vier Messwerten wird dann die Orientierung berechnet. Durch den Winkel, in dem das Objekt zum Radarsensor steht, wird ebenfalls abgeschätzt von welchem Punkt der zugewandten Seite das Radarecho stammt. In diesem Fall wird davon ausgegangen, dass die vordere linke Ecke für das Echo verantwortlich ist. Durch diese Werte wird nun ein Objekt mit den minimalen Abmessungen $3m$ mal $1m$ erzeugt.

An der Abweichung zwischen dem Radarobjekt und den Lidar-Messdaten sind auch die Probleme bei der Verwendung des Radarsensors zur Umfelderkennung zu erkennen. Sobald die Objekte aus der parallelen Bewegung ausscheren, wie hier beim Abbiegevorgang, weichen die ausgegebenen Werte von der tatsächlichen Position ab, da die interne Auswertung auf die Erfassung von Parallelverkehr spezialisiert ist. In diesem Fall folgen die Radarmessungen nicht dem engen Radius des Objekts, sondern driften nach außen weg. Dennoch kann über die abgeschätzte Orientierung die Lage des Objekts zumindest grob angenähert und mögliche Abbiegevorgänge detektiert werden.

Zusammenfassend lässt sich sagen, dass sich der Radarsensor aufgrund des geringen Öffnungswinkels sowie der Ungenauigkeiten im Bezug auf Position und Orientierung der Objekte nur bedingt zur Erfassung und Prädiktion des Umfelds eignet. Durch eine sinnvolle Kombination mit den Daten des Lidar-Sensors kann die Umfelderkennung jedoch deutlich verbessert werden, was im folgenden Kapitel beschrieben wird.

6 Sensordatenfusion

Der Begriff Sensordatenfusion beschreibt die Kombination (Fusion) von mehreren Sensoren mit dem Ziel, dass die fusionierten Daten genauer sind als die Daten der einzelnen Sensoren. Im folgenden Abschnitt wird zunächst allgemein auf die unterschiedlichen Fusionsarten und die entsprechenden Begriffe genauer eingegangen. Anschließend wird die hier nötige Sensorkalibrierung erläutert bevor abschließend die verwirklichte Fusion dargelegt wird.

6.1 Grundlagen der Datenfusion

Grundsätzlich beschreibt die Datenfusion die Kombination von Datensätzen aus unterschiedlichen Quellen. Da in diesem Anwendungsfall die Daten von Sensoren stammen, wird auch der Begriff Sensorfusion verwendet. Sensoren bezeichnen dabei Geräte, die durch die Messung physikalischer Größen Rückschlüsse auf die Umgebung zulassen. Wie durch diese sehr allgemeine Definition bereits angedeutet, gibt es zahlreiche, unterschiedliche Sensoren die grundsätzlich verschiedene Eigenschaften aufweisen können.

Datenfusion ist grundsätzlich in vielen unterschiedlichen Bereichen einsetzbar. Aufgrund der vorliegenden Problemstellung wird das Prinzip der Datenfusion hier anhand der Umfelderkennung erläutert. Außerdem eignet sich diese Anwendung auch sehr gut um das Prinzip der Sensorfusion zu verdeutlichen, da dieses räumliche Problem sowohl übersichtlich dargestellt werden kann als auch gut vorstellbar ist.

Da die ersten Ansätze zur Datenfusion bereits in den 80er Jahren in militärischen Projekten entwickelt wurden, gibt es zu diesem Thema umfassende Veröffentlichungen. Eine kompakte Einführung zur Datenfusion befindet sich in [RPL07], welche auch als Grundlage für diesen Abschnitt dient.

6.1.1 Fusionsarten

Im Allgemeinen gibt es drei unterschiedliche Arten Sensoren zu kombinieren. Es wird zwischen komplementärer, konkurrierender und kooperativer Fusion unterschieden, wobei die Grenzen zwischen diesen Bereichen fließend sind.

Komplementäre Fusion Wie aus der Bezeichnung komplementär bereits ersichtlich, geht es bei der komplementären Fusion darum, dass sich Sensoren ergänzen. Die Motivation für eine komplementäre Fusion ist die Erfassung einer größeren Messdatenmenge. Dabei können sich die Sensoren sowohl bei den Erfassungsbereichen als auch bei den Messgrößen ergänzen. Ein hier auftretendes Beispiel für komplementäre Fusionen ist die Kombination des Front- und des Rück-LIDAR. Diese tasten komplett unterschiedliche Bereiche ab.

Konkurrierende Fusion Bei der konkurrierenden Fusion werden von mehreren Sensoren identische physikalische Größen gemessen. Die Ziele sind dabei die Messgenauigkeit zu erhöhen und die Ausfallwahrscheinlichkeit zu reduzieren. Im einfachsten Fall werden zwei gleiche Sensoren verwendet, die den selben Bereich abtasten. Durch diese Redundanz können Sensorausfälle oder Übertragungsfehler erkannt und kompensiert werden. Für konkurrierende Fusionen werden oft auch Sensoren mit unterschiedlichen Messprinzipien verwendet. Da Messfehler häufig durch äußere Einflüsse

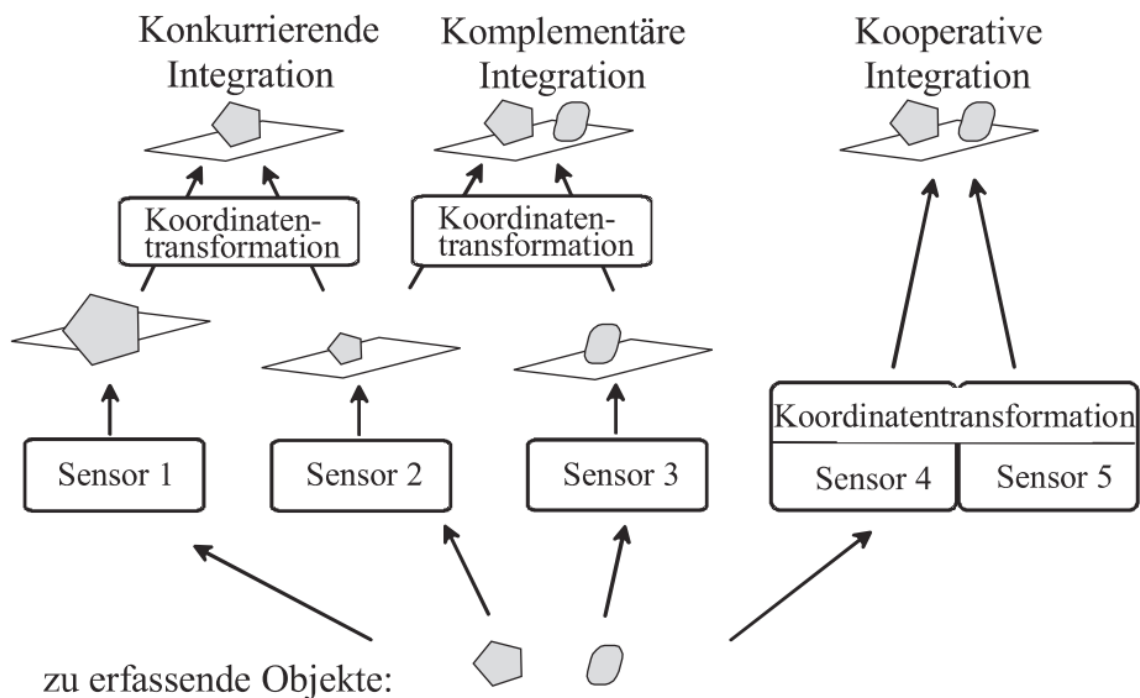


Abbildung 6.1: Fusionsarten [RPL07, S. 96]

verursacht werden, sind die Messfehler von unterschiedlichen Sensoren weniger stark korreliert als die von identischen Sensoren. So können Messungenauigkeiten verringert werden. Ein passendes Beispiel dafür liefern der in dieser Arbeit verwendete LIDAR- und der Radarsensor. Mit beiden Sensoren kann die Position eines Objekts bestimmt werden, wobei die Messung einmal auf der Reflexion eines Infrarotlasers und das andere mal auf der Reflexion von Radarwellen beruht. Auch wenn die Messprinzipien in diesem Fall ähnlich sind, unterscheiden sich die Umwelteinflüsse auf die Messergebnisse deutlich. So ist der LIDAR beispielsweise relativ anfällig gegenüber Niederschlag, wohingegen der Radar davon nahezu unbeeinflusst bleibt. Dieses Beispiel verdeutlicht auch die fließenden Grenzen zwischen den einzelnen Fusionsarten, da mit dem Radar im Gegensatz zum LIDAR auch die Relativgeschwindigkeit von Objekten gemessen werden kann. Somit kann diese Messgröße komplementär mit den LIDAR-Messungen fusioniert werden.

Kooperative Fusion Diese Fusionsart ist im Vergleich zu den beiden vorangegangenen komplexer und weniger intuitiv. Die Grundidee dabei ist, dass bei der kooperativen Fusion von Sensoren Messungen möglich sind, die mit einem einzelnen Sensor nicht erfasst werden können. Das klassische Beispiel für diese Art der Sensorfusion ist die Stereokamera. Dabei werden die Daten zweier Kameras verwendet, um die Entfernung einzelner Objekte zu ermitteln. Eine solche Messung ist ohne die Fusion der Datensätze nicht möglich, ausgenommen dem Fall einer bekannten Objektgröße.

In Abbildung 6.1 aus [RPL07] sind die drei verschiedenen Fusionsarten zur Veranschaulichung grafisch dargestellt. Bei der konkurrierenden Fusion wird ein Objekt von zwei Sensoren unterschiedlich erfasst und das Fusionsobjekt wird aus beiden Messungen erzeugt. Bei der komplementären Fusion erfassen die Sensoren unterschiedliche Objekte und bei der kooperativen Fusion entsteht erst durch die Fusion die entsprechende Information.

6.1.2 Fusions Ebenen

Neben den vorgestellten Fusionsarten kann die Datenfusion auch in verschiedenen Ebenen durchgeführt werden. Dabei beschreiben die Ebenen jeweils den Abschnitt in der Signalverarbeitungskette. Im Rahmen der Umfelderkennung besteht diese Kette aus drei Abstraktionsebenen.

Signalebene Die erste Ebene ist die Signalebene, bei welcher direkt die gemessenen Werte der Sensoren vorliegen. Diese Signale haben teilweise auch schon eine erste Vorverarbeitung im Sensor durchlaufen. Allerdings umfasst diese Bearbeitung nur Korrekturen die aufgrund des Messprinzips nötig sind, wie beispielsweise Belichtungsanpassungen bei Kameras. An dieser frühen Position in der Verarbeitungskette tritt noch kein Informationsverlust durch eine mögliche Bearbeitung auf. Allerdings ist die Verarbeitung aufgrund des geringen Abstraktionsgrads und der entsprechend großen Datenmenge auch am aufwendigsten. Für eine solche Fusion müssen die erfassten Signale synchronisiert und vergleichbar sein. Eine Fusion auf Signalebene wird aufgrund des Zeitpunktes der Zusammenführung auch als frühe Fusion bezeichnet.

Merkmalsebene Bei einer Fusion auf Merkmalsebene werden die Rohdaten der einzelnen Sensoren zunächst unabhängig voneinander verarbeitet. Die dabei erzeugten Merkmale wie beispielsweise Position oder Geschwindigkeit werden anschließend fusioniert. Durch den höheren Abstraktionsgrad sind in diesem Fall schon deutlich weniger Daten zu fusionieren als auf Signalebene.

Symbolebene Die höchste Fusionsebene ist die sogenannte Symbolebene oder Objektebene. Dabei werden die Daten von jedem Sensor getrennt verarbeitet und erst die verfolgten Objekte fusioniert. Eine solche Fusion wird auch als späte Fusion bezeichnet. Durch die weit fortgeschrittene Verarbeitung ist die Datenmenge für diese Fusion am geringsten, allerdings ist der mögliche Informationsverlust durch die Verarbeitung auch am größten. Ein großer Vorteil dieser Fusion ist außerdem die gute Modularität und Wiederverwendbarkeit, da die Fusion und die einzelnen Sensorvorverarbeitungen unabhängig voneinander sind. Dadurch können die Sensorverarbeitungen getrennt entwickelt und auch nachträglich noch Sensoren hinzugefügt oder entfernt werden.

6.2 Sensorsynchronisation und Transformation

Um Daten unterschiedlicher Sensoren fusionieren zu können, muss der Zusammenhang zwischen diesen Daten bekannt sein. Je nach Anwendungsfall und verwendeten Sensoren müssen dabei unterschiedliche Bereiche berücksichtigt werden. In dieser Arbeit sollen der LIDAR-Scanner (siehe Abschnitte 2.3.1 und 4) und der Radarsensor (siehe Abschnitte 2.3.2 und 5) fusioniert werden. Da mit diesen Sensoren dynamische Positionsdaten detektiert werden, müssen sowohl der räumliche als auch der zeitliche Zusammenhang zwischen den Daten bekannt sein. Der räumliche Zusammenhang zwischen den Sensoren ist notwendig, um die Positionsdaten in einem gemeinsamen Koordinatensystem darstellen zu können. Da auch dynamische Objekte erfasst werden, müssen auch die Zeitpunkte der jeweiligen Messungen berücksichtigt werden. Zusätzlich zu den beiden Sensoren zur Umfelderkennung wird für die Detektion der Eigenbewegung des Versuchsfahrzeug eine Inertialplattform verwendet.

6.2.1 Räumliche Kalibrierung

Für die räumliche Kalibrierung muss der Zusammenhang zwischen den unterschiedlichen Sensorkoordinatensystemen, in denen die jeweiligen Messdaten dargestellt werden, ermittelt werden. Dafür gibt es grundsätzlich zwei Möglichkeiten. Ist die Position und Orientierung der einzelnen Sensoren bekannt und kann die jeweilige Lage der verwendeten Sensorkoordinatensysteme aus den Spezifikationen entnommen werden, kann die gesuchte Transformation direkt gemessen werden. In dem hier vorliegenden Fall von Radar-Sensor und Front-LIDAR bietet sich dieses Vorgehen an. Die beiden Sensoren sind so ausgerichtet, dass die x-Achsen der Sensorkoordinatensysteme parallel zur Fahrzeugachse in Fahrtrichtung sind. Auch die y- und z-Achsen sind identisch orientiert. Die Lage und Position der Koordinatensysteme ist in Abbildung 2.3 dargestellt. Durch das Vermessen des Abstandes der beiden Sensoren in diesen drei Achsrichtungen kann der Translationsvektor zwischen

den beiden Koordinatensystemen ermittelt werden. Die gemessenen Werte sind in Tabelle 2.1 aufgeführt.

Die zweite Möglichkeit zur räumlichen Kalibrierung funktioniert über den Vergleich von den Messwerten. Wird ein Objekt von beiden Sensoren erfasst und an mehreren Stellen im Schnittbereich der beiden Sensoren gemessen, kann aus dem Vergleich dieser Messwerte auf die Lage der Sensorkoordinatensysteme relativ zueinander geschlossen werden. Da in diesem Fall bereits bekannt ist, dass die Orientierung der Systeme identisch ist, reicht theoretisch bereits ein Messwert um die Verschiebung in x- und y-Richtung zu ermitteln. Da der LIDAR-Scanner nur die xy-Ebene abtastet und der Radarsensor ebenfalls nicht in der z-Richtung differenziert, ist über die vertikale Verschiebung der Koordinatensysteme keine Aussage möglich. Da, aus den selben Gründen, ohnehin nur die Position auf der Fahrbahnebene interessant ist, ist diese Verschiebung bei der hier vorliegenden Anwendung auch nicht weiter von Interesse. Der Vorteil eines solchen Vorgehens ist, dass kein größeres Vorwissen über die Lage der Sensoren und deren verwendete Koordinatensysteme notwendig ist. Allerdings hängt die Güte der so vorgenommene Kalibrierung stark von den Sensorgenauigkeiten ab. Um die Einflüsse von Messfehlern zu reduzieren, ist es sinnvoll möglichst viele Messungen durchzuführen. Dadurch entsteht ein überbestimmtes Gleichungssystem, das mit Hilfe der Methode der kleinsten Quadrate gelöst und somit der quadratische Fehler minimiert werden kann.

Um die ausgemessenen Werte zu validieren, wurden spezielle Messwerte aufgezeichnet und ausgewertet. Dafür wurde eine metallische Rampe an verschiedenen Orten im gemischten Sensorbereich senkrecht positioniert. Durch die Auswertung der gemessenen Werte konnte eine Verschiebung um 30cm in x-Richtung bestätigt werden. Außerdem hat sich gezeigt, dass die Verschiebung in y-Richtung bereits bei der internen Radarauswertung berücksichtigt wird. Für die Transformation ergibt sich damit folgende Beziehung (vgl. Tabelle 2.1):

$$\mathbf{x}_{LIDAR} = \mathbf{x}_{RADAR} + \begin{bmatrix} -0,3m \\ 0m \\ 0,17m \end{bmatrix} \quad (6.1)$$

6.2.2 Temporale Kalibrierung

Sobald ein Sensorsystem in einer dynamischen Umgebung eingesetzt wird, spielen die Aufnahmezeitpunkte der Messwerte eine wichtige Rolle. Da sich die erfasste Umgebung mit jedem Zeitschritt ändert, können die Werte von unterschiedlichen Sensoren nur sinnvoll fusioniert werden, wenn die Zeitdifferenz zwischen den Messungen bekannt ist. Je größer die auftretende Dynamik im Vergleich zur Abtastrate der Sensoren ist, desto wichtiger ist die temporale Kalibrierung, bei der diese Latenzzeiten ermittelt werden.

6.2.2.1 Möglichkeiten zur temporalen Kalibrierung

Bei der Aufzeichnung und der Verarbeitung von Sensordaten treten grundsätzlich verschiedene Latenzzeiten auf, wie in Abbildung 6.2 dargestellt. Die Anteile der drei Blöcke Sensor, Kommunikation und Fusion können sich je nach Systemarchitektur und verwendeten Sensoren stark unterscheiden. Allerdings sind diese Anteile an der gesamten Latenzzeit im Allgemeinen nicht eindeutig bestimmbar ([Lin08]). Vor allem die Latenzzeit der Sensoren und deren Vorverarbeitung wird von den Herstellern oft nicht veröffentlicht. Deshalb muss zur temporalen Kalibrierung die Gesamtlatenzzeit abgeschätzt werden.

Grundsätzlich ist das über den Vergleich von Messwerten einer dynamischen Situation möglich. Eine ausführliche Beschreibung wie die einzelnen Gesamtlatenzzeiten von einem Laserscanner, einem Radarsensor und einigen Kameras ermittelt werden können, befindet sich in [Lin08]. Dabei

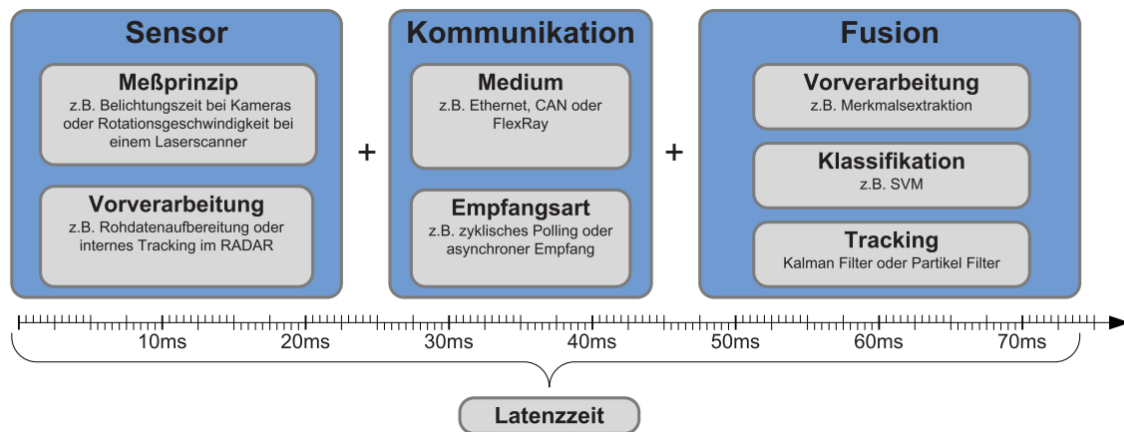


Abbildung 6.2: Latenzzeiten bei Datenfusion [Lin08, S.38]

werden zwei Fahrzeuge die jeweils mit der selben Inertialplattform (siehe Abschnitt 2.3.4) wie das Versuchsfahrzeug ausgestattet sind. Über die GPS-Zeit können diese Messeinheiten hoch genau synchronisiert werden und somit für die Gewinnung von Referenzdaten verwendet werden. Werden mit diesen Fahrzeugen nun dynamische Manöver ausgeführt, können die von den Sensoren erfassten Messwerte mit den Referenzdaten verglichen werden. Bewegen sich die Fahrzeuge beispielsweise aufeinander zu, wird der aktuell von einem Entfernungssensor ausgegebene Wert größer sein als der tatsächliche Wert, da die Messung aufgrund der Latenzzeit des Sensors bereits bei einem größeren Abstand durchgeführt wurde. Aus diesem Differenzwert kann anschließend die Gesamtlatenzzeit ermittelt werden. Diese Zeit umfasst alle Latenzen die zwischen der Messung des Sensors und der Aufzeichnung der ausgegebenen Daten liegen. Je nach Art der Sensoren bieten sich unterschiedliche Manöver an. So wird für die Kameras eine Querbewegung aufgezeichnet, da diese durch die hohe Winkelauflösung gut erfasst werden kann. Bei den entfernungsgebenden Sensoren wie dem LIDAR und Radar bieten sich aufeinander zufahrende oder sich entfernende Fahrzeuge an. Im Rahmen dieser Arbeit steht jedoch kein zweites Fahrzeug mit entsprechender Inertialsensorik zur Verfügung. Dementsprechend ist die Gewinnung von Referenzdaten ungleich schwieriger.

6.2.2.2 Abschätzung der Latenzzeiten über Sensoreigenschaften

Als erster Anhaltspunkt für die auftretenden Latenzzeiten können die von den Herstellern veröffentlichten Sensoreigenschaften dienen. In der Spezifikation des LIDAR-Scanners wird explizit auf den Zeitablauf der Messungen eingegangen. Bei einer Winkelauflösung von einem Grad benötigt der Scanner einen Spiegelumlauf pro Messung. Die Dauer für einen kompletten Umlauf wird mit 13,32ms angegeben [SIC03]. Die Daten werden in dieser Zeitspanne in einem Datenpuffer abgelegt und während dem nächsten Scan ausgegeben. Die Dauer dieser Ausgabe ist nicht genau beschrieben, muss aber kürzer wie der darauf folgende Scandurchlauf sein. Dadurch können die internen Latenzzeiten, je nach Winkel, mit 6,66ms bis 26,64ms abgeschätzt werden. Der niedrigste Wert ergibt sich aus der Messung bei 180° und der Annahme einer sofortigen Übertragung. Dementsprechend berechnet sich die Zeit lediglich aus der halben Umlaufzeit des Spiegels bis dieser wieder in der Ausgangsposition ist. Die angenommene maximale Latenzzeit ergibt sich aus der Betrachtung des ersten Messwerts bei 0° in Kombination mit der Annahme, dass die Ausgabe die gesamte Dauer des zweiten Scans in Anspruch nimmt. Zu diesen Zeiten kommt noch die Latenz durch die Übertragung zum Rechner hinzu.

Beim Radarsensor sind dagegen keine genaueren Angaben verfügbar. Den einzigen Anhaltspunkt stellt somit die Messrate von 10Hz dar (siehe Tabelle 2.2). Darüber hinaus findet in dem Sensor bereits die vollständige Messwertverarbeitung und Objektverfolgung statt, deren Dauer ebenfalls nicht

näher bekannt ist. Aufgrund der Abtastrate kann demnach von Latenzzeiten zwischen 100 und 200ms ausgegangen werden. Auch hier muss noch die Latenzzeit der Übertragung berücksichtigt werden, die in diesem Fall über den CAN-Bus läuft. Da diese Angaben jedoch deutlich zu ungenau für eine sinnvolle temporale Kalibrierung sind, werden diese Abschätzungen durch die Auswertung und den Vergleich von Messwerten präzisiert.

6.2.2.3 Temporale Kalibrierung durch Vergleich der Messdaten

Bei diesem Vorgehen wird der Umstand genutzt, dass sich die Sensorbereiche des LIDAR-Scanners und des Radarsensors überschneiden. Dadurch können Objekte in diesem Schnittbereich von beiden Sensoren gleichzeitig erfasst werden, falls die entsprechenden Remissionen der Objekte ausreichend stark sind. Befinden sich sowohl die Sensoren als auch das Objekt in Ruhe, kann durch den Vergleich der gemessenen Werte auch auf den räumlichen Zusammenhang zwischen den Sensoren geschlossen werden, wie bereits in Abschnitt 6.2.1 erläutert. Für die temporale Kalibrierung ist nun eine Relativbewegung zwischen Sensoren und Objekt nötig. Bewegt sich das Versuchsfahrzeug auf ein stehendes Objekt zu, sind die gemessenen Entfernungswerte abhängig vom Messzeitpunkt. Durch den Vergleich der gleichzeitig ausgegebenen Messwerte kann mit Hilfe der Relativgeschwindigkeit auf die Zeitdifferenz zwischen den Messungen geschlossen werden. Bei dem Beispiel sich annähernder Objekte korrespondiert eine größere Latenzzeit mit einem größeren Abstand, da die Distanz zum Messzeitpunkt noch größer als beim Ausgabezeitpunkt war. Um die Distanzmessungen adäquat vergleichen zu können, müssen die Werte in ein einheitliches Koordinatensystem transformiert werden, weshalb die räumliche Kalibrierung eine Voraussetzung für dieses Vorgehen ist.

Für die Abschätzung der Latenzzeiten mit dieser Methode wird ein Referenzobjekt gewählt. Dieses Objekt muss mehrere Voraussetzungen erfüllen. Zunächst einmal ist es notwendig, dass es sich im Messbereich beider Sensoren befindet und vor allem auch von beiden Sensoren detektiert wird. Dementsprechend muss es sowohl den Infrarotlaser des LIDAR-Scanners als auch die Radarwellen reflektieren. Außerdem muss zwischen Objekt und Versuchsfahrzeug eine Relativgeschwindigkeit auftreten, wobei größere Geschwindigkeiten zu bevorzugen sind, da die latenzzeitbedingten Unterschiede in der Entfernungsmessung proportional zur Relativgeschwindigkeit steigen. Außerdem sollte das Objekt nur eine geringe Tiefe haben, damit die Position der Radarreflexion bekannt ist. Dadurch kann die vom Radar gemessene Distanz direkt mit der des LIDAR-Scanners verglichen werden kann. Aus diesem Grund sind entgegenkommende Fahrzeuge eher ungeeignet, da das Radarecho zum Beispiel vom Motorblock stammen kann.

Aufgrund dieser Voraussetzungen wurde hier auf statische Objekte am Rande einer Landstraße zurückgegriffen. Dabei wurden sowohl Leitpfosten als auch Schilder verwendet. Die Leitpfosten können aufgrund der Reflektoren bereits in großen Entfernungen vom LIDAR-Scanner erkannt werden. Vor allem größere Schilder werden vom Radarsensor zuverlässig erkannt. Durch die sehr geringe Tiefe ist die longitudinale Position des Radarechos klar definiert. Um die beiden Sensoren kalibrieren zu können wurden nun die Schilder und Leitpfosten ausgewählt, welche von beiden Systemen erkannt wurden. Zusätzlich wird die Inertialplattform als Referenz herangezogen. Diese liefert neben den Eigenbewegungen des Versuchsfahrzeugs auch die korrigierte GPS-Position mit einer Genauigkeit unter 0,4m (siehe Abschnitt 2.3.4). Durch die über Google Maps gemessene GPS-Position der Objekte kann zusätzlich zu den Distanzmesswerten der beiden Sensoren auch die GPS-Distanz betrachtet werden. Allerdings beinhaltet diese auch weitere Ungenauigkeiten, da neben dem Messfehler der Inertialplattform vor allem die nachträglich bestimmten GPS-Werte unbekannte Fehler enthalten können.

In Abbildung 6.3 sind die Distanzmessungen der drei Sensoren zu einem am rechten Straßenrand befindlichen Leitpfosten sowie die von der Inertialplattform gemessene Geschwindigkeit dargestellt. Alle Werte beziehen sich auf das COG-Koordinatensystem des Versuchsfahrzeugs. Die blaue Kurve beschreibt die ermittelte GPS-Distanz zwischen dem Fahrzeug und dem Pfosten. Wie an dem Kurvenverlauf zu sehen ist, nähert sich das Fahrzeug mit relativ konstanter Geschwindigkeit an, pas-

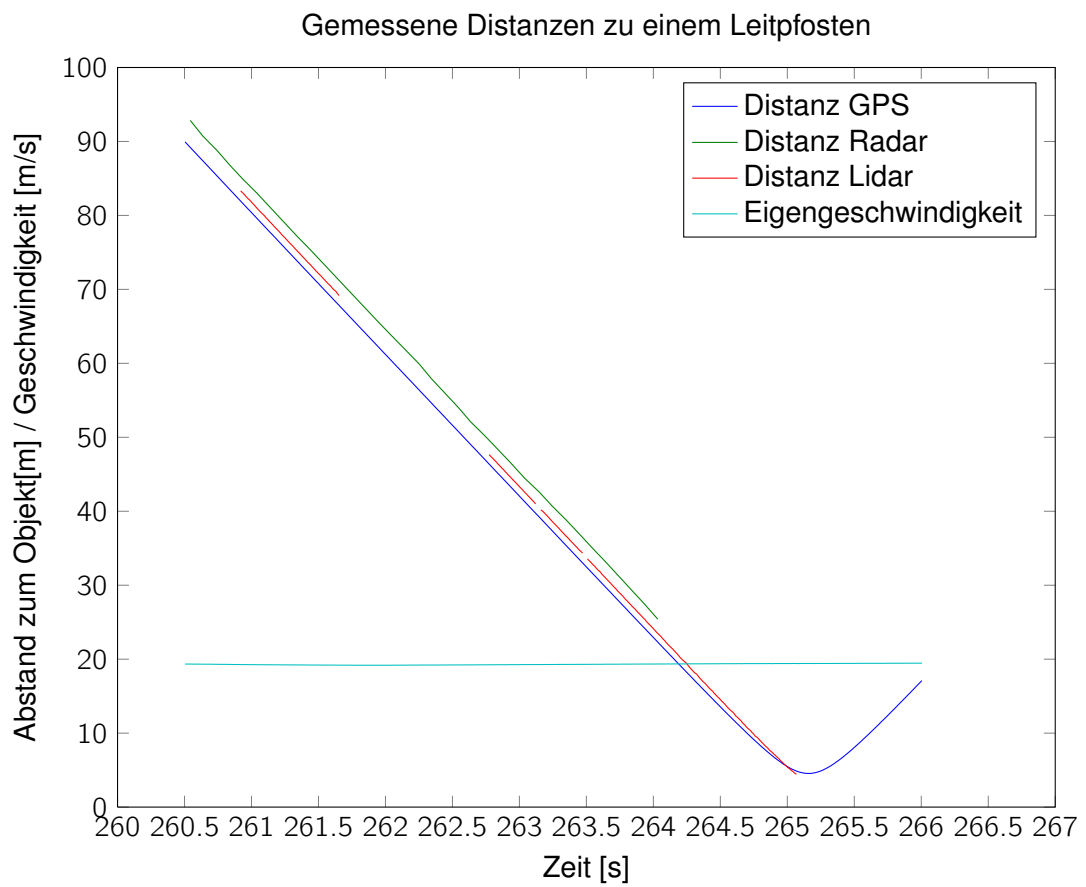


Abbildung 6.3: Distanzmessungen der unterschiedlichen Sensoren

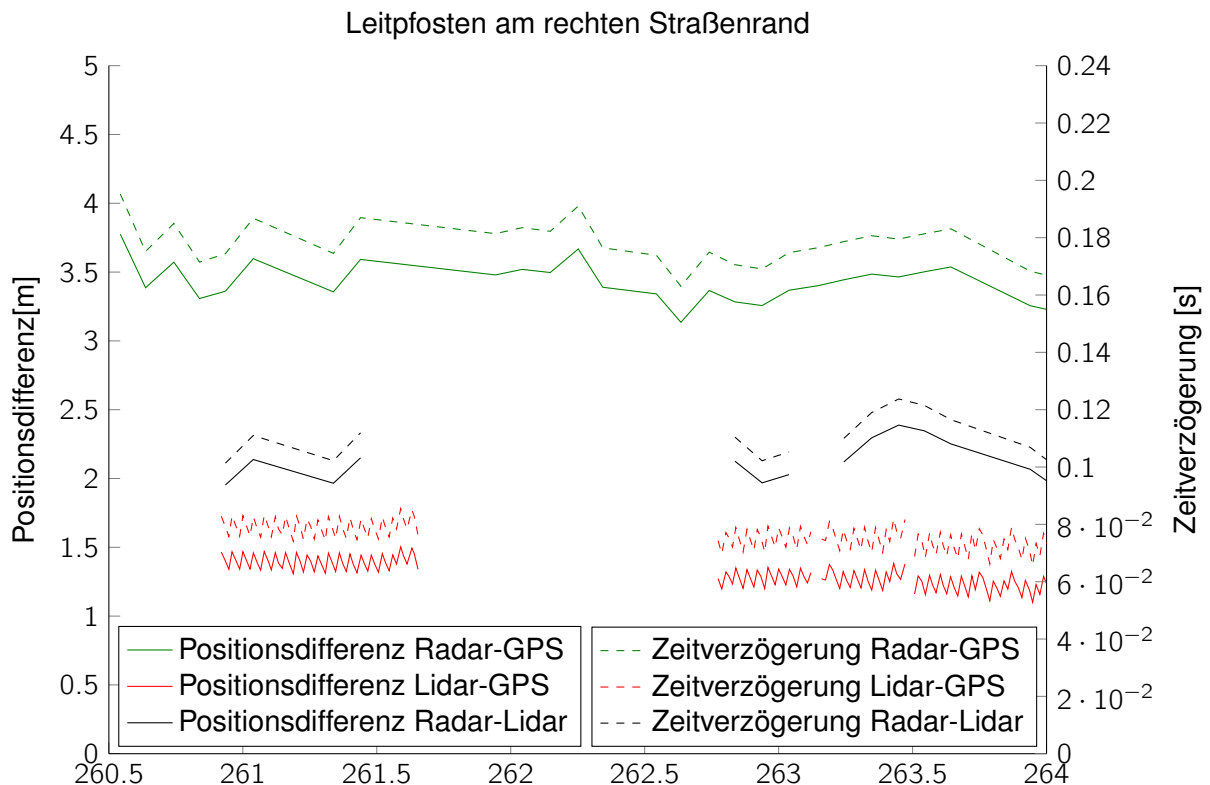


Abbildung 6.4: Bestimmung der Latenzzeiten

siert den Leitpfosten und entfernt sich anschließend wieder. Die grüne Kurve der Radarmesswerte endet bei ca. 25m, da der Pfosten dort den Radarmessbereich verlässt. Die Unterbrechungen im roten LIDAR-Graphen werden durch fehlende Detektion aufgrund zu geringer Remission verursacht. Die Eigengeschwindigkeit des Versuchsfahrzeugs liegt in dieser Zeitspanne zwischen $19,18$ und $19,46 \frac{m}{s}$ und ist somit annähernd konstant. Die dargestellten Kurven entsprechen dem erwarteten Verlauf. Alle drei Sensoren detektieren die Annäherung an das Objekt. Die identische Steigung der drei Kurven zeigt auch, dass diese die selbe Relativgeschwindigkeit wahrnehmen. Allerdings weisen die gemessenen Distanzkurven eine deutliche Verschiebung zueinander auf. Da die Verschiebung über die gesamte Erkennung relativ konstant und deutlich größer als die geringe Objektausdehnung ist, muss sie durch Latenzzeiten oder fehlerhafte räumliche Kalibrierung verursacht worden sein. Die räumliche Kalibrierung wurde jedoch experimentell überprüft (siehe Abschnitt 6.2.1), wodurch die Positionsdifferenz durch die Latenzzeiten der einzelnen Sensoren entsteht.

Um aus den gemessenen Kurven auf die Latenzzeiten zwischen den Sensoren zu schließen, gibt es zwei Möglichkeiten. Es können die Kurven entlang der Zeitachse verschoben werden bis sie deckungsgleich sind. D.h. es werden die Zeitpunkte ermittelt, an denen die Sensoren die selbe Entfernung gemessen haben. Dadurch können die Latenzzeiten direkt ohne weitere Informationen bestimmt werden. Oder es werden die Positionsdifferenzen an einem bestimmten Zeitpunkt verglichen und mit der Relativgeschwindigkeit die Zeitdifferenzen ermittelt. Da hier statische Referenzobjekte verwendet werden und die Eigengeschwindigkeit des Fahrzeugs von der Inertialen Messeinheit sehr genau ($0,02 \frac{m}{s}$ RMS [Oxf11]) erfasst wird, ist hier auch dieser Weg möglich. Beide Verfahren liefern dementsprechend auch übereinstimmende Ergebnisse. Aufgrund der besseren Vorstellbarkeit und den zusätzlichen Positionsdifferenzen werden hier die Ergebnisse des zweiten Verfahrens vorgestellt.

In Abbildung 6.4 sind die Positionsdifferenzen zwischen den drei Sensoren-Paarungen auf der linken y-Achse aufgetragen. Prinzipiell müssten diese Differenzen an identischen Zeitpunkten ermittelt werden, aber aufgrund der unterschiedlichen Abtastraten der Sensoren von ca. $10Hz$ beim Radar,

Tabelle 6.1: Positions- und Zeitdifferenzen zwischen den Messungen

Sensorpaar	mittlere Positionsdifferenz	mittlere Latenzzeit
RADAR - GPS	3,4281m	178,0ms
LIDAR - GPS	1,3398m	77,3ms
RADAR - LIDAR	2,1394m	111,0ms

75Hz beim LIDAR und 100Hz bei den GPS-Daten wurden stattdessen die Messungen mit der minimalen Differenz der Ausgabezeiten verglichen. Diese Unterschiede haben jedoch im Vergleich mit den Latenzzeiten kleine Maximalwerte von $6,6ms$ (halbe LIDAR-Abtastrate) sowie einen Mittelwert von nahezu Null und sind dementsprechend vernachlässigbar, da im Folgenden nur die Mittelwerte der Kurven von Interesse sind. Werden nun die ermittelten Positionsdifferenzen betrachtet, ist beim jeweiligen Vergleich mit den GPS-Daten ein geringes Messrauschen der beiden Sensoren zu erkennen. Da die GPS-Daten von der Inertialplattform gefiltert ausgegeben werden, weisen diese einen glatten Verlauf auf. Daraus kann geschlossen werden, dass das Rauschen allein von den Umgebungssensoren stammt. Abgesehen von dem Messrauschen bleiben die Positionsdifferenzen zwischen den Sensoren, wie erwartet, relativ konstant. Der Bildausschnitt wurde auf Distanzen, die größer sind als ca. $25m$, beschränkt, da der longitudinale Abstand bei dieser Auswertung dominieren soll. Die entsprechenden Latenzzeiten zwischen den Sensoren können dann über

$$t_{Latenz} = \frac{\Delta x}{v_{ego}} \quad (6.2)$$

mit Hilfe der Eigengeschwindigkeit v_{ego} zum jeweiligen Messzeitpunkt ermittelt werden. Die dadurch berechneten Zeitverzögerungen sind auf der rechten y-Achse der Abbildung 6.4 aufgetragen. Um den Einfluss des Messrauschens und des erwähnten geringen Zeitversatzes zwischen den Ausgabezeitpunkten der Sensoren zu reduzieren, werden die Mittelwerte der Differenzen gebildet. Diese sind sowohl für die Positionsdifferenzen als auch für die gesuchten Zeitverzögerungen in Tabelle 6.1 zusammengefasst. Die Werte beziehen sich nur auf die in Abbildung 6.4 dargestellte Messung, jedoch wurden vergleichbare Auswertungen mit mehreren Objekten bei unterschiedlichen Geschwindigkeiten durchgeführt. Dabei hat sich gezeigt, dass wie erwartet die Werte für die Kombinationen RADAR-GPS und LIDAR-GPS stark schwanken, die für die entscheidende Kombination RADAR-LIDAR aber relativ konstant sind. Dieses Verhalten wird durch die deutlich größere Ungenauigkeit der GPS-Referenz verursacht, da vor allem die Genauigkeit der Kartendaten nicht bekannt ist. Aus diesem Grund ist die Ermittlung der absoluten Latenzzeiten der beiden Sensoren aus diesen Daten nicht möglich. Dafür müsste eine eindeutige GPS-Referenz vorhanden sein. Ohnehin ist bei dieser Arbeit der Zeitunterschied zwischen den Messungen der beiden Sensoren von größerem Interesse. Da von der IMU nur die Drehrate und die Geschwindigkeit verwendet wird und sich diese Werte im Millisekundenbereich nur geringfügig ändern, ist auch keine genaue temporale Kalibrierung zur IMU notwendig. Bei der Auswertung von mehreren verschiedenen Objekten bei unterschiedlichen Relativgeschwindigkeiten wurden für die Latenzzeit zwischen Radar- und Lidar-Messungen Werte zwischen $100ms$ und $111ms$ ermittelt. Um die Kompensation der Latenzzeit auf das minimal Nötige zu beschränken, wird im weiteren Verlauf der Arbeit die Latenzzeit t_{Latenz} auf die untere Grenze von $100ms$ festgelegt. Die möglicherweise auftretende Abweichung von höchstens $10ms$ kann im Vergleich zu den Ungenauigkeiten der Radarmessung vernachlässigt werden.

6.2.2.4 Berücksichtigung der Latenzzeiten

Wie im vorhergehenden Abschnitt ermittelt, weisen die vom Radar gemessenen Werte im Vergleich zum LIDAR-Scanner eine Verzögerung von ca. $t_{Latenz} = 100ms$ auf. Um diese Messwerte kombinieren zu können, muss diese zeitliche Verzögerung berücksichtigt werden. Der naheliegendste Ansatz

wäre nun, die LIDAR-Daten künstlich zu verzögern, sodass die Daten synchron ausgewertet werden können. Allerdings steht das im Widerspruch zu dem Ziel dieser Arbeit, die Bewegung von anderen Verkehrsteilnehmer vorherzusagen. Wenn die Messwerte des LIDARs um die Latenzzeit des Radars verzögert werden, müsste bei der Prädiktion der Objektpositionen zusätzlich zur Übertragungszeit auch diese Latenzzeit berücksichtigt werden.

Aus diesem Grund wird bei dieser Arbeit der gegenteilige Ansatz verwendet. Die Lidardaten werden nicht verzögert, sondern die Radar-Messwerte werden prädiert. Dieses Vorgehen würde verhindern, dass die Umfelderkennung durch die Fusion mit den Radardaten um die Latenzzeit verzögert werden muss. Da diese pauschale Verzögerung von $100ms$ der 7,5-fachen LIDAR-Abtastrate entspricht und damit mindestens die aktuellsten sechs Messungen des LIDARs gar nicht berücksichtigt werden könnten, erscheint dieses Vorgehen sinnvoll.

Die Problematik bei diesem Ansatz liegt natürlich in der Prädiktion der Radar-Messwerte. Zum einen ist nicht bekannt, ob das Radarobjekt zum prädierten Zeitpunkt noch erfasst wird bzw. wo es sich dann befindet. Dementsprechend ist bei diesem Vorgehen auch mit Abstrichen bei den Messgenauigkeiten zu rechnen. Entscheidend ist nun, wie groß die Ungenauigkeiten der Prädiktion sind und ob diese in Kauf genommen werden können. Zur Evaluierung wurden zwei unterschiedliche Prädiktionsmodelle entwickelt und anhand realer Messwerte untersucht.

Für die Prädiktion der Radarobjekte muss die Relativbewegung zu dem Versuchsfahrzeug bestimmt werden. Bei dem ersten Modell werden die Radarobjekte durch

$$\begin{aligned}\hat{x} &= x + t_{Latenz} v_{rel} \\ \hat{y} &= y + \sin(\Delta\psi)\hat{x}\end{aligned}\tag{6.3}$$

prädiert. Dabei wird die x-Koordinate einfach über die Relativgeschwindigkeit angepasst und die y-Koordinate durch die Orientierungsänderung des Versuchsfahrzeugs.

Beim zweiten Ansatz wurde die Bewegung des Eigenfahrzeugs während des Prädiktionsschritts durch eine konstante Kreisbahn modelliert und die Objektgeschwindigkeit getrennt kompensiert:

$$\begin{aligned}\hat{x} &= x \cos(\Delta\psi) - y \sin(\Delta\psi) - v_{ego} t_{Latenz} \frac{\sin(\Delta\psi)}{\Delta\psi} + (v_{rel} + v_{ego}) t_{Latenz} \\ \hat{y} &= x \sin(\Delta\psi) + y \cos(\Delta\psi) - v_{ego} t_{Latenz} \frac{1 - \cos(\Delta\psi)}{\Delta\psi}\end{aligned}\tag{6.4}$$

Zur Evaluierung dieser Ansätze wurden bei realen Messdaten jeweils die prädierten Werte mit den Messwerten der nächsten Radarmessung verglichen. Da die Prädiktionszeit und die Abtastzeit des Radars übereinstimmen, kann die Positionsdifferenz zur Evaluierung herangezogen werden. Dafür wurden die aufgezeichneten Radardaten einer ca. 400s-langen Messfahrt verwendet. Wie in Kapitel 5 beschrieben, werden nur diejenigen Radarobjekte berücksichtigt, die über drei Radarmessungen eine Geschwindigkeit über $4 \frac{m}{s}$ aufweisen. Die knapp 3500 bewegten Radarmessungen wurden nun jeweils mit beiden Modellen prädiert und mit den tatsächlichen Messungen des nächsten Zeitschritts verglichen. Es hat sich gezeigt, dass die Ergebnisse der beiden Modelle sehr ähnlich sind und lediglich vernachlässigbare Unterschiede im Millimeterbereich aufweisen. In Tabelle 6.2 sind die Standardabweichungen jeweils in x- und y-Richtung sowie der Positionsfehler relativ zur Objektdistanz angegeben. Die Mittelwerte der entsprechenden Abweichungen sind nahezu Null. Aufgrund der Gleichwertigkeit der Ergebnisse wird hier der einfachere Ansatz verwendet, um unnötigen Rechenaufwand zu vermeiden.

Zur Veranschaulichung sind die x- und y-Abweichungen dieses Modells in Abbildung 6.5 dargestellt. Die meisten der 3500 prädierten Messwerte liegen sehr nah an den tatsächlichen Messwerten, jedoch gibt es einige größere Ausreisser. Diese können durch Messfehler des Radarsensors bzw. einen Wechsel des Radarechos verursacht werden.

Tabelle 6.2: Evaluierung der Radarprädiktion

Modell	$\sigma_x[m]$	$\sigma_y[m]$	rel. Fehler
Ansatz 1	0,1471	0,2241	0,00344
Ansatz 2	0,1494	0,2222	0,00343

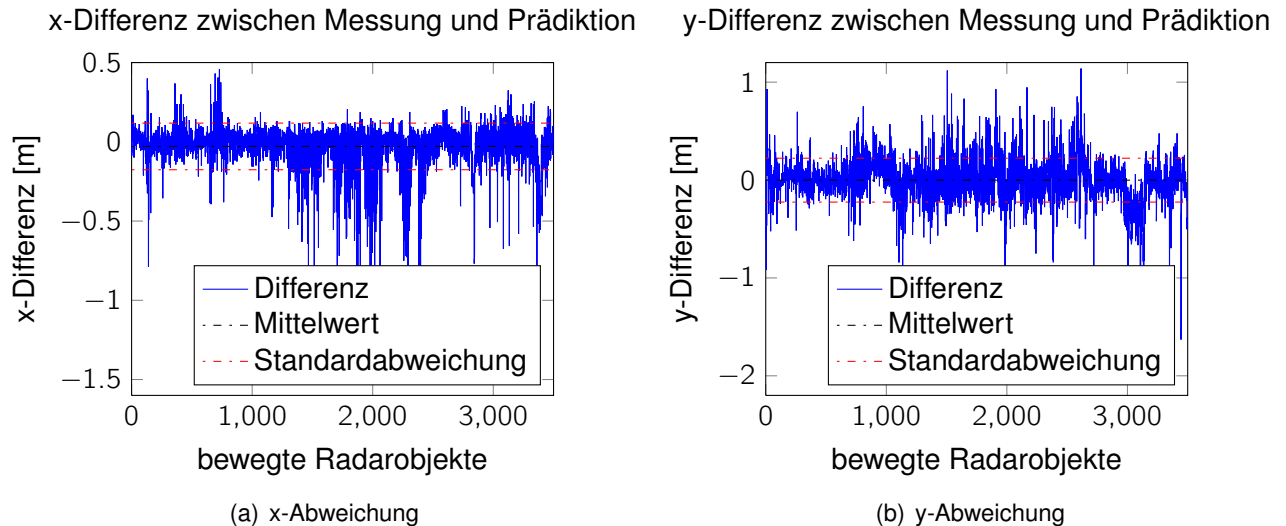


Abbildung 6.5: Abweichung zwischen prädizierten und gemessenen Radarwerten

Damit lässt sich abschließend sagen, dass die Prädiktion der Radardaten aufgrund der hier vorliegenden Sensorkombination und den erzielten Genauigkeiten sinnvoll ist. Zum einen sind die gemessenen Positionswerte des Radar-Sensors deutlich ungenauer als die des LIDAR-Sensors, wodurch diese ohnehin nicht so stark gewichtet werden. Zum anderen sind die durch die Prädiktion verursachten Ungenauigkeiten deutlich kleiner als z.B. der Sprung des Radarechos von einer Ecke zum Motorblock eines Fahrzeugs.

6.3 Umsetzung der Fusion

Nachdem in den vorangegangenen Kapiteln die Grundlagen und die nötigen Kalibrierungen für eine Sensorfusion erläutert wurden, geht es nun um die in dieser Arbeit umgesetzte Fusion. Wie aus Abschnitt 6.1.2 bekannt, gibt es unterschiedliche Ebenen auf denen eine Datenfusion durchgeführt werden kann. In diesem Fall ist jedoch bei dem Radarsensor kein Zugriff auf eine beliebige Stelle der Datenverarbeitungskette möglich. Es stehen lediglich die ausgegebenen Objekte zur Verfügung. Damit wäre von den vorgestellten Fusionsarten nur eine Fusion auf Symbolebene möglich. Diese zeichnet sich zwar durch den geringsten Aufwand aus, liefert aber auch den geringsten Mehrwert. Eine solchen Fusion ist es aber nur sinnvoll, wenn die Sensoren in ihrem Sensorbereich eigenständig gute Ergebnisse liefern. Allerdings treten beim LIDAR-Scanner häufig unvollständige Detektionen aufgrund zu geringer Remission der Objekte auf. Deshalb wird in dieser Arbeit eine komplexere Fusion verwendet, die nicht nur die Ergebnisse der einzelnen Objekterkennungen von LIDAR-Scanner (Kapitel 4) und Radarsensor (Kapitel 5) kombiniert, sondern auch die LIDAR-Datenverarbeitung unterstützt und verbessert. Dabei werden die Radardaten auf Symbolebene mit den Lidardaten auf Signalebene kombiniert. Aufgrund der Zugriffsbeschränkung auf die Radarverarbeitung ist diese Kombination die frühest mögliche. Da die Sensordaten von unterschiedlichen Abstraktionsebenen fusioniert werden, wird diese auch als gemischte Fusion bezeichnet.

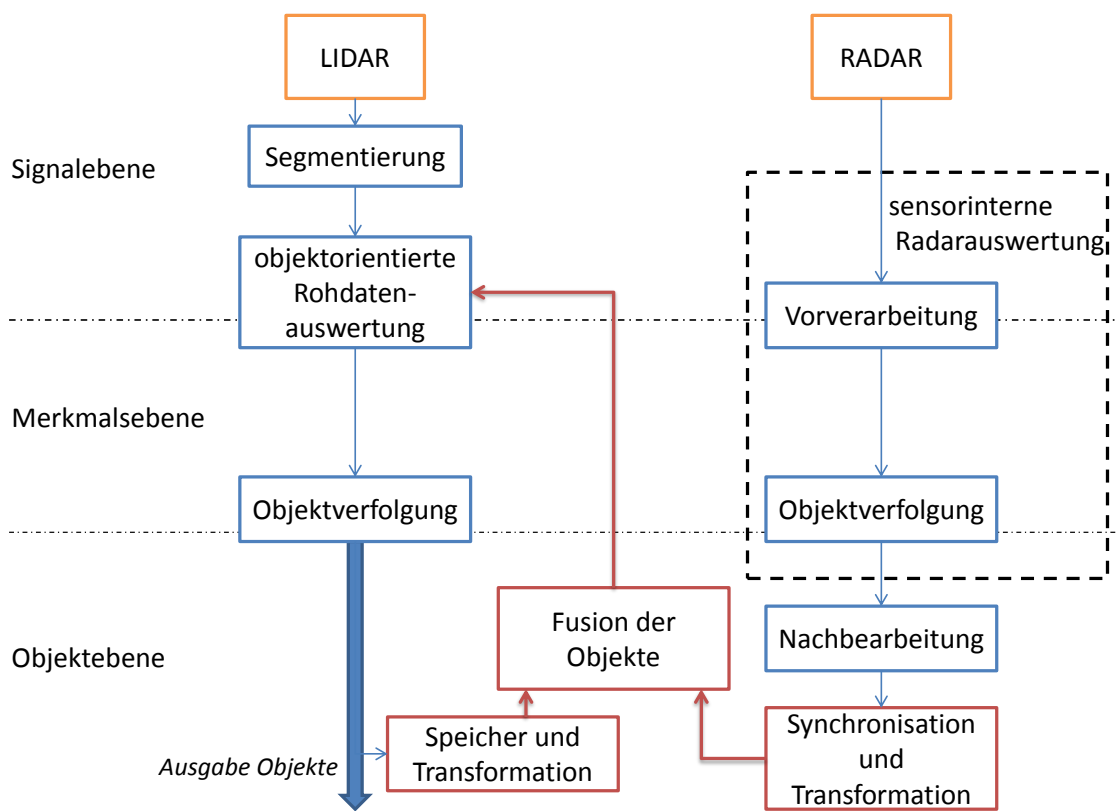


Abbildung 6.6: Funktionsweise der verwendeten Fusion

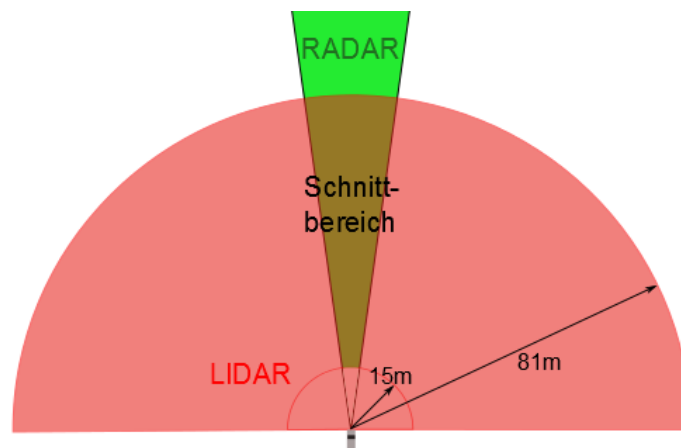


Abbildung 6.7: Sensorbereiche bei der Datenfusion in der Draufsicht

In Abbildung 6.6 ist die Funktionsweise der hier verwirklichten Fusion dargestellt. Darin sind die Signalverarbeitungsketten des LIDAR-Scanners und des Radarsensors eingezeichnet. Zusätzlich sind auch die in Abschnitt 6.1.2 eingeführten Fusionsebenen dargestellt. Da in der sensorinternen Radarauswertung keine Daten abgegriffen werden können, befindet sich der früheste Eingriffspunkt bereits in der Objektebene. Die durchgeführte Nachbearbeitung ist in Kapitel 5 erläutert. Im darauf folgenden Schritt, der Synchronisation und Transformation, werden die räumlichen und zeitlichen Unterschiede zwischen den Messwerten berücksichtigt, wie in Abschnitt 6.2 beschrieben. Die LIDAR-Auswertung besteht aus den aus Kapitel 4 bekannten Schritten.

Die jeweiligen Verarbeitungsschritte werden jedoch immer nur beim Erhalt neuer Messdaten ausgeführt. Aufgrund der kürzeren Abtastrate bestimmt der LIDAR-Scanner den Takt. Somit wird die in der Abbildung linke Verarbeitungskette, inklusive der Rückkopplung der Objekte, alle $13,32\text{ms}$ ausgeführt, während Nachbearbeitung, Synchronisation und Fusion der Radarobjekte nur bei einer neuen Radarmessung zusätzlich bearbeitet werden.

Wie in der Abbildung zu erkennen, handelt es sich grundsätzlich um eine Fusion auf Objektebene. Die Besonderheit ist jedoch, dass die fusionierten Objekte direkt in die LIDAR-Bearbeitung einfließen. Dabei werden die aufbereiteten und synchronisierten Radarobjekte mit den Objekten des letzten Zeitschritts fusioniert. Die Qualität der in Abschnitt 4.4 vorgestellten, objektorientierten Datenauswertung hängt schließlich von der Güte der bereits erkannten Objekte ab. Durch das zusätzliche Vorwissen aufgrund der Integration des Radarsensors, kann die LIDAR-Auswertung verbessert werden. Auf das genaue Vorgehen bei dieser Fusion wird in den folgenden Abschnitten eingegangen.

6.3.1 Sensorbereiche

Aus der Vorstellung der unterschiedlichen Fusionsarten in Abschnitt 6.1.1 ist bekannt, dass Sensoren sowohl konkurrierend als auch komplementär fusioniert werden können. Bei dieser Arbeit werden die Sensoren komplementär fusioniert. Eine konkurrierende Fusion wäre ohnehin nur im Schnittbereich der beiden Sensoren möglich. Dieser Bereich ist allerdings aufgrund des geringen Öffnungswinkels des Radarsensors relativ klein, wie in der Abbildung 6.7 zu sehen. Dort sind die Sensorbereiche des LIDAR-Scanners und des Radarsensors aus der Vogelperspektive dargestellt. Die komplementäre Fusion bietet sich außerdem für die hier vorliegende Problemstellung an, da fälschlicherweise erkannte Objekte bei einem Predictive Display deutlich unproblematischer als beispielsweise bei einem automatischen Bremssystem sind. Deshalb ist eine gegenseitige Absicherung der Messwerte durch unterschiedliche Sensoren nicht unbedingt notwendig und durch die vorliegenden Sensoren auch nicht sinnvoll umsetzbar.

In der Abbildung sind drei verschiedene Bereiche eingezeichnet, deren Färbung die dort verwendeten Sensoren abbildet. Im rot eingefärbten Bereich werden die LIDAR-Daten, im hellgrün gefärbten die Radar-Daten und im Schnittbereich die Messwerte beider Sensoren verwendet. Aufgrund des Messprinzips bzw. der Verarbeitung ist jedoch die Güte der Messwerte nicht auf dem ganzen Bereich konstant. So nimmt beim LIDAR mit zunehmendem Abstand zum Sensor die Detektionswahrscheinlichkeit stark ab, da immer weniger Infrarotlicht zum Sensor reflektiert wird. Aufgrund des Abtastprinzips steigt mit größerer Entfernung auch der Abstand zwischen den einzelnen Lichtimpulsen wodurch Objekte mit immer weniger Messpunkten erfasst werden. Beim Radarsensor ergeben sich durch die ACC-spezifische Datenverarbeitung vor allem Probleme bei Querverkehr oder abbiegendem Verkehr. Aus diesen Gründen werden im Nahbereich bis $15m$ vor dem Versuchsfahrzeug nur die LIDAR-Daten verwendet.

6.3.2 Fusion der Objekte

In diesem Arbeitsschritt müssen die Objekte der Radarauswertung mit denen der Lidarauswertung des letzten Zeitschritts kombiniert werden. Dadurch kann bereits die aktuelle Lidarauswertung unterstützt werden. Bei den vom Radarsensor erkannten Objekten werden zunächst nur die über mehrere Messungen bewegte Objekte verwendet und deren Orientierung abgeschätzt (siehe Kapitel 5). Somit stehen die gemessene Position und Relativgeschwindigkeit, sowie die vom Kalman-Filter der internen Radarsensorik ermittelte Relativbeschleunigung und die in der Nachbearbeitung berechnete Orientierung zur Verfügung. Wie in Abschnitt 6.2.2 beschrieben, werden diese Objektzustände noch prädiert, um die relative Sensorlatenz des Radars zu kompensieren. Die LIDAR-Auswertung ermittelt von den Objekten zusätzlich noch Länge, Breite und Drehrate. Im Gegensatz zum Radar wird die Orientierung gemessen und die Geschwindigkeit über einen Kalman-Filter ermittelt.

Um die Objekte der beiden Sensoren fusionieren zu können, muss zunächst untersucht werden, ob diese das selbe reale Objekt abbilden. Damit gibt es grundsätzlich drei verschiedene Möglichkeiten. Ein reales Objekt wird nur vom LIDAR-, nur vom Radar- oder von beiden Sensoren erfasst. In den Gebieten, in denen nur ein Sensor aktiv ist, ist die Zuordnung klar. Im Schnittbereich beider Sensoren muss jedoch ermittelt werden, ob zwei Sensorobjekte der unterschiedlichen Sensoren das selbe reale Objekt darstellen. Diese Problematik ähnelt sehr stark der Messwertassoziation in den Abschnitten 3.2 und 4.3.4. Auch hier muss die Assoziation so gewählt werden, dass zum einen zusammengehörende Objekte assoziiert werden, aber keine Objekte fälschlicherweise verschmelzen.

Die hier durchgeführte Assoziation läuft in zwei Schritten ab. Im ersten Schritt wird ein einfaches Nearest-Neighbour-Verfahren angewendet. Dabei wird um jedes Radarobjekt ein Toleranzbereich betrachtet und überprüft, ob ein Lidarobjekt innerhalb des Bereichs liegt. Die Radarobjekte werden durch Rechtecke mit $3 \times 1m$ repräsentiert (siehe Kapitel 5) und alle Lidarobjekte deren Abstand zum Rand kleiner als $1.5m$ ist, werden mit ihrem Abstand in eine Assoziationsmatrix eingetragen. In dieser Matrix steht jede Zeile für ein Lidarobjekt, jede Spalte für ein Radarobjekt und ein eingetragener Wert für ein mögliches Assoziationspaar. Auch wenn dieser Bearbeitungsschritt in einigen Fällen bereits für eine eindeutige Zuordnung ausreicht, entstehen bei nahe beieinander liegenden Objekten mehrere mögliche Zuordnungen. Da es sich hierbei um ein bekanntes Problem bei der Objektverfolgung handelt, gibt es relativ aufwendige mathematische Ansätze um dieses Problem zu lösen, wie beispielsweise die Ungarische Methode ([Kuh56]). Dabei wird die Positionsdifferenz jedoch nur für die erste Einteilung verwendet und somit muss daraus noch keine eindeutige Zuordnung berechnet werden.

Im zweiten Schritt wird die Assoziation, falls notwendig, mit weiteren Objektattributen verfeinert. Dafür wird zunächst überprüft, wie viele Lidarobjekte einem Radarobjekt zugeordnet wurden. Befindet sich mehr als ein aktuell gemessenes Lidarobjekt im Toleranzbereich, werden zwei Lidarobjekte $obj_{L,1}$ und $obj_{L,2}$ mit dem Radarobjekt obj_R verglichen. Als Vergleich wird jetzt jedoch nicht mehr die eukli-

dische Distanz herangezogen, sondern über

$$\varsigma(obj_1, obj_2) = ((x_1 - x_2)^2 + (y_1 - y_2)^2) \frac{1}{m^2} + 10(\psi_1 - \psi_2)^2 \frac{1}{rad^2} + (v_1 - v_2)^2 \frac{m^2}{s^2} \quad (6.5)$$

eine quadratische Distanz ς im Zustandsraum der Objekte berechnet. Dabei werden neben der Position auch die Orientierung und die Geschwindigkeit der Objekte berücksichtigt. Die Differenz der Orientierungen wird jedoch 10-fach gewichtet um den Unterschied der Größenordnung auszugleichen. Die zwei der drei Objekte mit dem kleinsten ς werden assoziiert. Falls

$$\varsigma(obj_{L,1}, obj_{L,2}) < \varsigma(obj_R, obj_{i,2}), \text{ mit } i = \{1, 2\} \quad (6.6)$$

gilt, also die beiden Lidarobjekte am ähnlichsten sind, werden diese zu einem Objekt verschmolzen und dann dem Radarobjekt zugeordnet. Befinden sich mehr als zwei Lidarobjekte im Toleranzbereich eines Radarobjekts, wird dieses Vorgehen wiederholt, bis nur noch ein Objekt übrig geblieben ist.

Ist einem Radarobjekt von Anfang an oder durch das beschriebene Ausschlussverfahren genau einem Lidarobjekt zugeordnet, wird diese Assoziation gespeichert. Wie in Abbildung 6.6 dargestellt, wird diese Information an die LIDAR-Auswertung übergeben und im nächsten Abschnitt näher erläutert.

Schließlich gibt es noch den Fall, dass einem Radarobjekt kein Lidarobjekt zugeordnet werden kann. Gerade in größeren Entfernungen in der Nähe der maximalen LIDAR-Reichweite tritt dieser Fall sehr häufig auf. Aus diesen Radarobjekten werden neue Lidarobjekte mit den entsprechenden Zustandsgrößen initialisiert. Da die LIDAR-Verarbeitung mit einer höheren Frequenz wie die Radarfusion läuft, ist es wichtig diese Objekte zu übergeben. Die zusätzlichen Objekte und Messdaten des Radars werden anschließend zur Verbesserung der LIDAR-Verarbeitung verwendet.

6.3.3 Integration in die Lidar-Objekterkennung

Die Grundidee bei der in dieser Arbeit entwickelten Umfelderkennung ist die objektorientierte Interpretation der verfügbaren Sensordaten. Wie bereits in Kapitel 4 erläutert, können vor allem LIDAR-Daten durch Verwendung von Vorwissen deutlich effektiver interpretiert werden, als lediglich auf Basis der Rohdaten. Deshalb werden bei der LIDAR-Auswertung die vom letzten Zeitschritt bekannten Objekte zur Verarbeitung herangezogen. Je genauer und je umfangreicher das verfügbare Vorwissen ist, desto besser können die Messdaten ausgewertet werden und dementsprechend viele Informationen aktualisiert werden. Deshalb werden die Informationen der Radarverarbeitung mit dem Vorwissen aus dem letzten Zeitschritt fusioniert. Damit wird die LIDAR-Verarbeitung durch die Fusion nicht nur ergänzt sondern auch deren Ergebnisse verbessert.

Wie im vorherigen Abschnitt beschrieben gibt es zwei verschiedene Fälle in denen Radarinformationen in die LIDAR-Verarbeitung einfließen. Zum einen bei den sogenannten radargestützten Objekten, die bereits aus dem letzten LIDAR-Zeitschritt bekannt sind und mit einem Radarobjekt fusioniert werden. Zum anderen bei den radarinitialisierten Objekten, die aufgrund von der Radarmessung neu initialisiert werden. Entscheidend ist nun, wie diese Informationen in die Verarbeitung integriert werden. Dafür wird die aus Kapitel 4 bekannte LIDAR-Verarbeitung um weitere Formen ergänzt. Es werden die radargestützten Formen IR-, LR- und ER- sowie die radarinitialisierte R-Form eingeführt. Aus Gründen der Konsistenz wird hier weiterhin von Formen gesprochen, obwohl diese Bezeichnungen eher für Repräsentationen mit bestimmten Eigenschaften stehen. Die Form eines Objekts legt fest wie danach in den Rohdaten gesucht wird und wie die Objektattribute interpretiert werden. Für diese neu eingeführten Formen stehen nun zusätzliche Informationen aufgrund der Radardaten zur Verfügung. Diese werden im Folgenden erläutert.

Zunächst einmal wird die Annahme getroffen, dass es sich bei Radarobjekten immer um Fahrzeuge oder Zweiräder handelt. Aufgrund den Voraussetzungen von einer Geschwindigkeit über $4 \frac{m}{s}$ und der

Erzeugung eines Radarechos durch elektrisch leitfähige Bestandteile können Fußgänger bzw. statische Objekte mit hoher Wahrscheinlichkeit ausgeschlossen werden. Durch diese Annahme können radarinitialisierte Objekte sofort klassifiziert werden und somit von Anfang an das entsprechende Bewegungsmodell für den Kalman-Filter verwendet werden. Wird ein Objekt radargestützt, so wird es ebenfalls als Fahrzeug klassifiziert. Die einzige Ausnahme sind bereits klassifizierte Zweiräder. Diese behalten ihre Klasse. Aufgrund des identischen Bewegungsmodells ist eine exakte Differenzierung zwischen Fahrzeugen und Zweirädern nicht unbedingt notwendig und mit den vorhandenen Sensoren auch nicht ohne weiteres möglich.

Die zweite wichtige Eigenschaft der Radarobjekte ist die Messbarkeit der Relativgeschwindigkeit. Deshalb wird bei Radarobjekten in den Lidarzeitschritten, in denen eine neue Radarmessung vorliegt, auch die gemessene Geschwindigkeit dem Kalman-Filter übergeben. Der größte Vorteil der Geschwindigkeitsmessung liegt aber vor allem bei der Initialisierung von neuen Objekten, da deren Geschwindigkeit von Anfang an berücksichtigt werden kann.

Abschließend stehen auch die vom Radar gemessenen Positionswerte zur Verfügung. Diese werden jedoch nur für den Fall einer ungenügenden LIDAR-Erkennung weiter verwendet. Der Grund für diese restriktive Fusion der Positionsdaten liegt in den vielen Ungenauigkeiten die mit der Radarposition verbunden sind. Neben der im Vergleich zum LIDAR deutlich schlechteren Winkelauflösung sowie der schlechteren Distanzgenauigkeit, kommen bei den verwendeten Radarmessungen noch zwei weitere Unbekannte dazu. Einerseits ist der genaue Ort des Radarechos nicht definiert, sondern kann theoretisch jedes größere elektrisch leitfähiges Bauteil des Objekts sein. Andererseits werden die Positionswerte um einen Abtastschritt des Sensors präzisiert, um mit den LIDAR-Werten verglichen werden zu können. Aus diesen Gründen werden die Positionsdaten des Radars nur bei den reinen Radarobjekten und für die Assoziation verwendet. Sobald mindestens zwei LIDAR-Messpunkte zur Verfügung stehen, wird die Position allein aufgrund dieser Daten ermittelt.

6.3.4 Fazit

Die hier verwendete Datenfusion kombiniert nicht nur die aus den Sensordaten berechneten Objekte, sondern verbessert durch sinnvolle Rückführung auch die LIDAR-Auswertung. Im Vergleich zur reinen Fusion auf Objektebene ist der zusätzliche Berechnungsaufwand überschaubar, aber trotzdem ein deutlicher Zusatznutzen erkennbar. Der zusätzliche Berechnungsaufwand wird lediglich durch die nun zusätzlich erkannten Objekte in der LIDAR-Verarbeitung verursacht. Die Assoziation kann bei beiden Fusionsarten identisch ausgeführt werden.

Durch die entwickelte Fusion entsteht der große Vorteil, dass vor allem Gegenverkehr frühzeitig vom Radar erkannt wird und durch die Einspeisung in die LIDAR-Verarbeitung beim Verlassen des Radarbereichs weiter verfolgt werden kann. Ein besseres Zusammenspiel der Sensoren wird leider durch die abgekapselte Radarverarbeitung verhindert. Dabei erschwert die ACC-Fokussierung der Sensorauswertung auch eine optimale Sensornutzung für die allgemeine Umfelderkennung.

7 Prädiktion

In den vorangegangenen Kapiteln wurde ausführlich auf die Erfassung des Umfelds eingegangen. Dabei wird die aktuelle Umgebung durch Sensoren erfasst, vorhandene Objekte identifiziert und deren Bewegungszustand durch Kalman-Filter angenähert. In dieser Arbeit soll jedoch die Zeitverzögerung der Übertragung beim teleoperierten Fahren kompensiert werden. Dafür sollen die zukünftigen Positionen der umgebenden Verkehrsteilnehmer vorausgesagt werden. Allerdings gilt dabei das berühmte Zitat

„Prognosen sind schwierig, besonders wenn sie die Zukunft betreffen.“

das dem Physik-Nobelpreisträger Niels Bohr (1885-1962) zugeschrieben wird.

7.1 Grundlagen der Prädiktion

Vorhersagen und Prädiktionen sind natürlich per Definition mit einer gewissen Unsicherheit behaftet. Es können jedoch viele Vorgänge durch Modelle beschrieben werden und von dem Verhalten dieser Modelle auf die zukünftige Entwicklung geschlossen werden. Je genauer die Abbildung des realen Vorgangs ist, desto genauer ist auch die vom Modell abgeleitete Prädiktion.

Die entscheidenden Größen sind die Komplexität des Systems, die Genauigkeit des Modells und die Zeitspanne der Prädiktion.

Komplexität und Systemverständnis Je genauer das Verhalten eines Systems beschrieben werden soll, desto besser müssen die Zusammenhänge und die Funktionsweise verstanden werden. Dabei spielt die Komplexität eine wichtige Rolle. Je einfacher ein System ist, desto eher können alle relevanten Vorgänge erkannt werden. So sind beispielsweise die Wechselwirkungen bei einem sehr einfachen mechanischem System wie einem Ein-Massen-Schwinger sehr genau bekannt. Wird dagegen das Wetter betrachtet, ein klassisches Beispiel für ein prognostiziertes System, sind die Zusammenhänge so komplex, dass selbst durch jahrhundert lange Forschung nur ein stark vereinfachtes Systemverständnis möglich ist. Da komplizierte Systeme selten bis ins letzte Detail verstanden werden können, ist es auch fundamental wichtig die verbleibenden Unsicherheiten adäquat zu berücksichtigen.

Genauigkeit des Modells Die Genauigkeit des Modells wird durch das Systemverständnis sowie die Modellierbarkeit der Vorgänge beschränkt. Wie bereits erwähnt, gehören zu dem Systemverständnis auch Unsicherheiten, die entsprechend modelliert werden können. Jedoch ist es keineswegs immer zielführend einen Vorgang in der maximal möglichen Detailtiefe abzubilden. Der sowohl für die Modellierung, als auch für die Berechnung nötige Aufwand ist oft der entscheidende Faktor bei der Wahl der Modellgenauigkeit. Dabei muss ein sinnvoller Kompromiss zwischen den Genauigkeitsanforderungen und dem Berechnungsaufwand gefunden werden.

Prädiktionszeit Die Prädiktionszeit beschreibt die Zeitspanne, für die das Systemverhalten vorhergesagt werden soll. Je größer diese Zeit ist, desto schwieriger wird die Prädiktion. Die geforderte Prädiktionszeit hat auch großen Einfluss auf die Detailtiefe und den Rechenaufwand des Modells.

7.2 Systembetrachtung

Die Definition sinnvoller Systemgrenzen ist für eine praktikable Prädiktion fundamental wichtig. Das Ziel dieser Arbeit ist die Prädiktion der Bewegung von Verkehrsteilnehmern die sich im Umfeld des Versuchsfahrzeugs befinden. Wird beispielsweise das gesamte Umfeld als System betrachtet, so wird schnell klar, dass dessen Komplexität nicht sinnvoll modelliert werden kann. Da sich das Versuchsfahrzeug bewegt, ändert sich laufend die umgebende Infrastruktur, in der sich die Verkehrsteilnehmer bewegen. Diese Verkehrsteilnehmer stehen zueinander in Wechselwirkungen und werden von unzähligen Faktoren beeinflusst. Das komplette Verhalten eines solchen Systems vorherzusagen ist aufgrund der vielen unbekannt Einflüsse schlichtweg nicht möglich. Das menschliche Gehirn ist nach vielen Jahren des Trainings im Straßenverkehr fähig, die Intentionen anderer Verkehrsteilnehmer zu erraten. Allerdings ist die Beurteilung einer solch komplexen Situation ein Gebiet, in dem es noch sehr lange dauern wird bis Maschinen ähnlich leistungsfähig werden wie der Mensch. Und selbst die menschliche Beurteilung liefert oft suboptimale Ergebnisse, wie wahrscheinlich die meisten aus der eigenen Erfahrung wissen. Aus diesem Grund muss die Komplexität des betrachteten Systems auf ein sinnvolles Maß reduziert werden.

Deshalb wird angenommen, dass alle Verkehrsteilnehmer während der Prädiktionszeit ihre momentane Aktion fortsetzen. Diese Vereinfachung ist zwar durch die fehlende Intensionserkennung erzwungen, aber bei kleineren Prädiktionszeiten durchaus angemessen, da bei der Übertragung Zeitverzögerungen von maximal $0,5s$ kompensiert werden sollen. Abrupte Ausweich- oder Bremsvorgänge können damit allerdings nicht abgebildet werden.

7.3 Modellierung

Die Modellierung des Systems muss zwei wichtige Bedingungen erfüllen. Einerseits soll das Systemverhalten möglichst gut nachgebildet werden und andererseits muss die Berechnung echtzeitfähig sein. Aufgrund der im vorhergehenden Abschnitt erläuterten Vereinfachung müssen lediglich die Bewegungen der Verkehrsteilnehmer modelliert werden. Dabei wird noch angenommen, dass diese Bewegungen durch Markov-Ketten beschrieben werden können. Für diese Markov-Ketten (auch als Markov-Prozesse bezeichnet) gilt, dass der nächste Zustand nur vom aktuellen Zustand abhängt. Das bedeutet insbesondere, dass eine Vorhersage nicht durch die zusätzliche Betrachtung der Historie verbessert werden kann. Der Vorteil dieser Annahme ist, dass der Rechen- und Speicheraufwand drastisch reduziert wird, da alle vergangenen Werte vernachlässigt werden können.

Die Beschreibung der Objektbewegungen ist nicht nur bei der Prädiktion notwendig, sondern auch bei der Konfiguration des Kalman-Filters zur Objektverfolgung. Dementsprechend sind mehrere mögliche Bewegungsmodelle bereits im Abschnitt 4.4.4 diskutiert worden. Der Unterschied zwischen der Prädiktion, die im Kalman-Filter der LIDAR-Verarbeitung stattfindet, und der hier gesuchten Prädiktion ist lediglich die Prädiktionszeit. Bei der Objektverfolgung wird der Objektzustand für den nächsten Messschritt in $13,32ms$ prädiziert. Für die Kompensation der Übertragungszeiten müssen allerdings Zeiten bis zu $500ms$ überbrückt werden. Wie bereits erwähnt, soll bei der Prädiktion lediglich der aktuelle Zustand fortgeführt werden, da mögliche Stelleingriffe nicht ausreichend vorhergesagt werden können. Weil die ermittelten Zustandsgrößen der Objekte vom Bewegungsmodell des Kalman-Filters abhängen, bietet es sich an, die selben Modelle für die Prädiktion zu verwenden.

Allerdings muss überprüft werden, ob die beiden Bewegungsmodelle aus Abschnitt 4.4.4 auch für die deutlich längere Prädiktionszeit noch gültig sind. Das Bewegungsmodell für Fußgänger und unbekannte Objekte in Gleichung (4.20) lässt sich direkt aus den beschreibenden Differentialgleichungen $\dot{x} = v_x$ und $\dot{y} = v_y$ herleiten. Aufgrund der Linearität des Modells hat die Zeitdauer keinen Einfluss auf die Genauigkeit. Diese hängt nur von der Güte der ermittelten Zustandsgrößen ab. Allerdings wird bei diesem Modell eine konstante Geschwindigkeit über den jeweiligen Zeitschritt angenommen. Diese Annahme wird mit zunehmender Dauer jedoch immer unrealistischer. Da beim Prädiktionsmodell

nur die erfassten Zustandsgrößen verwendet werden können, müsste auch das Kalman-Filter-Modell vergrößert werden. Deshalb gilt es abzuwägen, ob dieser deutlich höhere Aufwand sinnvoll ist. Da nur klassifizierte Objekte prädiziert werden sollen, gilt dieses Bewegungsmodell nur für Fußgänger. Bei der Betrachtung der Bewegungen von Fußgängern fällt auf, dass diese selten längere Beschleunigungsphasen durchlaufen. Im Normalfall werden Phasen annähernd konstanter Geschwindigkeit durch relativ abrupte Geschwindigkeitswechsel begrenzt. Aus diesem Grund ist die Annahme konstanter Beschleunigungen nicht unbedingt besser als die konstanter Geschwindigkeiten. Deshalb wurde hier das einfachere Modell verwendet. Da lediglich die Position prädiziert werden soll und Fußgänger keine Orientierung aufweisen, ergibt sich für die Fußgängerprädiktion

$$\begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_p = \begin{bmatrix} 1 & 0 & \tau & 0 \\ 0 & 1 & 0 & \tau \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}_k \quad (7.1)$$

Dabei steht der Index $_p$ für die prädizierten Zustände, $_k$ für die aktuell erfassten Zustände und τ für die Prädiktionszeit. Es hat sich gezeigt, dass das einfachere Modell für die Verfolgung und Prädiktion von Fußgängern bereits ausreichende Ergebnisse liefert. Dementsprechend wurde auf eine Evaluation der rechenaufwendigeren Variante mit zusätzlich verfolgter Beschleunigungen verzichtet.

Bei dem Bewegungsmodell für Fahrzeuge und Zweiräder ist die Ausgangslage jedoch anders. Um auf die bei der Objektverfolgung verwendete Differenzgleichung (4.21) zu kommen, wurde eine Linearisierung vorgenommen. Aufgrund der hohen Abtastrate des LIDAR-Scanners ist diese Vereinfachung auch sinnvoll. Bei den deutlich längeren Prädiktionszeiten sind die dadurch verursachten Fehler allerdings nicht mehr akzeptabel. Deshalb wird für die Prädiktion der Fahrzeug- und Zweirad-Zustände folgendes, von Lindl [Lin08] hergeleitetes Modell verwendet:

$$\begin{bmatrix} x \\ y \\ \psi \\ v \\ \omega \\ a \end{bmatrix}_p = \begin{bmatrix} x \\ y \\ \psi \\ v \\ \omega \\ a \end{bmatrix}_k + \begin{bmatrix} \frac{v+a\tau}{\omega} \sin(\psi + \omega\tau) + \frac{a}{\omega^2} \cos(\psi + \omega\tau) - \frac{v}{\omega} \sin(\psi) - \frac{a}{\omega^2} \cos(\psi) \\ -\frac{v+a\tau}{\omega} \cos(\psi + \omega\tau) + \frac{a}{\omega^2} \sin(\psi + \omega\tau) + \frac{v}{\omega} \cos(\psi) - \frac{a}{\omega^2} \sin(\psi) \\ \omega\tau \\ a\tau \\ 0 \\ 0 \end{bmatrix} \quad (7.2)$$

Für $\omega \approx 0$ wird die Näherung

$$\begin{bmatrix} x \\ y \\ \psi \\ v \\ \omega \\ a \end{bmatrix}_p = \begin{bmatrix} x \\ y \\ \psi \\ v \\ \omega \\ a \end{bmatrix}_k + \begin{bmatrix} \cos(\psi)v\tau + \frac{1}{2}\cos(\psi)a\tau^2 \\ \sin(\psi)v\tau + \frac{1}{2}\sin(\psi)a\tau^2 \\ \omega\tau \\ a\tau \\ 0 \\ 0 \end{bmatrix} \quad (7.3)$$

angewandt.

Durch diese Bewegungsmodelle kann der erfasste Bewegungszustand der Verkehrsteilnehmer über die geforderte Prädiktionszeit fortgesetzt werden. Da die Modelle die meisten Bewegungen sehr gut abbilden können, ist vor allem die Güte der Zustandsverfolgung für die Qualität der Prädiktion entscheidend.

8 Systembewertung

Die in den vorherigen Kapiteln vorgestellte Umfelderkennung soll nun abschließend bewertet werden. Es geht dabei allgemein um die Beurteilung der Leistungsfähigkeit des Systems. Dafür gibt es je nach Zielsetzung unterschiedliche Möglichkeiten. Eine Evaluierung der erzielten Ergebnisse kann sowohl für das Gesamtsystem als auch für einzelne Untersysteme durchgeführt werden. Auch die Art der Bewertung hängt stark von den Anforderungen ab.

So ist beispielsweise bei der Entwicklung des Lidar-Erkennungsalgorithmus ein Vergleich mit den Messwerten ausreichend. Für eine schnelle Überprüfung der Funktionalität genügt eine einfache optische Darstellung der Messwerte und der daraus extrahierten Objekte. In Abbildung 8.1 ist die hier verwendete Visualisierung gezeigt. Diese Art der Darstellung wurde bereits in Kapitel 4 zur Beschreibung des Algorithmus verwendet. Die Lidarmesswerte sind als rote \times und die erkannten Fahrzeuge als blaue Objekte eingezeichnet. Um neben der Erkennung auch die Verfolgung des Bewegungszustands und die Prädiktion der Objektposition bewerten zu können, sind zusätzlich auch die prädizierten Fahrzeugpositionen (Prädiktionszeit $\tau = 0,5s$) als farbige Objekte eingezeichnet. Dabei stellen die gestrichelten Rahmen die gemessene Objektposition zum prädizierten Zeitpunkt dar und erlauben somit eine Einschätzung der Fahrzeugbewegung und der Güte der Prädiktion. Um die Beurteilung noch zu erleichtern, wechselt die Farbe der prädizierten Objekte von grün zu rot, je schlechter die prädizierte Position ist. Eine schlechte Prädiktion deutet auf eine ungenügende Erfassung des Bewegungszustand eines Objekts hin. Die dargestellte Situation wurde an einer Kreuzung aufgezeichnet, wobei das Versuchsfahrzeug wieder im Koordinatenursprung zu finden ist und drei entgegenkommende Linksabbieger detektiert werden. Durch diese grafische Darstellung kann nun die Umfelderkennung in diesem Zeitpunkt offensichtlich sehr schnell evaluiert werden. Alleine durch die Betrachtung der Darstellung können mehrere Schlüsse gezogen werden. So passen die Positionen, Orientierung und Dimensionen der erkannten Objekte zu den entsprechenden Messpunkten des Scanners. Durch den Vergleich der prädizierten Positionen mit den später gemessenen Positionen, können auch Abweichungen beim erfassten Bewegungszustand erkannt werden. Bei dem gerade erst detektierten Objekt Nr. 5 ist beispielsweise bisher nur die Geschwindigkeit, nicht aber die Drehrate erkannt worden. Während die Kurvenfahrt des mittleren Fahrzeugs sehr gut abgebildet wird, geht der Kalman-Filter beim letzten Fahrzeug noch von einer Fortsetzung der Kurvenfahrt aus. Auch wenn diese Visualisierung äußerst praktikabel und effizient für die schnelle Überprüfung der Algorithmusergebnisse ist, kann damit keine unabhängige Systembewertung durchgeführt werden.

8.1 Grundlagen der Systembewertung

Für eine objektive Beurteilung sind deshalb mehrere Schritte nötig. In Abbildung 8.2 (aus [Lin08]) sind diese übersichtlich dargestellt. Der Algorithmus wertet die Sensordaten aus und dessen Ergebnisse sollen bewertet werden, weshalb diese als Evaluierungsdaten bezeichnet werden. Um die Evaluierung durchführen zu können, müssen Referenzdaten zur Verfügung stehen. Diese können auf unterschiedliche Weise erzeugt werden und teilweise auch auf den selben Sensordaten basieren. Diese beiden Datensätze werden anschließend über Metriken verglichen und durch aussagekräftige Gütemaße beschrieben. Im Folgenden wird zunächst auf die Möglichkeiten zur Referenzdatenerzeugung und sinnvolle Gütemaße bei der Umfelderkennung genauer eingegangen (nach [Lin08]). Anschließend wird die hier durchgeführte Systembewertung dargestellt.

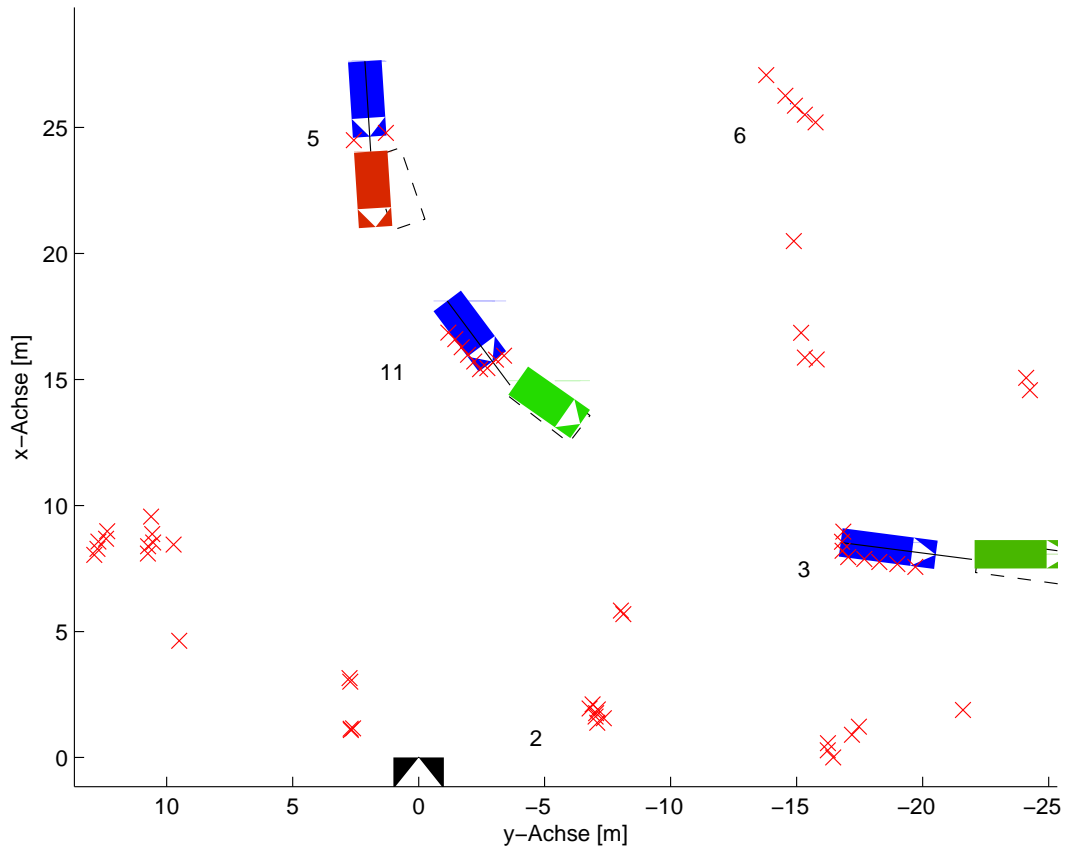


Abbildung 8.1: Grafische Darstellung zur optischen Evaluation

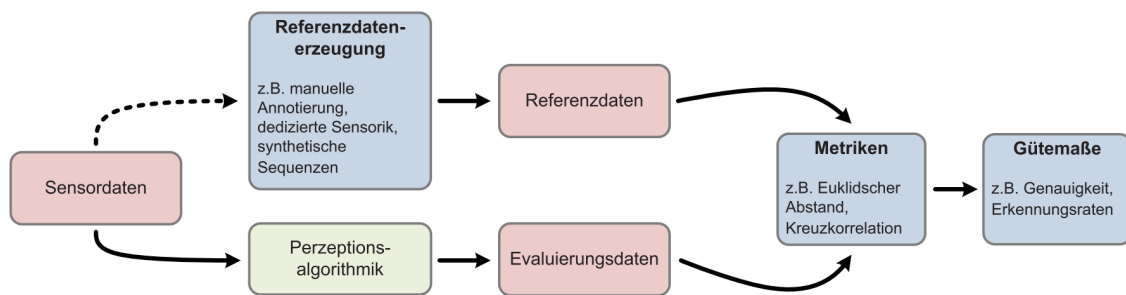


Abbildung 8.2: Übersicht über Systembewertung [Lin08, S. 139]

8.1.1 Erzeugung von Referenzdaten

Die Referenzdaten sollen dazu dienen, die vom System ermittelten Werte zu überprüfen. Dieses Vorgehen ist nur sinnvoll, wenn die Referenzdaten besser sind und im Idealfall auch die jeweiligen Unsicherheiten bekannt sind. Gerade bei komplexen Problemstellungen wie einer Umfelderkennung ist es schwierig hochgenaue Referenzdaten zu generieren. Lindl [Lin08] identifiziert dafür drei prinzipielle Möglichkeiten.

Simulation Bei diesem Ansatz wird das gesamte Testszenario simuliert. Da bei diesem Vorgehen alle Größen des Umfelds bekannt sind, stehen exakte Referenzdaten zur Verfügung. Allerdings hat dieser Ansatz auch gravierende Nachteile. Vor allem ist der Aufwand für eine solche Simulation sehr groß, da neben allen Objekten und deren Bewegungen auch die entsprechenden Sensordaten berechnet werden müssen. Je mehr Sensoren von dem zu prüfenden System ausgewertet werden, desto aufwendiger wird die Simulation. Außerdem können mit einer Simulation nicht alle real auftretenden Effekte berücksichtigt werden.

Referenzsensorik Die zweite Möglichkeit ist Objekte im Umfeld mit einer Referenzsensorik auszustatten. Wird beispielsweise ein anderes Fahrzeug ebenfalls mit einer IMU (siehe Abschnitt 2.3.4) ausgestattet und diese mit der IMU des Versuchsfahrzeugs synchronisiert, können die Bewegungen beider Fahrzeuge präzise aufgezeichnet werden. Durch die genaue Erfassung des Bewegungszustands inklusive der Beschleunigungen und Drehraten können diese Referenzdaten auch zur Beurteilung der Objektverfolgung verwendet werden. Mit Hilfe der bekannten Abmessungen des Referenzobjekts kann auch die Umfeldsensorik bewertet werden. Der große Nachteil liegt in dem hohen finanziellen Aufwand für die Referenzsensorik. Außerdem können jeweils nur Referenzdaten für einzelne Objekte generiert werden.

Attributierung Bei der dritten Variante werden die Referenzdaten nachträglich aus aufgezeichneten Messdaten gewonnen. So können beispielsweise mit Hilfe von Kamerainformationen auftretende Objekte von einem Benutzer klassifiziert werden. Während die Klassifizierung gut möglich ist, wird die Bestimmung der Objektpositionen schon deutlich aufwendiger, da diese für mehrere Zeitschritte nötig ist. Deshalb wird in [Lin08] ein Vorgehen vorgestellt, mit dem die Attributierung größtenteils von einem Algorithmus ausgeführt wird und der Benutzer lediglich die Ergebnisse überwacht. Der große Vorteil der nachträglichen Attributierung ist die einfache Anwendbarkeit, da weder Referenzsensoren noch genaue Sensormodelle zur Verfügung stehen müssen. Allerdings werden die Referenzdaten lediglich auf Basis der aufgezeichneten Messwerte generiert, wodurch die Bewertung nicht sensorunabhängig erfolgen kann.

Alle drei Möglichkeiten weisen jeweils spezifische Vor- und Nachteile auf und die Erzeugung von Referenzdaten ist, unabhängig von der gewählten Methode, mit großem Aufwand verbunden [Lin08]. Da bei dieser Arbeit keine Referenzsensorik verfügbar ist und die Simulation des gesamten Fahrumfelds den zeitlichen Rahmen sprengen würde, wurde hier die Methode der Attributierung verwendet. Bei dieser korreliert der Aufwand direkt mit der gewünschten Anzahl an Referenzgrößen und deren Genauigkeit. Damit kann die Evaluierung und der Aufwand gut an die Gegebenheiten angepasst werden.

8.1.2 Definition geeigneter Gütemaße

Gütemaße werden benötigt, um die Abweichungen der Evaluierungsdaten von den Referenzdaten anschaulich zu darzustellen. Das Ziel ist es dabei, die gesamte Leistungsfähigkeit eines Systems mit wenigen Kenngrößen zu beschreiben. Nach Lindl [Lin08] gibt es bei einer Umfelderkennung die drei Bereiche Genauigkeiten, Erkennungsraten und Laufzeiten, welche durch Gütemaße beschrieben werden müssen.

8.1.2.1 Erkennungsraten

Mit Hilfe von Erkennungsraten soll die Verlässlichkeit der Objektdetektion beschrieben werden. Werden die unterschiedlichen Objektklassen und mögliche räumliche Ungenauigkeiten vernachlässigt, liegt eine binäre Entscheidung vor. Entweder der Algorithmus detektiert an dieser Stelle ein Objekt oder nicht. Demnach können zwei Fehler auftreten. Ein existierendes Objekt wird nicht erkannt (Nicht-Erkennung) oder es wird ein Objekt erkannt, obwohl es in der Realität nicht existiert (Fehlerkennung). Da beide Fehlertypen bei der Bewertung des Systems von Interesse sind, werden dafür wie in [Lin08] zwei Gütemaße eingeführt.

Richtig-Positiv-Rate (RPR): Die Richtig-Positiv-Rate beschreibt, wie viele reale Objekte richtig erkannt werden. Sie wird durch

$$\text{RPR} = \frac{t_p}{n_{obj}} = \frac{\text{richtig erkannte Objekte}}{\text{Anzahl Objekte}} \quad (8.1)$$

berechnet. Darin ist n_{obj} die Gesamtanzahl an Verkehrsteilnehmern, die im betrachteten Zeitfenster im Sensorbereich auftreten. Um zu den richtig erkannten Verkehrsteilnehmern t_p zu gehören, muss ein Objekt innerhalb des Zeitfensters richtig klassifiziert worden sein und sich näher als $\delta = 1,5m$ am Referenzobjekt befinden. Durch die Forderung der korrekten Klassifizierung wird diese Rate nicht nur durch Nicht-Erkennung, sondern auch durch mögliche Falschklassifizierungen reduziert.

Fehlererkennungsrate (FER): Mit diesem Gütemaß soll die Anzahl der Fehlerkennungen f_n charakterisiert werden. Eine Fehlerkennung ist in diesem Fall ein Hintergrundobjekt, das fälschlicherweise als Fahrzeug, Fußgänger oder Fahrradfahrer klassifiziert wird. Allerdings gibt es in diesem Fall keine Referenzgröße wie bei der RPR, mit der dieser Wert normiert werden könnte. Aus diesem Grund wird f_n zeitlich normiert, sodass

$$\text{FER} = \frac{f_n}{Std} = \frac{\text{Fehlerkennungen}}{\text{Stunde}} \quad (8.2)$$

definiert wird.

Falsch-Klassifizierungs-Rate (FKR): Da bei der RPR auch die korrekte Klassifizierung der Objekte gefordert wird, ist es sinnvoll auch die Rate der Falsch-Klassifizierungen anzugeben. In Kombination mit der RPR kann nun auch auf die klassenunabhängige Objekterkennungsrate geschlossen werden.

8.1.2.2 Genauigkeiten

Während im vorherigen Abschnitt nur bewertet wurde, ob Objekte überhaupt erkannt und klassifiziert werden, soll mit den folgenden Gütemaßen beschrieben werden, wie genau diese Erkennung ist. Diese Betrachtung beinhaltet neben der Genauigkeit der Zustandsgrößen auch eine zeitliche Komponente.

Zeitliche Abdeckung (ZA): Zur Beurteilung dieser zeitlichen Komponente verwendet Lindl [Lin08] die Zeitliche Abdeckung. Diese berechnet sich wie folgt

$$\text{ZA} = \frac{\sum_i t_{kl,i}}{\sum_i l_i} = \frac{\text{klassifizierte Zeit}}{\text{Lebensdauer}} \quad (8.3)$$

Dabei ist $t_{kl,i}$ die Zeitdauer, die das Objekt i richtig erkannt und klassifiziert wurde. l_i ist die Lebensdauer des Objekts, d.h. die Zeit die sich das Objekt im Sensorbereich befindet.

Root Mean Square Error (RMSE) der Zustandsgrößen: Um die Abweichungen der Zustandsgrößen x von den Referenzgrößen \bar{x} zu beurteilen, bietet sich der Root Mean Square Error an (Wurzel

des mittleren quadratischen Fehlers):

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (\tilde{x}_t - x_t)^2} \quad (8.4)$$

Ist die mittlere Differenz der beiden Größen gleich Null, entspricht der RMSE der Standardabweichung. Durch diese allgemeinere Formel können mit dem RMSE beliebige Größen verglichen werden. Somit kann mit dem RMSE für alle verfügbaren Referenzgrößen der mittlere Fehler der korrespondierenden Objektzustände berechnet werden.

8.1.2.3 Erfassungsbereich und Rechenzeit

Abschließend müssen noch zwei weitere Faktoren berücksichtigt werden. Damit die bisher definierten Gütemaße nicht verfälscht werden, muss die Evaluierung auf den von den Sensoren erfassten Bereich beschränkt bleiben. Da Objekte außerhalb des Erfassungsbereichs nicht detektiert werden können, würden die Erkennungsraten andernfalls negativ beeinflusst. Bei Systemen, die auf mehrere Sensoren zugreifen, sind auch mehrere unabhängige Bewertungen von unterschiedlichen Bereichen denkbar. So können beispielsweise der Schnitt- und der Vereinigungs-Bereich einzeln untersucht werden.

Als letztes Gütemaß gilt es noch die Rechenzeit zu berücksichtigen. Da das Umfelderkennungssystem in Echtzeit ablaufen soll, ist die benötigte Rechenzeit von entscheidender Bedeutung. Im Idealfall ist die Rechenzeit des Systems so lang wie die kürzeste Abtastzeit der verwendeten Sensoren. Ein System zur Umfelderkennung kann als echtzeitfähig bezeichnet werden, wenn es etwa 15 Zyklusdurchläufe pro Sekunde erreicht ([Lin08]). Da die Rechenzeit stark von der verwendeten Hardware abhängt, wird diese immer in Kombination mit der entsprechenden Hardwarekonfiguration angegeben.

8.2 Bewertung der entwickelten Umfelderkennung

8.2.1 Optische Beurteilung

Wie bereits in der Einleitung dieses Kapitels erwähnt, wurde bei der Entwicklung des vorliegenden Algorithmus vor allem eine optische Systembewertung durchgeführt. Dabei wurde neben der in Abbildung 8.1 dargestellten Übersicht auch ein Kamerabild der Umgebung verwendet. Mit Hilfe der Kameraparameter und der daraus abgeleiteten Transformation können die detektierten Objekte in das Kamerabild eingeblendet werden. Die zusätzliche Betrachtung des Videobilds liefert einerseits eine gewohnte Repräsentation der Situation, wodurch die Klassen, Positionen und in gewissem Maße auch die Bewegungen der Verkehrsteilnehmer ersichtlich sind. Andererseits stammen diese Informationen von einem unabhängigen Sensor und somit können auch fehlende Detektionen des Lidar- oder Radar-Sensors erkannt werden.

In Abbildung 8.3 ist die grafische Darstellung der Erkennungs- und Prädiktions-Werte während einer Überlandfahrt abgebildet. Im Videobild (a) sind, neben den leicht angedeuteten Lidar-Messwerten, die erkannten Objektpositionen durch einen blauen Rahmen und die prädierten Positionen durch einen roten Rahmen gekennzeichnet. Um die Bewegung der Objekte noch zu verdeutlichen, sind die beiden Positionen durch einen Pfeil verbunden. In Abbildung 8.3 (b) ist die Situation wieder schematisch dargestellt. Offensichtlich können durch die Kombination beider Darstellungen die momentanen Ergebnisse der Umfelderkennung gut beurteilt werden. Allerdings ist diese Beurteilung sehr subjek-

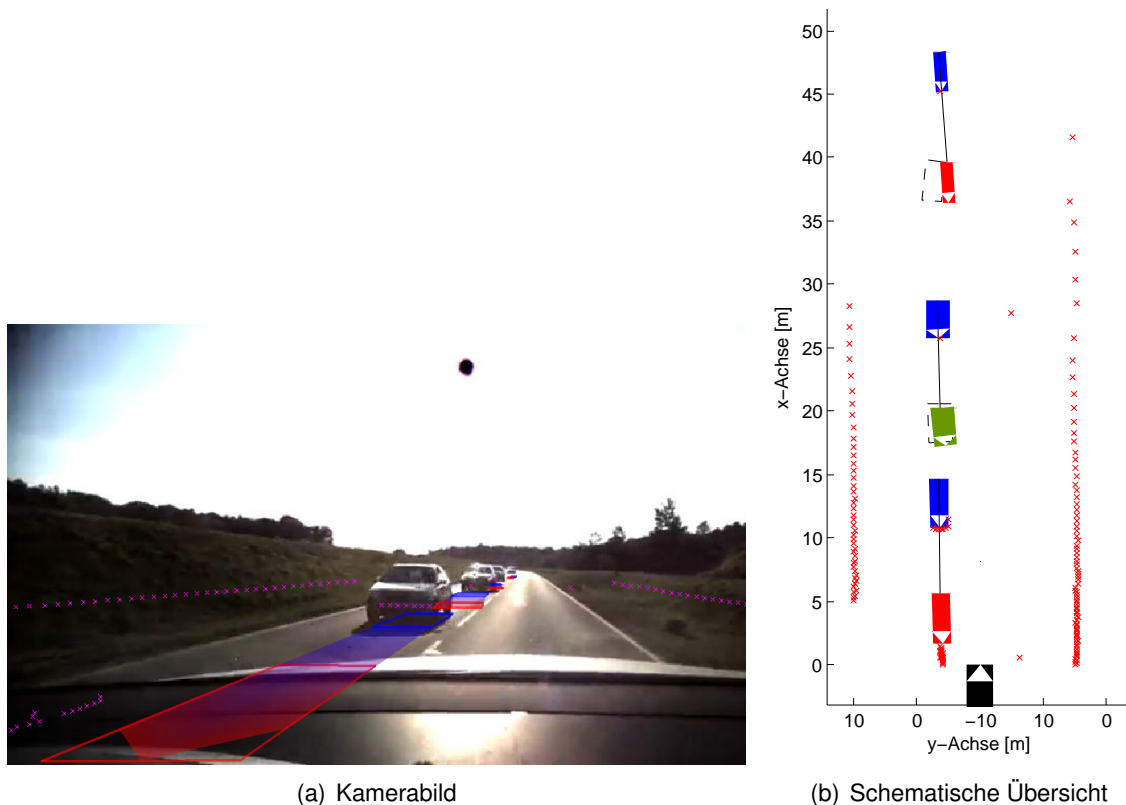


Abbildung 8.3: optische Evaluierung einer Landstraßensituation

tiv, weshalb im Folgenden eine Systembewertung mit den in Abschnitt 8.1 beschriebenen Vorgehen durchgeführt wird.

8.2.2 Verwendete Referenzdaten

Die Generierung von Referenzdaten sollte ursprünglich im Rahmen einer anderen studentischen Arbeit stattfinden. Allerdings ist zum momentanen Zeitpunkt noch keine automatisierte Erstellung von verlässlichen Referenzdaten möglich. Durch eine manuelle Überprüfung und Auswahl stehen Referenzdaten einer sehr kurzen Sequenz zur Verfügung. Die daraus ermittelten Gütemaße sind aufgrund der kurzen Zeitspanne jedoch nicht aussagekräftig. Zur Veranschaulichung werden diese Ergebnisse dennoch am Ende dieses Kapitels vorgestellt.

Um trotzdem eine Evaluierung des Systems zu ermöglichen, wurden eigene Referenzdaten mit Hilfe einer einfachen Attributierung von Messfahrten erzeugt (siehe Abschnitt 8.1.1). Da keine Referenzsensorik zur Verfügung steht und eine ausreichend realitätsnahe Simulation in diesem Rahmen zu aufwendig ist, bleibt die Attributierung die einzige Möglichkeit zur Referenzdatengenerierung. Um den Aufwand der Attributierung gering zu halten, wurden nur die Größen betrachtet, bei denen ein gutes Aufwand-Nutzen-Verhältnis zu erwarten ist. Deshalb wurden in den Messesequenzen lediglich die tatsächlich auftretenden Objekte mit ihrer Lebensdauer und Klassenzugehörigkeit ermittelt und die Objekte der Umfelderkennung zugeordnet. Da die Bewegungszustände nicht berücksichtigt werden, ist der Attributierungsaufwand überschaubar. Dennoch können mit diesen Werten bereits die Erkennungsraten Richtig-Positiv-Rate (RPR), Fehlererkennungsrate (FER), Falsch-Klassifizierungs-Rate (FKR), sowie die Zeitliche Abdeckung (ZA) ermittelt werden (siehe Abschnitt 8.1.2). Lediglich die Fehler der Zustandsgrößen können nicht abgebildet werden. Eine ausführlichere Attributierung

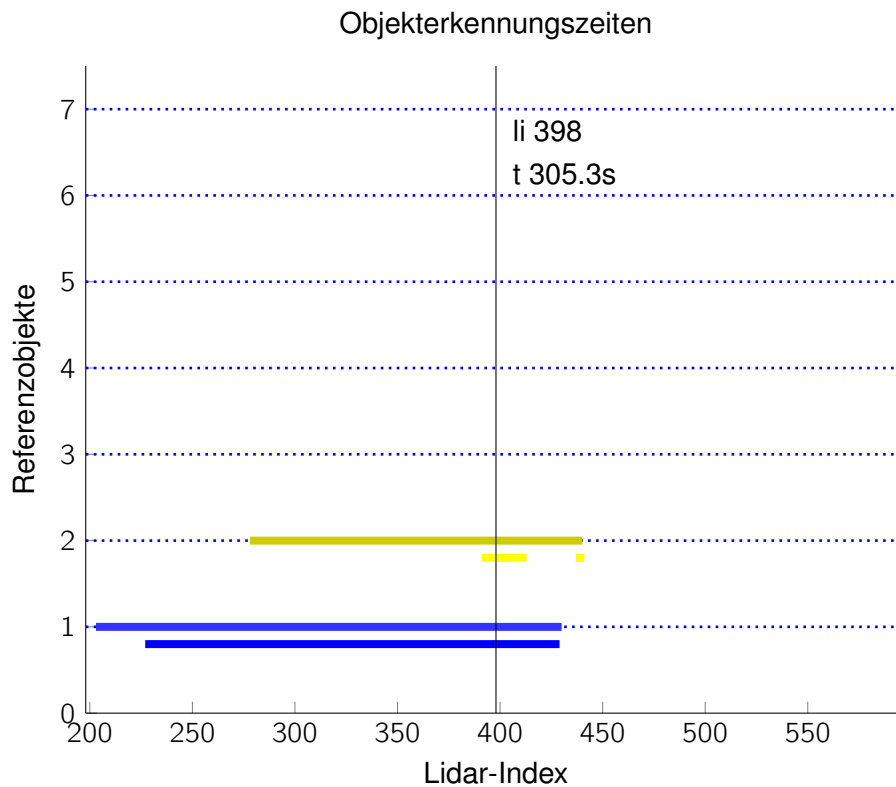


Abbildung 8.4: Grafische Darstellung der Objekterkennungszeiten

wäre an dieser Stelle jedoch wenig zielführend, da die Berechnung der Zustandsgenauigkeiten nur bei genauen Referenzdaten Sinn macht und deren Erzeugung um ein Vielfaches aufwändiger ist.

Attributierung Auch wenn die hier durchgeführte Attributierung auf sehr wenige, relativ einfach erkennbare Größen beschränkt ist, darf der notwendige Aufwand nicht unterschätzt werden. Vor allem weil die Aussagekraft der ermittelten Gütemaße mit der Länge der betrachteten Sequenzen zunimmt. Bei der Attributierung müssen mehrere Arbeitsschritte durchgeführt werden. Zunächst müssen alle auftretenden Verkehrsteilnehmer als solche erkannt werden. Dafür bietet sich eine Videoaufzeichnung der Messsequenz an. Durch die kombinierte Betrachtung der Videoaufzeichnungen und der Messwerte können dann die Zeitpunkte ermittelt werden, in denen diese Objekte den relevanten Bereich erreichen oder verlassen. Außerdem kann festgestellt werden, ob es sich um ein Fahrzeug, ein Fahrradfahrer oder einen Fußgänger handelt. Im letzten und aufwendigsten Schritt müssen alle von der Umfelderkennung erkannten Objekte überprüft werden und entweder dem entsprechenden Referenzobjekt zugeordnet oder als Fehlerkennung markiert werden.

Für eine aussagekräftige Systembewertung ist es entscheidend, dass möglichst viele Verkehrssituationen berücksichtigt werden. Dafür ist eine möglichst effiziente Attributierung notwendig. Aus diesem Grund wurde eine vorhandene Benutzeroberfläche so angepasst, dass dem Benutzer alle benötigten Informationen übersichtlich dargestellt werden und die entsprechenden Eingaben auf das Nötigste reduziert werden. Zur Darstellung der Verkehrssituation und der erkannten Objekte werden wieder das aufbereitete Videobild und eine schematische Übersicht, wie in Abbildung 8.3, verwendet. Durch eine entsprechende Zeitsteuerung können unterschiedlichen Aufnahmezeitpunkte beliebig gewählt und abgespielt werden. Zu diesen bereits vorhandenen Funktionalitäten wurden noch eine Grafik sowie die benötigten Eingabefelder hinzugefügt. Da bei dieser Attributierung bestimmt werden soll, zu welchen Zeitpunkten welche Objekte erfasst wurden, werden in der Grafik die Erkennungszeiten und Klassen der einzelnen Objekte dargestellt.

8 Systembewertung

In Abbildung 8.4 ist eine solche Darstellung der Objektlebenszeiten gezeigt. Die x-Achse entspricht dabei der Zeitachse, jedoch sind die Werte nicht in Sekunden sondern in Lidar-Abtastschritten angegeben, da das untersuchte System mit der Abtastrate des Lidars ausgeführt wird. Demnach entspricht jeder Lidarschritt $13,3\text{ms}$. Auftretende Objekte werden als horizontale Linien dargestellt. Der Anfangs- bzw. End-Punkt der Linien stellen somit den Zeitpunkt der Erkennung bzw. des Verlierens dar. Da bei diesem System nur klassifizierte Objekte als Verkehrsteilnehmer betrachtet werden, gilt ein Objekt als erkannt, sobald es einer Klasse zugeordnet werden konnte. Die unterschiedlichen Klassen der Objekte werden über die Farbe der Linien modelliert. Dabei werden die selben Farben wie bei der Einzeichnung ins Videobild verwendet: Blau für Fahrzeuge, Grün für Radfahrer und Gelb für Fußgänger. In dieser Grafik werden nun die erzeugten Referenzobjekte sowie die erkannten Objekte dargestellt. Über den y-Wert der Linien können die Objekte unterschieden werden. So wird jedem Referenzobjekt seine Nummer als y-Wert zugewiesen. Demnach sind in Abbildung 8.4 als Referenzobjekte ein Fahrzeug und ein Fußgänger abgebildet. Es handelt sich dabei um die in Tabelle 8.1 bei Szene 2 abgebildete Situation. Die erkannten und zugeordneten Objekte werden dann jeweils leicht nach unten versetzt dargestellt. Wie zu erkennen ist, wird das Fahrzeug bereits nach kurzer Verzögerung, der Fußgänger jedoch nur sporadisch erkannt. Durch die Markierung des aktuellen betrachteten Zeitschritts mit der horizontalen Linie ist auf einen Blick ersichtlich, welche Objekte momentan existieren und welche einander zugeordnet sind.

Die benötigten Schritte der Attributierung sind nun durch wenige Eingaben möglich. Durch die Betrachtung des Videobilds und der grafischen Übersicht kann nun der Eintritts- bzw. Austritts-Zeitpunkt sowie die Klasse aller auftretenden Verkehrsteilnehmer in eine entsprechende Tabelle eingetragen werden. Durch die unmittelbare Darstellung der eingetragenen Werte in die Grafik, können die eingegebenen Werte auch gut überprüft werden. Um die Zuordnung zwischen den Referenzobjekten und den erkannten Objekten zu erleichtern, werden die Lebensdauern und Klassen dieser Objekte ebenfalls dargestellt. Mit Hilfe der in Videobild, Übersicht und Lebenszeit-Grafik angezeigten Informationen, muss nun für jedes erkannte Objekt die Nummer des entsprechenden Referenzobjekts ermittelt und eingetragen werden. Stellt das Objekt keinen Verkehrsteilnehmer dar, wird die Fehlerkennung durch den Eintrag -1 gekennzeichnet. Da dieser Schritt sehr zeitaufwendig ist, wäre eine Automatisierung wünschenswert. Allerdings müssten für die Zuordnung auch mindestens Positionswerte der Referenzobjekte vorliegen und deren manuelle Erzeugung würde einen deutlich höheren Aufwand bedeuten.

8.2.3 Testsznarien

Für eine aussagekräftige Systembewertung ist es sinnvoll, möglichst viele Situationen des Anwendungsfalls zu überprüfen. Allerdings treten bei einer automobilen Umfelderkennung unzählige unterschiedlichste Situationen auf. Die große Anzahl ergibt sich schlichtweg aus der hohen Komplexität und den vielen Einflussfaktoren auf Verkehrssituationen. So spielen beispielsweise die Ortslage (Stadt, Landstraße, Autobahn), die Anzahl der Verkehrsteilnehmer sowie deren Klassen und Verhalten eine entscheidende Rolle. Wird die Situation auf einer Autobahn bei wenig Verkehr mit der einer vielbefahrenen, städtischen Kreuzung verglichen, werden sich die jeweils ermittelten Gütemaße stark unterscheiden. Zusätzlich spielen im Bereich der Umfelderkennung, je nach verwendeten Sensoren, die Witterungsbedingungen eine Rolle. Deshalb wurden zur Bewertung solche Messfahrten ausgewählt, die diese Vielzahl möglichst gut repräsentieren. Die verwendeten Messsequenzen sind in Tabelle 8.1 zusammengefasst. Auf die Besonderheiten der einzelnen Messfahrten wird bei der Diskussion der Ergebnisse in Abschnitt 8.2.6 noch genauer eingegangen.

8.2.4 Systemkonfigurationen

Grundsätzlich soll bei der Systembewertung das entwickelte Gesamtsystem getestet werden. Jedoch ist die Leistungsfähigkeit des hier entwickelten Algorithmus von einer Vielzahl von Grenzwerten und

Tabelle 8.1: Testscenarien

Beschreibung	Bildausschnitt
Szene 1: Stehendes Fahrzeug, 1 Fahrrad, 1 Fußgänger; Dauer 58,6s	
Szene 2: Stadtszenario, Abbiegung und Kurven, 11 Fahrzeuge, 5 Fußgänger; Dauer 65s	
Szene 3: Stadtszenario, 12 Fahrzeuge, 2 Fahrräder; Dauer 45,6s	
Szene 4: Landstraße, 20 Fahrzeuge; Begrenzung durch Erdwälle; Dauer 79s	
Szene 5: Landstraße, 16 Fahrzeuge; leichter Regen; passierte Abzweigungen; Dauer 60s	
Szene 6: Autobahn, 12 Fahrzeuge davon 2 LKWs; Dauer 58,6s	

Parametern abhängig. Mit Hilfe einer umfangreichen und automatisierten Systembewertung könnten mehrere Systemkonfigurationen verglichen und damit die Systemleistung optimiert werden. Da hier aufgrund der fehlenden Referenzdaten eine solche Auswertung nicht möglich ist, wird lediglich die in den vorherigen Kapiteln beschriebene Parametrierung verwendet. Diese Werte wurden größtenteils durch Überlegungen und anhand kurzer Messsequenzen festgelegt.

Eine Optimierung der Parameter ist aufgrund der vielen Freiheitsgrade nur bei einer automatisierten Bewertung sinnvoll, aber auch mit der hier vorliegenden manuellen Auswertung können unterschiedliche Systemkonfigurationen getestet werden. Um den Nutzen der durchgeführten Sensorfusion bewerten zu können, werden die Ergebnisse der Lidar-Erkennung und der Fusions-Erkennung getrennt evaluiert. Die beiden Systemkonfigurationen werden im Folgenden mit (K_{Lid}) und (K_{Fus}) abgekürzt. Auf die getrennte Betrachtung des Radar-Sensors wird hier verzichtet, da der Sensorbereich keine sinnvolle Umfelderkennung erlaubt.

8.2.5 Referenzbereich

Eine entscheidende Größe bei der Systembewertung ist der Referenzbereich. Dieser beschreibt das Gebiet, in dem Verkehrsteilnehmer als Referenzobjekte deklariert werden und hat somit direkten Einfluss auf die erzielten Gütemaße. Deshalb sind die Werte einer Systembewertung nur bei einer eindeutigen Definition des Referenzbereichs sinnvoll interpretierbar. Um eine vergleichbare Bewertung der beiden Systemkonfigurationen (K_{Lid}) und (K_{Fus}) zu erzielen, wird nur der Lidar-Sensorbereich betrachtet, da dieser von beiden Konfigurationen erfasst wird. Auch aus diesem Grund wäre eine zusätzliche Betrachtung einer reinen Radar-Konfiguration problematisch, da in diesem Fall nur noch der kleine Schnittbereich der beiden Sensoren zur Verfügung steht.

Mit der Wahl des maximal möglichen Lidar-Bereichs wird dementsprechend die Objekterkennung in einem vor dem Fahrzeug liegenden Halbkreis mit dem Radius von ca. $81m$ bewertet (siehe Abbildung 6.7). Allerdings erscheint dieser Bereich nicht für alle Objektklassen sinnvoll. Während Fahrzeuge durchaus in einer so großen Entfernung vom Lidar erfasst und auf Basis der Messwerte klassifiziert werden können, ist das für Fußgänger und Fahrradfahrer nicht möglich. Aufgrund der geringeren Objektausdehnungen werden für diese Klassen Referenzradien von $40m$ für Fahrradfahrer und $25m$ für Fußgänger gewählt. Einerseits wird dadurch eine unnötige Verschlechterung der erzielten Erkennungsraten durch unrealistisch große Referenzbereiche vermieden. Und andererseits die Erzeugung der Referenzobjekte erleichtert, da Fußgänger nicht bereits in $80m$ Entfernung auf dem Kamerabild erkannt werden müssen.

Außerdem wird die Identifizierung von Referenzobjekten noch weiter eingeschränkt. In dieser Arbeit sollen die Bewegungen von Verkehrsteilnehmern prädiziert werden und somit sind nur bewegte Verkehrsteilnehmer von Interesse. Deshalb werden Objekte erst ab der ersten Bewegung innerhalb des Referenzbereichs berücksichtigt. Sobald ein Verkehrsteilnehmer einmal als Referenzobjekt markiert ist, wird dieser auch über mögliche Haltephasen bis zum Verlassen des Referenzbereichs berücksichtigt.

8.2.6 Ergebnisse

Mit Hilfe der Referenzdaten wurden die aus Abschnitt 8.1.2 bekannten Gütemaße RPR, FKR, FER und ZA berechnet. Um die Umfelderkennung differenzierter beurteilen zu können, wurden diese Gütemaße sowohl klassenübergreifend als auch für jede Klasse ermittelt. In Tabelle 8.2 sind die Gütemaße für alle Testszenarien (siehe Tabelle 8.1) und beide Systemkonfigurationen zusammengefasst. Dabei sind die klassenspezifischen Gütemaße durch die Indizes $_p$ für Fußgänger, $_c$ für Fahrzeug und $_b$ für Fahrrad gekennzeichnet. Falls eine Verkehrsteilnehmerklasse in einer Messfahrt nicht auftritt, kann das entsprechende Gütemaß auch nicht berechnet werden, was in der Tabelle durch ein „-“

8 Systembewertung

Gütemaß	Szene 1 Stehend		Szene 2 Stadt		Szene 3 Stadt		Szene 4 Land- straße		Szene 5 Land- straße Regen		Szene 6 Auto- bahn	
	K_{Lid}	K_{Fus}	K_{Lid}	K_{Fus}	K_{Lid}	K_{Fus}	K_{Lid}	K_{Fus}	K_{Lid}	K_{Fus}	K_{Lid}	K_{Fus}
RPR [%]	1.00	1.00	0.81	0.81	0.93	1.00	0.65	1.00	0.63	0.94	0.83	0.92
RPR_p [%]	1.00	1.00	0.80	0.80	-	-	-	-	-	-	-	-
RPR_c [%]	-	-	0.82	0.82	0.92	1.00	0.65	1.00	0.63	0.94	0.83	0.92
RPR_b [%]	1.00	1.00	-	-	1.00	1.00	-	-	-	-	-	-
FKR [%]	0.00	0.00	0.13	0.13	0.00	0.00	0.05	0.00	0.06	0.06	0.00	0.00
FKR_p [%]	0.00	0.00	0.00	0.00	-	-	-	-	-	-	-	-
FKR_c [%]	-	-	0.18	0.18	0.00	0.00	0.05	0.00	0.06	0.06	0.00	0.00
FKR_b [%]	0.00	0.00	-	-	0.00	0.00	-	-	-	-	-	-
FER [$\frac{1}{s}$]	0.17	0.17	0.69	0.72	0.77	0.77	0.18	0.19	0.73	0.78	0.12	0.14
FER_p [$\frac{1}{s}$]	0.02	0.02	0.34	0.34	0.48	0.48	0.16	0.18	0.48	0.52	0.12	0.12
FER_c [$\frac{1}{s}$]	0.14	0.14	0.29	0.32	0.26	0.26	0.01	0.01	0.23	0.25	0.00	0.02
FER_b [$\frac{1}{s}$]	0.02	0.02	0.06	0.06	0.02	0.02	0.00	0.00	0.02	0.02	0.00	0.00
ZA [%]	0.73	0.73	0.36	0.37	0.42	0.65	0.26	0.87	0.19	0.67	0.36	0.87
ZA_p [%]	0.99	0.99	0.27	0.26	-	-	-	-	-	-	-	-
ZA_c [%]	-	-	0.41	0.43	0.44	0.71	0.26	0.87	0.19	0.67	0.36	0.87
ZA_b [%]	0.64	0.63	-	-	0.31	0.28	-	-	-	-	-	-

Tabelle 8.2: Berechnete Gütemaße der Systembewertung

gekennzeichnet ist. Im Folgenden werden zunächst die Ergebnisse der sechs Testszenarien einzeln diskutiert, bevor zusammenfassend das Gesamtsystem bewertet wird.

Szene 1 Diese Szene wurde ursprünglich aufgezeichnet, um die Erfassung von Radfahrern mit Hilfe des Lidar-Scanners zu untersuchen. Dementsprechend steht das Versuchsfahrzeug und es bewegen sich lediglich ein Radfahrer und ein Fußgänger im Referenzbereich. Diese beiden Verkehrsteilnehmer werden bei jedem Durchqueren des Bereichs richtig klassifiziert, woraus Richtig-Positiv-Raten von 100% folgen und die Falsch-Klassifizierungs-Raten dementsprechend gleich Null sind. Die Qualität der Erkennung von realen Objekten wird noch durch die Zeitliche-Abdeckung beschrieben. Dabei sind deutliche Unterschiede zwischen der Erfassung des Fußgängers und der des Radfahrers erkennbar. Der Fußgänger wird fast während der gesamten Aufenthaltsdauer im Referenzbereich (99% der Zeit) erkannt. Der deutlich geringere Wert des Fahrradfahrers ($ZA_b = 0.64$) wird durch Phasen verursacht, in denen nur ein Teil des Fahrrads erkannt wird. Die geringe Ausdehnung des erkannten Objekts führt dann teilweise zu einer phasenweisen Klassifizierung als Fußgänger. Abgesehen von den kurzen Falschklassifizierungen des Fahrradfahrers liefert die Erfassung der Referenzobjekte gute Ergebnisse. Allerdings treten auch einige Fehlerkennungen auf. In diesem konkreten Fall wird die relativ hohe Fehlerkennungsrate FER_c durch eine bewegliche Begrenzungskette verursacht. Durch diese treten trotz stehendem Versuchsfahrzeug charakteristische Bewegungen in den Messwerten auf. Beim Vergleich der beiden Systemkonfigurationen fällt auf, dass die erzielten Ergebnisse nahezu identisch sind. Das entspricht den Erwartungen, da der Radarsensor Fußgängern und Radfahrern meistens nur im Nahbereich detektiert und solche Messungen aufgrund der ungenauen Winkelmessung nicht verwendet werden.

Szene 2 Die zweite Testsequenz wurde, wie alle weiteren, im realen Straßenverkehr aufgezeichnet. Das Versuchsfahrzeug fährt dabei durch ein städtisches Umfeld mit breiten Grünstreifen, wobei die Strecke, neben einer Abbiegung, aus einer lang gezogenen Kurve besteht. Die deutliche Verschlechterung der berechneten Gütemaße im Vergleich zu dem ruhenden, ersten Szenario lässt sich auf mehrere Ursachen zurückführen. So wird einer der Fußgänger nicht erkannt, da er beim Vorbeifahren des Fahrzeugs stehen bleibt und dadurch nicht klassifiziert wird. Die nicht erkannten Fahrzeuge tauchen nur kurz im Referenzbereich auf. Aus diesem Grund werden sie nur unvollständig detektiert, wodurch sie falsch klassifiziert werden. Was auch an der Falschklassifizierungsrate FKR festgestellt werden kann. Die schlechten ZA-Werte lassen sich dadurch erklären, dass die Objekte erst in der näheren Umgebung richtig eingeordnet werden können, die Referenzobjekte davor aber bereits lange im Referenzbereich waren. Bei der Bewertung dieser Rate muss berücksichtigt werden, dass die Referenzbereiche im Vergleich zur Erkennungsleistung des Lidar-Scanners sehr großzügig gewählt wurden. Als Letztes ist noch eine deutlich größere Fehlerkennungsrate zu erkennen. Diese wird vor allem durch die Vegetation am Straßenrand verursacht, welche in Kombination mit der Eigenbewegung des Fahrzeugs oft zu kurzzeitigen Fehldetektionen führt. Auch in diesem Szenario lässt sich kein relevanter Unterschied zwischen den beiden Sensorkonfigurationen feststellen.

Szene 3 Im zweiten Stadtszenario treten neben mehreren Fahrzeugen auch zwei Fahrradfahrer auf. Das Versuchsfahrzeug legt dabei annähernd gerade Straßenabschnitte zurück, die durch einen kurzen Abbiegevorgang verbunden sind. Bei der Betrachtung der berechneten Gütemaße sind vor allem zwei Punkte interessant. Erstens sind die Werte der Lidar-Konfiguration vergleichbar mit denen des ersten Stadtszenarios, was für die Aussagekraft der ermittelten Gütemaße spricht. Zweitens lässt sich durch die Fusion mit dem Radarsensor die Leistungsfähigkeit der Umfelderkennung in dieser Situation deutlich steigern, wie an den Werten der Fusions-Konfiguration ersichtlich. Die zurückgelegten Streckenabschnitte sind im Vergleich zum Szenario 2 relativ gerade und liegen dadurch größtenteils in dem schmalen Erfassungsbereich des Radars. Dadurch ermöglicht der Radarsensor nun eine deutlich frühere Erkennung von Fahrzeugen und damit eine deutlich Steigerung der zeitlichen Abdeckung.

Szene 4 In der Szene 4 bewegt sich das Versuchsfahrzeug über eine Landstraße mit einer lang gezogenen Kurve, wobei sich über weite Teile der Strecke an beiden Straßenseiten leichte Erdwälle befinden. Dadurch ist die vom Lidar erfasste Umgebung sehr homogen. Diese Homogenität bewirkt die geringe Fehlerkennungsrate bei diesem Szenario. Lediglich vereinzelte Büsche oder Schilder werden kurzzeitig als Fußgänger klassifiziert. Die Verbesserung durch den Fusionsansatz zeigt sich hier noch deutlicher als in der vorherigen Szene. Während die Lidar-Konfiguration entgegenkommende Fahrzeuge meist sehr spät und teilweise gar nicht erkennt, werden mit der Fusions-Konfiguration alle Fahrzeuge richtig klassifiziert und auch fast über den gesamten Referenzbereich detektiert.

Szene 5 In Szene 5 handelt es sich wieder um eine Verkehrs

Szene 6 Als letztes TestszENARIO wird noch eine Autobahnfahrt untersucht. Durch die definierten Fahrbahnbegrenzungen wird auch in dieser Szene wieder eine relativ geringe FER erzielt. Bei dieser Messfahrt wurden neben mehreren PKWs auch zwei Lastwagen aufgenommen. Deren Erkennung und Verfolgung anhand von Lidar-Messwerten ist relativ schwierig, da LKWs häufig keine definierte Form aufweisen. Auch die Bewegungen eines Gespanns sind schwieriger abzubilden als die eines PKWs. Aus diesen Gründen wäre es sinnvoll, Lastwagen durch eine eigene Klasse zu repräsentieren, was aus zeitlichen Gründen nicht mehr möglich war. Auch bei dieser Messfahrt liefert das Fusionsmodell deutlich bessere Ergebnisse. Das entspricht den Erwartungen, da das ACC-System vor allem für Autobahnszenarien konzipiert wurde.

Fazit Der positive Eindruck der bei der optischen Evaluierung (siehe Abschnitt 8.2.1) konnte mit den ermittelten Gütemaßen in Tabelle 8.2 bestätigt werden. Das Gesamtsystem mit der Fusion von Lidar- und Radarsensor liefert für alle TestszENARIEN vielversprechende Erkennungsraten. Auch die hohe zeitliche Abdeckung deutet auf eine verlässliche Objektverfolgung hin. Die Konfiguration K_{Lidar} fällt dagegen vor allem bei Überland- und Autobahnsituationen deutlich ab.

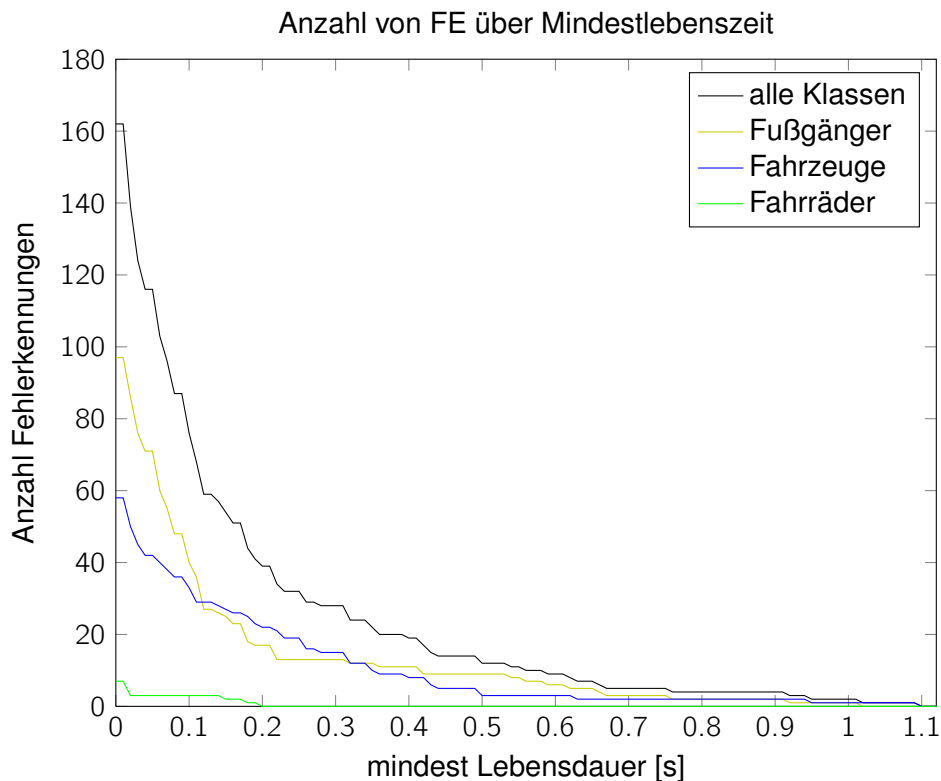


Abbildung 8.5: Lebensdauer von Fehlerkennungen

Die relativ hohen Fehlerkennungsraten werden vor allem durch Vegetation am Straßenrand verursacht. Durch die Bewegung des Versuchsfahrzeugs und Wind treten in den Lidarmesswerten von Grünstreifen verkehrsteilnehmertypische Formen auf, welche gelegentlich zu Fehlerkennungen führen können. Allerdings wird der Großteil der Fehlerkennungen nur sehr kurz detektiert. Zur Verdeutlichung sind in Abbildung 8.5 die Anzahl aller Fehlerkennungen der sechs Testszenarien über ihrer Erkennungsdauer aufgetragen.

Demnach werden insgesamt 162 falsche Objekte detektiert. Jedoch haben die meisten dieser Fehlerkennungen nur eine sehr kurze Lebensdauer. Mehr als die Hälfte wird kürzer als eine Zehntel Sekunde erkannt. Somit könnte die Anzahl der falsch eingeblendeten Objekte durch eine zeitliche Filterung drastisch reduziert werden.

Als letztes Kriterium wird die benötigte Rechenzeit berücksichtigt. Da der entwickelte Algorithmus bisher nur als Simulinkmodell vorliegt und noch nicht in C-Code umgewandelt bzw. in die ADTF-Umgebung integriert wurde, wird nur eine grobe Abschätzung durchgeführt. Dafür wird die für die Auswertung einer Messfahrt benötigte Simulationszeit mit der Dauer der Messfahrt verglichen.

Dabei wurden bei den untersuchten Testszenarien weniger als ein Drittel der Realzeit für die gesamte Simulation benötigt. Die Simulation wurde auf dem studentischen Laptop mit einem Intel Core 2 Duo T6400 Prozessor mit 2 GHz und 4 GB RAM durchgeführt. Somit konnte selbst in der aufwendigeren Simulinkumgebung und mit dieser Hardware dreifache Echtzeitfähigkeit erreicht werden. Der zusätzliche Rechenaufwand des Fusionskonfiguration beträgt lediglich zwischen 1,8% und 7,5% der Rechenzeit. Aufgrund der deutlich besseren Ergebnisse und dem nur geringfügig höheren Rechenaufwand ist diese Systemkonfiguration eindeutig zu bevorzugen. Abschließend lässt sich sagen, dass die hier entwickelte Umfelderkennung im Hinblick auf die Anwendung in einem Predictive Display gute Ergebnisse liefert.

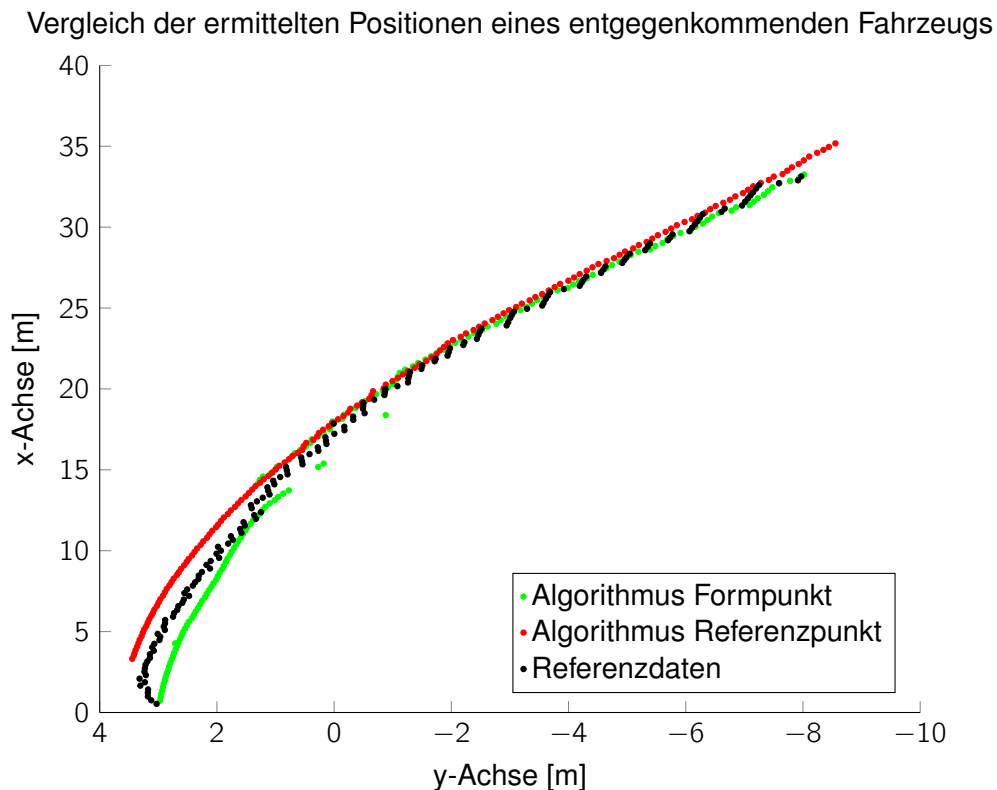


Abbildung 8.6: Vergleich mit Positions-Referenzdaten

8.2.7 Bewertung durch externe Referenzdaten

In einem anderen studentischen Projekt sollen aus den aufgezeichneten Messwerten möglichst vollständige Referenzdaten gewonnen werden. Im Vergleich zu dem hier vorgestellten Algorithmus müssen dabei keine Echtzeitanforderungen erfüllt werden und außerdem kann auf die gesamte Messsequenz und damit auch auf zukünftige Messwerte zugegriffen werden. Durch die Betrachtung des gesamten Zeitbereichs können die Bewegungszustände von Objekten auch über Verdeckungsphasen oder bis zum Eintritt in den Referenzbereich interpoliert werden. Vor allem in Kombination mit einer abschließenden Kontrolle durch einen Benutzer lassen sich so verlässliche und umfangreiche Referenzdaten generieren.

Allerdings ist diese Entwicklung noch nicht abgeschlossen und es steht zum momentanen Zeitpunkt nur eine sehr kurze Sequenz an Referenzdaten zu Verfügung. Diese Sequenz stammt noch von einer älteren Messfahrt und beschreibt ein in einer Rechtskurve entgegenkommendes Fahrzeug. In Abbildung 8.6 sind die ermittelten Positionsdaten des hier entwickelten Algorithmus und die Referenzdaten eingezeichnet. Dabei ist zu beachten, dass hier erstmals auf eine gleiche Achsenormierung verzichtet wurde, um die Abweichungen besser sichtbar zu machen. Dabei wurden die Referenzpositionen in schwarz, die beiden im Algorithmus auftretenden Positionen in grün und rot dargestellt. Bei den grünen Formpunkten handelt es sich jeweils um den repräsentativen Punkt der verwendeten Formen (siehe Abschnitt 4.3.1.2). Die roten Referenzpunkte sind jeweils die angenommenen Referenzpunkte des Objekts. Da es sich um ein Fahrzeug handelt wird versucht, dafür den Mittelpunkt der Hinterachse aus den Abmessungen zu schätzen und zu verwenden. Wie in der Abbildung zu erkennen ist, ergibt sich für diesen Fahrzeug-Referenzpunkt eine relativ gleichmäßige Bewegung. Der Formpunkt weist dagegen einige Sprünge auf, die durch einen Wechsel in der erkannten Form verursacht werden. Bei der Betrachtung der schwarzen Referenzdaten fällt auf, dass diese ebenfalls einen unruhigen Verlauf aufweisen. Vor allem der charakteristische Verlauf der vielen abgesetzten Striche am Anfang der Erkennung, bei x-Werten über 20m, wird durch die rohdatenbasierte Lidar-

8 Systembewertung

auswertung verursacht. Jedes Mal wenn das Objekt mit einem neuen Lidarstrahl erfasst wird, macht der ermittelte Punkt einen Sprung. Aufgrund der Probleme, die bei den Referenzdaten noch auftreten, erscheint ein ausführlicher Vergleich momentan nicht sehr sinnvoll. Wenn künftig verlässliche Referenzdaten zur Verfügung stehen, können dann auch die Genauigkeiten der ermittelten Objektzustände überprüft und durch RMSE-Werte abstrakt repräsentiert werden.

9 Zusammenfassung und Ausblick

9.1 Zusammenfassung

In dieser Arbeit wurde ein Algorithmus zur Erkennung und Prädiktion von Verkehrsteilnehmern im Umfeld eines Versuchsfahrzeugs entwickelt. Die prädizierten Positionen sollen in ein Predictive Display zur Steuerung des teleoperierten Fahrzeugs eingeblendet werden. Zur Umfelderkennung konnte dabei auf einen Lidar-Scanner und einen Radarsensor zurückgegriffen werden. Um mit den zur Verfügung stehenden Daten das bestmögliche Ergebnis zu erzielen, wurde ein innovativer Ansatz zur Lidardatenauswertung entwickelt. Im Gegensatz zur weit verbreiteten, rohdatenbasierten Merkmalsextraktion werden die Messwerte objektorientiert ausgewertet. Durch die Berücksichtigung von Vorwissen bei der Datenauswertung können einerseits höhere Erkennungsraten erzielt werden und andererseits Objekte ohne definierte Formen, wie beispielsweise Radfahrer, verfolgt werden. Außerdem wurde auch eine Routine zur Klassifizierung von Verkehrsteilnehmer vorgestellt. Diese unterscheidet, unter Berücksichtigung der Objektdynamik und den Abmessungen, zwischen Fahrzeugen, Fahrrädern und Fußgängern. Allerdings kann das Verfahren problemlos um weitere Klassen ergänzt werden. Um die Umfelderkennung noch weiter zu verbessern, wurde der serienmäßige Radarsensor mit Hilfe einer Datenfusion integriert. Der hier vorgestellte Fusionsansatz zielt dabei ebenfalls auf eine möglichst effektive Datenauswertung ab. Deshalb werden die erkannten Radarobjekte nicht nur mit den Lidarobjekten fusioniert, sondern auch zur Verbesserung der Lidardatenauswertung verwendet. Die Verfolgung und Ermittlung der Bewegungszustände der Verkehrsteilnehmer wird durch Kalman-Filter verwirklicht. Dabei werden die verschiedenen Klassen teilweise mit unterschiedlichen Bewegungsmodellen modelliert, um die Filterleistung zu verbessern. Für die nichtlinearen Bewegungsmodelle bei Fahrzeugen und Fahrradfahrern kommen Extended-Kalman-Filter zum Einsatz, da diese bei dem vorliegenden linearisierbaren Problem ein sehr gutes Leistungs-/Aufwandverhältnis aufweisen. Die für das Predictive Display benötigte Prädiktion der Positionen beruht auf der Fortsetzung des aktuell erfassten Bewegungszustands der Verkehrsteilnehmer.

Die Evaluation des entwickelten Systems wurde anhand aufgezeichneter Messfahrten im realen Straßenverkehr durchgeführt. Dabei wurden durch die manuelle Attributierung der Datensätze grundsätzliche Gütemaße zur Systembewertung ermittelt. Die erreichten Werte sind äußerst vielversprechend. Das hier vorliegende System kann somit zur Erfassung und Prädiktion des umgebenden Verkehrsgeschehens verwendet werden und ermöglicht durch die Klassifizierung der Objekte auch eine differenzierte Darstellung im Predictive Display.

9.2 Ausblick

Auch wenn die entwickelte Umfelderkennung bereits gute Ergebnisse liefert und für das Predictive Display verwendet werden kann, so gibt es dennoch einige Erweiterungs- und Verbesserungsmöglichkeiten. Zunächst einmal könnte das vorliegende Modell um eine eigene Klasse für LKWs erweitert werden. Dabei dürften vor allem die unklare Form der Lidarmesswerte und ein geeignetes Bewegungsmodell problematisch sein. Jedoch können durch den hier vorliegenden, objektorientierten Ansatz auch undefinierte Formen verfolgt werden, wie bereits bei der Klasse der Radfahrer demonstriert. Mit der Ergänzung der LKWs sind dann alle wichtigen Verkehrsteilnehmer abgedeckt.

Um die Verlässlichkeit der Objekterkennung zu erhöhen, könnten die Informationen der installierten Kameras verwendet werden. Für die Fusion der Kameras mit dem vorliegenden System sind mehrere Varianten denkbar. In der Einleitung dieser Arbeit wird auf einige Forschungsarbeiten verwiesen, die sich mit möglichen Fusionen dieser Sensoren befassen (siehe Abschnitt 1.2.2). Die große Herausforderung bei der Bildverarbeitung ist die Erfüllung der Echtzeitbedingung, da sehr große Datenmengen bearbeitet werden müssen. Deshalb werden die Ergebnisse von anderen Sensoren häufig zur Einschränkung des untersuchten Bereichs genutzt. Aufgrund der großen Möglichkeiten die dieser Bereich bietet, sollte eine solche Fusion untersucht werden.

Bisher konnte auch noch keine ausführliche Evaluation der Systemleistung durchgeführt werden, da dafür präzise Referenzdaten benötigt werden. Die Erstellung solcher Referenzdaten ist zwar mit großem Aufwand verbunden, kann aber neben der reinen Systembewertung auch zur Optimierung dienen. So können die verwendeten Parameter und Grenzwerte mit Hilfe der Systemevaluierung angepasst werden und damit die Leistung der Umfelderkennung gesteigert werden. Außerdem kann damit der Nutzen von entwickelten Erweiterungen bewertet werden.

Abschließend muss das vorliegende Modell noch in die Software-Umgebung des Versuchsträgers integriert werden. Da der Algorithmus aber bereits in Form eines Simulinkmodells vorliegt, kann daraus automatisiert C-Code generiert werden und lediglich die benötigten Schnittstellen müssen angepasst werden.

Abbildungsverzeichnis

2.1	Versuchsfahrzeug mit Sensoren	6
2.2	Systemarchitektur des Versuchsträgers	7
2.3	Verwendete Koordinatensysteme	7
2.4	Abdeckungsbereiche der Sensoren	8
2.5	SICK Lidar-Sensor [SIC03]	9
2.6	Reichweite des LIDARs in Abhängigkeit des Remissionswerts des Objekts [SIC03, S.14]	10
3.1	Komponenten eines Trackingsystems [BSF88, S.3]	14
3.2	Grundprinzip eines Abtastverfahrens [JU04, S. 5]	20
4.1	Lidar Rohdaten	24
4.2	Übersicht über Trackingalgorithmus	25
4.3	Veranschaulichung des Segmentierungsgrenzwerts	26
4.4	Segmentierte Lidar-Messung	27
4.5	Verbinden von Segmenten	29
4.6	charakteristische Segmentformen	31
4.7	Überprüfung der Koordinatentransformation	35
4.8	Messwertassoziation	36
4.9	Übersicht über Trackingalgorithmus	38
4.10	Rohdatenauswertung	39
4.11	Gültigkeitsbereiche für Messpunkte	41
4.12	Gewichtungsfunktionen	45
5.1	Radarauswertung	52
6.1	Fusionsarten [RPL07, S. 96]	55
6.2	Latenzzeiten bei Datenfusion [Lin08, S.38]	58
6.3	Distanzmessungen der unterschiedlichen Sensoren	60
6.4	Bestimmung der Latenzzeiten	61
6.5	Abweichung zwischen prädierten und gemessenen Radarwerten	64
6.6	Funktionsweise der verwendeten Fusion	65
6.7	Sensorbereiche bei der Datenfusion in der Draufsicht	66
8.1	Grafische Darstellung zur optischen Evaluation	74
8.2	Übersicht über Systembewertung [Lin08, S. 139]	74
8.3	optische Evaluierung einer Landstraßensituation	78
8.4	Grafische Darstellung der Objekterkennungszeiten	79
8.5	Lebensdauer von Fehlerkennungen	85
8.6	Vergleich mit Positions-Referenzdaten	86

Tabellenverzeichnis

2.1	Translationsvektoren zwischen den Koordinatensystemen	8
2.2	Kenndaten der Radar-Einheit [200, S.160]	11
2.3	Genauigkeiten der inertialen Messeinheit [Oxf11]	12
4.1	erwartete Abmessungen der Verkehrsteilnehmer	33
6.1	Positions- und Zeitdifferenzen zwischen den Messungen	62
6.2	Evaluierung der Radarprädiktion	64
8.1	Testszenarien	81
8.2	Berechnete Gütemaße der Systembewertung	83

Literaturverzeichnis

- [200] *Autoelektrik, Autoelektronik: [neu: Vernetzung, Hybridantrieb ; Systeme und Komponenten]*. 5. Aufl. Studium und Praxis. Wiesbaden: Vieweg, 2007. ISBN: 3-528-13875-0.
- [AB63] J. E. ARNOLD und P. W. BRAISTED. *Design and evaluation of a predictor for remote control systems operating with signal transmission delays*. Bd. 2229. NASA Technical note. Washington, DC National Aeronautics und Space Administration, 1963.
- [Amd+13] A. AMDITIS et al. „Multi-sensor tracking and lane estimation in highly automated vehicles“. In: *IET Intelligent Transport Systems 7.1* (2013), S. 160–169. ISSN: 1751-956X. DOI: 10.1049/iet-its.2011.0166.
- [BAB] J. BURLET, O. AYCARD und Q. BAIG. „Interacting multiple models based classification of moving objects“. In: *Vision (ICARCV 2010)*, S. 2472–2478. DOI: 10.1109/ICARCV.2010.5707920.
- [BSF88] Y. BAR-SHALOM und T. E. FORTMANN. *Tracking and data association*. Bd. 179. Mathematics in science and engineering. Boston: Academic Press, 1988. ISBN: 0-12-079760-7.
- [FC] F. FAYAD und V. CHERFAOUI. „Tracking objects using a laser scanner in driving situation based on modeling target shape“. In: *2007 IEEE Intelligent Vehicles Symposium*, S. 44–49. DOI: 10.1109/IVS.2007.4290089.
- [FLD] K. C. FUERSTENBERG, D. T. LINZMEIER und K. C. DIETMAYER. „Pedestrian recognition and tracking of vehicles using a vehicle based multilayer laserscanner“. In: *Proceedings of ITS 2003, 10th World Congress on Intelligent Transport Systems*.
- [FLN12] B. FORTIN, R. LHERBIER und J.-C. NOYER. „Feature Extraction in Scanning Laser Range Data Using Invariant Parameters: Application to Vehicle Detection“. In: *IEEE Transactions on Vehicular Technology* 61.9 (2012), S. 3838–3850. ISSN: 0018-9545. DOI: 10.1109/TVT.2012.2211630.
- [Fre+13] FREDERIC CHUCHOLOWSKI et al., Hrsg. *Prediction Methods for Teleoperated Road Vehicles*. 2013.
- [GA08] M. S. GREWAL und A. P. ANDREWS. *Kalman filtering: Theory and practice using MATLAB*. 3. Aufl. Hoboken: John Wiley, 2008. ISBN: 978-0-470-17366-4.

- [GSL] S. GNATZIG, F. SCHULLER und M. LIENKAMP. „Human-machine interaction as key technology for driverless driving - A trajectory-based shared autonomy control approach“. In: *2012 RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, S. 913–918. DOI: 10.1109/ROMAN.2012.6343867.
- [JU04] S. JULIER und J. UHLMANN. „Unscented Filtering and Nonlinear Estimation“. In: *Proceedings of the IEEE* 92.3 (2004), S. 401–422. ISSN: 0018-9219. DOI: 10.1109/JPROC.2003.823141.
- [KBD] N. KAEMPCHEN, M. BUEHLER und K. DIETMAYER. „Feature-level fusion for free-form object tracking using laserscanner and video“. In: *IEEE Proceedings. Intelligent Vehicles Symposium, 2005*, S. 453–458. DOI: 10.1109/IVS.2005.1505145.
- [KKne] N. KAWASAKI und U. KIENCKE. „Standard platform for sensor fusion on advanced driver assistance system using bayesian network“. In: *IEEE Intelligent Vehicles Symposium, 2004*. June 14-17, 2004, S. 250–255. DOI: 10.1109/IVS.2004.1336390.
- [Kae+ne] N. KAEMPCHEN et al. „IMM object tracking for high dynamic driving maneuvers“. In: *IEEE Intelligent Vehicles Symposium, 2004*. June 14-17, 2004, S. 825–830. DOI: 10.1109/IVS.2004.1336491.
- [Kal] R. E. KALMAN. „A New Approach to Linear Filtering and Prediction Problems“. In: *Tans. ASME J. Basic Engineering*. Bd. 82, S. 34–45.
- [Kuh56] H. W. KUHN. „Variants of the hungarian method for assignment problems“. In: *Naval Research Logistics Quarterly* 3.4 (1956), S. 253–258. ISSN: 00281441. DOI: 10.1002/nav.3800030404.
- [LH] Y.-C. LEE und T. HSIAO. „Object tracking via the probability-based segmentation using laser range images“. In: *2010 IEEE Intelligent Vehicles Symposium (IV)*, S. 197–202. DOI: 10.1109/IVS.2010.5548081.
- [LW] P. LINDNER und G. WANIELIK. „Multi level fusion for an automotive pre-crash safety system“. In: *2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, S. 143–146. DOI: 10.1109/MFI.2008.4648124.
- [Lab+05] R. LABAYRADE et al. „Cooperative Fusion for Multi-Obstacles Detection With Use of Stereovision and Laser Scanner“. In: *Autonomous Robots* 19.2 (2005), S. 117–140. ISSN: 0929-5593. DOI: 10.1007/s10514-005-0611-7.
- [Lin08] R. LINDL. „Tracking von Verkehrsteilnehmern im Kontext von Multisensorsystemen“. Diss. München: Technische Universität München, 2008.
- [MBNne] A. MENDES, L. BENTO und U. NUNES. „Multi-target detection and tracking with a laserscanner“. In: *IEEE Intelligent Vehicles Symposium, 2004*. June 14-17, 2004, S. 796–801. DOI: 10.1109/IVS.2004.1336486.

- [Oxf11] OXFORD TECHNICAL SOLUTIONS LIMITED. *RT Inertial and GPS Measurement Systems: User Manual*. Hrsg. von OXFORD TECHNICAL SOLUTIONS LIMITED. Oxfordshire, 2011.
- [RPL07] H. RUSER und F. PUENTE LEÓN. „Informationsfusion – Eine Übersicht (Information Fusion – An Overview)“. In: *tm – Technisches Messen* 74.3 (2007). ISSN: 0171-8096. DOI: 10.1524/teme.2007.74.3.93.
- [SDS 2] J. SPARBERT, K. DIETMAYER und D. STRELLER. „Lane detection and street type classification using laser range images“. In: *2001 IEEE Intelligent Transportation Systems. Proceedings*. 25-29 Aug. 2001, S. 454–459. DOI: 10.1109/ITSC.2001.948700.
- [SIC03] SICK AG. *Lasermesssysteme LMS 200/ LMS 211/ LMS 220/ LMS 221/ LMS 291: Technische Beschreibung*. Hrsg. von SICK AG. 2003.
- [SRS] SONGHWAI OH, S. RUSSELL und S. SASTRY. „Markov chain Monte Carlo data association for general multiple-target tracking problems“. In: *2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, 735–742 Vol.1. DOI: 10.1109/CDC.2004.1428740.
- [Sta12] STATISTISCHES BUNDESAMT. *Verkehr: Verkehrsunfälle*. Wiesbaden, 2012.
- [TA] TRUNG-DUNG VU und O. AYCARD. „Laser-based detection and tracking moving objects using data-driven Markov chain Monte Carlo“. In: *2009 IEEE International Conference on Robotics and Automation (ICRA)*, S. 3800–3806. DOI: 10.1109/ROBOT.2009.5152805.
- [TN+] H. TEHRANI NIKNEJAD et al. „Embedded multi-sensors objects detection and tracking for urban autonomous driving“. In: *2011 IEEE Intelligent Vehicles Symposium (IV)*, S. 1128–1135. DOI: 10.1109/IVS.2011.5940563.
- [Wal08] L. WALCHSHÄUSL. „Maschinelle Erkennung von Verkehrsteilnehmern mittels heterogener Sensorik“. Diss. München: Technische Universität München, 2008.
- [Win12] H. WINNER. „Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort ; mit 45 Tabellen“. In: *Handbuch Fahrerassistenzsysteme* (2012, 2012).
- [Wol12] M. WOLTER. „Erkennung und Verfolgung von statischen und dynamischen Objekten auf Basis von Laserdaten“. Diss. München: Technische Universität München, 2012.
- [Wu+] B. WU et al. „Fast Pedestrian Detection with Laser and Image Data Fusion“. In: *Graphics (ICIG)*, S. 605–608. DOI: 10.1109/ICIG.2011.107.