ulm university universität
uulm

Multisensorsystem zur erweiterten Fahrzeugumfelderfassung

DISSERTATION

zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS
(DR.-ING.)

der Fakultät für Ingenieurwissenschaften
und Informatik der Universität Ulm

von

STEFAN WENDER
aus Mühlhausen / Thür.

1. Gutachter:	Prof. Dr.-Ing. K. C. J. Dietmayer
2. Gutachter:	Prof. Dr.-Ing. H.-M. Groß
Amtierender Dekan:	Prof. Dr.-Ing. M. Weber

Datum der mündlichen Prüfung: 24.10.2008

Vorwort

Die vorliegende Arbeit entstand im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Mess-, Regel- und Mikrotechnik der Universität Ulm.

Mein Dank gilt Herrn Prof. Klaus Dietmayer für die Förderung meiner Arbeiten und die exzellenten Möglichkeiten in der ARGOS-Gruppe. Insbesondere bin ich dankbar für die eingeräumten wissenschaftlichen Freiheiten bei der Erschaffung des in dieser Arbeit vorgestellten Gesamtsystems inklusiver seiner einzelnen Ansätze.

Bei Herrn Prof. H.-M. Groß möchte ich mich für die Übernahme des Koreferats, die förderliche Diskussion und die hilfreichen Anmerkungen und Vorschläge bedanken.

Die Arbeiten meines Kollegen Thorsten Weiss bildeten eine wichtige Grundlage für verschiedene Aspekte meiner Forschungstätigkeit. Vielen Dank für die exzellente Teamarbeit insbesondere im Rahmen der gemeinsam bearbeiteten Projekte und die zahlreichen Diskussionen.

Meinem ehemaligen Kollegen Nico Kämpchen danke ich für die Zusammenarbeit sowie das von ihm gesammelte Know-how, die Erfahrungen und Ergebnisse im Bereich der laserscannerbasierten Objektverfolgung. Danke auch für die zahlreichen Tipps zu dieser Arbeit.

Meinem Kollegen Holger Berndt danke ich für viele hilfreiche Diskussionen und die nützlichen Tipps zu verschiedenen, von mir verfassten Arbeiten.

Weiterhin danke ich auch allen Kollegen am Institut für Mess-, Regel- und Mikrotechnik für die Diskussionen, das angenehme Arbeitsklima und die willkommene Ablenkung in regelmäßigen Freizeitaktivitäten. Insbesondere möchte ich hier Mirko Mählisch, Claudia Gnahm, Christine Hartung, Andreas Wimmer, Michael Munz und Marcel Schenderlein nennen.

Mein ganz besonderer Dank gilt Martin Nieß, Franz Degenhard, Oliver Betz und Thomas Löffler für die perfekte technische Unterstützung und insbesondere für die zahlreichen Arbeiten bei der Integration der Mess- und Rechentechnik im Versuchsfahrzeug.

Ich danke auch den zahlreichen Studenten, die meine Arbeiten an zahlreichen Stellen unterstützt haben. Dies sind: Michael Schönherr, Simon Clemen, Christof Rehbach, Huo Jing, Ju Ziyang, Yo Xiao, Kent Chen, Rustam Mursaliew, Hadiuzzaman Syed, Tobias Bund, Sebastian Wiehr und Carolin Hank

Mein Dank gilt auch der Firma IBEO Automobile Sensor GmbH und insbesondere meinem ehemaligen Kollegen und jetzigem Leiter der dortigen Forschungsabteilung Kay Fürstenberg für zahlreiche interessante Projekte und Herausforderungen, die sehr gute Unterstützung bezüglich der Messtechnik sowie die Gelegenheit, meine Algorithmen erfolgreich in Onlineanwendungen zu testen.

Schließlich möchte ich meiner Familie für die regelmäßige Unterstützung bei privaten und beruflichen Fragen danken.

Inhaltsverzeichnis

Symbolverzeichnis **ix**

1 Einleitung **1**

2 Stand der Technik und Motivation **4**
2.1 Sensoren . . . 4
2.1.1 LIDAR . . . 4
2.1.2 RADAR . . . 5
2.1.3 Videosensoren . . . 6
2.2 Sensorfusion . . . 7
2.3 Erfassung der Infrastruktur . . . 9
2.4 Drahtlose Kommunikation . . . 10
2.5 Objektverfolgung . . . 11
2.6 Klassifikation . . . 13
2.7 Fahrerassistenzsysteme . . . 14
2.7.1 Überblick . . . 14
2.7.2 Kundenakzeptanz und rechtliche Aspekte . . . 16
2.7.3 Forschungsprojekte . . . 17
2.7.4 EU-Projekte . . . 18
2.8 Motivation . . . 20

3 Systembeschreibung **22**
3.1 Sensoren und Versuchsträger . . . 22
3.2 Systemüberblick . . . 23

4 Vorverarbeitung der Laserscannerdaten **28**
4.1 Sensorinterne Vorverarbeitung . . . 28
4.2 Reflektordetektion . . . 28
4.3 Ausblenden von Entfernungsmessungen . . . 29
4.4 Segmentierung . . . 30
4.5 Verdeckungserkennung . . . 30
4.6 Zusammenfassung . . . 31

5 Laserscannerbasierte Objektverfolgung **32**
5.1 Kalman-Filter . . . 32
5.2 Ablauf der Objektverfolgung . . . 35
5.3 Verwendetes Prozessmodell . . . 37
5.4 Rechteckiges Objektmodell . . . 38
5.5 Reduktion der Messvektoren . . . 40
5.6 Formbestimmung . . . 41
5.7 Segmentfusion . . . 44
5.8 Anwendungsspezifische Ausgabe . . . 46
5.9 Zusammenfassung . . . 47

6 Laserscannerbasierte Klassifikation 48
6.1 Merkmalsextraktion . . . 49
6.1.1 Merkmale der Laserscannermessdaten . . . 50
6.1.2 Merkmale der Objektverfolgung . . . 53
6.2 Künstliche neuronale Netze . . . 55
6.2.1 Neuronen . . . 55
6.2.2 Strukturen neuronaler Netze . . . 56
6.2.3 Lernverfahren . . . 58
6.2.4 Mehrfachinitialisierung . . . 61
6.2.5 Automatische Code-Generierung . . . 61
6.3 Analyse neuronaler Netze . . . 61
6.3.1 Blackbox-Charakter neuronaler Netze . . . 61
6.3.2 Stückweise lineare Netze . . . 63
6.4 Mustererkennung . . . 65
6.4.1 One-VS-All-Methode . . . 65
6.4.2 Optimierung der Merkmale . . . 65
6.4.3 Gesamtklassifikator . . . 70
6.4.4 Signifikanz der Merkmale . . . 71
6.5 Regelbasierte Klassifikation . . . 71
6.6 Anwendungsspezifische Ausgabe . . . 72
6.7 Zusammenfassung . . . 73

7 Fusionsansätze mit Videomessdaten 74
7.1 Kaskadenklassifikatoren . . . 75
7.2 Fusionsebene . . . 79
7.3 Aufmerksamkeitssteuerung . . . 80
7.4 Detektion von PKWs . . . 83
7.4.1 Erscheinungsbild von PKWs . . . 83
7.4.2 Objektauswahl . . . 84
7.4.3 Berechnung der ROI . . . 84
7.4.4 Perspektivische Rekonstruktion . . . 86
7.4.5 Anwendung des Kaskadenklassifikators . . . 86
7.5 Detektion von Fußgängern . . . 87
7.5.1 Erscheinungsbild von Fußgängern . . . 87
7.5.2 Objektauswahl . . . 88
7.5.3 Berechnung der ROI . . . 88
7.5.4 Anwendung des Kaskadenklassifikators . . . 89
7.6 Merkmalsextraktion . . . 90
7.6.1 Mehrfachdetektionen . . . 90
7.6.2 Merkmalsberechnung aus dem Kaskadenklassifikator . . . 90
7.6.3 Verwendete Merkmale . . . 92
7.7 Zusammenfassung . . . 92

8 Fusionsansätze mit präzisen digitalen Karten 93
8.1 Karteninhalt . . . 93
8.2 Positionierung . . . 95
8.3 Unsicherheiten . . . 96

8.3.1 Positionsunsicherheit . . . 96
8.3.2 Rotationsunsicherheit . . . 96
8.4 Kombination der Unsicherheiten . . . 97
8.5 Anwendung der Karteninformationen . . . 98
8.6 Zusammenfassung . . . 99

9 Nutzung drahtloser Datenübertragung 100
9.1 Motivation . . . 100
9.2 Übertragene Daten . . . 102
9.3 Ansatz zur Integration der übertragenen Daten . . . 102
9.4 Prädiktion des dynamischen Zustands . . . 103
9.4.1 Motivation . . . 103
9.4.2 Einspurmodell in WGS84-Koordinaten . . . 104
9.4.3 Prozessrauschen . . . 106
9.4.4 Versuchsaufbau . . . 107
9.4.5 Zeitliche Realisierung des Filters . . . 109
9.4.6 Prädiktion auf den Zeitpunkt der Sensormessdaten . . . 110
9.5 Assoziation . . . 110
9.5.1 Transformation in das Fahrzeugkoordinatensystem . . . 111
9.5.2 Gating . . . 112
9.5.3 Assoziationswahrscheinlichkeiten . . . 112
9.5.4 Zeitliche Filterung . . . 115
9.6 Offsetschätzung . . . 115
9.7 Anwendungsspezifische Ausgabe . . . 118
9.8 Zusammenfassung . . . 119

10 Experimentelle Untersuchungen 120
10.1 Beispieldaten . . . 120
10.1.1 Verwendete Datensätze . . . 120
10.1.2 Laserscannermessdaten . . . 121
10.1.3 Videomessdaten . . . 123
10.2 Zeitmessungen . . . 125
10.3 Objektverfolgung . . . 128
10.3.1 Rechenzeitreduktion durch die Optimierung der Objektverfolgung . 128
10.3.2 Auswertung der Orientierung . . . 128
10.3.3 Segmentfusion . . . 130
10.3.4 Rechenzeit . . . 130
10.4 Klassifikation . . . 131
10.4.1 Verwendete Bewertungsmethode . . . 131
10.4.2 Laserscannerbasierte Klassifikation . . . 136
10.4.3 Anpassung an Anwendungsanforderungen . . . 137
10.4.4 Rechenzeit . . . 141
10.5 Fusion von Laserscanner- und Videomessdaten . . . 141
10.5.1 Detektion von PKWs . . . 142
10.5.2 Detektion von Fußgängern . . . 144
10.5.3 Auswertung aller Objektklassen . . . 146
10.5.4 Rechenzeit . . . 146

10.6 Verwendung präziser Karten 147
10.6.1 Verbesserung der Klassifikation 147
10.6.2 Rechenzeit 148
10.7 Drahtloskommunikation 149
10.7.1 Simulation der drahtlosen Übertragung 149
10.7.2 Assoziation 149
10.7.3 Einfluss der Übertragungszeit 151
10.7.4 Analyse der Geschwindigkeitsschätzung 151
10.7.5 Rechenzeit 155

11 Zusammenfassung und Ausblick **156**

A Liste der Merkmale für die Mustererkennung **159**

B Kombination der Positions- und Rotationsunsicherheit **161**
B.1 Bereich der Straße 161
B.2 Bereich außerhalb der Straße 162
B.3 Ignorierte Gebiete 163
B.4 Verkehrsinseln 163

Literatur **164**

Betreute Studien- und Diplomarbeiten **177**

Liste der Publikationen **178**

Symbolverzeichnis

Generelle Notationsform

a	Skalare Variable
$\mathbf{a}$	Vektor
$\mathbf{A}$	Matrix
$\mathbf{A}^T$	Transponierte der Matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	Inverse der Matrix $\mathbf{A}$
$\mathbf{J}_{\mathbf{f(x)}}$	Jacobi-Matrix der Funktion $\mathbf{f}(\mathbf{x})$
$\mathbf{H}_{f(\mathbf{x})}$	Hesse-Matrix der skalaren Funktion $f(\mathbf{x})$
$\mathbf{I}$	Einheitsmatrix

Objektverfolgung

d_{min}	Geschwindigkeitsabhängiger Mindestabstand bei der Segmentfusion
$\mathbf{e}_i$	i-ter kartesischer Basisvektor
$\mathbf{f}(\ldots)$	Nichtlineares Prozessmodell
$\mathbf{F}(k)$	Systemmatrix für den Zeitschritt k
$\mathbf{G}(k)$	Steuermatrix für den Zeitschritt k
$\mathbf{h}(\ldots)$	Nichtlineares Messmodell
$\mathbf{H}(k)$	Messmatrix für den Zeitschritt k
l	Länge
$\mathbf{P}(k)$	Fehlerkovarianz des Zustands zum Zeitpunkt k
$\hat{\mathbf{P}}(k)$	Geschätzte Fehlerkovarianz des Zustands zum Zeitpunkt k
$\hat{\mathbf{P}}^-(k)$	Zeitliche Prädiktion der Fehlerkovarianz des Zustands zum Zeitpunkt k
$\mathbf{Q}(k)$	Kovarianz des Prozessrauschens für den Zeitschritt k
$\mathbf{R}(k)$	Kovarianz des Messrauschens für den Zeitschritt k
s_{Form}	Gütekriterium zur Auswahl der besten Form für gemessene Segmente
$\mathbf{S}(k)$	Innovationskovarianz für den Zeitschritt k
T	Länge des aktuellen Zeitschritts
$\mathbf{u}(k)$	Steuervektor für den Zeitschritt k
v	Betrag der Geschwindigkeit
v_x	Geschwindigkeit in Richtung der x-Achse
v_y	Geschwindigkeit in Richtung der y-Achse
$\mathbf{v}(k)$	Prozessrauschen für den Zeitschritt k
w	Breite
$\mathbf{w}(k)$	Messrauschen für den Zeitschritt k
x	x-Koordinate einer kartesischen Position
$\mathbf{x}(k)$	Realer Zustandsvektor zum Zeitpunkt k
$\hat{\mathbf{x}}(k)$	Geschätzter Zustandsvektor zum Zeitpunkt k
$\hat{\mathbf{x}}^-(k)$	Zeitliche Prädiktion des Zustandsvektors zum Zeitpunkt k
y	y-Koordinate einer kartesischen Position
$\mathbf{z}(k)$	Messvektor zum Zeitpunkt k
$\hat{\mathbf{z}}(k)$	Prädizierter Messvektor zum Zeitpunkt k
ψ	Orientierung
ω	Gierrate

Klassifikation

b_i	Anzahl der Bedingungen für Teilbereich i
$b_{max,\ Aktivierung}$	Maximale Anzahl von Bedingungen pro Teilbereich in der Aktivierung
$b_{max,\ Ausgabe}$	Maximale Anzahl von Bedingungen pro Teilbereich in der Ausgabe
$c_{Hintergrund}$	Zugehörigkeitswert für die Klasse „Hintergrund“
$c_{Fußgänger}$	Zugehörigkeitswert für die Klasse „Fußgänger“
$c_{Zweirad}$	Zugehörigkeitswert für die Klasse „Zweirad“
c_{PKW}	Zugehörigkeitswert für die Klasse „PKW“
c_{LKW}	Zugehörigkeitswert für die Klasse „LKW“
$\mathbf{c}_{\mathbf{Klassifikation}}$	Ausgabe der Mustererkennung
$\mathbf{c}_{\mathbf{a\ priori}}$	A priori Gewichtungsvektor
$\mathbf{c}_{\mathbf{Ausgabe}}$	Vektor der ausgegebenen Zugehörigkeitswerte
E	Ausgabefehler
$E(\mathbf{w})$	Ausgabefehler als Funktion aller Gewichte
$E'(\mathbf{w})$	Gradient der Fehlerfunktion
$E''(\mathbf{w})$	Hesse-Matrix der Fehlerfunktion
$f(z)$	Ausgabefunktion eines Neurons
$F(A)$	Kumulative Verteilungsfunktion für Ereignis A
l_i	Anzahl der Teilbereiche in Partition i
$l_{max,\ Aktivierung}$	Maximale Anzahl von Teilbereichen in der Aktivierung
$l_{max,\ Ausgabe}$	Maximale Anzahl von Teilbereichen bei der Ausgabe
$l_{max,\ Skalierung}$	Maximale Anzahl von Teilbereichen bei der automatischen Skalierung
n	Anzahl der Merkmale
M	Anzahl der Neuronen in der Ausgabeschicht
$M(L)$	Anzahl der Neuronen der Schicht L
$NEES_{v,stationär}$	*NEES* der Geschwindigkeit, Annahme eines stationären Objekts
$NEES(k)$	Aktueller *NEES* der Geschwindigkeit des betrachteten Objekts
$P(A)$	Wahrscheinlichkeitsdichtefunktion für Ereignis A
r	Anzahl der kombinierten Partitionen
$\mathbf{r}$	Richtung der Gewichtsanpassung
$s_{i,j}$	Stützstelle für die manuelle Skalierung
$\mathbf{t}$	Vorgabe für den Ausgabevektor beim Training eines Netzes
v_j	Koeffizient der linearen Ausgabe eines Teilbereiches
$v_{i,j}$	Koeffizient der linearen Ausgabe des i-ten Teilbereiches
$u_{i,j}$	Koeffizient der linearen Begrenzungen eines Teilbereiches
w_j	Gewicht des j-ten Eingangs eines Neurons
$\mathbf{w}$	Gewichtsvektor
$\Delta\mathbf{w}$	Update des Gewichtsvektors
$w_{i,j}^L$	Gewicht des j-ten Eingangs des i-ten Neurons der Schicht L
$\Delta w_{i,j}^L$	Gewichtsupdate für den j-ten Eingangs des i-ten Neurons der Schicht L
x_i	i-tes Merkmal des Merkmalsvektors
$\mathbf{x}$	Eingangsvektor der Mustererkennung mit $\dim(\mathbf{x}) = n$ Merkmalen
y	skalarer Ausgabewert eines neuronalen Netzes oder eines Neurons
y_i	i-ter Ausgabewert des Ausgabevektors
y_i^L	Ausgabe des i-ten Neurons der Schicht L
$\mathbf{y}$	Ausgabevektor der Mustererkennung

z	Aktivierung eines Neurons
z_i^L	Aktivierung des i-ten Neurons der Schicht L
α	Schrittweite der Gewichtsanpassung
η	Lernrate

Fusionsansätze mit Videomessdaten

A	Pixelwert zur Berechnung der Pixelsumme eines Rechtecks
B	Pixelwert zur Berechnung der Pixelsumme eines Rechtecks
C	Pixelwert zur Berechnung der Pixelsumme eines Rechtecks
d_{Objekt}	Distanz des Objekts zur Kamera
D	Pixelwert zur Berechnung der Pixelsumme eines Rechtecks
f	Brennweite der Kamera
h_{Bild}	Objektgröße im Videobild
h_{Objekt}	Reale Objektgröße
$h_j\left(k, x_{i,j}\right)$	Primitiver Klassifikator für Merkmal j zum Zeitschritt k
$H(\mathbf{x})$	Gesamtklassifikator für eine Kaskadenstufe
$j\,(k)$	Bestes Merkmal zum Zeitschritt k
l	Anzahl der Trainingsbeispiele für Hintergrund
m	Anzahl der Trainingsbeispiele für den gesuchten Objekttyp
R	Summe der Pixel unter einem achsenparallelen Rechteck
w_i	Gewicht des Trainingsbeispiels i
x	x-Koordinate eines Pixels
$\mathbf{x_i}$	Merkmalsvektor des Trainingsbeispiels i
y	y-Koordinate einer kartesischen Position
y_i	Klasse des Trainingsbeispiels i
$\beta_{j(k)}$	Gewicht für den primitiven Klassifikator $h_{j(k)}\left(k, x_{i,j(k)}\right)$
ϵ_j	Klassifikationsleistung des primitiven Klassifikators $h_j\left(k, x_{i,j}\right)$

Nutzung drahtloser Datenübertragung

a	Beschleunigung
a_{WGS84}	Große Halbachse des Erdellipsoids
$A\,(r, \alpha)$	Flächeninhalt des Gitterelements an der Stelle (r, α)
b_{WGS84}	Kleine Halbachse des Erdellipsoids
e_{WGS84}	Exzentrizität
$\mathbf{f_d}$	zeitdiskrete Zustandsübergangsfunktion
$\mathbf{f_r}\left(\mathbf{x_g}\right)$	Transformation von globalen Koordinaten in Fahrzeugkoordinaten
$\mathbf{f_{offset}}$	Funktion zur Berechnung des kartesischen Offsets
$g\,(r, \alpha)$	Funktionswert der Gaußglocke an der Stelle (r, α)
h	Höhe
k	Zeitschritt
$P\,(i)$	Assoziationswahrscheinlichkeit für Objekt i
$P\,(i\|M)$	Assoziationswahrscheinlichkeit für Objekt i, falls eine Messung vorliegt
$P\,(M)$	Messwahrscheinlichkeit für das empfangene Objekt
$P\,(r, \alpha)$	Verdeckungswahrscheinlichkeit an der Stelle (r, α)
$\mathbf{P_e}$	Kovarianz der globalen Zustandsschätzung des Versuchsfahrzeugs
$\mathbf{P_g}$	Kombinierte Kovarianz für Versuchsfahrzeug und Objekt

$\mathbf{P_o}$	Kovarianz der globalen Zustandsschätzung des empfangenen Objekts
$\mathbf{P_{pos}}$	Kovarianz für die Position im Fahrzeugkoordinatensystem
$\mathbf{P_r}$	Kovarianz für den Zustandsvektor im Fahrzeugkoordinatensystem
q_a	Rauschterm für die Beschleunigung
q_ω	Rauschterm für die Gierrate
$\mathbf{Q}\,(k)$	Zeitdiskretes Prozessrauschen
$\mathbf{Q_c}\,(k)$	Zeitkontinuierliches Prozessrauschen
$\mathbf{Q_d}\,(t_1, t_2)$	Auswirkung des Rauschens zum Zeitpunkt t_1 auf den Zeitpunkt t_2
$\mathbf{Q_{offset}}\,(k)$	Prozessrauschen für die Offsetschätzung
$\mathbf{Q}_{a,a}\,(k)$	Element aus $\mathbf{Q}\,(k)$ für die Varianz der Beschleunigung
$\mathbf{Q}_{\omega,\omega}\,(k)$	Element aus $\mathbf{Q}\,(k)$ für die Varianz der Gierrate
r	Radius im Polarkoordinatensystem bei der Verdeckungserkennung
r_λ	Ellipsenradius zur Berechnung der Änderung des Längengrades
r_φ	Ellipsenradius zur Berechnung der Änderung des Breitengrades
$\mathbf{R_m}$	Messfehlerkovarianz des Laserscanners für das assoziierte Objekt
$\mathbf{R_k}$	Kombinierte Kovarianz zur Bestimmung des kartesischen Offsets
$\mathbf{R_{offset}}$	Kovarianz des gemessenen Offsets in kartesischen Koordinaten
$s_{Assoziation}$	Schwellwert zur Ausgabe einer Assoziation
t	Zeitpunkt
Δt	Länge eines Zeitschrittes
υ	Geschwindigkeit
x_m	x-Koordinate der sensorbasiert gemessenen Position
x_r	x-Koordinate der empfangenen Position
$\mathbf{x}$	Zustandsvektor
$\mathbf{x_g}$	Kombinierter Zustandsvektor für Versuchsfahrzeug und Objekt
$\mathbf{x_i}$	Zustandsvektor des sensorbasiert erfassten Objekts i
$\mathbf{x_k}$	Kombinierter Zustandsvektor zur Bestimmung des kartesischen Offsets
$\mathbf{x_{offset}}$	Zustandsvektor zur Schätzung des globalen Offsets
$\mathbf{x_r}$	Relativer, empfangener Zustandsvektor im Fahrzeugkoordinatensystem
$\Delta\mathbf{x}\,(r, \alpha)$	Abstand des Gitterelements (r, α) von der empfangenen Position
y_m	y-Koordinate der sensorbasiert gemessenen Position
y_r	y-Koordinate der empfangenen Position
$\mathbf{z_{offset}}$	Gemessener Offset in kartesischen Koordinaten
α	Winkel im Polarkoordinatensystem bei der Verdeckungserkennung
γ	Schwelle für die 3σ-Ellipse
λ	Längengrad
σ	Standardabweichung
τ	Zeitraum
φ	Breitengrad
ψ	Orientierung
ω	Gierrate
$\mathcal{N}\,(\mathbf{x}, \boldsymbol{\mu}, \mathbf{P})$	Multivariate Gaußverteilung

Experimentelle Untersuchungen

m	Anzahl der Testbeispiele
p_0^{ist}	Klassifikationswahrscheinlichkeit für die Klasse „Hintergrund“
p_1^{ist}	Klassifikationswahrscheinlichkeit für die Klasse „Fußgänger“
p_2^{ist}	Klassifikationswahrscheinlichkeit für die Klasse „Zweirad“
p_3^{ist}	Klassifikationswahrscheinlichkeit für die Klasse „PKW“
p_4^{ist}	Klassifikationswahrscheinlichkeit für die Klasse „LKW“
p_0^{soll}	A priori Wahrscheinlichkeit für das Auftreten der Klasse „Hintergrund“
p_1^{soll}	A priori Wahrscheinlichkeit für das Auftreten der Klasse „Fußgänger“
p_2^{soll}	A priori Wahrscheinlichkeit für das Auftreten der Klasse „Zweirad“
p_3^{soll}	A priori Wahrscheinlichkeit für das Auftreten der Klasse „PKW“
p_4^{soll}	A priori Wahrscheinlichkeit für das Auftreten der Klasse „LKW“
s_{Form}	Gütekriterium zur Auswahl der besten Form für gemessene Segmente
$\alpha_{Fußgänger}$	Öffnungswinkel der ROI für den Fußgängerschutz
σ_{ψ}	Standardabweichung des Messrauschens der Orientierung

1 Einleitung

Der Individualverkehr ist heute bereits ein fester Bestandteil unserer modernen Gesellschaft und gilt als Ausdruck von Freiheit und Flexibilität. Trotz erheblicher Verteuerung der Mineralölpreise und der dadurch zu erwartenden bewussteren Nutzung von Fahrzeugen nahm die Anzahl der in Deutschland zugelassenen Fahrzeuge auch im Jahr 2006 wieder zu [127]. Die kontinuierliche Erhöhung der Verkehrsdichte führt zu einer immer stärkeren Belastung der Verkehrsteilnehmer.

Während unter dem Aspekt des Umweltschutzes effizientere und sauberere Antriebskonzepte entwickelt werden, wird unter dem Gesichtspunkt der Verkehrssicherheit eine Reduktion von Unfällen mit Verletzungen oder Todesfolge angestrebt. Moderne Fahrzeuge besitzen bereits eine Vielzahl von Sicherheitssystemen. Aktive Systeme wie das ABS (Antiblockiersystem) oder das ESP (Elektronisches Stabilitätsprogramm) verbessern die Fahrstabilität. Passive Systeme wie Sicherheitsgurte, Airbags oder Seitenaufprallschutz reduzieren die Auswirkungen eines Unfalls auf die Fahrzeuginsassen.

Infolge der zunehmenden Ausstattung des Fahrzeugbestandes mit entsprechenden Sicherheitssystemen nahm die Anzahl der im Straßenverkehr tödlich verunglückten Personen in den letzten Jahren kontinuierlich ab [127]. Trotz dieser positiven Tendenz starben im Jahr 2005 allein in Deutschland 5.361 Personen infolge schwerer Verkehrsunfälle. Betrachtet man alle Unfälle mit Verletzten und Getöteten im Jahr 2005, dann wird in einem überwiegenden Anteil von 86,4% als Ursache ein Fehlverhalten der Fahrzeugführer genannt [127].

Um eine weitere Verbesserung der Verkehrssicherheit zu erreichen, wurde mit der Entwicklung und Integration neuer innovativer Systeme begonnen, die den Mensch beim Fahren entlasten und unterstützen können. Diese modernen Fahrerassistenzsysteme werden bereits in gefährlichen Situationen aktiv, bevor ein Unfall eintritt. Dafür werden Informationen über andere Verkehrsteilnehmer und das weitere Fahrzeugumfeld benötigt. Zu diesem Zweck wird mit Hilfe verschiedener Sensoren das Fahrzeugumfeld vermessen. Eine anschließende Sensordatenverarbeitung versorgt die Assistenzsysteme mit den benötigten Informationen.

Erste, auf dem Markt erhältliche Systeme, wie zum Beispiel Nachtsichtsysteme zur Verbesserung der Sicht bei eingeschaltetem Abblendlicht, sind meist sehr eng mit einem bestimmten Sensor verkoppelt. Parallel dazu werden in die Fahrzeuge andere Systeme integriert, die nicht selten unabhängig davon andere Sensoren verwenden. So wird beispielsweise die Spurverlassenswarnung in einigen Fahrzeugen mit einer zusätzlichen Kamera realisiert. Auch im Bereich der Komfortsysteme werden unterschiedliche Sensoren für unterschiedliche Assistenten eingesetzt. So verwenden abstandsgeregelte Tempomaten (*engl. ACC = adaptive cruise control*) in der Regel RADAR-Sensoren, während für Einparkassistenten oft Ultraschall-Sensoren integriert werden.

Zukünftige Fahrerassistenzsysteme können von einer gemeinsamen Nutzung der Sensoren und der Sensordatenverarbeitung profitieren. Hierfür wird eine zusätzliche Abstraktionsschicht benötigt, die die gesammelten Informationen aus dem Fahrzeugumfeld repräsentiert und den einzelnen Fahrerassistenzsystemen zur Verfügung stellt.

Im Rahmen dieser Arbeit wird ein solches Fahrzeugumfelderkennungssystem vorgestellt. Das System führt basierend auf Messdaten verschiedener Sensoren eine zeitliche Verfolgung

und Klassifikation der im Fahrzeugumfeld detektierten Objekte durch. Zusätzlich können Informationen aus drahtloser Kommunikation berücksichtigt werden. Das resultierende Fahrzeugumfeldmodell kann dann simultan verschiedene Fahrerassistenzsysteme unter unterschiedlichen Vorgaben mit Informationen versorgen.

Während sich viele aktuelle Systeme vorerst auf Funktionalitäten für Autobahnen und Landstraßen beschränken, ist das hier vorgestellte System für eine Nutzung in allen Bereichen des Straßenverkehrs konzipiert. Dies schließt auch eine Anwendung in städtischen Gebieten und auf Kreuzungen ein.

Für das Design des vorgestellten Systems war dessen Echtzeitfähigkeit von besonderer Bedeutung, da die entsprechenden Algorithmen zur Fahrzeugumfelderkennung ihre Berechnungen während des Fahrens durchführen müssen, um eine geeignete Unterstützung und Entlastung des Fahrers zu ermöglichen. Viele der hier vorgestellten Algorithmen wurden im Rahmen von Unteraufträgen für Forschungsprojekte der Europäischen Union entwickelt. So wurde die laserscannerbasierte Objektverfolgung sowie die auf Scanner- und Kartendaten basierende Klassifikation für das Projekt INTERSAFE [62] entwickelt, in einem Demonstrationsfahrzeug des Projektes integriert und erfolgreich in Onlineanwendungen getestet. Die Algorithmen zur Integration drahtloser Kommunikation wurden für das Projekt SAFESPOT [118] entwickelt. Eine kurze Vorstellung der von der Europäischen Union geförderten Projekte wird in Abschnitt 2.7.4 gegeben.

Diese Arbeit ist folgendermaßen gegliedert:

Im folgenden Kapitel wird ein Überblick über den aktuellen Stand der Technik im Bereich Fahrerassistenzsysteme gegeben. Das Kapitel schließt mit der Motivation für diese Arbeit.

In Kapitel 3 erfolgt eine Beschreibung der verwendeten Hardware. Weiterhin gibt dieses Kapitel einen Überblick über die entwickelten Softwaremodule. Schließlich wird die Integration in einem verteilten System im Fahrzeug diskutiert.

Kapitel 4 stellt verschiedene Algorithmen zur Vorverarbeitung der vom Laserscanner aufgenommenen Entfernungsmessungen vor. Dies beinhaltet sensorinterne Vorverarbeitungsschritte, die Erkennung von Reflektoren und die Ausblendung von störenden Entfernungsmessungen. Algorithmen zur Segmentierung gruppieren einzelne Entfernungsmessungen, die voraussichtlich vom selben Objekt stammen. Weiterhin werden unter Berücksichtigung des radialen Messprinzips des Laserscanners Informationen über potenzielle Verdeckungen gesammelt.

Eine Beschreibung der laserscannerbasierten Objektverfolgung erfolgt in Kapitel 5. Im Rahmen der Objektverfolgung wurden bestehende Arbeiten erweitert und für die Echtzeitanwendung in Demonstrationsfahrzeugen optimiert.

Im Anschluss wird in Kapitel 6 das Gesamtsystem zur Objektklassifikation vorgestellt. Die Mustererkennung basiert auf einer speziellen Struktur neuronaler Netze, die automatisch die Merkmale in geeigneter Weise skaliert und darüber hinaus nach dem Training noch analysiert werden kann. Zusätzlich werden die Methoden der Mustererkennung mit regelbasiertem a priori Wissen kombiniert. Ein zeitlicher Filter kombiniert die Klassifikationsergebnisse des aktuellen Zeitschrittes mit den vorherigen Ergebnissen. Schließlich erfolgt eine anwendungsspezifische Selektion der zugeordneten Objektklasse.

In Kapitel 7 wird die Objektklassifikation um Fusionsansätze mit Videodaten erweitert. Hier dienen die Messwerte des Laserscanners zur Aufmerksamkeitssteuerung für die Bildverarbeitung. Die Algorithmen der Bildverarbeitung ermitteln mit Hilfe eines dreidimensionalen Objektmodells zusätzliche Merkmale für die Klassifikation von Fußgängern und PKWs.

Kartenbasierte Fusionsansätze werden in Kapitel 8 diskutiert. Hier kommen die Ergebnisse einer Arbeit zur präzisen Positionierung in digitalen Karten zum Einsatz. Basierend auf diesen bestehenden Arbeiten verbessern neue Methoden die Klassifikation durch die Auswertung der in der Karte enhaltenen Zusatzinformationen.

Kapitel 9 beinhaltet die Beschreibung neuer Methoden zur Erweiterung des sensorbasierten Fahrzeugumfeldmodells durch drahtlos übertragene Informationen. Dabei liegt der Schwerpunkt auf der Zuordnung empfangener Informationen zu Objekten, die mit den Sensoren des Fahrzeugs erfasst und verfolgt wurden.

Schließlich erfolgt in Kapitel 10 eine detaillierte Auswertung der entwickelten Algorithmen. Diese basieren auf einer zu diesem Zweck erstellten Testdatenbank. Nach der Beschreibung neuer Methoden zur Erstellung von Beispieldaten erfolgt die Auswertung der neuen Algorithmen in der Reihenfolge ihrer Vorstellung.

Die Arbeit schließt mit der Zusammenfassung und einem Ausblick.

2 Stand der Technik und Motivation

2.1 Sensoren

Zur Realisierung vieler moderner Fahrerassistenzsysteme werden detaillierte Informationen über das aktuelle Fahrzeugumfeld benötigt. Zur Fahrzeugumfeldvermessung werden typischerweise Abstandssensoren wie LIDAR oder RADAR sowie Videosensoren eingesetzt. Ultraschallsensoren werden aufgrund ihres niedrigen Preises für Einparkhilfen verwendet, sind jedoch aufgrund ihrer sehr begrenzten Reichweite nicht für eine allgemeine Fahrzeugumfelderkennung geeignet. Im Folgenden werden die verschiedenen Messprinzipien der für die Fahrzeugumfelderkennung typischen Sensoren näher beschrieben.

2.1.1 LIDAR

LIDAR-Sensoren (*engl. LIDAR = light detection and ranging*) basieren auf der Entfernungsmessung nach dem Prinzip der Lichtlaufzeit. Im Bereich der Fahrzeugumfelderkennung werden in der Regel Laserpulse im nicht sichtbaren Infrarotspektrum von etwa 900 nm verwendet. Wenn die vom Sensor ausgesendeten Laserpulse auf einen Gegenstand treffen, wird ein Teil der ausgesendeten Leistung zurück zur Empfangseinheit reflektiert. Durch Auswertung der Laufzeit des Laserpulses lässt sich mit Hilfe der bekannten Lichtgeschwindigkeit die Entfernung des reflektierenden Gegenstands bestimmen. Im Wesentlichen werden zwei Bauformen für LIDAR-Sensoren unterschieden, die in den folgenden Absätzen beschrieben werden.

Beim Mehrstrahl-LIDAR (*engl. multibeam LIDAR*) werden mehrere fest verbaute, horizontal angeordnete Sende- und Empfangseinheiten verwendet. Jede Einheit liefert entsprechend ihrer Ausrichtung Entfernungsmessungen. Der Vorteil dieser Bauweise ist der Verzicht auf bewegliche Bauteile. Aufgrund der begrenzten Anzahl von Sende- und Empfangseinheiten ist jedoch die Anzahl der verfügbaren Entfernungsmessungen stark begrenzt. Das bedeutet, dass entsprechende Sensoren bei hoher Auflösung nur einen kleinen Winkelbereich abdecken können, wie beispielsweise der LIDAR-Sensor der Firma Hella [135]. Alternativ sind natürlich auch Sensoren mit großem Öffnungswinkel und niedriger Auflösung realisierbar. Ein Beispiel für einen solchen Sensor bietet Continental Automotive Systems mit einem dreistrahligen Nahbereichssensor [53, 111].

Das auch als Laserscanner bezeichnete, scannende LIDAR lenkt den gesendeten Laserpuls über einen Spiegel oder ein Prisma in die Messrichtung um. Durch das Rotieren des Spiegels bzw. des Prismas können nacheinander radiale Entfernungsmessungen in verschiedene Richtungen durchgeführt werden. Um auch in Fällen von Bodenunebenheiten oder dem Einnicken des Fahrzeugs Messungen liefern zu können, werden oft mehrere Sende-/Empfangseinheiten verwendet, die in unterschiedlichen vertikalen Winkeln arbeiten. Hierdurch werden Messungen in mehreren übereinander liegenden Messebenen erzeugt.

Ein scannender LIDAR-Sensor erfordert durch die beweglichen Teile einen aufwendigeren Aufbau als ein Mehrstrahl-LIDAR. Da eine einzelne Sende- und Empfangseinheit nacheinander in verschiedene Richtungen messen muss, ist die erreichbare maximale Messfrequenz deutlich niedriger als bei Mehrstrahl-Sensoren. Der Vorteil scannender

LIDAR-Sensoren liegt im extrem großen horizontalen Winkelbereich, in dem Messungen mit hoher Winkelauflösung durchgeführt werden können. Scannende LIDAR-Sensoren liefern beispielsweise die Firmen Denso [20, 131] und IBEO AS [45, 60].

2.1.2 RADAR

RADAR-Sensoren (*engl. RADAR = radio detection and ranging*) führen Entfernungsmessungen mit Radarstahlen im Wellenlängenbereich von wenigen Millimetern durch. Zur Entfernungsmessung existieren zwei gängige Messprinzipien.

Das pulsmodulierte RADAR arbeitet vergleichbar mit dem Prinzip der LIDAR-Sensoren mit kurzen Wellenpulsen, deren Laufzeit zwischen Sender, Objekt und Empfänger ausgewertet wird. Zur Bestimmung der Relativgeschwindigkeit kann die durch den Dopplereffekt bedingte Frequenzänderung des gesendeten Pulses ausgewertet werden. Entsprechende Geräte werden als Puls-Doppler-Radar bezeichnet.

Das frequenzmodulierte Dauerstrichradar (*engl. FMCW = frequency modulated continuous wave*) verändert zeitlich die Frequenz einer kontinuierlich ausgesendeten Millimeterwelle. Über die zeitliche Frequenzverschiebung des empfangenen Signals kann dann die Entfernung ermittelt werden. Auch bei diesem Verfahren kann die Frequenzänderung durch den Dopplereffekt zur Bestimmung der Relativgeschwindigkeit ausgenutzt werden. Diese Geschwindigkeit äußert sich in einem Frequenzoffset des empfangenen Signals [160]. Zur Bestimmung der Richtung des gemessenen Objektes kommen wie beim LIDAR Mehrstrahl- und scannende Ansätze zur Anwendung [158]. Weiterhin werden die Sensoren nach der Anordnung von Sende- und Empfangseinheiten unterschieden. Bei monostatischen Verfahren wird eine Antenne für das Senden und Empfangen benutzt. Bistatische Verfahren verwenden jeweils eine Sende- und eine Empfangsantenne. Bei multistatischen Verfahren sendet eine Antenne, während mehrere verteilte Antennen die reflektierte Welle empfangen [115].

Für den KFZ-Bereich werden zwei Frequenzbänder benutzt. Das Band bei 24 GHz wird für den Nahbereich eingesetzt und liefert Messungen in einem relativ großen Winkelbereich. Das Frequenzband bei 76 GHz wird für den Fernbereich bis 200 m genutzt, deckt aber nur einen sehr kleinen Winkelbereich ab [37].

RADAR-Sensoren für den Automobilbereich werden von verschiedenen Herstellern angeboten. Bosch stellt einen FMCW-Radar im 76 GHz Band mit vier Strahlen her [11, 111]. Continental Automotive Systems produziert verschiedene Puls-Doppler-Radar-Sensoren im 76 GHz Band [111, 17] sowie ein Puls-Doppler-Radar für den Nahbereich im 24 GHz Band [111]. Weitere RADAR-Sensoren werden beispielsweise von den Firmen Denso, Delphi, Hella, TRW und Valeo hergestellt [111].

Der Vorteil von RADAR-Sensoren liegt in der Eigenschaft, Relativgeschwindigkeiten direkt zu messen. LIDAR-Sensoren müssen hier eine Objektverfolgung durchführen und die Geschwindigkeit durch Ableitung der Positionsmessung schätzen. Aufgrund der beim RADAR verwendeten, großen Wellenlänge werden die Messungen deutlich weniger von Nebel und Regen gestört. Erst bei größeren Regentropfen tritt eine deutliche Reichweitenreduktion durch Dämpfung auf [158]. Die große Wellenlänge wirkt sich aber auch gleichzeitig nachteilig aus: Die Strahlen lassen sich deutlich schlechter fokussieren als beim

LIDAR. Damit kann die Richtung des gemessenen Objekts vergleichsweise schlecht bestimmt werden. Die große Strahlaufweitung führt zusätzlich dazu, dass mit einem Strahl unterschiedliche Teile eines Objektes gemessen werden. Die ausgegebene Entfernungsmessung ist dann ein Mittelwert der einzelnen Abstände und kann bei der zeitlichen Änderung der relativen Lage stark schwanken. LIDAR-Sensoren erzeugen hier meist deutlich konstantere Entfernungsmessungen und bieten eine deutlich höhere Winkelauflösung. Ein weiterer Vorteil der LIDAR-Sensoren liegt in den fehlenden Frequenzbandeinschränkungen. Für RADAR-Sensoren stehen lediglich zwei Frequenzbänder zur Verfügung, wobei das Frequenzband bei 24 GHz mit anderen Anwendungen wie Funkastronomie und der Satelliten-Erderkundung geteilt werden muss und die Nutzung momentan zeitlich bis zum Jahr 2013 limitiert ist. Als alternative Frequenz ist für die Zeit danach ein Band bei 79 GHz festgelegt worden [30].

2.1.3 Videosensoren

Videokameras bilden neben entfernungsmessenden Sensoren eine weitere wichtige Sensorgruppe für die Fahrzeugumfelderfassung. Durch ihre hohe Auflösung bieten Kameras erheblich mehr Informationen über beobachtete Objekte als Abstandssensoren. Aufgrund der fehlenden Abstandsinformation ist die Trennung zwischen unterschiedlichen Objekten jedoch deutlich schwieriger. Dementsprechend ist die Sensordatenverarbeitung für Videokameras auch aufwendiger und benötigt vergleichsweise viel Rechenzeit.

Im Bereich des für das menschliche Auge sichtbaren Lichts werden zwei Technologien für Videosensoren unterschieden. Beide Varianten beruhen auf dem lichtelektrischen Effekt, das heißt, Photonen bestimmter Wellenlänge werden von einem dotierten Siliziumkristall absorbiert und erzeugen dort elektrische Ladungen.

CCD-Sensoren (*engl. CCD = charge coupled device*) sammeln die erzeugten Ladungen während der Belichtungszeit in der Fotozelle. Anschließend werden diese über den gesamten Chip verschoben und am Rand seriell ausgelesen. Der Vorteil dieser Technik ist eine hohe Lichtempfindlichkeit und entsprechend niedriges Bildrauschen, da wenig zusätzliche Schaltungselemente pro Pixel auf dem Chip untergebracht werden müssen. Aufgrund des Messprinzips sind entsprechende Sensoren empfindlich für Überlaufen der erzeugten Ladungen in benachbarte Pixel (*engl. blooming*) und unerwünschte Belichtungseffekte beim Verschieben der Ladung über den Chip (*engl. smear*).

CMOS-Sensoren (*engl. CMOS = complementary metal oxide semiconductor*) wandeln die erzeugten Ladungen direkt in jedem Pixel in einen messbaren Strom um. Hierzu sind einige zusätzliche Transistoren zur Wandlung und Verstärkung in jedem Bildelement notwendig. Da diese Transistoren Teile der Chipoberfläche belegen, sinkt der Anteil der lichtempfindlichen Chipfläche. Die Folge ist ein höheres Bildrauschen. Der Vorteil dieser Technik ist eine deutlich geringere Anfälligkeit für Überbelichtungseffekte und die Möglichkeit, Teile der Vorverarbeitung direkt im Chip durchführen zu können [54, 110].

Um Farbbilder aufzunehmen, werden Kameras pixelweise direkt mit Farbfiltern versehen, so dass jedes Pixel nur das Licht einer der Farben rot, grün oder blau messen kann [110]. Dies führt zu einer Reduktion der effektiven Auflösung und einer niedrigeren Lichtempfindlichkeit.

Die Empfindlichkeit der vorgestellten Messprinzipien erlaubt auch eine Anwendung im Nahinfrarotbereich. In diesem Fall wird häufig eine aktive Beleuchtung in diesem Spektrum verwendet. Da das menschliche Auge im Nahinfrarotbereich nicht sehen kann, wird durch die aktive Beleuchtung ein großer Sichtbereich bei Nacht erreicht, ohne andere Verkehrsteilnehmer zu blenden.

Auch im Bereich der Wärmestrahlung mit einer Wellenlänge von $8\,\mu$m bis $12\,\mu$m existieren bildgebende Sensoren. Die auch als Ferninfrarot bezeichnete Strahlung wird in der Fahrzeugumfelderfassung oft mit ungekühlten Mikrobolometern erfasst [76, 111]. Diese Sensortechnologie kann Objekte wie Fußgänger oder Wild messen, deren Temperatur sich deutlich von der Umgebungstemperatur unterscheidet.

Neuere 3D-Kameras kombinieren die Aufnahme von Textur- und Abstandsinformationen [7, 57, 100, 119]. Durch Auswertung der Laufzeit von abgestrahltem, moduliertem Laserlicht wird für jedes Pixel zusätzlich zum Helligkeitswert eine Abstandsinformation bestimmt. Es wurden sowohl Sensoren mit kontinuierlicher, amplitudenmodulierter Beleuchtung, die auch als PMD-Sensoren (*engl. PMD = photonic mixer device*) bezeichnet werden [7], als auch Sensoren mit gepulstem Laserlicht entwickelt [100].

Der bisherige Nachteil von 3D-Kameras liegt in der kurzen Reichweite und einer relativ niedrigen Auflösung. So existieren zwar Sensoren mit einer Auflösung von 160×120 Pixeln, die maximal messbare Entfernung liegt hier aber bei nur 7,5 m [57].

2.2 Sensorfusion

Für eine umfassendere Fahrzeugumfelderkennung werden häufig mehrere Sensoren kombiniert eingesetzt. Die Nutzung von Sensorinformationen aus verschiedenen Sensoren wird als Sensorfusion bezeichnet. Hierfür existieren je nach Anwendungsfall verschiedene Kombinationsmöglichkeiten.

Da Videosensoren üblicherweise keine Abstandsinformationen liefern können, werden Fusionssysteme mit zwei Kameras benutzt, um ähnlich zum menschlichen Sehvermögen Tiefeninformationen bestimmen zu können. Bei dem als Stereo-Vision bezeichneten Verfahren werden in beiden Kamerabildern korrespondierende Bildausschnitte gesucht. Aufgrund der bekannten Anbaupositionen und -winkel der Kameras kann dann mit den Positionen der korrespondierenden Bildausschnitte in den Einzelbildern ein Abstand zum Kamerasystem bestimmt werden [34]. Ähnliche Berechnungen sind auch durch Auswertung zeitlich versetzter Bilder einer bewegten Kamera möglich [39].

Eine weiterer Ansatz zur Fusion kombiniert mehrere Entfernungssensoren zur Erweiterung des Messbereichs. Hierbei werden mit den Sensoren unterschiedliche Bereiche des Fahrzeugumfelds vermessen [104, 112, 124]. Die Hauptaufgabe der Fusionsalgorithmen besteht hier in der Wahrung der Konsistenz beim Übergang vom Messbereich eines Sensors in den eines anderen Sensors.

Die meisten Vorteile verspricht man sich von der Fusion heterogener Sensoren. Aus diesem Grund werden häufig Abstandssensoren mit Videosensoren fusioniert [69, 94, 131, 145]. Die Informationen des Abstandssensors können zur Identifikation von Objekthypothesen im Videobild verwendet werden. Dabei kann sowohl die Objektposition als auch die Objektgröße eingeschränkt werden. Anschließende Bildverarbeitungsalgorithmen können dann zur Verbesserung der Positionsschätzung [69] oder der Objektklassifikation [94] beitragen.

Eine wichtige Entscheidung beim Design eines Fusionssystems besteht in der Wahl der Abstraktionsebene, auf der die Sensordaten fusioniert werden sollen. Häufig wird auf der Objektebene (*engl. high level fusion*) fusioniert [28, 36, 78, 104]. Hierbei wird die komplette Sensordatenverarbeitung bis zur Liste verfolgter Objekte getrennt für jeden einzelnen Sensor durchgeführt. Anschließend werden die Objektdaten aus den unterschiedlichen Sensoren fusioniert. Dieser Ansatz hat den Vorteil, dass die komplette Sensordatenverarbeitung direkt im Steuergerät des Sensors durchgeführt werden kann. Zur Fusion müssen dann nur die berechneten Objektdaten an das zentrale Fusionssystem übertragen werden. Zusätzlich sind Fusionssysteme auf der Objektebene leicht skalierbar, da bei einer einheitlichen Objektschnittstelle prinzipiell Sensoren inklusive der Datenverarbeitung ausgetauscht werden können.

Der schwerwiegendste Nachteil der Fusion auf der Objektebene ist die geringe Informationsmenge, die zur Fusion zur Verfügung steht. Es sind nur noch die Objektinformationen verfügbar, die teilweise auch auf unterschiedlichen Objektmodellen basieren. Falls Mehrdeutigkeiten bei der Fusion auftreten, stehen die Messdaten zur Auflösung nicht mehr zur Verfügung. Ein informationstechnisch günstigerer Ansatz ist aus diesem Grund die Fusion auf Rohdatenebene (*engl. low level fusion*). Hier werden unter Verwendung eines gemeinsamen Objektmodells alle Sensorrohdaten zentral fusioniert. Mit Hilfe der kompletten Sensormessdaten lassen sich eventuelle Mehrdeutigkeiten einfacher auflösen. Der Nachteil dieses Ansatzes liegt im hohen Übertragungsaufwand für die Sensorrohdaten und der schlechten Skalierbarkeit des Systems.

Um die Vorteile beider Ansätze vereinen zu können, wurde der Fusionsansatz auf der Merkmalsebene (*engl. feature level fusion*) vorgestellt [66]. Hierbei werden vorverarbeitete Sensormessdaten zentral fusioniert. Dafür werden geeignete Merkmale für jeden einzelnen Sensor berechnet, die alle notwendigen Sensorinformationen enthalten. Anstatt die gesamten Rohmessdaten zu übertragen, müssen dann nur die Merkmale an die zentrale Fusionseinheit übermittelt werden. Hier erfolgt wie bei der Fusion auf Rohdatenebene eine Objektbildung mit einem einheitlichen Objektmodell. Bei geeigneter Wahl der Merkmale besteht auch weiterhin eine Skalierbarkeit des Systems [66].

In der Literatur werden Fusionsansätze, die auf Rohdaten- oder Merkmalsbasis arbeiten, auch als frühe Fusion (*engl. early fusion*) bezeichnet. Beispiele hierfür finden sich in [84, 133, 145]. Unterschiedliche Fusionsebenen werden auch in [134] diskutiert.

Falls Sensorinformationen auf Rohdaten- oder Merkmalsebene fusioniert werden, ist nach [66] eine synchrone Aufnahme der unterschiedlichen Sensordaten empfehlenswert. Dies kann mit einem entsprechenden Aufbau realisiert werden, wenn triggerbare Sensoren vorliegen bzw. maximal ein Sensor nicht triggerbar ist, aber seinen Messzeitpunkt entsprechend signalisiert. Werden die Sensordaten nicht synchron aufgenommen, kann sich die Perspektive aufgrund der bewegten Sensorplattform zwischen den Aufnahmen ändern. Zusätzlich kann sich die aufgenommene Szene durch bewegte Objekte verändern. Die daraus

resultierenden Änderungen in den Messdaten lassen sich oft nicht kompensieren und führen zu größeren Ungenauigkeiten bei der Fusion.

Wenn Sensorinformationen aus verschiedenen Sensoren fusioniert werden, müssen diese vorher in ein gemeinsames Koordinatensystem transformiert werden. Hierfür bietet sich das nach DIN 70000 genormte Fahrzeugkoordinatensystem an [66]. Für eine derartige Transformation müssen die Anbaupositionen und -orientierungen der Sensoren relativ zu dem gemeinsamen Koordinatensystem bekannt sein. Diese Information wird aus der Sensorkalibrierung gewonnen. Hierbei wird die intrinsische Kalibrierung und die extrinsische Kalibrierung unterschieden.

Die intrinsische Kalibrierung bestimmt die internen Sensorparameter zur Umrechnung der Sensordaten in ein dreidimensionales, kartesisches Koordinatensystem. Sie erfolgt in der Regel einmalig mit einem speziellen Kalibrieraufbau. Die intrinsische Kalibrierung einer Kamera kann zum Beispiel mit Hilfe eines Schachbrettmusters [12] durchgeführt werden.

Die extrinsische Kalibrierung ermittelt die Anbauposition und -orientierung der Sensoren. Sie kann beim Aufbau offline mit entsprechenden Kalibrierobjekten [66, 94] oder online anhand markanter Punkte in den Sensordaten [107, 162] erfolgen.

Mit Hilfe der Kalibrierung können neben der Transformation der Sensordaten in das gemeinsame Fahrzeugkoordinatensystem auch Transformationen aus dem gemeinsamen Fahrzeugkoordinatensystem in die dreidimensionalen Sensorkoordinatensysteme oder in die Bildebene der Kamera durchgeführt werden.

2.3 Erfassung der Infrastruktur

Für viele Anwendungen sind Informationen über die das Fahrzeug umgebende Infrastruktur wie Fahrbahnverlauf, Fahrspur, Kreuzungen, Ampeln, Schilder usw. hilfreich. In diesem Fall ist jeweils die Position der Infrastrukturelemente relativ zum eigenen Fahrzeug von Interesse. So existieren bereits verschiedene Arbeiten zur videobasierten Fahrspurerkennung [3, 33, 77]. Laserscannerbasierte Arbeiten beschäftigen sich mit der Fahrbahnverlaufserkennung [22, 73, 125] sowie ebenfalls mit der Fahrspurerkennung [21, 32]. Auch radarbasierte Ansätze zu Fahrbahnverlaufserkennung werden verfolgt [102]. In [14] werden Fusionsansätze für RADAR- und Videodaten zur Detektion von Leitplanken beschrieben.

Deutlich detailliertere Infrastrukturinformationen lassen sich durch Verwendung von digitalem Kartenmaterial als zusätzlichem „Sensor“ erzielen [9]. Sämtliche in der Karte gespeicherten Infrastrukturinformationen werden durch eine Positionierung des eigenen Fahrzeugs relativ zur Karte in Bezug zum Fahrzeugumfeld gebracht. Die Positionierung kann wie in den üblichen Navigationssystemen über GPS-Sensoren (*engl. GPS = global positioning system*) beziehungsweise DGPS-Sensoren (*engl. DGPS = differential global positioning system*) erfolgen [89]. Das „NAVSTAR Global Positioning System“ (GPS) ist das satellitengestützte, globale Positionierungssystem des US-Verteidigungsministeriums. Entsprechende Sensoren empfangen von Satelliten ausgestrahlte Signale. Aufgrund der Signallaufzeiten bestimmt der Empfänger seine Position [106].

Bei Standard-GPS-Empfängern und DGPS-Empfängern erfolgt die Positionierung mit relativ hoher Unsicherheit von einigen Metern. Die Positionsunsicherheit kann durch die

Korrelation von Karte und zurückgelegter Strecke verbessert werden. Es existieren aber auch Verfahren mit deutlich präziserer Positionsbestimmung durch Fusion der Karteninformationen mit Videodaten [123] oder mit LIDAR-Messdaten [148, 150, 152].

Schließlich existieren Verfahren, die online während des Betriebs aufgrund der Sensorinformationen Karten erzeugen und sich gleichzeitig darin lokalisieren [144, 149].

2.4 Drahtlose Kommunikation

Fahrzeuge können mit der Infrastruktur und auch mit anderen Fahrzeugen über drahtlose Kommunikation Informationen austauschen. Die drahtlose Kommunikation kann also ebenfalls als zusätzliche Informationsquelle bzw. als zusätzlicher „Sensor“ interpretiert werden.

Zur Datenübertragung wird häufig WLAN-Technologie nach dem Standard IEEE 802.11 a/b/g verwendet [2, 24, 40, 56]. Einige Arbeiten beschäftigen sich auch schon mit einer speziellen Weiterentwicklung für die Fahrzeug-Fahrzeug-Kommunikation, dem IEEE 802.11p Standard [15, 40]. In manchen Fällen wird auch die Verwendung der digitalen oder der analogen Radiofrequenzen zum Empfang von Zusatzinformationen im Fahrzeug untersucht [114].

Der Aspekt der Datensicherheit bei der Übertragung persönlicher Informationen liegt ebenfalls im Fokus aktueller Forschungsaktivitäten [40, 51].

Bei der Entwicklung von Fahrerassistenzsystemen unter Verwendung von Drahtloskommunikation muss die jeweilige Ausstattungsrate der im Straßenverkehr aktiven Fahrzeuge berücksichtigt werden. So werden anfangs nur sehr wenige Fahrzeuge mit Kommunikationstechnologie ausgerüstet sein. Eine komplette Ausstattung aller Fahrzeuge wird sich in absehbarer Zeit nur schwer realisieren lassen, da aktuelle Fahrzeuge ohne diese Technologie voraussichtlich noch lange im Einsatz sein werden. Demzufolge wird drahtlose Kommunikation in absehbarer Zeit eine Fahrzeugumfelderfassung mit Sensoren wohl nicht komplett ersetzen können. Zur Verbesserung der Fahrerassistenz kann jedoch durch zusätzliche, drahtlos übertragene Informationen durchaus beigetragen werden.

Aufgrund der anfangs geringen Ausstattungsraten werden anfangs Anwendungen der drahtlosen Kommunikation zwischen Fahrzeug und Infrastruktur für den Käufer attraktiv, da hier bei entsprechender Infrastruktur sofort ein Vorteil für diesen wahrnehmbar ist. Anwendungen zur Fahrzeug-Fahrzeug-Kommunikation können erst später bei höheren Ausstattungsraten Verbesserungen erreichen, da hier jeweils mindestens zwei beteiligte Fahrzeuge ausgestattet sein müssen. Demzufolge kann eine höhere Ausstattungsrate nach der Festlegung eines einheitlichen Standards am besten durch die Bereitstellung geeigneter Infrastruktur gefördert werden.

Aktuelle Entwicklungen zur Anwendung von drahtloser Kommunikation umfassen die Übertragung von Verkehrsfunk und Kartenupdates an Navigationsgeräte [114], die Erfassung des Verkehrsflusses [96, 163] sowie die Sichtbereichserweiterung für Assistenzssysteme über die Sensorreichweite hinaus.

Für die Erweiterung des Sichtbereiches wurde in verschiedenen Arbeiten die Übertragung von Informationen über Schlechtwettergebiete, Staus oder liegengebliebene Fahrzeuge beschrieben [40, 113, 137, 138]. Im von der Europäischen Union geförderten Projekt INTERSAFE wurde ein Kreuzungsassistent vorgestellt, der unter anderem Informationen

über Ampelphasen an das Fahrzeug überträgt [24]. Zur Bestimmung der Position relativ zur Ampel wurde hier ein präziser Positionierungsansatz mit Landmarken verwendet [148]. In [10] wurde ein System zur Verkehrszeichenerkennung vorgeschlagen, das sich per DGPS positioniert.

Informationen über andere Fahrzeuge werden für automatisches, kooperatives Fahren zur Erweiterung des abstandsgeregelten Tempomats und zum automatischen Einfädeln an Engpässen [40, 113] verwendet. In [132] werden sogar Ansätze zur Kollisionsvermeidung basierend auf drahtloser Kommunikation untersucht. Hier wurde allerdings bereits ein Nachteil einer reinen Positionierung mit üblichen DGPS-Sensoren erläutert. Hohe Unsicherheiten und Messfehler verhindern eine präzise Bestimmung der relativen Lage zweier Fahrzeuge. Als Ausweg wurde hier auf aufwendigere, präzisere GPS-Sensoren mit einer Positionsunsicherheit von etwa 0,5 m zurückgegriffen. Eine weitere Idee zur Verbesserung der relativ ungenauen Positionsinformationen ist die Fusion der drahtlos empfangenen Objektdaten mit Sensormessdaten, sobald die Objekte im Sichtbereich der Sensoren liegen [68, 137, 138].

2.5 Objektverfolgung

In der Fahrzeugumfelderfassung wird die zeitliche Objektverfolgung (*engl. Tracking*) in der Regel auf die Messwerte angewendet, die die eingesetzen Sensoren bei der Vermessung anderer Verkehrsteilnehmer liefern. Da der Objekttyp bei der Messung oft unklar ist, muss die Objektverfolgung häufig auf alle gemessenen Gegenstände angewendet werden.

In einem ersten Schritt müssen hier mit geeigneten Methoden in der Objektverwaltung aus den Sensormessdaten Objekte gebildet werden. Diesen Objekten werden dabei aus den Sensormessdaten ermittelte Messgrößen wie Position oder Größe zugeordnet. In jedem weiteren Messzyklus werden dann bestehenden Objekten in der Datenassoziation jeweils neue Messgrößen zugeordnet. Die so erzeugte Objektliste muss in jedem Messzyklus weiterhin überprüft, korrigiert und erweitert werden, da Objekte den Messbereich verlassen und neue Objekte im Messbereich auftauchen können.

Aus den Messgrößen werden iterativ die Objektparameter wie Position, Größe, und Geschwindigkeit geschätzt. Das Ziel der Objektverfolgung ist die optimale Schätzung der Objektparameter, auch wenn diese Größen verrauscht, nur teilweise oder nicht direkt messbar sind. Ein Beispiel für die Schätzung eines nicht direkt messbaren Objektzustandes ist die Schätzung der Geschwindigkeit aus zeitlich versetzten Positionsmessungen. Verschiedene Methoden der Objektverfolgung werden in verschiedenen bestehenden Arbeiten diskutiert [5, 6, 8, 66, 129, 153, 157].

Arbeiten zur Objektverfolgung im Fahrzeugumfeld basieren häufig auf dem Kalman-Filter. Entsprechende Ansätze dienen der Objektverfolgung mit RADAR-Sensoren [55, 103] mit Videokameras [39, 109], LIDAR-Sensoren [44, 66, 73, 87, 128] und der Sensorfusion [66, 88, 129]. Neuere Arbeiten verwenden auch Partikelfilter zur videobasierten Fahrspurverfolgung [3, 123] und zur Objektverfolgung basierend auf verschiedenen Sensoren [52, 61].

Neuere Ansätze verwenden das Integrated-Probabilistic-Data-Association-Filter beziehungsweise auch das Joint-Integrated-Probabilistic-Data-Association-Filter zur Objektverfolgung [92, 93].

Neben unterschiedlichen Filteransätzen liegen wesentliche Unterschiede in der Bildung der Messwerte aus den Rohdaten. Da die in dieser Arbeit verwendete Objektverfolgung auf Laserscannermessdaten beruht, soll nun ein Überblick über bestehende Ansätze der Messwertgenerierung gegeben werden.

Um aus den Rohmessdaten des Laserscanners Messvektoren für verfolgte Objekte berechnen zu können, wird eine Segmentierung durchgeführt, die abstandsbasiert Messpunkte gruppiert. Dies geschieht unter der Annahme, dass dicht beieinander liegende Punkte zum selben Objekt gehören. Details über Segmentierungsverfahren können in [43, 73, 155] gefunden werden. Anschließend werden die Segmente bestehenden Objekten zugeordnet. Verschiedene Methoden für diese Datenassoziation werden in [87] verglichen. Basierend auf den zugeordneten Segmenten werden nun die Messvektoren der Objekte berechnet.

In [73] wird als Position der Schwerpunkt aller Laserscannermesspunkte verwendet. Die Messwerte für die Objektgröße werden aus einem achsenparallelen Rechteck ermittelt, dessen Größe so gewählt wird, dass es gerade alle Entfernungsmessungen des Laserscanners enthält.

Andere Arbeiten [44, 87] schlagen einen klassenabhängigen Ansatz vor. Bei PKWs wird ausgenutzt, dass der Laserscanner von einem Objekt mit rechteckiger Grundfläche nur eine oder zwei Seiten vermessen kann. Zusätzlich zeichnet sich im Fall von zwei Seiten der Eckpunkt durch die geringste Distanz zum Sensor aus. Auf diese Weise lassen sich die Laserscannermesspunkte der zwei Seiten trennen. Gemessene Seiten eines Objektes werden nun getrennt mit je einem Kalman-Filter als Linie verfolgt. Die Objektposition und -größe wird anschließend aus den einzelnen Kalman-Filtern berechnet. Auf Objekte anderer Klassen wird die bereits beschriebene Schwerpunktmethode angewendet.

In [128] wird mit Hilfe von Regressionsgeraden die Orientierung eines Objektes bestimmt. Mit der berechneten Orientierung wird ein Rechteck mit minimaler Ausdehnung berechnet, das alle Laserscannermesspunkte enthält. Diese sogenannten Bounding-Boxen beschreiben die Objektgröße. Ihr Mittelpunkt wird als Referenzpunkt genutzt. Darüber hinaus wird ein Multi-Hypothesen-Ansatz vorgeschlagen, um den möglichen Zerfall von Objekten in mehrere Segmente und Mehrdeutigkeiten in der Datenassoziation zu berücksichtigen. Hierbei werden redundante Objekte gebildet und verfolgt, die jeweils eine Möglichkeit der Datenassoziation und eine Objektklasse abdecken. Eine Auflösung der widersprüchlichen Hypothesen erfolgt später durch die Auswertung der Objekthistorie.

Ein neues Objektmodell wird in [66] beschrieben. Hierfür werden verschiedene Objektformen definiert, die den möglichen Messungen eines Objektes mit rechteckiger Grundfläche entsprechen. Es wird versucht, jede der Objektformen mit Hilfe des ICP-Algorithmus (*engl. ICP = iterative closest point*) in die zugehörigen Entfernungsmessungen des Laserscanners einzupassen. Falls die am besten passende Form einem vorgegebenen Gütekriterium entspricht, kann das Objekt als Rechteck verfolgt werden. Wenn kein Rechteck zugeordnet werden kann, dann wird versucht, die Entfernungsmessungen aufeinander folgender Zeitschritte durch den ICP-Algorithmus aufeinander abzubilden, um dadurch zumindest die Objektbewegung bestimmen zu können. Die Arbeit [66] berücksichtigt weiterhin, dass die gemessenen Seiten eines Objektes mit rechteckiger Grundfläche häufig nicht komplett vermessen werden. Als Gründe hierfür kommen Verdeckungen, der Austritt aus dem Messbereich oder eine ungünstige Orientierung zum Sensor in Frage. Um diese wechselnde Messbarkeit der Objektkontur zu berücksichtigen, wird als Positionsmesswert für das

Kalman-Filter eine geeignete Ecke des Rechtecks oder der Mittelpunkt einer der Seiten gewählt.

2.6 Klassifikation

Allgemein ist die Klassifikation eine Unterteilung von untersuchten Mustern in verschiedene Kategorien [23]. In der Fahrzeugumfelderkennung bezieht sich dies auf Objekte, die mit Sensoren vom Fahrzeug gemessen werden und in Klassen unterteilt werden sollen. Zu diesem Zweck werden verschiedene Merkmale aus den Sensordaten extrahiert, die den sogenannten Merkmalsraum aufspannen. Objekte einer Klasse werden nun durch Abgrenzung bestimmter Merkmalskombinationen von Objekten anderer Klassen unterschieden. Diese Abgrenzung entspricht einer Unterteilung des Merkmalsraums in verschiedene Bereiche mit unterschiedlichen Objektklassen.

Der einfachste Anwendungsfall der Klassifikation ist die Objektdetektion. Hier wird versucht, eine bestimmte Objektart von allen anderen gemessenen Objekten zu trennen. Beispiele für derartige gesuchte Vordergrundobjekte können zum Beispiel Ampeln [86] oder Verkehrsschilder [38, 101] im Videobild oder aber bewegte Objekte wie Fußgänger [47, 49, 91, 131, 159] oder PKWs [108, 130, 164] sein. In manchen Fällen wird auch eine Bewertung der Objektverfolgung durchgeführt, um korrekt verfolgte Objekte von verrauschten, unsicheren Objekten zu trennen [66]. Es sind aber auch Unterteilungen in mehrere Objektklassen wie zum Beispiel „Fußgänger“, „Zweirad“, „PKW“, „LKW“ und „Hintergrund“ denkbar [43, 64].

Prinzipiell kann die Klassifikation in regelbasierte Ansätze und anhand von Beispieldaten optimierte Klassifikatoren unterschieden werden. Regelbasierte Ansätze unterteilen die gemessenen Objekte aufgrund von a priori Wissen der Systementwickler. Der Vorteil derartiger Verfahren besteht in der guten Nachvollziehbarkeit des Klassifikationsprozesses. Beispiele für auf a priori Wissen basierende Systeme finden sich in [14, 43, 47, 69].

In vielen Fällen ist jedoch eine Klassifikation mit Hilfe von a priori Wissen wegen hochdimensionaler Merkmalsräume, komplexer Klassengrenzen, Messrauschen und wechselnder Beobachtbarkeit der Merkmale nur unzureichend oder überhaupt nicht realisierbar. In diesen Fällen ist es vorteilhaft, die Klassengrenzen durch einen automatischen Prozess statistisch optimal festzulegen.

Entsprechende Verfahren beruhen auf Stichproben von Beispieldaten. Diese Daten bestehen aus Paaren von berechneten Merkmalen und einer Vorgabe für die Objektklasse, die vom Systementwickler festgelegt wurde. In einer Trainingsphase werden nun die Klassengrenzen innerhalb des Merkmalsraums aufgrund der Beispieldaten festgelegt und statistisch optimiert. Anschließend können die ermittelten Klassengrenzen zur Klassifikation beliebiger Daten verwendet werden.

Bei statistisch optimierten Verfahren ist die Generalisierungsfähigkeit des trainierten Systems von Bedeutung. Diese gibt an, wie gut das System aufgrund der Optimierungen auch Objekte klassifizieren kann, die nicht im Trainingsprozess verfügbar waren. Gerade verrauschte Daten können häufig zu einer zu starken Adaption der Klassengrenzen an die Beispieldaten führen, die dann oft eine Fehlklassifikation bisher unbekannter Daten nach sich zieht.

Um der Überadaption entgegenzuwirken, werden die Beispieldaten in Trainings- und Validierungsdaten unterteilt. Eine iterative Trainingsprozedur verwendet nun nur die Trainingsdaten zur Optimierung. Nach jeder Iteration wird die Generalisierungsfähigkeit mit Hilfe der Validierungsdaten überprüft. Schließlich werden die Klassifikationsparameter mit den besten erreichten Generalisierungseigenschaften gewählt.

Zur Fahrzeugumfelderfassung wird bereits eine Vielzahl von Ansätzen für statistisch optimierte Klassifikatoren eingesetzt. So werden unter anderem Gaussian Mixture Modells [1, 80], Support Vector Machines [122, 130], Bayes'sche Netze [71] und Hyperpermutationsnetzwerke [91, 98] verwendet. Aber auch Neuronale Netze [59, 122] inklusive spezieller Vertreter, wie Time Delay Neural Networks [159] und Convolutional Neural Networks [131] kommen zum Einsatz.

Für die Videobildverarbeitung werden zusätzlich Verfahren wie das konturbasierte Chamfer Contour Matching [49, 52, 91] oder die Kaskadenklassifikatoren von P. Viola und M. Jones [141] eingesetzt [64, 72, 86, 91, 94, 108, 145, 164].

2.7 Fahrerassistenzsysteme

2.7.1 Überblick

Als Fahrerassistenzsysteme kann man alle Systeme im Fahrzeug zusammenfassen, die den Fahrer durch die Übernahme automatisierbarer Tätigkeiten entlasten oder ihn durch gezielte Versorgung mit relevanten Zusatzinformationen unterstützen. Eine Einteilung der Fahrerassistenzsysteme bezüglich ihrer Anwendung in den einzelnen Phasen eines Unfalls wird in [41] und [139] gegeben. Abbildung 2.1 zeigt eine Einteilung bisher vorgestellter Fahrerassistenzsysteme nach diesem Schema.

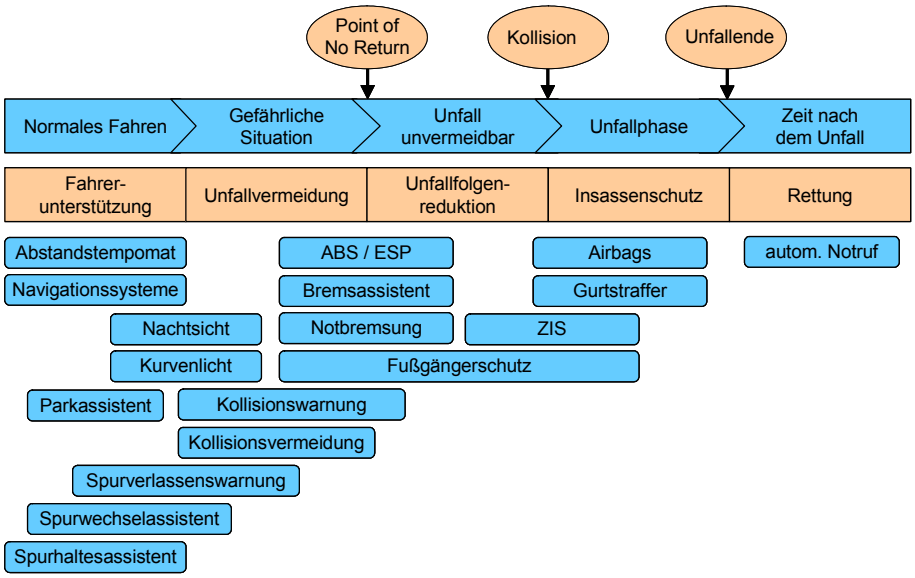

Abbildung 2.1: Einteilung aktueller Fahrerassistenzsysteme bezüglich der Anwendung in den einzelnen Phasen des Fahrens bis hin zu einem Unfall.

Nach diesem Schema ist die erste Phase vor einem eventuellen Unfall das normale Fahren. In dieser Zeit wird der Fahrer durch verschiedene Systeme in der Führungstätigkeit unterstützt. Abstandsgeregelte Tempomaten (*engl. ACC = adaptive cruise control*) übernehmen die Längsführung des Fahrzeugs. Bei neueren Ausführungen geschieht dies bis zum Stillstand (*engl. ACC Stop & Go*). Zur Regelung der Querführung existieren ebenfalls Arbeiten für Spurhalteassistenten, die das Fahrzeug durch Lenkeingriffe in der aktuellen Spur halten. Navigationssysteme informieren den Fahrer über geeignete Routen. Hier werden sogar dynamische Informationen über die Verkehrslage mit einbezogen, die von verschiedenen Radiosendern per TMC (*engl. traffic message channel*) ausgestrahlt werden. Beim Einparken unterstützen Parkassistenten. Während frühere Systeme hier nur durch Abstandsmessung mit Ultraschallsensoren informierten, übernehmen neueste Systeme sogar bereits das Lenken beim Einparken [143].

In der zweiten Phase vor dem Unfall ist die Situation bereits gefährlich, ein Unfall kann jedoch noch vermieden werden. Hier müssen Assistenzsysteme versuchen, die kritische Situation durch Information, Warnungen oder aktives Eingreifen zu entschärfen. Viele moderne Assistenzsysteme unterstützen sowohl die erste als auch die zweite Phase, da sie die tägliche Führung des Fahrzeugs vereinfachen und gleichzeitig den Fahrer gezielt über auftretende gefährliche Situationen informieren. Hierzu zählen Nachtsichtsysteme, die durch spezielle Kameras im Nah- oder Ferninfrarotspektrum den Bereich vor dem Fahrzeug aufnehmen und visualisieren, der vom Fahrer bei Abblendlicht mit bloßem Auge nicht mehr wahrgenommen werden kann. Weitere Sichtverbesserungen werden durch adaptives Kurvenlicht erreicht. Diese Systeme leuchten das Kurveninnere abhängig vom prädizierten Kurs des Fahrzeugs stärker aus. Systeme zur Spurverlassenswarnung informieren den Fahrer bei unbeabsichtigtem Überfahren der Fahrspurmarkierungen. Zusätzlich unterstützen Spurwechselassistenten den Fahrer beim gewünschten Spurwechsel. Dies geschieht üblicherweise durch Informationen über eventuell vorhandene Fahrzeuge auf benachbarten Fahrspuren. Schließlich existieren Arbeiten zur Kollisionswarnung und zur automatischen Kollisionsvermeidung.

Ab einem gewissen kritischen Zeitpunkt ist der Unfall nicht mehr zu verhindern (*engl. point of no return*). Assistenzsysteme können hier immer noch z.B. durch Verringerung der Aufprallgeschwindigkeit die Schwere des Unfalls und somit die Unfallfolgen reduzieren. Auch hier ist eine eindeutige Zuordnung oft nicht möglich, da die meisten hier angewendeten Systeme auch in der Phase der Unfallvermeidung erfolgreich agieren können. Das Antiblockiersystem (ABS) oder das Elektronische Stabilitätsprogramm (ESP) gehören inzwischen in einem Großteil der Neuwagen zur Serienausstattung. Neuere Fahrzeuge verfügen darüber hinaus über Bremsassistenten, die den Fahrer bei Vollbremsungen durch verbesserten und schnelleren Aufbau des Bremsdruckes unterstützen. Auch eine automatisch durchgeführte Notbremsung ist denkbar. Aus rechtlichen Gründen wird in [79] eine Vollbremsung nur in Situationen empfohlen, in denen der Unfall nicht mehr zu vermeiden ist. Aber auch ein automatisches Abbremsen mit reduzierter Bremskraft [18] kann dem Fahrer helfen. Betrachtet man die Reaktion des Fahrers als Teil der Kollisionswarnsysteme, können diese auch in der späten Phase noch einen Beitrag leisten, selbst wenn die Warnung bereits in der vorhergehenden Phase erfolgte. Falls die Reaktion des Fahrers nicht ausreicht, um den Unfall zu vermeiden, so ist möglicherweise noch eine Reduktion des Aufpralls zu erreichen.

Eine besondere Stellung hat der Fußgängerschutz. Während die bisher beschriebenen Sicherheitssysteme als primäres Ziel den Schutz der Fahrzeuginsassen verfolgen, besteht auch die

Notwendigkeit, andere verletzliche Verkehrsteilnehmer (*engl. VRU = vulnerable road user*) zu schützen. Speziell für Fußgänger existieren hier bereits verschiedene Ansätze. Warnsysteme können den Fahrer über Fußgänger in kritischen Bereichen des Fahrzeugumfeldes informieren.

Ab dem Zeitpunkt der Kollision bis zum Unfallende versuchen Sicherheitssysteme, die Fahrzeuginsassen bestmöglich zu schützen. Hierfür werden bereits seit Jahren Seitenaufprallschutzsysteme, Airbags und Gurtstraffer integriert. Diese werden momentan in der Regel erst beim Kontakt mit einem anderen Objekt aktiviert. Zukünftige Versionen, die auch das Fahrzeugumfeld beobachten, könnten auch bereits kurz vor dem ersten Kontakt reversibel initialisiert oder sogar komplett ausgelöst werden. Moderne zusätzliche Insassenschutzsysteme (ZIS) optimieren darüber hinaus sogar die Sitzposition der Passagiere oder schließen eventuell geöffnete Fenster oder Schiebedächer.

Auch der Fußgängerschutz ist in diesem Zeitraum von Bedeutung. Bei Kollisionen können Verletzungen durch Dämpfung des Aufpralls deutlich reduziert werden. Dieses Ziel wird durch spezielle Karosserieformen, Außenairbags oder Motorhauben verfolgt. Letztere können sich selbständig ankippen, um sich beim Aufprall durch den erhöhten Abstand zum Motorblock besser verformen zu können [4].

Der eventuell notwendige Einsatz von Rettungskräften nach einem Unfall kann ebenfalls durch automatische Notrufsysteme unterstützt werden. Mit dem sogenannten eCall wird nach einer Initiative der Europäischen Union eine Serienausstattung bis 2010 angestrebt [31, 83].

2.7.2 Kundenakzeptanz und rechtliche Aspekte

Systeme, die automatisch in die Fahrzeugführung eingreifen, bieten den enormen Vorteil, dass sie extrem schnell agieren können. Wird der Fahrer mit in die Entscheidung einbezogen, erhöht sich die zur Ausführung der Aktion benötigte Zeit um dessen Reaktionszeit. Gleichzeitig stehen Entwickler derartiger Systeme aber auch vor einer immensen Herausforderung. Wenn das System vollständig selbstständig handelt, kann man die Verantwortung nur schwer dem Fahrer übertragen. Demzufolge müssen sämtliche Aktionen gerechtfertigt sein und dürfen auf keinen Fall zusätzliche Gefährdungen auslösen. Aufgrund dieser Aspekte und der enormen Komplexität der täglichen Situationen im Fahrzeugumfeld wird nach heutigem Wissensstand ein komplett automatisches Fahren in absehbarer Zeit nicht erwartet. Der Fokus richtet sich vielmehr auf eine Optimierung der vom Fahrer durchgeführten Aktionen [41]. Alternativ ist auch ein nur geringfügiger Eingriff wie z.B. ein schwaches Abbremsen denkbar, das für andere Verkehrsteilnehmer ungefährlich ist.

Ein wichtiger Punkt beim Design von Fahrerassistenzsystemen ist die Kundenakzeptanz. Der Käufer muss einen deutlichen Vorteil des Systems erkennen können, um dafür zusätzlich zu investieren. Zählt das System bereits zur Ausstattung des Fahrzeugs, muss die Mensch-Maschine-Schnittstelle (*engl. HMI = human machine interface*) übersichtlich und die Aktionen des Systems nachvollziehbar sein. Führt das System nicht nachvollziehbare oder falsche Aktionen durch, wird der Kunde das System irgendwann abstellen. Laut einer Studie in [74] ist zum Beispiel bei Warnsystemen eine Fehlwarnung pro 10 korrekter Warnungen für die meisten Kunden gerade noch akzeptabel. Erhöht sich die Zahl

der Fehlwarnungen, wird das System eher als störend empfunden. Weiterhin ist zu berücksichtigen, dass die meisten Fahrer die Entscheidungsgewalt über ihr Fahrzeug behalten möchten. Sie lassen sich durch Assistenzsysteme bei Bedarf gern unterstützen, möchten aber nicht bevormundet werden [41].

Ein letzter Aspekt ist die scheinbare Sicherheit, die der Fahrer bei guter Unterstützung durch Assistenzsysteme empfindet. Dieser Eindruck kann beim Fahrer zu nachlassender Aufmerksamkeit und damit in Extremfällen auch zu einer Reduktion der Verkehrssicherheit durch das Assistenzsystem führen. Um dieser Entwicklung entgegenzuwirken, muss das Systemdesign so gewählt werden, dass dem Fahrer seine Aufgabe und Verantwortung auch weiterhin bewusst bleibt [41].

2.7.3 Forschungsprojekte

Im Bereich der Forschung und Entwicklung befassen sich die Arbeiten sowohl mit der Verbesserung bestehender als auch mit der Entwicklung neuer Fahrerassistenzsysteme. So befassen sich einige Arbeiten mit neuen Konzepten zum Fußgängerschutz [4, 47, 91, 131]. Weitere Aktivitäten konzentrieren sich auf automatische Notbremssysteme zur Unfallfolgenminderung [66, 85, 120]. Darüber hinaus existieren Untersuchungen zur automatischen Kollisionsvermeidung [132], zur automatischen Durchführung bestimmter Fahrmanöver [40, 113] und zum vollständig autonomen Fahren [75].

Ein weiterer Trend ist die Einbeziehung komplexerer Szenarien. Während erste Fahrerassistenzsysteme sich primär auf die Anwendung in einfachen Szenarien auf Autobahnen oder Landstraßen konzentrierten, versuchen neuere Systeme auch Lösungen für komplexere Szenarien im städtischen Verkehr anzubieten. Ein besonderer Fokus liegt dabei auf Kreuzungen, da hier ein hoher Anteil der Unfälle auftritt [46].

Für Kreuzungen wurden verschiedene Assistenzsysteme entworfen, bei denen verschiedene Sensoren, Karteninformationen und auch Drahtloskommunikation zum Einsatz kamen [24, 46, 50, 62, 116]. Ein wesentlicher Schwerpunkt für Kreuzungsassistenten ist die Unterstützung des Fahrers bei der Beachtung der Vorfahrtsregelungen. Die Herausforderung besteht hier zum einen in der Vielfalt der möglichen Situationen, in denen die Vorfahrt gewährt werden muss. Manche Assistentzsysteme unterstützen den Fahrer darum vorerst nur bei ausgewählten Manövern auf Kreuzungen, wie zum Beispiel dem Linksabbiegen [99]. Zum anderen muss für entsprechende Assistenzsysteme ein sehr großer Bereich des Fahrzeugumfelds erfasst werden, da beispielsweise auch Querverkehr erkannt werden muss. Bisherige Arbeiten bewältigen diese Herausforderung entweder über drahtlose Datenverarbeitung oder über Laserscanner mit einem extrem großen Messbereich. Ein zweiter Anwendungsbereich ist die Ampelassistenz. Entsprechende Systeme versuchen primär das unbeabsichtigte Überfahren einer auf rot geschalteten Ampel zu vermeiden. Darüber hinaus sind aber auch Systeme mit Geschwindigkeitsvorschlägen für das optimale Heranfahren an die Ampel realisiert worden.

Auch Fahrerassistenzsysteme für neue Anwendungsfälle werden untersucht. Beispielsweise wird das Einparken von LKWs mit Anhängern unter Wechselaufbauten automatisiert. Für dieses zentimetergenaue Einparken zum Ankoppeln der Wechselaufbauten wurde ein laserscannerbasiertes System vorgestellt [126]. Verschiedene Systeme zur Verkehrszeichenerkennung [13, 38, 101] können den Fahrer über aktuelle Vorschriften wie zum Beispiel das

aktuelle Geschwindigkeitslimit informieren. In [90] wird schließlich ein System beschrieben, dass dem Fahrer kartenbasiert gefährliche Überholpassagen anzeigt.

Die aktuellste Herausforderung auf dem Gebiet der automatisierten Fahrzeugführung war die von der US-Militärforschungsbehörde DARPA (*engl. DARPA = Defence Advanced Research Projects Agency*) ausgeschriebene Urban Challenge [19]. Hier mussten die Teilnehmer mit vollständig autonom fahrenden Fahrzeugen antreten. Die Fahrzeuge sollten innerhalb einer abgesperrten Kleinstadt verschiedene Aufgaben erfüllen. So waren vorgegebene Ziele abzufahren, wobei das Fahrzeug entsprechend der Verkehrsregeln agieren und andere Verkehrsteilnehmer berücksichtigen musste. Aktionen wie beispielsweise das Einfädeln in fließenden Verkehr mussten dabei vom Fahrzeug komplett selbständig bewältigt werden.

2.7.4 EU-Projekte

Die Europäische Union fördert diverse Aktivitäten zur Verbesserung der Sicherheit und der Effizienz des Straßenverkehrs [29]. Koordinierende Projekte wie eSAFETYSUPPORT, ESCOPE und SAFETY TECHNOPRO sorgen dabei für Wissenstransfer und Öffentlichkeitsarbeit. Zur Bestimmung wichtiger Gefahrenquellen im Straßenverkehr wird im Projekt TRACE eine globale Unfallforschung durchgeführt.

Die Akzeptanz neuer Sicherheitssysteme in der Gesellschaft und beim Nutzer ist ein wesentlicher Aspekt bei der Verbesserung der Verkehrssicherheit durch den Einsatz dieser neuen Systeme. Gesellschaftliche und wirtschaftliche Auswirkungen neuer Assistenzsysteme werden in eIMPACT und SEiSS betrachtet. Die Projekte AIDE und HUMANIST untersuchen Nutzeraspekte und geeignete Mensch-Maschine-Schnittstellen.

Die rapide Entwicklung der Fahrzeugelektronik zieht gleichzeitig eine große Menge neuer Fehlerquellen im Automobil nach sich. Zur Reduktion möglicher Fehlerquellen und Entwicklung neuer, sicherer Komponenten werden verschiedene Aktivitäten unterstützt. Die Standardisierung von Softwaretechnologie und Architekturen für Assistenzsysteme werden in ATESST und EASIS vorangetrieben. Die Weiterentwickelung elektrischer Komponenten und die Verbesserung der Ausfallsicherheit wird in SPARC und SAFETEL verfolgt.

Die Effizienz des Straßenverkehrs wirkt sich auch gleichzeitig auf die Belastung der Infrastruktur aus. Häufige Staus stellen eine zusätzliche Belastung für Infrastruktur und Umwelt dar. Gleichzeitig kommt der effiziente, flüssige Verkehr jedem einzelnen Teilnehmer zugute. Aus diesem Grund wird in verschiedenen Projekten wie EURAMP, TRACKSS und MORYNE versucht, den Verkehrsfluss zu optimieren.

Einige bereits identifizierte Gefahrenquellen im Straßenverkehr sollen direkt in verschiedenen Projekten beseitigt werden. So wird im Projekt FRICTION Straßenglätte als wichtige Gefahrenquelle untersucht. Informationen hierüber sollen mit geeigneten Sensoren von Fahrzeugen oder der Infrastruktur gemessen und an andere Fahrzeuge weitergeleitet werden. Eine bessere Sicherung von Gefahrguttransporten soll in GOODROUTE erreicht werden.

Ein großer Fokus liegt auch auf der Entwicklung kooperativer Systeme, die durch drahtlosen Datenaustausch mit der Infrastruktur oder anderen Verkehrsteilnehmern die Sicherheit des Straßenverkehrs oder dessen Effizienz erhöhen. Die Europäische Union fördert hier eine Reihe von Projekten, die sich mit verschiedenen Aspekten beschäftigen. Dazu gehören

COMeSAFETY, COOPERS, COVER, CVIS, WATCH-OVER, SEVECOM, SAFESPOT, REPOSIT, I-WAY, FEEDMAP, GST, IM@GINE-IT und CYBERCARS2.

Schließlich wurden im integrierten Projekt PReVENT verschiedene aktive und preventive Sicherheitssysteme zur Verbesserung der Verkehrssicherheit entwickelt und untersucht [27]. Das Projekt wurde im Herbst 2007 mit einer öffentlichen Präsentation der entwickelten Assistenzsysteme in den aufgebauten Demonstrationsfahrzeugen beendet.

PReVENT war ein Projekt mit sehr großem Umfang, das sich mit verschiedenen Aspekten von Assistenzsystemen und unterschiedlichen Situationen und Szenarien befasste. Zur Untersuchung einzelner Themen war das Projekt hierfür in verschiedenen Unterprojekten organisiert (Abbildung 2.2).

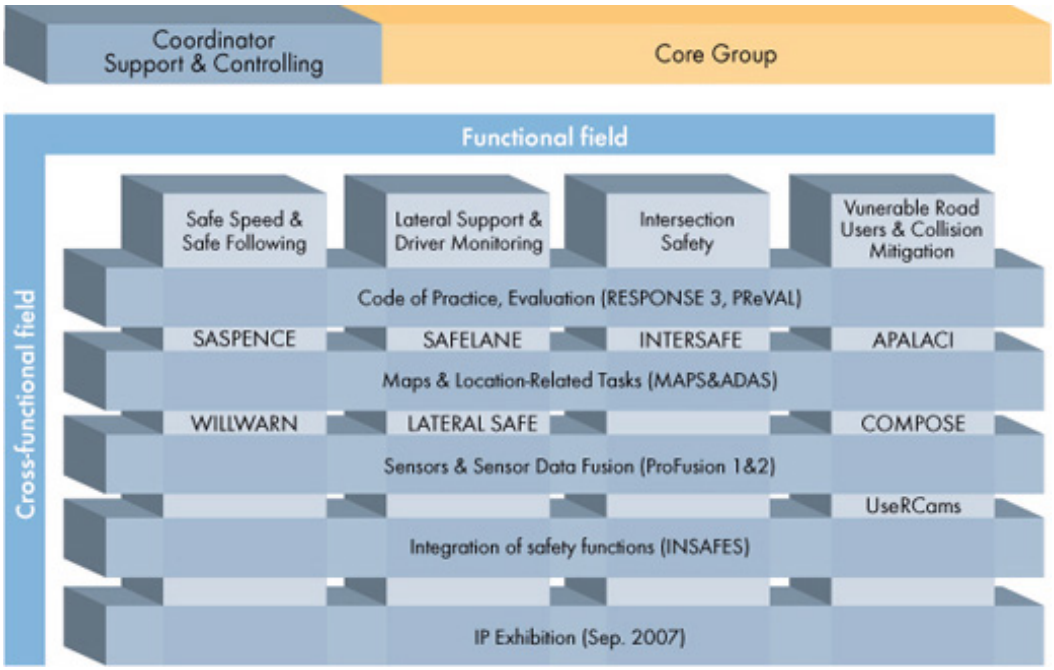

Abbildung 2.2: Organisation der Aktivitäten im Projekt PReVENT: Einzelne Assistenzsysteme für die Unterstützung unterschiedlicher Situationen werden in vertikalen Unterprojekten bearbeitet. Übergreifende Aktivitäten sind in horizontalen Unterprojekten organisiert [27].

Verschiedene Szenarien der Fahrerassistenz wurden in vertikalen Unterprojekten behandelt. So wurden die Themen sichere Geschwindigkeit und sichere Kolonnenfahrt in SASPENCE und WILLWARN bearbeitet. Die Unterstützung beim Halten der Spur und für sichere Spurwechsel bildete den Inhalt von SAFELANE und LATERAL SAFE. Im Rahmen des Projektes INTERSAFE wurden verschiedene Konzepte und Systeme zur Assistenz in Kreuzungsszenarien entwickelt. Schließlich wurden diverse Konzepte und Assistenzsysteme zum Schutz von Personen in APALACI, COMPOSE und UseRCams untersucht. Zum einen bezieht sich dies auf den Schutz besonders gefährdeter Verkehrsteilnehmer, wie Fußgänger und Zweiradfahrer. Zum anderen wurde hier auch die Reduktion der Aufprallgeschwindigkeit bei Unfällen adressiert, um das Verletzungsrisiko der Fahrzeuginsassen zu reduzieren.

Szenarienübergreifende Aspekte wurden in horizontalen Unterprojekten bearbeitet. Mit dem „Code of Practice“ wurden geeignete Randbedingungen für die Realisierung von Assistenzsystemen unter Berücksichtigung rechtlicher Aspekte festgelegt und evaluiert. Weiterhin wurde die Weiterentwicklung von Karten und deren Integration in Assistenzsysteme vorangetrieben. Verfügbare Sensoren und Ansätze zur Sensorfusion wurden ebenfalls in einem horizontalen Projekt untersucht. Ein weiteres Projekt beschäftigte sich mit der Integration verschiedener Assistenzsysteme in ein Gesamtkonzept. Den Abschluss bildete die gemeinsame Ausstellung der Ergebnisse aller Unterprojekte im Herbst 2007.

2.8 Motivation

Viele aktuell verfügbare Fahrerassistenzsysteme beruhen auf der exklusiven Nutzung eines Sensors. In manchen Fällen wird auch bereits eine Sensorfusion betrieben. Darüber hinaus existieren schon erste zusätzliche Fahrerassistenzsysteme, die auf einem Basissystem wie dem abstandsgeregelten Tempomat und seinem Sensor aufbauen. Da in Zukunft voraussichtlich eine deutlich höhere Anzahl von Fahrerassistenzsystemen im Fahrzeug zu erwarten ist, wird die gemeinsame Nutzung von Sensormessdaten in ihrer Bedeutung deutlich steigen. Gemeinsame Algorithmen für die Sensormessdatenverarbeitung können darüber hinaus zum strukturierten Aufbau und zur Kostenersparnis beitragen. Das Ziel dieser Arbeit war die Entwicklung eines derartigen gemeinsamen Systems zur Sensormessdatenverarbeitung.

Um verschiedene Assistenten parallel mit Informationen versorgen zu können, ist eine zusätzliche Abstraktionsebene notwendig. Dieses sogenannte Fahrzeugumfeldmodell enthält alle Ergebnisse der Sensormessdatenverarbeitung. Während sich [66] auf die Objektverfolgung zur Fahrzeugumfeldmodellierung konzentrierte, sollten nun auch Klasseninformationen in ein erweitertes Fahrzeugumfeldmodell eingetragen werden (Abbildung 2.3).

Bisherige Klassifikationsansätze basieren meist auf der anwendungsspezifischen Detektion einer ausgewählten Objektklasse. Unterschiedliche Fahrerassistenzsysteme konzentrieren sich allerdings durchaus auf unterschiedliche Objekttypen. So sind Fußgänger für den Fußgängerschutz von Bedeutung, abstandsgeregelte Tempomaten benötigen zur Auswahl eines Zielobjektes aber Informationen über vorausfahrende Fahrzeuge. Dieser Sachverhalt begründet Klassifikationsansätze mit mehreren Objektklassen. Diese Arbeit konzentriert sich auf die Erkennung und Unterscheidung der im Fahrzeugumfeld befindlichen Verkehrsteilnehmer. Die Wahl der unterschiedenen Verkehrsteilnehmer orientiert sich an den in [43] vorgeschlagenen Objektklassen. Somit werden Objekte der Klassen „Fußgänger“, „Zweirad“, „PKW“ und „LKW“ von allen restlichen Objekten unterschieden, die die Klasse „Hintergrund“ zugewiesen bekommen. Eine weitere Unterscheidung der Hintergrundobjekte, wie zum Beispiel eine Verkehrszeichenerkennung, wird hier nicht vorgenommen.

In [43] existieren bereits einfache, regelbasierte Ansätze für eine derartige Klassifikationsaufgabe. Für bessere Klassifikationsergebnisse und die Anwendung in komplexeren Szenarien ist jedoch ein leistungsfähigeres System mit diversen verschiedenen Merkmalen notwendig, das ebenfalls Aspekte wie Objekthistorie und Sensorfusion berücksichtigen kann und zusätzlich die Klassifikationsergebnisse aus unterschiedlichen Zeitschritten auswertet.

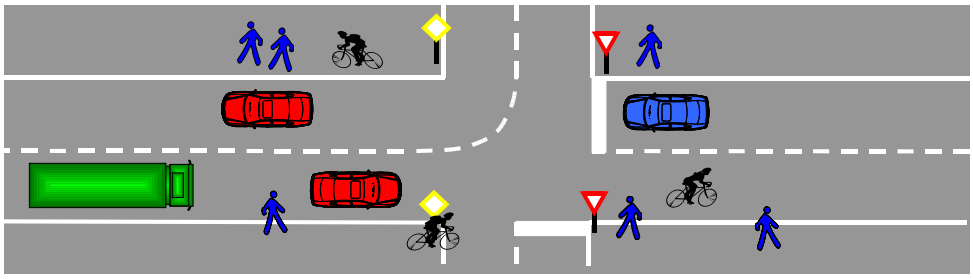

Abbildung 2.3: Für die parallele Unterstützung verschiedener Assistenzsysteme wird das Fahrzeugumfeldmodell als zusätzliche Abstraktionsebene eingeführt. In dieses Modell werden alle gesammelten Informationen über das Fahrzeugumfeld gespeichert.

Eine weitere Zielstellung war die Integration drahtlos empfangener Informationen in das erweiterte Fahrzeugumfeldmodell. Bisherige Anwendungen der Drahtloskommunikation konzentrieren sich auf die Übertragung globaler Informationen über Gefährdungen wie beispielsweise Staus oder Schlechtwettergebiete. Diese lassen sich leicht in das Fahrzeugumfeldmodell eintragen. Eine größere Herausforderung besteht darin, Informationen über bewegte Objekte zu übertragen und in das sensorbasierte Umfeldmodell zu integrieren. Hierzu war ein allgemeiner Ansatz zu verfolgen, der sowohl Sichtbereichserweiterungen als auch die Zuordnung von Zusatzinformationen zu Objekten realisiert, die mit den Onboard-Sensoren vermessen wurden. Die übertragenen Objektdaten können dafür aus verschiedenen Quellen stammen: Verkehrsteilnehmer können die eigenen Objektdaten verschicken oder es können Daten von Objekten übertragen werden, die von anderen Verkehrsteilnehmern oder Infrastruktursensoren vermessen wurden.

Unterschiedliche Fahrerassistenzsysteme können durchaus unterschiedliche Anforderungen an die Sensormessdatenverarbeitung stellen. Da das Fahrzeugumfeldmodell als Basis für verschiedene Assistenzsysteme dienen soll, waren auch Methoden zur Optimierung der Ausgabe dieses gemeinsamen Fahrzeugumfeldmodells zu untersuchen.

Bei der Auswahl der zu verwendenden Methoden war deren Echtzeitfähigkeit zu berücksichtigen, da die Algorithmen im Fahrzeug online funktionieren müssen, um Fahrerassistenzsysteme dort mit Informationen versorgen zu können.

3 Systembeschreibung

In diesem Kapitel soll ein allgemeiner Überblick über das entwickelte System zur Fahrzeugumfelderfassung gegeben werden. Im ersten Abschnitt wird der benötigte Hardwareaufbau beschrieben. Anschließend wird das Softwaresystem in einer Gesamtübersicht vorgestellt und seine einzelnen Komponenten kurz erläutert. Eine detaillierte Beschreibung der Softwarekomponenten erfolgt in den folgenden Kapiteln.

3.1 Sensoren und Versuchsträger

Zur Fahrzeugumfelderfassung wurde ein Versuchsträger mit verschiedenen Sensorsystemen ausgestattet. Das Versuchsfahrzeug, die Sensoren und ihre Anbaupositionen sind in Abbildung 3.1 dargestellt.

Abbildung 3.1: Versuchsträger zur Fahrzeugumfelderfassung: In der vorderen Stoßstange wurde ein Laserscanner zur Abstandsmessung integriert. Eine Videokamera wurde neben dem Rückspiegel angebracht. Ein DGPS-Empfänger liefert eine grobe globale Position.

Als Abstandssensor wurde ein ALASCA XT Laserscanner der Firma IBEO Automobile Sensor GmbH in der vorderen Stoßstange integriert. Dieser liefert Entfernungsmessungen bis zu einer Reichweite von etwa 200 m. Um auch bei Unebenheiten oder Einnicken des Fahrzeugs weiterhin nutzbare Messdaten erzeugen zu können, werden Entfernungsmessungen in vier horizontalen Messebenen zur Verfügung gestellt. Bei der verwendeten Frequenz von 12,5 Hz liefert der Sensor Messdaten in Winkelschritten von 0,25°. Neben der präzisen Winkelauflösung bietet der verwendete Sensor ebenfalls einen extrem breiten Messbereich. Aufgrund der zentralen Integration wird ein horizontaler Winkelbereich von 140° abgedeckt. Die in dieser Arbeit vorgestellten Algorithmen können darüber hinaus auch auf Messdaten von zwei an den vorderen Fahrzeugecken integrierten Laserscannern angewendet werden [46]. In diesem Fall erhöht sich der abgedeckte Winkelbereich noch einmal deutlich.

Als Videokamera kommt eine PCO Pixelfly CCD-Kamera zum Einsatz, die auch bereits in [66] verwendet wurde. Sie wurde neben dem Rückspiegel hinter der Frontscheibe angebracht. Diese Kamera liefert Bilder in einer Auflösung von 640 × 480 Pixeln. Zwei verschiedene Objektive wurden verwendet. Für Szenarien an Kreuzungen und in städtischen

Gebieten wird ein breiter Sichtbereich benötigt. Aus diesem Grund kam hier ein Weitwinkelobjektiv mit einer Brennweite von 4,8 mm und einem Öffnungswinkel von etwa 71° zum Einsatz. Für Autobahnszenarien sind Messdaten in höherer Entfernung vor dem Fahrzeug von Bedeutung. Diese lassen sich durch die höhere Auflösung eines Teleobjektivs mit kleinerem Öffnungswinkel verbessern. Zu diesem Zweck wurde ein Objektiv der Brennweite 12 mm mit einem Öffnungswinkel von etwa 30° eingesetzt.

Die Videokamera und der Laserscanner wurden zum Fahrzeug bezüglich Anbauposition und Ausrichtung kalibriert. Die Messungen beider Sensoren erfolgen synchron. Zur Kalibrierung und Synchronisierung wurden Ansätze aus [66] verwendet und auf den neuen Laserscanner angepasst.

Für eine globale Positionsbestimmung wurde ein Afusoft Raven 6 DGPS-Empfänger in das Fahrzeug integriert. Dieser Sensor wird zusammen mit dem Laserscanner auch für die präzise Positionierung in digitalen Karten genutzt [150, 152]. Die Ergebnisse dieser Arbeiten zur Positionierung liefern zusammen mit den präzisen digitalen Karten wertvolle Zusatzinformationen und werden in Kapitel 8 als zusätzlicher Sensor zur Objektklassifikation hinzugezogen.

Bei der Anwendung der Drahtloskommunikation standen Methoden zur Nutzung der übertragenen Objektinformationen im Vordergrund. Da geeignete Techniken zur drahtlosen Datenübertragung im Fahrzeugumfeld im Rahmen einer Kooperation von der Firma IBEO Automobile Sensor GmbH untersucht werden, wurde hier auf Untersuchungen in diese Richtung verzichtet. Da bisher keine geeignete Übertragungshardware zur Verfügung stand, musste die drahtlose Kommunikation inklusive Übertragunszeiten simuliert werden. Diese Simulation wird in Abschnitt 10.7.1 näher beschrieben.

Der aufgebaute Versuchsträger ermöglichte die Aufzeichnung diverser Sequenzen von Messdaten der unterschiedlichsten Szenarien. In einem Teil dieser Sequenzen wurden zum Training der entwickelten Klassifikatoren die unterschiedenen Objektklassen manuell markiert. Zur Analyse der entwickelten Algorithmen wurde in einem weiteren Teil ebenfalls eine manuelle Markierung der Objekte durchgeführt. Eine detaillierte Beschreibung der Trainings- und Testdaten erfolgt in Kapitel 10. Bei der Entwicklung der Algorithmen wurde insbesondere auf die Echtzeitfähigkeit geachtet, da eine sinnvolle Anwendung in Verbindung mit Fahrerassistenzsystemen online erfolgen muss. Diese Echtzeitfähigkeit konnte im Rahmen von Drittmittelprojekten überprüft werden. So wurde ein Großteil der Algorithmen im Rahmen dieser Projektarbeiten in andere Versuchsträger integriert und dort online getestet.

3.2 Systemüberblick

In diesem Abschnitt wird ein Überblick über das erstellte Softwaresystem zur Fahrzeugumfelderfassung gegeben. Es werden die im vorherigen Abschnitt beschriebenen Sensoren eingesetzt. Grundsätzlich sind in der verwendeten Softwarearchitektur aber auch Fusionsansätze mit anderen Abstands- und Videosensoren denkbar. Auch der Einsatz zusätzlicher Sensoren in der Fusion ist prinzipiell möglich.

Abbildung 3.2 zeigt eine Übersicht über das Softwaresystem. Das vorgestellte System ist modular aufgebaut. Das blau dargestellte Basissystem führt mit Hilfe der Laserscannermess-

daten nach einer Vorverarbeitungsstufe eine Objektverfolgung und -klassifikation durch. Zur Schätzung absoluter Geschwindigkeiten der gemessenen Objekte werden aus den ESP-Sensoren des Fahrzeugs Messdaten über dessen Bewegung ermittelt und in die Objektverfolgung einbezogen. Die Ergebnisse der Objektverfolgung und -klassifikation werden im erweiterten Fahrzeugumfeldmodell gespeichert.

Ein spezielles Modul optimiert die Ausgabe des Fahrzeugumfeldmodells für jedes Assistenzsystem entsprechend dessen Anforderungen. Die Grundlage für diese Optimierung besteht darin, notwendige, harte Schwellwertoperationen so spät wie möglich durchzuführen. So werden zum Beispiel prinzipiell alle Objekte verfolgt. Eine Einschränkung der Ausgabe auf bestimmte Objekte erfolgt erst in diesem letzten Modul. Ähnliche Ansätze werden bei der Klassifikation verfolgt. Der Klassifikator bestimmt für jedes Objekt Zugehörigkeitswerte für alle Objektklassen. Die harte Zuordnung der ausgegebenen Klasse erfolgt dann erst bei der Übergabe an das jeweilige Assistenzsystem. In die durchgeführten Entscheidungen können anwendungsspezifische Einschränkungen und Anforderungen eingehen, die zu unterschiedlichen Ausgaben für unterschiedliche Assistenzsysteme führen können. So kann für den Fußgängerschutz zum Beispiel die Klassifikation auf „Fußgänger“ und „Hintergrund“ eingeschränkt werden, während ein abstandsgeregelter Tempomat Ausgaben mit den Objektklassen „Fahrzeug“ und „Hintergrund“ erhält. Die Voraussetzung für ein derartiges Vorgehen ist das unabhängige Agieren der so versorgten Assistenzsysteme.

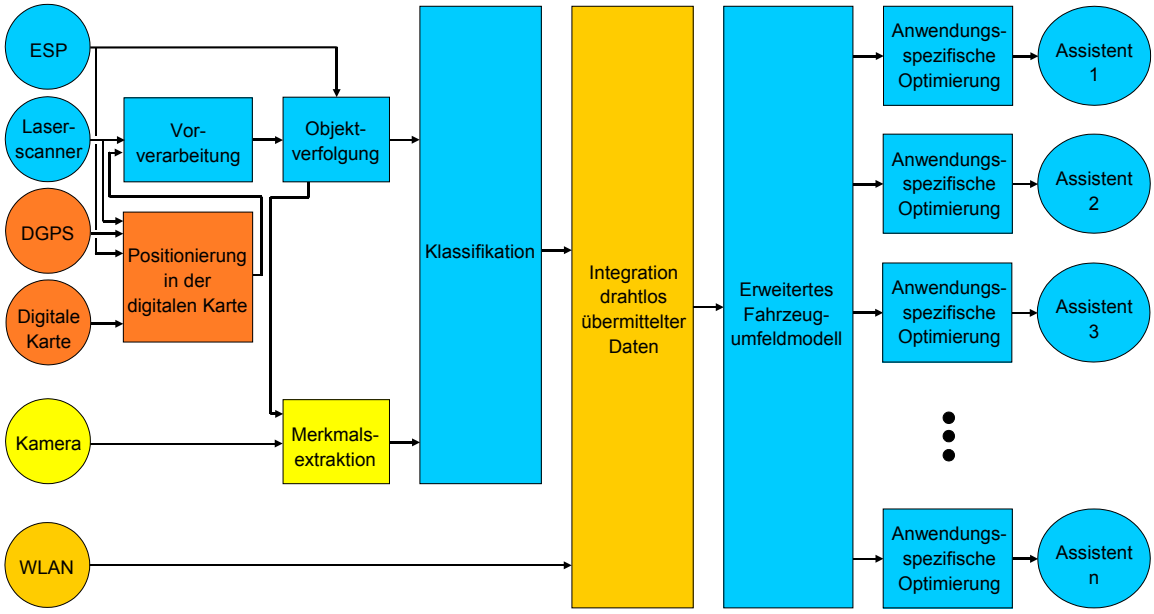

Abbildung 3.2: Überblick über das modulare Softwaresystem: Das Basissystem (blau) führt mit Hilfe von Laserscannermessdaten eine Objektverfolgung und -klassifikation durch, deren Ergebnisse im erweiterten Fahrzeugumfeldmodell gespeichert werden. Spezielle Module optimieren die Ausgabe an unterschiedliche Assistenzsysteme. Fusionsansätze mit präzisen digitalen Karten (rot) verbessern die Objektverfolgung und die Klassifikation. Die Fusion mit Videomessdaten (gelb) führt zu einer leistungsfähigeren Klassifikation. Schließlich werden drahtlos empfangene Objektdaten in das erweiterte Fahrzeugumfeldmodell integriert (orange).

Das Basissystem kann um verschiedene Fusionsansätze erweitert werden. Die erste, in rot dargestellte Erweiterung verwendet die Ergebnisse der Arbeiten zur präzisen Positionierung in digitalen Karten [150, 152]. Die zusätzlichen Informationen verbessern sowohl die Objektverfolgung als auch die Klassifikation.

Ein zweiter Fusionsansatz nutzt Videomessdaten zur Verbesserung der Objektklassifikation. Hier werden zusätzliche Merkmale aus Videobildern berechnet. In diese Berechnung gehen auch Positions- und Orientierungsmessungen ein, die für jedes Objekt aus Laserscannermessdaten gewonnen werden. Fusionsansätze mit Videomessdaten zur Verbesserung der Objektverfolgung werden hier nicht berücksichtigt, da entsprechende Untersuchungen bereits in [66] durchgeführt wurden. Die dort vorgestellten Ansätze lassen sich aber prinzipiell in das verwendete Gesamtsystem integrieren.

Schließlich kann das System zur Fahrzeugumfelderfassung noch durch Integration von drahtlos übermittelten Objektdaten erweitert werden. Diese Daten werden konsistent in das erweiterte Fahrzeugumfeldmodell integriert.

Die einzelnen Komponenten des Systems werden in den folgenden Kapiteln getrennt vorgestellt. Eine Ausnahme bildet dabei die anwendungsspezifische Optimierung der Ausgabe des erweiterten Fahrzeugumfeldmodells. Die dort durchgeführten harten Entscheidungen beziehen sich jeweils auf die Ergebnisse eines der Module Objektverfolgung, Klassifikation oder Integration der drahtlos übermittelten Daten. Aus diesem Grund werden die entsprechenden Optimierungsansätze jeweils in dem Kapitel erläutert, das das zugehörige Modul beschreibt.

Im Rahmen dieser Arbeit wurden alle Softwaremodule in einem zentralen Computer ausgeführt. Alle Sensormessdaten wurden an diesen Rechner übertragen. Diese Übertragung der kompletten Messdaten des Laserscanners und der Kamera führt zu einem enormen Übertragungsaufwand, der in der Serienproduktion aufgrund der notwendigen Übertragungsmedien mit erhebliche Kosten verbunden wäre. Eine effiziente Realisierung im Fahrzeug sollte Teile der Berechnungen direkt beim Sensor durchführen, um die Menge der zu übertragenden Daten zu reduzieren.

Eine mögliche Realisierung besteht darin, lediglich Objektmerkmale zu übertragen [66]. Dafür sollte die Objektverfolgung nach [66] in zwei Teile zerlegt werden (Abbildung 3.3). Die Schätzung des Objektzustands und die zeitliche Prädiktion wird weiterhin in einer zentralen Recheneinheit durchgeführt. Diese sendet die prädizierten Objektmerkmale an die Recheneinheiten der Sensoren. Dort wird die Merkmalsprädiktion, die Merkmalsextraktion und die Datenassoziation realisiert. Anschließend werden die aus den aktuellen Sensormessdaten extrahierten Merkmale wieder an die zentrale Recheneinheit übertragen, in der eine Fusion auf der Merkmalsebene vorgenommen wird.

Werden mehrere Sensoren verwendet, eignet sich dieser Ansatz der Objektverfolgung zur Sensorfusion auf der Merkmalsebene. Verfahren der Sensorfusion zur Objektverfolgung wurden bereits in bestehenden Arbeiten untersucht [66, 69]. In dieser Arbeit wird auf derartige Ansätze verzichtet. Der Schwerpunkt der Sensorfusion liegt hier in der Verbesserung der Algorithmen zur Klassifikation.

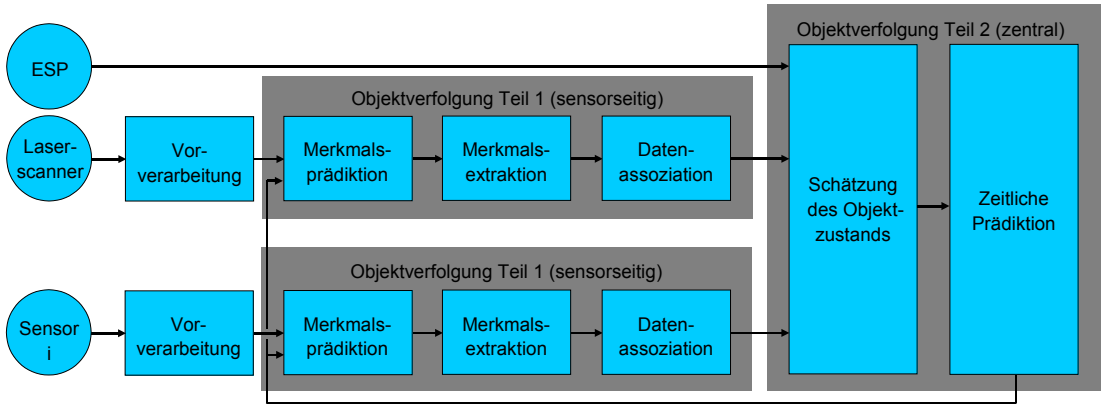

Abbildung 3.3: Aufteilung der Objektverfolgung auf verschiedene Recheneinheiten nach [66]: Die Schätzung des Objektzustands und die zeitliche Prädiktion erfolgen zentral, während die restlichen Schritte sensorseitig realisiert sind. Zwischen den einzelnen Recheneinheiten müssen nur Objektmerkmale übertragen werden.

Eine sinnvolle Aufteilung der Module des kompletten Systems zur erweiterten Fahrzeugumfelderfassung ist in Abbildung 3.4 gegeben. Hier wurde die Objektverfolgung in einen sensorseitigen und einen zentralen Teil zerlegt.

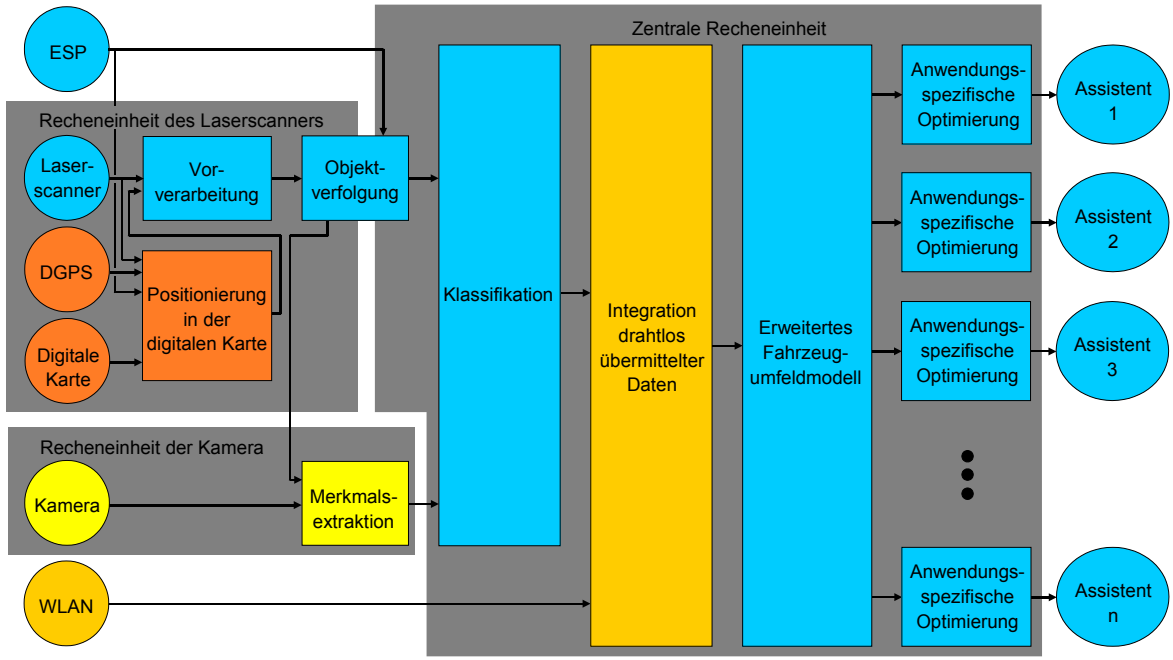

Abbildung 3.4: Unterteilung der Module für eine sinnvolle Verteilung auf Rechnern im Fahrzeug: Die Messdaten werden direkt am Sensor vorverarbeitet. So werden statt der kompletten Messdaten nur noch Objektlisten zwischen den einzelnen Rechnern übertragen.

Die Recheneinheit des Laserscanners bearbeitet in diesem Fall zusätzlich zu der Vorverarbeitung der Messdaten und einem Teil der Objektverfolgung auch die Algorithmen zur

präzisen Positionierung in der digitalen Karte. Eine weitere Unterteilung ist denkbar. Da hier jedoch nur die Ergebnisse der Arbeiten zur Positionierung genutzt werden, wird eine weitere Unterteilung an dieser Stelle nicht weiter verfolgt.

Die videobasierten Merkmale für die Klassifikation werden mit Hilfe von Positions- und Orientierungsmessungen des Laserscanners in der Recheneinheit der Videokamera berechnet. So müssen anstatt der kompletten Videobilder nur wenige Merkmale pro Objekt übertragen werden.

Der zweite Teil der Objektverfolgung, die Objektklassifikation, die Integration der drahtlos übermittelten Daten und die Ausgabeoptimierung werden weiterhin zentral ausgeführt. Dies kann in einem oder mehreren Rechnern realisiert werden.

Aufgrund dieser Anordnung müssen die zentralen Module statt mit den gesamten Sensormessdaten nur noch mit Objektdaten versorgt werden. Zur Durchführung der Klassifikation enthalten die Objekte eine Reihe von Merkmalen, die bereits bei dem Laserscanner oder der Videokamera aus den Messdaten berechnet wurden.

4 Vorverarbeitung der Laserscannerdaten

In diesem Kapitel werden einige Schritte zur Vorverarbeitung der Laserscannermessdaten erläutert. Die Beschreibung umfasst die Berechnung von Zusatzinformationen für einzelne Entfernungsmessungen, Methoden zum Ausblenden unerwünschter Entfernungsmessungen und die Segmentierung der Messdaten. Schließlich wird ein Verfahren zur Erkennung von verdeckten Bereichen präsentiert.

4.1 Sensorinterne Vorverarbeitung

Von der Software des Sensorherstellers werden verschiedene Berechnungen durchgeführt. Aus den Entfernungsmessungen werden durch Auswertung der Winkelinformationen kartesische Koordinaten für jeden Messpunkt berechnet. Zusätzlich ermittelt die Software Entfernungsmessungen, die aus Reflexionen von der Straßenoberfläche resultieren. Diese Messungen werden entsprechend markiert und dadurch von den folgenden Berechnungen ausgeschlossen.

4.2 Reflektordetektion

Zusätzlich zur Entfernungsmessung liefert der verwendete Laserscanner mit der Echopulsbreite auch ein Maß für die vom vermessenen Gegenstand zurückgestrahlte Energie. Die Auswertung von Echopulsbreiten wurde bereits für ältere Laserscanner in [43, 66] untersucht. Ergebnisse dieser Arbeiten zeigen, dass mit Hilfe der Echopulsbreite Retroreflektoren erkannt werden können. Diese befinden sich zum Beispiel in den Rückleuchten von Fahrzeugen, in Nummernschildern, auf Verkehrsschildern oder auf Leitpfosten. In diesem Abschnitt wird die Erkennung von Entfernungsmessungen beschrieben, die von einem Reflektor stammen. Die dabei identifizierten Entfernungsmessungen werden später in Kapitel 6 zur Unterstützung der Klassifikation von PKWs eingesetzt.

In einem ersten Schritt wurde das Messverhalten des verwendeten ALASCA XT Laserscanners untersucht. Auch für diesen Sensor ist die von Reflektoren zurückgestrahlte Energie deutlich größer als bei anderen Oberflächen. Da die vom Laserscanner empfangene Energie auch von der Distanz zum Objekt abhängt, muss diese bei der Erkennung von Reflektoren berücksichtigt werden. Reflektoren werden darum über einen distanzabhängigen Schwellwert für die Echopulsbreite charakterisiert.

Sinnvolle Schwellwerte werden mit Hilfe einer Versuchsmessung festgelegt. Dazu wird ein Reflektor gewählt, der den Querschnitt des Laserstrahls komplett abdeckt. Somit wird die maximal mögliche Echopulsbreite gemessen. Der Reflektor wird nun von verschiedenen Entfernungen vermessen. Die Echopulsbreiten werden über der Entfernung abgetragen. Ähnliche Messungen werden für gut reflektierende andere Gegenstände durchgeführt. Hier bieten sich weiße Flächen oder silberne Fahrzeuglackierungen an. Die resultierenden Messwerte des verwendeten Sensors sind in Abbildung 4.1 dargestellt.

Der Schwellwert wird nun so festgelegt, dass er dicht an den Messwerten der normal reflektierenden Gegenstände liegt. Dies ist dadurch begründet, dass hier Maximalwerte für

Echopulsbreiten gemessen wurden. Wenn ein Objekt den Laserstrahl nur teilweise abdeckt, ist die reflektierte Energie entsprechend geringer. Somit kann davon ausgegangen werden, dass Reflektoren auch deutlich geringere Echopulsbreiten erzeugen können. Andere Objekte werden jedoch nicht wesentlich höhere Echopulsbreiten erreichen, als im Versuch ermittelt wurde.

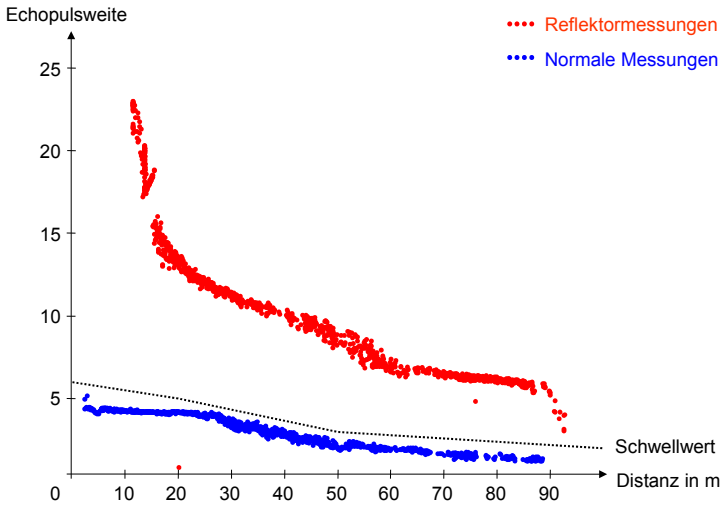

Abbildung 4.1: Entfernungsabhängiger Echopulsbreitenverlauf: Reflektoren erreichen deutlich höhere Echopulsbreiten (rot) als andere Gegenstände (blau). Ein entfernungsabhängiger Schwellwert (schwarz) kann somit zur Detektion von Reflektoren verwendet werden.

In [66] wurde eine Gruppierung von dichten Reflektormessungen vorgeschlagen, um einzelne Reflektoren zu charakterisieren. Dieses Verfahren wird auch hier angewendet.

4.3 Ausblenden von Entfernungsmessungen

Aufgrund der Integration des Laserscanners in einem Gehäuse in der vorderen Stoßstange des Versuchsträgers treten an bestimmten Stellen unerwünschte Entfernungsmessungen auf. So schränkt das Sichtfenster des Gehäuses den maximalen Winkelbereich ein, in dem sinnvolle Entfernungen gemessen werden können. Aus diesem Grund ist die Sensorausgabe auf den entsprechenden Winkelbereich zu begrenzen.

Zusätzlich kann eine Verschmutzung der Frontscheibe des Einbaugehäuses zu Reflexionen auf dieser Scheibe und damit zu Entfernungsmessungen führen. Der Sensor liefert in solchen Fällen meist auch hinter diesen Reflexionen noch Entfernungsmessungen, da nur ein Teil der Energie von der Scheibe reflektiert wird und vom Sensor pro ausgesendetem Laserpuls mehrere Entfernungsmessungen ermittelt werden. Verwendet man jedoch die Entfernungsmessungen von der Frontscheibe in der Objektverfolgung, entsteht das unerwünschte Trugbild eines Objektes, das direkt vor der Front des Versuchsträgers fährt. Um das Auftreten derartiger Scheinobjekte zu unterbinden, werden Messungen erst ab einer festgelegten minimalen Entfernung verwendet. Alle dichteren Entfernungsmessungen werden ausgeblendet.

4.4 Segmentierung

Die Segmentierung gruppiert Entfernungsmessungen, die voraussichtlich zum gleichen Objekt im Fahrzeugumfeld gehören. Als Kriterium wird hier der Abstand zwischen den einzelnen Messungen gewählt [43]. Für diesen Abstand wird eine obere Schranke gewählt, die zusätzlich linear mit der Distanz zum Sensor ansteigt, um das radiale Messprinzip des Laserscanners zu berücksichtigen.

4.5 Verdeckungserkennung

Aufgrund des radialen Messprinzips des Laserscanners lassen sich basierend auf vorhandenen Entfernungsmessungen Rückschlüsse über verdeckte Bereiche ziehen, in denen keine Messungen auf eventuell vorhandene Objekte erwartet werden können. Diese Information wird später beim Kombinieren drahtlos empfangener Daten mit gemessenen Objekten verwendet, aber auch Anwendungen in der Objektverfolgung sind denkbar.

Die Erkennung von Verdeckungen wird für diskrete Bereiche in Polarkoordinaten durchgeführt. In einem ersten Schritt werden hierfür alle Einzelmessungen des Laserscanners, die während eines Zeitschrittes aufgenommen wurden, in einem entsprechenden Raster eingetragen (Abbildung 4.2 links).

Nun wird für jede Rasterzelle eine Verdeckungswahrscheinlichkeit berechnet. Dabei werden die vier horizontalen Messebenen des Laserscanners einzeln berücksichtigt. Für jeden diskreten Winkelschritt werden dabei, beim Sensor beginnend, alle Zellen nacheinander durchlaufen. Solange dabei keine Messungen in den Zellen vorliegen, ist die Verdeckungswahrscheinlichkeit 0 Prozent. Für jede Messebene, von der Entfernungsmessungen gefunden werden, wird die Wahrscheinlichkeit um 25 Prozent erhöht. Damit werden fünf unterschiedliche Werte für die Verdeckungswahrscheinlichkeit unterschieden: nicht verdeckt und verdeckt in einer, zwei, drei und vier Messebenen (Abbildung 4.2 rechts).

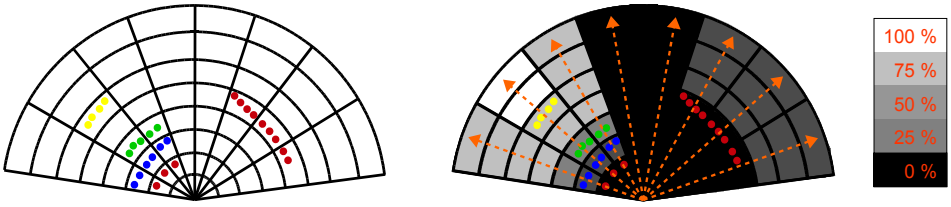

Abbildung 4.2: Berechnung der Verdeckungswahrscheinlichkeiten: Die Einzelmessungen des Laserscanners werden in einem Raster in Polarkoordinaten eingetragen. Vom Sensor aus werden, für jeden Winkelschritt getrennt, alle Zellen durchlaufen. Für jede Messebene, von der Entfernungsmessungen gefunden werden, wird die Verdeckungswahrscheinlichkeit um 25 Prozent erhöht.

In Abbildung 4.3 ist ein Beispiel für die berechneten Verdeckungswahrscheinlichkeiten basierend auf realen Messdaten dargestellt. Im Bereich vor dem Sensor ist die Verdeckungswahrscheinlichkeit 0 Prozent. Hinter den gemessenen Fahrzeugen ist der Wert 100 Prozent, da alle vier Messebenen die Fahrzeuge treffen. Die Leitplanke wird in höheren Distanzen nur

noch in zwei Ebenen vermessen. Aus diesem Grund ist die Verdeckungswahrscheinlichkeit dahinter geringer.

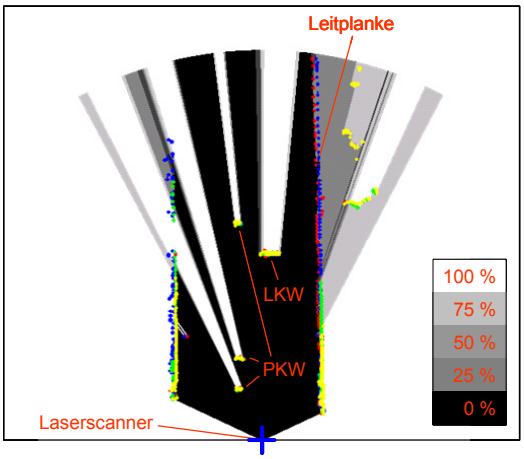

Abbildung 4.3: Verdeckungswahrscheinlichkeiten für reale Messdaten: Zwischen dem Sensor und den ersten Entfernungsmessungen liegt keine Verdeckung vor. Hinter den Fahrzeugen ist die Wahrscheinlichkeit 100 Prozent. Die Wahrscheinlichkeit hinter der Leitplanke ist in hohen Distanzen geringer, da einige Messebenen darüber hinweg messen.

Aufgrund der Berechnungen in Polarkoordinaten kann die Verdeckungserkennung auch für ein relativ feines Raster mit einer Auflösung von 0,2 m und 0,5° über den gesamten Winkelbereich und bis zu einer Distanz von 150 m sehr effizient durchgeführt werden. Auf einem normalen Desktop PC mit 3,4 GHz benötigte die Berechnung pro Scan deutlich weniger als eine Millisekunde.

4.6 Zusammenfassung

Verschiedene Algorithmen zur Vorverarbeitung der Laserscannermessdaten wurden vorgestellt. Nach sensorinternen Vorverarbeitungsschritten ermöglicht die Auswertung der Echopulsbreite die Detektion von Retroreflektoren. Einzelne unerwünschte Messungen des Laserscanners können gezielt ausgeblendet werden. Algorithmen zur Segmentierung bilden Gruppierungen einzelner Entfernungsmessungen, die entsprechend der Modellannahme unterschiedlichen Gegenständen im Fahrzeugumfeld entsprechen. Schließlich führt ein neuer Algorithmus eine Analyse der vorhandenen Messwerte durch, mit der sich Bereiche des Fahrzeugumfelds identifizieren lassen, die aufgrund des radialen Messprinzips des Laserscanners durch vorhandene Gegenstände verdeckt werden. Die Ergebnisse der einzelnen Vorverarbeitungsschritte bilden die Grundlage für die in den folgenden Kapiteln beschriebenen Algorithmen.

5 Laserscannerbasierte Objektverfolgung

Während der Laserscanner nur Entfernungen messen kann, sind für Fahrerassistenzsysteme auch weitere Informationen über die Dynamik der im Fahrzeugumfeld gemessenen Objekte von Interesse. Durch eine zeitliche Verfolgung der gemessenen Objekte lassen sich auch abgeleitete Größen wie beispielsweise deren Geschwindigkeiten ermitteln. Die verwendete Objektverfolgung benutzt für jedes Objekt einen separaten Kalman-Filter zur Schätzung der Position, der Orientierung, der Geschwindigkeit, der Gierrate und der Größe. Mit Hilfe des Kalman-Filters lässt sich zudem der Einfluss des Messrauschens auf die Objektparameter verringern, die direkt gemessen werden. Die vorgestellte Objektverfolgung basiert auf vorherigen Arbeiten [66], wurde jedoch an einigen Stellen erweitert und zur Echtzeitanwendung in Versuchsfahrzeugen optimiert.

Im folgenden Abschnitt wird kurz die verwendete Notation des Kalman-Filters vorgestellt. Danach wird der aus [66] übernommene Ablauf der Objektverfolgung erläutert. Anschließend werden das verwendete Bewegungsmodell und das übernommene Objektmodell beschrieben. Verschiedene neue Erweiterungen und Optimierungen der Objektverfolgung werden in den folgenden Abschnitten vorgestellt.

5.1 Kalman-Filter

Im Jahr 1960 veröffentlichte R. E. Kalman [65] sein Filter zur optimalen Schätzung des dynamischen Objektzustands $\mathbf{x}(k)$ für lineare Systeme unter der Bedingung, dass das Mess- und das Prozessrauschen additiv, mittelwertfrei, weiß, normalverteilt und unkorelliert sind. Für diesen Ansatz werden zwei lineare Modelle verwendet. Das Prozessmodell beschreibt die Veränderung der Objektparameter vom diskreten Zeitschritt k zum nächsten Zeitschritt $k+1$:

$$\mathbf{x}(k+1) = \mathbf{F}(k) \cdot \mathbf{x}(k) + \mathbf{G}(k) \cdot \mathbf{u}(k) + \mathbf{v}(k). \tag{5.1}$$

Hierbei ist $\mathbf{F}(k)$ die Systemmatrix, $\mathbf{G}(k)$ die Steuermatrix, $\mathbf{u}(k)$ der Steuervektor und $\mathbf{v}(k)$ das weiße normalverteilte Prozessrauschen, das durch die Kovarianzmatrix $\mathbf{Q}(k)$ beschrieben wird.

Das Messmodell beschreibt den linearen Zusammenhang zwischen dem Zustandsvektor und den im Messvektor $\mathbf{z}(k)$ zusammengefassten Messgrößen:

$$\mathbf{z}(k) = \mathbf{H}(k) \cdot \mathbf{x}(k) + \mathbf{w}(k). \tag{5.2}$$

Der Messvektor ergibt sich basierend auf der Messmatrix $\mathbf{H}(k)$ und dem weißen normalverteilten Messrauschen $\mathbf{w}(k)$, das durch die Kovarianzmatrix $\mathbf{R}(k)$ charakterisiert wird. Zusätzlich zum Zustandsvektor schätzt das Kalman-Filter auch die Fehlerkovarianz $\mathbf{P}(k)$ der Zustandsgrößen.

Das Kalman-Filter führt eine iterative Schätzung des Objektzustandes $\mathbf{x}(k)$ und der Fehlerkovarianz $\mathbf{P}(k)$ für diskrete Zeitschritte durch. Dabei wird der geschätzte Objektzustand mit $\hat{\mathbf{x}}(k)$ und die geschätzte Kovarianz mit $\hat{\mathbf{P}}(k)$ bezeichnet.

Nach einer einmaligen Initialisierung des Zustandsvektors erlaubt das Kalman-Filter eine Schätzung der Objektdynamik. Der Prädiktionsschritt bestimmt dabei eine Schätzung des dynamischen Zustands (zeitliche Prädiktion) und der zu erwartenden Messwerte (Messwertprädiktion) für den folgenden Zeitschritt. Die zeitliche Prädiktion verwendet das lineare Prozessmodell:

$$\hat{\mathbf{x}}^{-}(k+1) = \mathbf{F}(k) \cdot \hat{\mathbf{x}}(k) + \mathbf{G}(k) \cdot \mathbf{u}(k). \tag{5.3}$$

Das Prozessrauschen wird in der Schätzung der Fehlerkovarianz berücksichtigt:

$$\hat{\mathbf{P}}^{-}(k+1) = \mathbf{F}(k) \cdot \hat{\mathbf{P}}(k) \cdot \mathbf{F}(k)^{T} + \mathbf{Q}(k). \tag{5.4}$$

Mit Hilfe des Messmodells wird eine Messwertprädiktion durchgeführt:

$$\hat{\mathbf{z}}(k+1) = \mathbf{H}(k+1) \cdot \hat{\mathbf{x}}^{-}(k+1). \tag{5.5}$$

Hier wird das Messrauschen in der Prädiktion der Innovationskovarianzmatrix berücksichtigt:

$$\mathbf{S}(k+1) = \mathbf{H}(k+1) \cdot \hat{\mathbf{P}}^{-}(k+1) \cdot \mathbf{H}(k+1)^{T} + \mathbf{R}(k+1). \tag{5.6}$$

Falls neue Messwerte vorliegen, führt der Innovationsschritt damit ein Update des geschätzten Objektzustands durch. Hierfür wird zunächst die Filterverstärkung berechnet:

$$\mathbf{K}(k+1) = \hat{\mathbf{P}}^{-}(k+1) \cdot \mathbf{H}(k+1)^{T} \cdot \mathbf{S}(k+1)^{-1}. \tag{5.7}$$

Mit dem neuen Messvektor $\mathbf{z}(k+1)$ kann dann das Update für den Objektzustand bestimmt werden:

$$\hat{\mathbf{x}}(k+1) = \hat{\mathbf{x}}^{-}(k+1) + \mathbf{K}(k+1) \cdot \left[\mathbf{z}(k+1) - \hat{\mathbf{z}}(k+1)\right]. \tag{5.8}$$

Zusätzlich wird auch das Update für die Kovarianz durchgeführt:

$$\hat{\mathbf{P}}(k+1) = \left[\mathbf{I} - \mathbf{K}(k+1) \cdot \mathbf{H}(k+1)\right] \hat{\mathbf{P}}^{-}(k+1). \tag{5.9}$$

Das Kalman-Filter wird in vielen Bereichen zur optimalen Schätzung angewendet. Die Annahme linearer Systeme ist in der Praxis jedoch oft nicht erfüllt. So können oft auch nichtlineare Prozess- oder Messmodelle auftreten. Ein nichtlineares Prozessmodell hat die allgemeine Form:

$$\mathbf{x}(k+1) = \mathbf{f}(k, \mathbf{x}(k), \mathbf{u}(k), \mathbf{v}(k)). \tag{5.10}$$

Ein nichtlineares Messmodell hat die allgemeine Form:

$$\mathbf{z}(k) = \mathbf{h}(k, \mathbf{x}(k), \mathbf{w}(k)). \tag{5.11}$$

In nichtlinearen Fällen werden Abwandlungen des Kalman-Filters eingesetzt, die die optimale Lösung approximieren. Ein Beispiel hierfür ist das Unscented Kalman-Filter. Es propagiert eine Reihe von konkreten Punkten der geschätzten Zufallsvariable direkt mit Hilfe der nichtlinearen Funktionen. Anschließend wird aus diesen Punkten die Schätzung und ihre Varianz rekonstruiert [129].

In [66] wird mit dem Extended Kalman-Filter eine andere Alternative verwendet. Im Extended Kalman-Filter wird von einem additiven Mess- und Prozessrauschen ausgegangen. Für

die folgenden Betrachtungen wird weiterhin auf die Notation des Steuervektors verzichtet, da dieser in der vorliegen Arbeit bei der Verfolgung der durch die Sensoren vermessenen Objekte nicht bekannt ist und nicht modelliert wird. Somit ergibt sich als Prozessmodell:

$$\mathbf{x}(k+1) = \mathbf{f}(k, \mathbf{x}(k)) + \mathbf{v}(k). \tag{5.12}$$

Als Messmodell ergibt sich:

$$\mathbf{z}(k) = \mathbf{h}(k, \mathbf{x}(k)) + \mathbf{w}(k). \tag{5.13}$$

Das Prozessmodell ist für die Prädiktion nur exakt, wenn man als Parameter den exakten Zustand $\mathbf{x}(k)$ übergibt. Da nur der Schätzwert $\hat{\mathbf{x}}(k)$ bekannt ist, wird der exakte Zustand durch eine Taylor-Entwicklung um den Schätzwert $\hat{\mathbf{x}}(k)$ approximiert.

$$\mathbf{x}(k+1) = \mathbf{f}\left(k, \hat{\mathbf{x}}(k)\right) + \mathbf{J}_{\mathbf{f(x)}} \Delta\mathbf{x} + \frac{1}{2} \sum_{i=1}^{\dim(\mathbf{x})} \mathbf{e_i} \Delta\mathbf{x}^T \mathbf{H}_{\mathbf{f_i(x)}} \Delta\mathbf{x} + \ldots + \mathbf{v}(k) \tag{5.14}$$

mit

$$\Delta\mathbf{x} = \mathbf{x}(k) - \hat{\mathbf{x}}(k). \tag{5.15}$$

Hierbei bezeichnet $\mathbf{e_i}$ den i-ten kartesischen Basisvektor mit der gleichen Dimension wie der Zustandsvektor $\mathbf{x}(k)$. Das bedeutet, alle Werte aus $\mathbf{e_i}$ sind Null, lediglich der i-te Wert ist Eins. Weiterhin bezeichnen $\mathbf{J_{f(x)}}$ die Jacobi-Matrix der Funktion $\mathbf{f}(k, \mathbf{x}(k))$ und $\mathbf{H_{f_i(x)}}$ die Hesse-Matrix der i-ten Komponente von $\mathbf{f}(k, \mathbf{x}(k))$. Bricht man die Taylor-Entwicklung nach dem quadratischen Glied ab, ergibt sich für die Prädiktion des Extended Kalman-Filters [5]:

$$\hat{\mathbf{x}}^-(k+1) = \mathbf{f}\left(k, \hat{\mathbf{x}}(k)\right) + \frac{1}{2} \sum_{i=1}^{\dim(\mathbf{x})} \mathbf{e_i} \,\text{Spur}\left(\mathbf{H}_{\mathbf{f_i(x)}} \hat{\mathbf{P}}(k)\right). \tag{5.16}$$

$$\begin{aligned} \hat{\mathbf{P}}^-(k+1) \;=\; & \mathbf{J}_{\mathbf{f(x)}} \hat{\mathbf{P}}(k) \, \mathbf{J}_{\mathbf{f(x)}}{}^T + \\ & \frac{1}{2} \sum_{i=1}^{\dim(\mathbf{x})} \sum_{j=1}^{\dim(\mathbf{x})} \mathbf{e_i} \mathbf{e_j}^T \,\text{Spur}\left(\mathbf{H}_{\mathbf{f_i(x)}} \hat{\mathbf{P}}(k) \, \mathbf{H}_{\mathbf{f_j(x)}} \hat{\mathbf{P}}(k)\right) + \\ & \mathbf{Q}(k). \end{aligned} \tag{5.17}$$

In der Praxis wird die Taylor-Entwicklung oft auch bereits nach dem linearen Glied abgebrochen. Dann entfallen jeweils die zweiten Summanden im Prädiktionsschritt.

Eine analoge Taylor-Entwicklung ist für die nichtlineare Messfunktion notwendig:

$$\begin{aligned} \mathbf{z}(k+1) \;=\; & \mathbf{h}\left(k, \hat{\mathbf{x}}^-(k+1)\right) + \mathbf{J}_{\mathbf{h(x)}}(k+1) \, \Delta\mathbf{x} + \\ & \frac{1}{2} \sum_{i=1}^{\dim(\mathbf{x})} \mathbf{e_i} \Delta\mathbf{x}^T \mathbf{H}_{\mathbf{h_i(x)}} \Delta\mathbf{x} + \ldots + \\ & \mathbf{w}(k) \end{aligned} \tag{5.18}$$

mit

$$\Delta\mathbf{x} = \mathbf{x}^-(k+1) - \hat{\mathbf{x}}^-(k+1). \tag{5.19}$$

Wieder bezeichnen $\mathbf{J_{h(x)}}$ die Jacobi-Matrix und $\mathbf{H_{h_i(x)}}$ die Hesse-Matrix der i-ten Komponente der Funktion $\mathbf{h}\left(k, \hat{\mathbf{x}}^{-}(k+1)\right)$.

Für die Prädiktion der Messung ergibt sich beim Abbruch der Taylor-Entwicklung nach dem quadratischen Glied [5]:

$$\hat{\mathbf{z}}(k+1) = \mathbf{h}\left(k, \hat{\mathbf{x}}^{-}(k+1)\right) + \frac{1}{2}\sum_{i=1}^{\dim(\mathbf{x})} \mathbf{e_i}\,\text{Spur}\left(\mathbf{H_{h_i(x)}}\hat{\mathbf{P}}^{-}(k+1)\right) . \tag{5.20}$$

Entsprechend kann die Innovationskovarianz nach der folgenden Formel bestimmt werden:

$$\begin{aligned}\mathbf{S}(k+1) \;=\; & \mathbf{J_{h(x)}}\hat{\mathbf{P}}(k)\,\mathbf{J_{h(x)}}^{T} + \\ & \frac{1}{2}\sum_{i=1}^{\dim(\mathbf{x})}\sum_{j=1}^{\dim(\mathbf{x})} \mathbf{e_i}\mathbf{e_j}^{T}\,\text{Spur}\left(\mathbf{H_{h_i(x)}}\hat{\mathbf{P}}^{-}(k+1)\,\mathbf{H_{h_j(x)}}\hat{\mathbf{P}}^{-}(k+1)\right) + \\ & \mathbf{R}(k) .\end{aligned} \tag{5.21}$$

Auch bei der Messwertprädiktion entfallen jeweils die zweiten Summanden, wenn die Taylor-Entwicklung bereits nach dem linearen Glied abgebrochen wird.

5.2 Ablauf der Objektverfolgung

In diesem Abschnitt wird ein Überblick über die Anwendung des Kalman-Filters zur laserscannerbasierten Objektverfolgung im Fahrzeugumfeld gegeben. Der grundlegende Ablauf basiert auf dem in [66] vorgestellten Verfahren. Abbildung 5.1 zeigt einen Überblick über die notwendigen Schritte für die laserscannerbasierte Objektverfolgung mit einem Kalman-Filter nach [67]. Ähnliche Systemgestaltungen finden sich auch in [43, 87, 128].

Das gesamte System wird für jeden Messzyklus durchlaufen. Die Messdaten des Laserscanners durchlaufen zuerst die Vorverarbeitung. Hier werden die im Kapitel 4 beschriebenen Vorverarbeitungsschritte durchgeführt. Das Ergebnis sind segmentierte Messdaten.

Nun müssen die Objektmerkmale wie Position, Objektgröße und -orientierung bestimmt werden. Zu diesem Zweck wird das in [66] vorgestellte, rechteckige Objektmodell verwendet. Dieser Ansatz versucht, rechteckige Objekte in die aktuellen Entfernungsmessungen zu legen. In [66] wird hierfür das iterative Verfahren WICP (*engl. WICP = weighted iterative closest point*) angewendet. Anschließend können aus dem eingepassten Rechteck die Messwerte für Orientierung, Größe und Position ermittelt werden. Eine detaillierte Beschreibung des rechteckigen Objektmodells erfolgt in Abschnitt 5.4. Da die Objektverfolgung parallel mit anderen Algorithmen in Echtzeit ablaufen muss, wurde hier auf den Einsatz des rechenaufwändigen WICP verzichtet. Eine Alternative zur Formbestimmung wird in Abschnitt 5.6 präsentiert.

Die Prädiktion sorgt für die Vergleichbarkeit der gespeicherten Objekthistorien aus vorherigen Zeitschritten mit den aktuellen Messwerten. Hierzu wird als Erstes der Zustand der gespeicherten Objekte auf den Zeitpunkt der Messdaten prädiziert. Anschließend muss die Bewegung des eigenen Fahrzeugs kompensiert werden, da die Objektverfolgung absolute Geschwindigkeiten schätzen soll, während der Laserscanner die Objektpositionen nur

relativ zum eigenen Fahrzeug messen kann. Schließlich berechnet die Messwertprädiktion die erwarteten Messwerte.

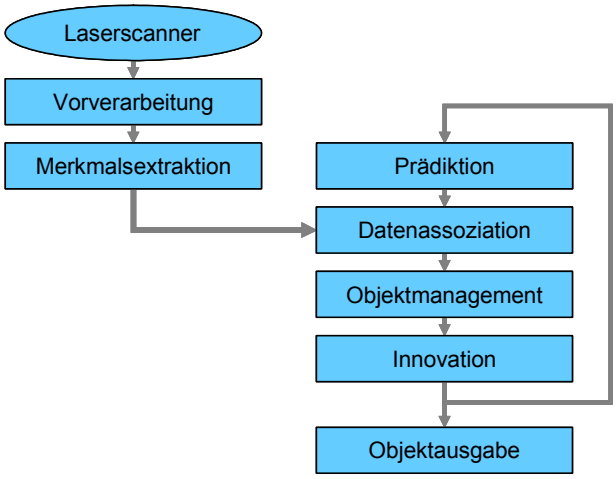

Abbildung 5.1: Überblick über die laserscannerbasierte Objektverfolgung mit dem Kalman-Filter: Nach der Vorverarbeitung der Messdaten werden die Messvektoren für das Kalman-Filter extrahiert. Objekte aus vorherigen Zeitschritten werden auf den aktuellen Zeitpunkt prädiziert. Mit Hilfe prädizierter Messvektoren wird die Zuordnung der Messwerte zu bekannten Objekten durchgeführt. Das Objektmanagement bearbeitet nicht zugeordnete Messungen und Objekte, die den Messbereich der Sensoren verlassen haben. Schließlich wird in der Innovation das Filterupdate durchgeführt. Die Objektausgabe wählt die an das Assistenzsystem übergebenen Objekte aus.

Nun erfolgt die Datenassoziation, die den prädizierten Objekten passende gemessene Segmente zuordnet. Hierbei wird unter Berücksichtigung der Unsicherheiten die Mahalanobis-Distanz zwischen prädizierter und gemessener Position ausgewertet. Verfahren zur Datenassoziation werden in [66, 87, 129] beschrieben.

In bisherigen Arbeiten wurde hier jedem Objekt nur ein Segment zugeordnet. In der Praxis zerfallen Objekte teilweise auch in mehrere Segmente. Ein Ansatz zur Zuordnung mehrerer Segmente zu einem prädizierten Objekt wird in Abschnitt 5.7 vorgestellt.

Für nicht zugeordnete Segmente legt das Objektmanagement neue Objekte an. Gleichzeitig werden Objekte aus dem Speicher entfernt, die den Erfassungsbereich des Laserscanners verlassen haben.

Die Innovation führt das Update des Kalman-Filters mit den aktuellen Messwerten durch.

Schließlich erfolgt die Objektausgabe. Hier werden die Objekte ausgewählt, die dem Fahrerassistenzsystem übergeben werden. Da das hier vorgestellte System mehrere Assistenzsysteme mit Informationen versorgen kann, wird in Abschnitt 5.8 auf die anwendungsabhängige Objektauswahl eingegangen.

5.3 Verwendetes Prozessmodell

Das Prozessmodell beschreibt bei der Verfolgung von Objekten im Fahrzeugumfeld im Wesentlichen deren Bewegung. Die Objektgröße wird sinnvollerweise als konstant angenommen.

Ein allgemeingültiges Bewegungsmodell ist das Freie-Masse-Modell. Hier werden die Bewegungsanteile v_x in x-Richtung und v_y in y-Richtung unabhängig voneinander geschätzt. Zusätzlich kann unabhängig davon eine Orientierung des Objektes geschätzt werden [66].

Wenn es sich bei dem verfolgten Objekt um ein Fahrzeug handelt, kann dessen Bewegung allerdings besser durch ein sogenanntes Einspurmodell beschrieben werden. Hier wird die Bewegung nur in Richtung der Orientierung des Objektes modelliert. Eine Richtungsänderung der Bewegung wird durch eine Orientierungsänderung des Objektes erreicht. Das Einspurmodell eignet sich für Fußgänger jedoch nur bedingt, da sich diese auch seitlich bewegen können. Diese Flexibilität müsste dann durch hohe Prozessunsicherheiten in der Orientierung modelliert werden [5, 66].

Die in den folgenden Abschnitten beschriebene Objektverfolgung erfolgt unabhängig vom Typ des Objektes. Sie orientiert sich an [66] und verwendet ein Freie-Masse-Modell. Da der Laserscanner nur Positionsmessungen liefert, lässt sich die Beschleunigung nur indirekt durch zweimalige Ableitung nach der Zeit ermitteln. Eine Schätzung der Beschleunigung durch das Kalman-Filter zeigt darum ein starkes Nachlaufen bezüglich der realen Beschleunigungswerte. Experimentellen Untersuchungen ergaben darüber hinaus ein stärkeres Rauschen der geschätzten Objektgeschwindigkeiten. Aus diesem Grund erfolgt in der vorliegenden Arbeit eine Modellierung der Objektbewegung durch eine konstante Geschwindigkeit. Das Prozessrauschen modelliert dabei eventuell auftretende Beschleunigungsvorgänge. Zusätzlich werden unabhängig von der Bewegung eine Orientierung ψ und eine Gierrate ω geschätzt. Das Kalman-Filter schätzt ebenfalls die Breite w und die Länge l des Objektes, wobei diese als konstant modelliert sind.

Die Objektverfolgung verwendet somit folgenden Zustandsvektor:

$$\mathbf{x} = \begin{pmatrix} x & v_x & y & v_y & \psi & \omega & w & l \end{pmatrix}. \tag{5.22}$$

Das zugehörige Prozessmodell für einen Zeitschritt der Länge T ist durch folgende Gleichung beschrieben:

$$\mathbf{x}(k+1) = \begin{bmatrix} 1 & T & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & T & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & T & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{x}(k) + \mathbf{v}(k). \tag{5.23}$$

Der Steuervektor $\mathbf{u}(k)$ ist in diesem Modell nicht bekannt und wird dementsprechend auf Null gesetzt. Das Prozessrauschen wird entspechend des direkt diskreten Prozessmodells für konstante Beschleunigung [5] gewählt und wurde aus [66] übernommen.

5.4 Rechteckiges Objektmodell

Die Messdaten des Laserscanners bestehen für jedes Objekt aus einzelnen Entfernungsmessungen. Für Fahrerassistenzsysteme sind jedoch Informationen über Objekte im Fahrzeugumfeld von Bedeutung. Aus diesem Grund müssen aus den segmentierten Entfernungsmessungen sinnvolle Objektinformationen gewonnen werden. Diese beinhalten die Position, die Größe und die Orientierung des vermessenen Objektes.

Zur Bestimmung der gemessenen Position muss ein geeigneter Referenzpunkt aus den einzelnen Entfernungsmessungen des Objektes berechnet werden. Dies ist speziell bei Objekten von Bedeutung, die ihre Orientierung zum Sensor ändern. Bereits in [43] wird erläutert, dass bei ungünstiger Bestimmung des Referenzpunktes bei größeren Objekten Inkonsistenzen auftreten können, die zu nicht vernachlässigbaren Geschwindigkeitsoffsets bei der Schätzung führen. Aus diesem Grund führt die Verwendung des Schwerpunktes der Entfernungsmessungen als Referenzpunkt in verschiedenen Fällen nur zu ungenauen Geschwindigkeitsschätzungen.

Viele Objekte können im Messmodell mit einer rechteckigen Grundfläche approximiert werden. Hierzu zählen Fahrzeuge, Leitplanken, Häuser oder auch Zäune und Wände. Diese Objekte werden vom Laserscanner in der Regel von einer oder von zwei Seiten vermessen. Die entsprechenden Seiten können mit der Orientierung zum Sensor wechseln.

Die Verfolgung der Objekte erfolgt mit dem in [66] vorgeschlagenen rechteckigen Objektmodell. Dabei wird aus den vorhandenen Entfernungsmessungen des Laserscanners eine rechteckige Form extrahiert. Daraus lässt sich die gemessene Orientierung des Objektes festlegen. Der verwendete Referenzpunkt ergibt sich in Abhängigkeit von den vermessenen Seiten. Werden zwei Seiten vermessen, ist die eingeschlossene Ecke ein markanter Referenzpunkt. Im Fall einer vermessenen Seite erfolgt eine Verdeckungsanalyse für deren Anfang und Ende. Ergeben sich lediglich Unsicherheiten bezüglich der Position einer Ecke aufgrund von Verdeckungen oder der Messbereichsgrenzen, stellt die andere Ecke der Seite einen geeigneten Referenzpunkt dar. Falls die Position beider Ecken in vergleichbarer Unsicherheit vorliegt, verwendet die Objektverfolgung die Seitenmitte als Referenzpunkt (Abbildung 5.2).

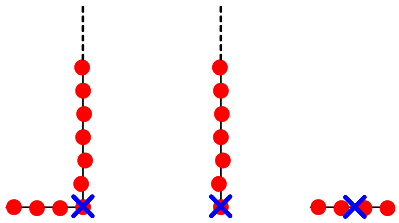

Abbildung 5.2: Referenzpunkte beim rechteckigen Objektmodell: Liegen Messnungen von zwei der Objektseiten vor, ist die eingeschlossene Ecke der Referenzpunkt (links). Falls der Sensor nur eine Seite erfasst, muss untersucht werden, ob diese komplett vermessen wurde. In manchen Fällen liefert der Laserscanner keine Messwerte von einem Ende der Seite. In diesen Fällen stellt die andere Ecke einen geeigneten Referenzpunkt dar (Mitte). Ansonsten ist die Seitenmitte der Referenzpunkt (rechts).

Insgesamt existieren somit acht mögliche Referenzpunkte. Während der Messwertprädiktion ist nicht bekannt, welcher Referenzpunkt am besten vom Laserscanner gemessen wird, da noch keine Messwerte zugeordnet wurden. Aus diesem Grund müssen alle Referenzpunkte prädiziert werden. Bei dieser Prädiktion gehen die Größe und die Orientierung des Objektes mit in das Messmodell ein, da das Kalman-Filter als Position die Objektmitte schätzt. Abbildung 5.3 zeigt die möglichen Referenzpunkte.

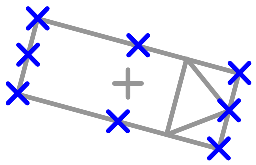

Abbildung 5.3: Prädizierte Referenzpunkte: Bei der Messwertprädiktion werden alle Referenzpunkte prädiziert. Dies wird durch eine geeignete Messmatrix erreicht, die zur Berechnung der Referenzpunkte die Objektposition (geometrischer Mittelpunkt), die Orientierung und die Objektgröße berücksichtigt.

In der Datenassoziation werden dem Objekt aktuelle Messdaten zugeordnet. Dabei wird auch aufgrund der Messdaten der passende Referenzpunkt ausgewählt. In die Innovation geht anschließend nur der Abstand zwischen diesem prädizierten und dem gemessenen Referenzpunkt ein (Abbildung 5.4).

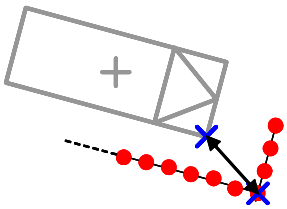

Abbildung 5.4: Nach der Datenassoziation steht der gemessene Referenzpunkt fest. In die Innovation geht nur die Distanz zwischen prädiziertem und gemessenem Referenzpunkt ein.

Für die Objektgröße wird die Länge und die Breite des Rechteckes als Messwert benutzt. Die Größen gehen nur in das Filter ein, wenn die entsprechende Seite bis zu beiden Enden vermessen wurde.

Einige Objekte passen nicht in dieses Schema und müssen bei der Verfolgung getrennt betrachtet werden. Wenn aufgrund der geringen horizontalen Ausdehnung im Messbereich keine Form erkennbar ist, lässt sich ohne Probleme der Schwerpunkt als Referenzpunkt verwenden. Zu diesen Fällen zählen Fußgänger, Leitpfosten, Schilder oder Bäume. Die Orientierung wird hier mit 0° festgelegt oder ergibt sich aus einer eventuellen Bewegung des Objektes.

In manchen Fällen weisen Objekte eine sehr komplexe Form auf, die stark von dem rechteckigen Objektmodell abweicht. Dementsprechend lässt sich hier ebenfalls keine Orientierung aus den Entfernungsmessungen bestimmen. In diesen Fällen schlägt [66] eine Objektverfolgung durch den direkten Vergleich der einzelnen Entfernungsmessungen aufeinanderfolgender Zeitschritte vor. Da dieses Verfahren jedoch sehr zeitaufwendig ist, wurde hier darauf verzichtet. Für entsprechende Objekte wird die Orientierung ebenfalls auf 0° festgelegt. Die achsenparallele Ausdehnung wird als Größe interpretiert. Als Referenzpunkt wird der geometrische Mittelpunkt gewählt.

5.5 Reduktion der Messvektoren

Zum Zeitpunkt der Messwertprädiktion ist unklar, welcher der Referenzpunkte des rechteckigen Objektmodells am besten gemessen werden kann. Zur Assoziation der Messdaten müssen darum alle Referenzpunkte prädiziert werden. Bisherige Arbeiten führten diese Prädiktion direkt im Kalman-Filter durch. Der Messvektor enthielt damit außer den Werten für Länge, Breite und Orientierung pro Referenzpunkt zwei Einträge. Das sind bei acht Referenzpunkten insgesamt 19 statt der sonst notwendigen 5 Einträge im Messvektor. Bei der üblichen Prädiktion im Kalman Filter wirkt sich das direkt auf die Dimension der Messmatrix $\mathbf{H}$ aus. Wie in Gleichung 5.6 ersichtlich ist, hat die dort berechnete Innovationskovarianz $\mathbf{S}$ somit die Dimension 19×19. Diese muss bei der Innovation in Gleichung 5.7 zur Berechnung der Filterverstärkung invertiert werden. Da die Rechenzeit von Matrixinversionen im Allgemeinen überproportional mit der Dimension steigt, ist der notwendige Mehraufwand bei der neuen Matrixgröße nicht zu vernachlässigen.

In der Innovation wird lediglich ein Referenzpunkt zum Filter-Update verwendet. Aus diesem Grund ist hier ein Messvektor mit fünf Einträgen ausreichend. Eine entsprechende Reduktion der Dimension der invertierten Matrix bedingt jedoch eine zweistufige Messwertprädiktion (Abbildung 5.5).

Die Prädiktion des Messvektors erfolgt nur für den geometrischen Mittelpunkt des Rechteckes. Ein zweiter Schritt berechnet anschließend die einzelnen Referenzpunkte für die Datenassoziation direkt aus diesem Punkt, der Größe und der Orientierung. Die Datenassoziation wählt den verwendeten Referenzpunk dann später in Abhängigkeit von den aktuellen Messdaten aus.

Damit zur Innovation der korrekte prädizierte Referenzpunkt und die korrespondierende Messmatrix verwendet werden können, wird vor der Innovation eine zweite Merkmalsprädiktion durchgeführt, die eine Schätzung des benötigten Referenzpunktes ermittelt. Die Ergebnisse der ersten Messwertprädiktion werden dabei überschrieben.

Durch die vorgestellte Änderung konnte der Rechenaufwand der Matrixinvertierung deutlich reduziert werden. Eine genaue Auswertung der eingesparten Rechenzeit befindet sich in Kapitel 10.

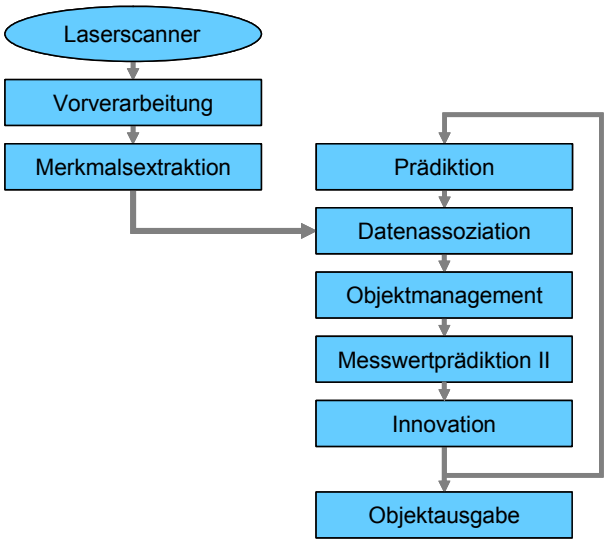

Abbildung 5.5: Neuer Ablauf der Objektverfolgung: Zur Dimensionsreduktion der Matrixinvertierung im Kalman-Filter wird in der Prädiktion nur noch der geometrische Mittelpunkt als Messwert prädiziert. Die anderen Referenzpunkte werden daraus berechnet. Für eine korrekte Innovation muss in einer zweiten Merkmalsprädiktion der gemessene Referenzpunkt mit der korrespondierenden Messmatrix prädiziert werden.

5.6 Formbestimmung

Zur Extraktion der Messwerte für das Kalman-Filter muss aus den einzelnen Entfernungsmessungen des Laserscanners eine Objektform bestimmt werden. Als Ersatz für das recht rechenintensive WICP wird hier ein neuer Ansatz vorgestellt, der mit relativ einfachen Methoden die benötigten Konturinformationen bestimmt. Da das Verfahren pro Zeitschritt auf viele Segmente angewendet werden muss, wurde hier auf aufwändige Algorithmen wie zum Beispiel lineare Regression verzichtet. Bei ausreichender Rechenkapazität kann das Verfahren mit derartigen Ansätzen verbessert werden.

Auch hier kommt das rechteckige Objektmodell zum Einsatz. Der Laserscanner liefert je nach Perspektive von rechteckigen Objekten Messungen von einer oder von zwei Objektseiten. Die Verteilung der Messpunkte wird entsprechend als L-Form beziehungsweise als I-Form bezeichnet. Objekten mit komplexeren Konturen oder rechteckige Objekte, von denen nicht ausreichend Entfernungsmessungen vorliegen, liefern keine derartig charakteristische Messwertverteilung. Diese Verteilungen werden mit der O-Form zusammengefasst. Abbildung 5.6 zeigt Messdaten des Laserscanners in der Draufsicht und ein Videobild der entsprechenden Szene. Beispiele für die beschriebenen Messpunktverteilungen wurden in der Abbildung hervorgehoben.

Den einzelnen Entfernungsmessungen eines Objektes muss nun eine entsprechende Form zugeordnet werden. Informationen über Position, Größe und Orientierung können dann aus der Form abgeleitet werden.

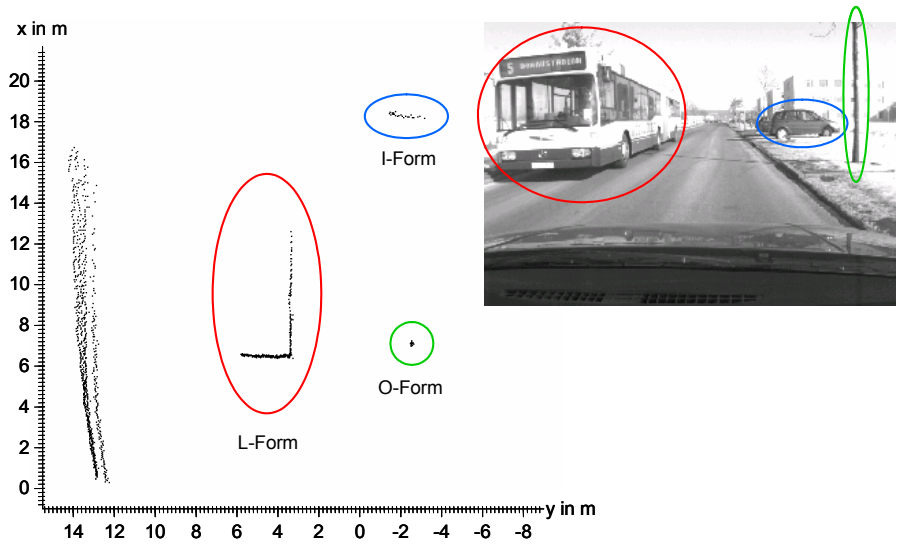

Abbildung 5.6: Mögliche Messpunktverteilungen: Bei rechteckigen Objekten können Entfernungsmessungen von einer Seite (I-Form) oder von zwei Seiten (L-Form) vorliegen. Alle weiteren Messpunktverteilungen werden als O-Form zusammengefasst.

Ein einfacher Ansatz zur Formberechnung wird bereits in [43] beschrieben. Bei I-Formen werden einfach die Punkte am linken und am rechten Ende des Objektes verbunden. Bei L-Formen zeichnet sich die Ecke dadurch aus, dass die entsprechende Entfernung von allen Messungen des Objektes die kleinste ist. Die beiden Seiten der L-Form entstehen durch Verbindung des entsprechenden Eckpunktes mit den äußeren Punkten. Ein Beispiel für die verwendeten Punkte ist im linken Teil von Abbildung 5.7 dargestellt.

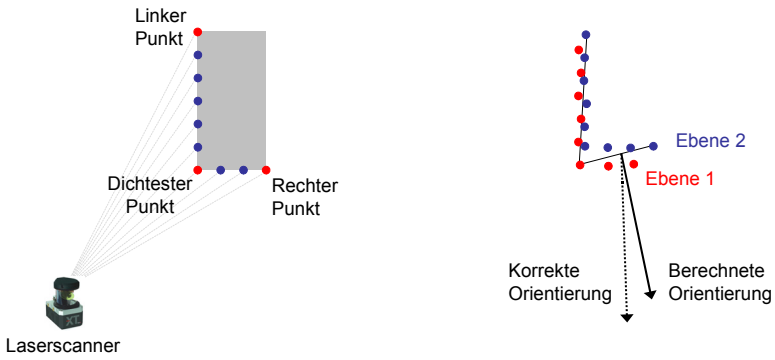

Abbildung 5.7: In einem ersten Ansatz werden zur Bestimmung einer L-Form der dichteste Messpunkt am Sensor sowie die beiden äußeren Punkte der Kontur verwendet (links). Die Anwendung auf reale Daten zeigt Probleme bei Entfernungsmessungen unterschiedlicher Messebenen (rechts).

Wird dieser Ansatz jedoch auf reale Messdaten angewendet, ergeben sich verschiedene Modellabweichungen (Abbildung 5.7 rechts). Durch das vorhandene Messrauschen liegen die Messpunkte nicht wie erwartet auf einer Geraden. Noch deutlicher wirken sich Entfernungsunterschiede zwischen einzelnen Messebenen aus. Zum einen werden hier Objekte in unterschiedlichen Höhen vermessen. Da Fahrzeuge oberhalb von Motorhaube und Kofferraum deutlich kleiner sind, können hier unterschiedliche Entfernungsmessungen durchaus der Realität entsprechen. Zum anderen traten bei manchen Sensoren älteren Typs für Abstandsmessungen unterschiedliche absolute Messfehler in den einzelnen Messebenen auf, die ähnliche Effekte erzeugten.

Die Unterschiede in den Entfernungsmessungen verschiedener Messebenen werden in einem neuen Ansatz berücksichtigt. In einem ersten Schritt werden hierzu verschiedene Hypothesen für die Objektform berechnet. Diese Hypothesen beinhalten je eine L-Form und eine I-Form, die aus den Messpunkten einer einzelnen Ebene berechnet wurden. Zusätzlich werden eine L-Form und eine I-Form basierend auf allen Messpunkten erstellt. Dies ist sinnvoll, wenn in den einzelnen Messebenen nicht ausreichend Messungen vorhanden sind, um sicher eine Form bestimmen zu können. Die aufgestellten Hypothesen für die Form eines Objektes sind in Abbildung 5.8 an einem Beispiel veranschaulicht.

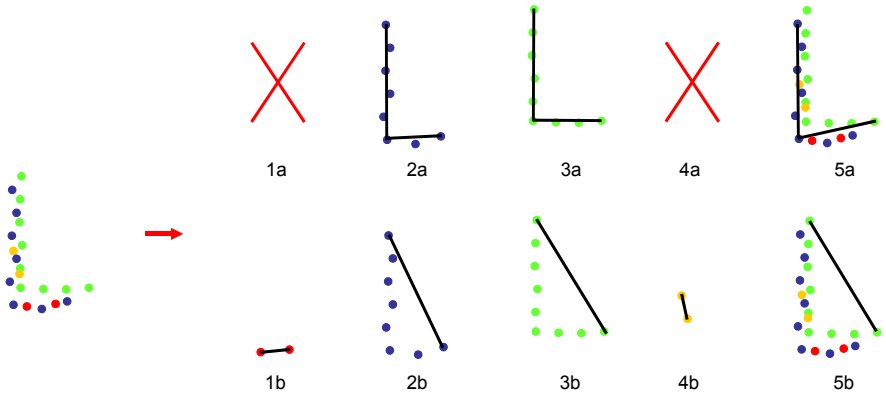

Abbildung 5.8: Für die Formbestimmung werden verschieden Formhypothesen berechnet. Je eine L-Form (1a - 4a) und eine I-Form (1b - 4b) werden aus den Messpunkten der einzelnen Messebenen bestimmt. Eine zusätzliche L-Form (5a) und eine zusätzliche I-Form (5b) basieren auf allen Messpunkten.

Verschiedene Formhypothesen können möglicherweise wegen zu geringer Messpunktanzahl nicht aufgestellt werden (1a und 4a). Andere Hypothesen werden aufgrund verschiedener Restriktionen verworfen. So werden die Hypothesen 1b und 4b aufgrund ihrer Größe verworfen. Die Hypothese 5a scheidet aufgrund des ungünstigen Winkels zwischen den Seiten der L-Form aus.

Aus allen verfügbaren Hypothesen muss nun die Beste ausgewählt werden. Dieser Auswahl basiert auf dem mittleren Abstand der Messpunkte von der bestimmten Form. Da viele Messpunkte in der Regel eine Form mit längeren Kanten erzeugen, auf die das Messrauschen des Laserscanners bezüglich der Orientierung einen geringeren Einfluss hat, berücksichtigt die Auswahl ebenfalls die Anzahl der Messpunkte. In statistischen Experimenten wurden verschiedene Methoden zur Kombination beider Kriterien untersucht. Die besten Ergebnisse

erzielte dabei die Division des mittleren Abstands durch den Logarithmus der Messpunktanzahl. Das Maß zur Bewertung der Formen kann also beschrieben werden durch:

$$s_{Form} = \frac{\textit{mittlerer Abstand}}{\log\left(\textit{Anzahl der Messpunkte}\right)} \,. \tag{5.24}$$

Es wird die Form mit dem geringsten gewichteten Abstand ausgewählt. Überschreitet dieser Wert eine vorgegebene Schranke, konnte keine geeignete Form gefunden werden. In solchen Fällen wird eine O-Form zugeordnet. In Abbildung 5.9 ist die Auswahl der Hypothese 3a des obigen Beispiels als beste Form dargestellt.

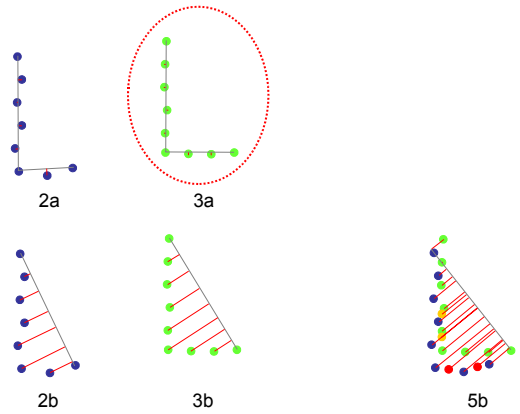

Abbildung 5.9: Auswahl der besten Formhypothese: Nach dem vorgestellten Kriterium ist Hypothese 3a die beste berechnete Form.

Die Bestimmung des Referenzpunktes und der Messwerte für die Objektgröße und -orientierung erfolgt bei L-Formen und I-Formen entsprechend der Beschreibungen in Abschnitt 5.4.

Für O-Formen existieren zwei Ansätze. Falls aus vorherigen Zeitschritten für das Objekt eine Orientierung bekannt ist, wird diese verwendet. Für diese Orientierung wird ein Rechteck berechnet, das alle Entfernungsmessungen enthält. Von allen acht Referenzpunkten wird unter Berücksichtigung von Verdeckungen und der Sensorcharakteristik derjenige Referenzpunkt gewählt, der die beste Orientierung zum Laserscanner hat. Ist auch in vorherigen Zeitschritten keine Rechteckform anwendbar gewesen, dann wird eine Orientierung von 0° angenommen. Die Objektgröße wird durch ein achsenparalleles Rechteck beschrieben. Als Referenzpunkt wird der geometrische Mittelpunkt gewählt.

5.7 Segmentfusion

In der Segmentierung werden die einzelnen Entfernungsmessungen des Laserscanners gruppiert. Dies geschieht auf Basis des Abstands zwischen den Entfernungsmessungen unter der Annahme, dass Messungen, die zum selben Objekt gehören, dicht beieinander liegen. Leider existieren auch Fälle, in denen es der Segmentierung nicht gelingt, alle Messungen eines Objektes in einem Segment zu gruppieren. Das Objekt zerfällt also in mehrere Segmente. Dies kann zum Beispiel bei LKWs auftreten, die häufig nur im Bereich der

Räder und des Unterbaus vermessen werden. Die dabei vermessene Kontur weist oft größere Lücken auf und die abstandsbasierte Segmentierung ist dadurch nicht in der Lage, alle Entfernungsmessungen in einem Segment zu gruppieren (Abbildung 5.10).

Bisherige Verfahren zur Datenassoziation weisen einem Objekt in der Regel nur maximal ein Segment zu. Eine Ausnahme bildet der in [128] vorgestellte Ansatz mit diversen Assoziationshypothesen. Dieser Ansatz ist hier jedoch aufgrund seiner Komplexität und des enormen Bedarfs an Rechenzeit ungeeignet. Um Sonderfälle mit mehreren Segmenten besser verarbeiten zu können, wird hier eine neue Methode zur Gruppierung von Segmenten vorgestellt.

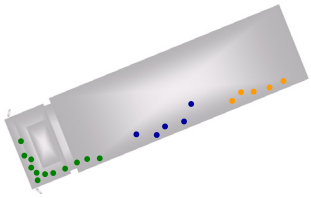

Abbildung 5.10: In einigen Fällen können Objekte in mehrere Segmente zerfallen. Ein Beispiel hierfür sind dicht vorbeifahrende LKWs, die im Bereich der Räder und des Unterbaus vermessen werden und damit keine durchgängige Kontur liefern.

In einem ersten Schritt wird jedem Objekt maximal ein Segment zugeordnet. Dieser Schritt entspricht den bisherigen Verfahren zur Datenassoziation und beruht auf der Mahalanobis-Distanz. In einem zweiten Schritt werden weitere Segmente zum Objekt hinzugefügt. Hier werden sogar verschiedene Objekte zu einem Objekt verschmolzen, wenn davon ausgegangen werden kann, dass die zugehörigen Messdaten vom gleichen Objekt im Fahrzeugumfeld stammen.

Als Grundlage für das Zusammenfügen von Messdaten wird der im Straßenverkehr übliche Abstand bei bewegten Objekten verwendet. In Deutschland drohen bei einem Abstand von weniger als einem Viertel des Tachowertes in Metern bereits deutliche Sanktionen inklusive mehrerer Punkte im Verkehrszentralregister in Flensburg. Beispielaufnahmen mit dem Laserscanner zeigten allerdings, dass dieser Abstand in der Praxis trotzdem oft nicht eingehalten wird. Aus diesem Grund wird hier die Hälfte, also ein Achtel des Tachowertes in Metern als minimaler Abstand zwischen nacheinander fahrenden Fahrzeugen angenommen. Verwendet man zur Berechnung die von der Objektverfolgung geschätzte Geschwindigkeit v in $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$, wird der Mindestabstand d_{min} folgendermaßen ermittelt:

$$d_{min} = \frac{v \cdot 3,6\,\mathrm{s}}{8}. \tag{5.25}$$

Da der Mindestabstand von der Objektgeschwindigkeit abhängt, muss eine verlässliche Geschwindigkeitsschätzung vorliegen, damit die Fusion der Messdaten durchgeführt werden kann. Dies wird durch Überprüfung der Konsistenz der Objekthistorie bezüglich Mess- und Prozessunsicherheiten sichergestellt. Ein entsprechendes Verfahren beruht auf dem NIS (*engl. NIS = normalized innovation squared*) und wurde in [66] und [8] vorgestellt.

Falls die Objekthistorie und damit die geschätzte Geschwindigkeit als verlässlich angenommen werden kann, wird der Bereich vor und hinter dem zugeordneten Segment auf der Länge des Sicherheitsabstandes nach weiteren Messwerten durchsucht. Wenn neue Messwerte zum Objekt hinzugefügt wurden, wird für das so konstruierte Objekt wiederum der Bereich davor und dahinter untersucht. Abbildung 5.11 illustriert die beschriebene Vorgehensweise. Das mittlere Segment wird als Erstes dem Objekt zugeordnet. Da die Objekthistorie als verlässlich eingestuft wurde, wird der Bereich vor und hinter dem Segment durchsucht. Dabei werden Entfernungsmessungen der äußeren Segmente gefunden. Diese Segmente werden darum zum Objekt hinzugefügt. Als Ergebnis entsteht die rechts dargestellte Objektgröße.

Sobald sich das Objekt nicht mehr bewegt, kann mit der hier vorgestellten Methode keine Segmentfusion mehr durchgeführt werden, da stehende Objekte deutlich dichter beieinander auftreten können. Aufgrund der Objekthistorie bleibt jedoch die geschätzte Objektgröße erhalten. Diese wird dann am zugeordneten Segment ausgerichtet. In der Regel werden dann die übrigen Segmente von der rechteckigen Objektbox überdeckt. Derartig überdeckte Objekte können aus Konsistenzgründen in der Objektausgabe unterdrückt werden. Auf diese Weise erhält das unterstützte Fahrerassistenzsystem eine konsistente Umfeldbeschreibung mit einem Objekt der korrekten Objektgröße und ohne störende Zusatzobjekte.

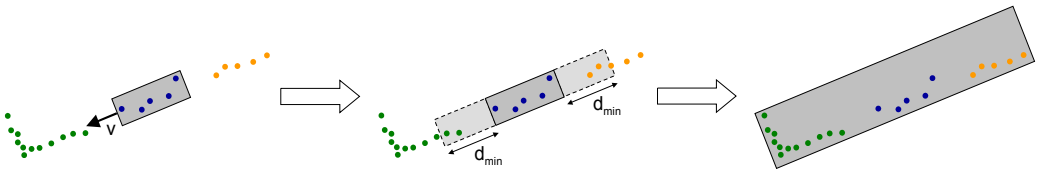

Abbildung 5.11: Segmentfusion mit Hilfe des Sicherheitsabstands: Bei einer verlässlichen Geschwindigkeitsschätzung wird in Bewegungsrichtung vor und hinter dem Objekt ein minimaler Abstand zu anderen Objekten vorausgesetzt. Befinden sich innerhalb dieses Abstands weitere Entfernungsmessungen, muss das zugehörige Segment ein Teil des gleichen Objektes sein.

5.8 Anwendungsspezifische Ausgabe

Über die Ausgabe der verfolgten Objekte an das Assistenzsystem wird nach verschiedenen Kriterien entschieden. Diese Entscheidung kann parallel für verschiedene Assistenzsysteme mit unterschiedlichen Parametern durchgeführt werden.

Als Erstes wird die in [66] vorgestellte Konsistenzprüfung der Objekthistorie berücksichtigt. Aus dem bei der Konsistenzprüfung berechneten Gütemaß wird hier für jedes Objekt mit einer Schwellwertoperation entschieden, ob es ausgegeben werden kann. Der Schwellwert kann anwendungsspezifisch gewählt werden.

Aus Konsistenzgründen werden an dieser Stelle auch Überdeckungen zwischen verschiedenen Objekten überprüft (Abschnitt 5.7). Kleinere Objekte, die annähernd komplett von großen Objekten überdeckt werden, werden hier bei der Ausgabe unterdrückt, da davon ausgegangen wird, dass die entsprechenden Messdaten zum größeren Objekt gehören.

Weiterhin kann über die Ausgabe bestimmter Objektparameter aufgrund der korrespondierenden Unsicherheiten entschieden werden. So ist für verschiedene Anwendungen die Ausgabe einer Objektgeschwindigkeit nur bei relativ sicheren Schätzwerten sinnvoll [66].

Schließlich kann die Objektausgabe auch entsprechend der Anwendungsanforderungen auf eine bestimmte ROI (*engl. ROI = region of interest*) reduziert werden. Diese ROI kann Einschränkungen bezüglich Position, Bewegungsrichtung, Größe und Geschwindigkeit enthalten. So ist für den Fußgängerschutz beispielsweise eine deutlich geringere Reichweite der Fahrzeugumfelderfassung notwendig, als für einen geschwindigkeitsabhängigen Tempomaten, der auch bei hohen Geschwindigkeiten Objekte vor dem Fahrzeug ausreichend frühzeitig erkennen muss. Detaillierte Angaben über die verwendeten Kriterien beinhaltet das Kapitel 10.

5.9 Zusammenfassung

In diesem Kapitel wurden die verwendeten Algorithmen zur laserscannerbasierten Objektverfolgung vorgestellt. Diese Algorithmen basieren auf bestehenden Arbeiten. Es wurden jedoch verschiedene Optimierungen und Erweiterungen zur Reduktion der benötigten Rechenzeit vorgenommen. Darüber hinaus wurden neue Ansätze zur Segmentfusion vorgestellt. Diese Ansätze wirken dem Zerfall von Verkehrsteilnehmern in mehrere Objekte entgegen, der durch die abstandsbasierte Segmentierung entsteht. Die Leistungsfähigkeit der vorgestellten Verbesserungen wird in Kapitel 6 quantitativ untersucht.

6 Laserscannerbasierte Klassifikation

Neben Positions-, Größen- und Bewegungsinformationen ist für verschiedene Assistenzsysteme auch die Unterscheidung bestimmter Objekttypen notwendig. Diese kann zum Beispiel zur Einschätzung der Masse und des Kollisionswiderstands für die Auslösung des Airbags verwendet werden. Weitere Beispiele zur sinnvollen Nutzung von Objekttypeninformationen sind der Fußgängerschutz oder die Identifikation von Fahrzeugen als Zielobjekte für abstandsgeregelte Tempomaten. Da für verschiedene Assistenzsysteme unterschiedliche Objekte von Bedeutung sind, werden die Objekte hier in fünf verschiedene Klassen unterteilt. Die Klassifikation soll Objekte der Klassen „Fußgänger“, „Zweirad“, „PKW“ und „LKW“ von den verbleibenden Objekten der Klasse „Hintergrund“ unterscheiden.

Dieses Kapitel beschreibt das Konzept der hier vorgestellten Klassifikation unter Verwendung von Laserscannermessdaten. Erweiterungen mit zusätzlichen Sensormessdaten werden in den folgenden Kapiteln 7 und 8 beschrieben. In Abbildung 6.1 ist der Ablauf der Klassifikation skizziert.

Als Grundlage der Klassifikation werden verschiedene Merkmale aus den Messdaten des Laserscanners und aus der in der Objektverfolgung geschätzten Objektdynamik extrahiert. Die Merkmalsextraktion erfolgt bei der Sensordatenvorverarbeitung beziehungsweise nach der Objektverfolgung.

Aufgrund der hohen Dimension des Merkmalsraums wurde ein Klassifikator gewählt, der die optimalen Klassifikationsgrenzen aufgrund von Beispieldaten erlernt. In ersten Untersuchungen wurden künstliche neuronale Netze und Support-Vektor-Maschinen im Rahmen einer Diplomarbeit [122] bezüglich ihrer Leistungsfähigkeit zur Klassifikation mit Hilfe der extrahierten Merkmale verglichen. Eine große Herausforderung stellte dabei die unbekannte Datenverteilung dar. Während neuronale Netze beliebige Verteilungen direkt aufgrund der Beispieldaten approximieren können, wird für Support-Vektor-Maschinen ein geeigneter Kernel zur Transformation der Daten in einen Hyperraum benötigt, in dem die Daten linear separierbar sind. Aufgrund der unbekannten Datenverteilung und der vorliegenden Erfahrungen fiel die Wahl auf neuronale Netze als ein Standardverfahren zur Mustererkennung. In Abschnitt 6.2 erfolgt darum eine kurze Einführung in das Themengebiet der künstlichen neuronalen Netze.

Anschließend werden in Abschnitt 6.3 Möglichkeiten diskutiert, wie die vom Netzwerk erlernte Abbildung vom hochdimensionalen Merkmalsraum auf die unterschiedenen Objektklassen nachvollziehbar analysiert werden kann.

Zur Mustererkennung wird ein Klassifikator aus verschiedenen neuronalen Netzen aufgebaut. Dieser wird in Abschnitt 6.4 beschrieben. Als Ergebnis liefert der Klassifikator für jedes Objekt einen Zugehörigkeitswert zwischen 0 und 1 für jede Klasse.

Die Ergebnisse des Klassifikators können mit regelbasiertem Wissen überprüft, korrigiert und in einem vorgegebenen Rahmen gehalten werden. Dies geschieht durch Veränderung der berechneten Zugehörigkeitswerte. Dieser Teil der Klassifikation wird in Abschnitt 6.5 näher erläutert.

Ein zeitliches Mittelwertfilter glättet das Ergebnis der Klassifikation durch Einbeziehung einer festen Anzahl von Klassifikationsergebnissen aus vorherigen Zeitschritten.

In dem Modul „anwendungsspezifische Optimierung“ wird dann letztendlich die für die Ausgabe zugewiesene Klasse gewählt. Diese Auswahl kann applikationsspezifisch durchgeführt werden und wird bei der Objektausgabe für das Assistenzsystem realisiert. Details der Klassenauswahl werden in Abschnitt 6.6 vorgestellt.

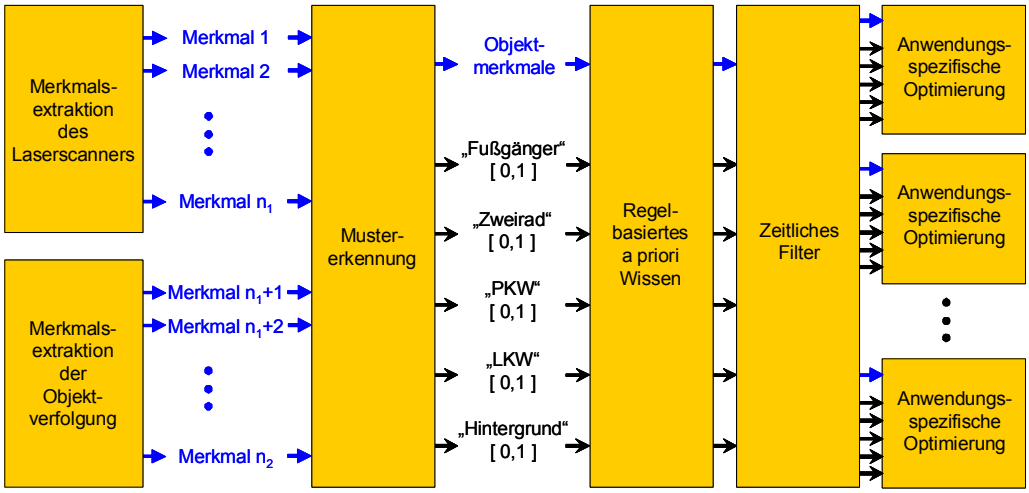

Abbildung 6.1: Ablauf der Klassifikation für jedes Objekt: Die Klassifikation beruht auf einer Vielzahl von Merkmalen, die aus Laserscannermessdaten und der Objektverfolgung extrahiert werden. Die Mustererkennung berechnet Zugehörigkeitswerte für jede Klasse. Diese werden mit regelbasiertem Vorwissen modifiziert. Ein zeitlicher Mittelwertfilter berücksichtigt Klassifikationsergebnisse vorheriger Zeitschritte. Die ausgegebene Klasse wird in der anwendungsspezifischen Optimierung der Ausgabe aus den Zugehörigkeitswerten ermittelt.

6.1 Merkmalsextraktion

Als Grundlage der Klassifikation extrahiert eine erster Schritt aus den Sensormessdaten diverse Merkmale. Dieser Abschnitt enthält eine Beschreibung aller Merkmale, die bei einer ausschließlichen Nutzung des Laserscanners zur Erfassung des Fahrzeugumfelds zur Anwendung kommen. Im Wesentlichen werden hier zwei Gruppen von Merkmalen unterschieden.

Ein Teil der Merkmale basiert direkt auf den Laserscannermessdaten. Bei einer Aufteilung der Algorithmen auf verschiedene Recheneinheiten gemäß Abschnitt 3.2 erfolgt die Berechnung dieser Merkmale nach der Datenassoziation in der Recheneinheit des Laserscanners.

Wenn nur der Laserscanner zur Fahrzeugumfelderfassung verwendet wird, ist eine leistungsfähige Klassifikation nicht aufgrund eines einzelnen Zeitschrittes möglich. Die Messungen eines Baumes oder eines Schildes können beispielsweise einem Fußgänger sehr ähnlich sein. Ebenso sind Hintergrundobjekte mit Messeigenschaften anderer Objektklassen denkbar. Aus diesem Grund wird die Objektbewegung als wichtiges Merkmal zur Klassifikation hinzuge-

nommen. Entsprechende Merkmale werden nach dem Filterupdate am Ende der Objektverfolgung in der zentralen Recheneinheit ermittelt.

Die Klassifikation kann durch zusätzliche Merkmale aus weiteren Sensoren deutlich verbessert werden. In Kapitel 7 werden zusätzliche Merkmale beschrieben, die aus Videomessdaten extrahiert wurden.

Eine komplette Liste aller Merkmale, die in dieser Arbeit verwendet wurden, befindet sich in Anhang A.

6.1.1 Merkmale der Laserscannermessdaten

Als primäre Information liefert der Laserscanner Messdaten über die horizontale Objektkontur. Als Merkmale werden die gefilterte Länge und Breite inklusive der entsprechenden Unsicherheiten verwendet (Abbildung 6.2 links). Zusätzlich werden auch die korrespondierenden zeitlichen Maximalwerte aus allen Messungen verwendet. Da der Laserscanner Messdaten in mehreren Ebenen liefert, können auch Informationen über die Objekthöhe gewonnen werden. Das vermessene Objekt muss mindestens die Höhe zwischen oberster und unterster Entfernungsmessung aufweisen (Abbildung 6.2 rechts). Als Merkmal wird der zeitliche Maximalwert dieser Differenz für die Klassifikation bestimmt.

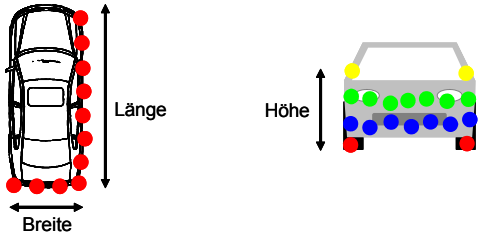

Abbildung 6.2: Objektgröße: Länge und Breite des Fahrzeugs werden gemessen und gefiltert. Der Abstand zwischen oberster und unterster Entfernungsmessung des Laserscanners gibt Informationen über die minimale Objekthöhe.

Weitere Merkmale werden aus der Form berechnet, die aus den einzelnen Entfernungsmessungen bestimmt wurde. Die entsprechenden Merkmale sollen speziell PKWs besser charakterisieren. Diese weisen in der Regel eine ausgeprägt rechteckige Grundform auf, während andere Objekte meist stärker von diesem Modell abweichen. Als erstes Merkmal geht die Anzahl der vermessenen Objektseiten ein. Dies sind zwei bei der L-Form, eine bei der I-Form und keine bei der O-Form. Zusätzlich wird der ermittelte Winkel zwischen den Seiten von L-Formen als Merkmal verwendet.

Mit einem weiteren Merkmal sollen unregelmäßige Objekte wie Büsche oder Messungen auf Böschungen am Straßenrand als Hintergrund aussortiert werden. Dieses Merkmal ist der mittlere Abstand zweier bezüglich des Winkels benachbarter Entfernungsmessungen im Segment. In Abbildung 6.3 sind beispielhafte Messungen auf verschiedene Objekte dargestellt.

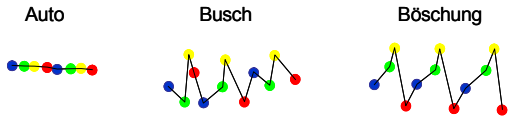

Abbildung 6.3: Mittlerer Abstand zweier benachbarter Entfernungsmessungen: Bei Büschen treten starke Variation der Entfernungsmessungen auf. Bei Böschungen variieren die Entfernungsmessungen unterschiedlicher Messebenen. Aus diesem Grund können beide Objekttypen aufgrund des deutlich höheren mittleren Abstands zwischen benachbarten Entfernungsmessungen von Objekten mit glatten, vertikalen Oberflächen getrennt werden.

An glatten Objekten mit vertikalen Seitenflächen liegen benachbarte Entfernungsmessungen sehr dicht beieinander. Dies gilt sowohl für PKWs als auch für Häuser oder Leitplanken. Messungen von Büschen oder Böschungen weichen jedoch charakteristisch von dieser Anordnung ab. Bei Büschen sind die Entfernungsmessungen stark verteilt, da viele der Messungen auch von Blättern aus dem Inneren stammen. Die Entfernungsmessungen von Böschungen verhalten sich etwas anders. Messungen der gleichen Messebene liegen nach wie vor dicht beieinander. Da unterschiedliche Ebenen die Böschung jedoch in unterschiedlichen Höhen vermessen, variiert die Distanz von einer Ebene zur nächsten. In der Praxis stammen benachbarte Entfernungsmessungen aus unterschiedlichen Messebenen, wenn die Messdaten nach dem Winkel sortiert betrachtet werden. Darum ist auch hier der mittlere Abstand zwischen zwei benachbarten Entfernungsmessungen deutlich höher als bei glatten, vertikalen Oberflächen.

Eine weitere Eigenschaft des Laserscanners ist die Möglichkeit, Reflektoren zu messen. Bei Fahrzeugen sind in der Regel in den Nummernschildern und in den Rückleuchten Reflektoren integriert. Abbildung 6.4 zeigt die erwarteten Positionen der Reflektoren bei Fahrzeugen. Als Merkmale für Reflektoren werden für jeden gemessenen Reflektor die Breite und der Abstand zur erwarteten Position berechnet.

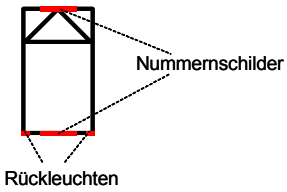

Abbildung 6.4: Detektierbare Reflektoren an Fahrzeugen: Die Nummernschilder und die Rückleuchten von Fahrzeugen enthalten Reflektoren, die mit dem Laserscanner detektiert werden können.

Eine letzte Gruppe von Merkmalen beschäftigt sich mit der Frage, inwieweit die Länge und die Breite des Objektes messbar ist. Im Wesentlichen existieren drei Fälle, in denen die Objektkontur nicht bis zum Ende vermessen wird. Die offensichtlichen Fälle sind die Verdeckung durch ein anderes Objekt oder der Austritt des Objekts aus dem Messbereich des Laserscanners.

Es ist aber auch möglich, dass aufgrund der Reflektionscharakteristik nicht genügend Energie zum Sensor reflektiert wird, obwohl der Raum zwischen Objekt und Sensor frei von anderen Objekten ist. Dies ist der Fall, wenn der Laserpuls das Objekt in einem sehr flachen Winkel trifft. Dann ist der in Richtung des Sensors reflektierte Anteil sehr gering. Auch stark streuende Oberflächen können einen ähnlichen Effekt auslösen.

Mit zwei Merkmalen wird angegeben, ob das Objekt am Rand des Winkelbereiches liegt, in dem der Laserscanner Messdaten liefert. Darüber hinaus werden die Winkel berechnet, in denen die Laserpulse auf die Seite und das Heck beziehungsweise die Front des Objektes treffen (Abbildung 6.5).

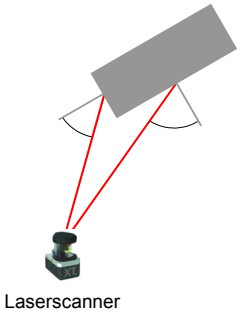

Abbildung 6.5: Bei flachem Einfallswinkel kann eine Objektseite häufig nicht auf der kompletten Länge vermessen werden. Darum wird der Einfallswinkel auf die vermessenen Objektseiten als Merkmal benutzt.

Weiterhin werden Verdeckungen und fehlende Messwerte ausgewertet, um die maximale Objektgröße zu bestimmen. So können zum Beispiel LKWs und PKWs ausgeschlossen werden, falls die maximal mögliche Objektgröße unter den dafür üblichen Größen liegt.

Abbildung 6.6 zeigt links eine Beispielszene, in der das dunkelgraue Objekt teilweise verdeckt wird. Dadurch kann die Objektlänge nicht gemessen werden. Im mittleren und im rechten Teil der Abbildung werden zwei Methoden zur Bestimmung der maximalen Objektgröße illustriert. Als Grundlage dient die gemessene Objektform und die dabei bestimmten Seiten des Objektes. Diese Seiten werden so weit verlängert, bis sie einen ermittelten Grenzwinkel schneiden. Der Grenzwinkel wird durch Auswertung des radialen Messprinzips des Laserscanners bestimmt. Hierbei werden nur Messebenen berücksichtigt, die Messungen auf dem Objekt enthalten.

In einem ersten Schritt wird die mögliche Größe bestimmt, die durch andere Objekte verdeckt wird (Abbildung 6.6, Mitte). Hier werden die Objektseiten so weit vergrößert, bis sie einen Messwinkel schneiden, in dem keine Verdeckung mehr vorliegt. Es wird davon ausgegangen, dass das Objekt in diesem Messwinkel eine Entfernungsmessung erzeugen würde, falls es sich bis in diesen Winkelbereich ausdehnt. Wenn keine Messung vorliegt oder hinter dem betrachteten Objekt ein anderes Objekt gemessen wurde, wird der Messwinkel als Begrenzung der verdeckten Größe angenommen.

In der Praxis erzeugen Objekte jedoch nicht immer eine Entfernungsmessung, wenn ein Laserpuls auf sie trifft. Wie bereits erläutert wurde, wird unter Umständen nur ein sehr geringer Anteil der Energie zurück zum Sensor reflektiert, der dann nicht für eine Entfernungsmessung ausreicht. Berücksichtigt man diesen Effekt, dann ist die maximal mögliche Objektgröße erst durch Messungen auf Objekte beschränkt, die sich hinter dem betrachteten Objekt befinden. Diese maximal mögliche Objektgröße wird ebenfalls in Form von Merkmalen für Länge und Breite zur Klassifikation verwendet (Abbildung 6.6, rechts).

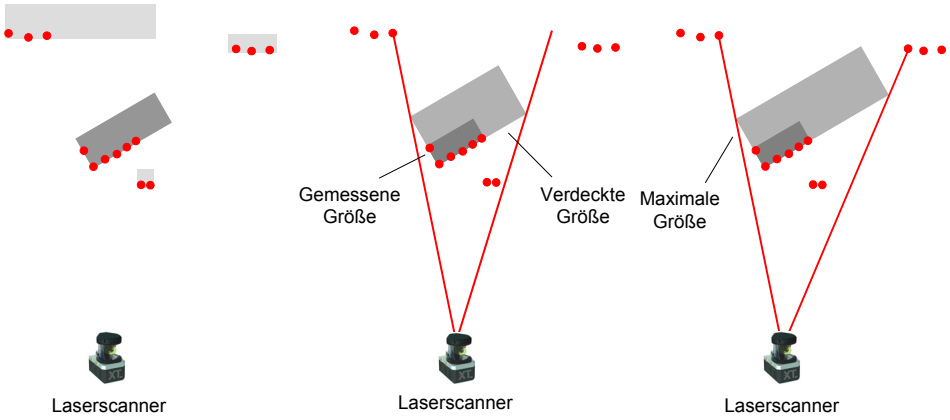

Abbildung 6.6: Bestimmung der maximalen Objektgröße: In einigen Fällen wird die Objektgröße nicht komplett vermessen (links). Vergrößert man das Objekt um den Bereich, der durch andere Messungen verdeckt ist, erhält man die verdeckte Objektgröße (Mitte). Die maximal mögliche Objektgröße ist durch Messungen hinter dem betrachteten Objekt eingeschränkt (rechts).

6.1.2 Merkmale der Objektverfolgung

Zur Klassifikation bewegter Objekte werden aus der Objektverfolgung signifikante Merkmale gewonnen. Die Objektgeschwindigkeit wird redundant in Merkmalen für momentane, mittlere, Median- und Maximalgeschwindigkeit verwendet.

Um Fahrzeuge signifikant von Fußgängern und stehenden Objekten trennen zu können, wird zusätzlich ein Zähler eingeführt, der die Anzahl der Zyklen zählt, in denen sich das Objekt schneller als $4\frac{m}{s}$ bewegt hat.

Die Unsicherheit der Geschwindigkeitsschätzung wird in einem Merkmal berücksichtigt, das basierend auf Arbeiten aus [66] ein Maß für signifikante Bewegungen liefert. Hierzu wird von der Hypothese ausgegangen, dass das Objekt stationär ist. Dementsprechend ist der Schätzfehler gleich der geschätzten Geschwindigkeit. Die Auswertung basiert auf dem mit der Kovarianz normalisierten Schätzfehlerquadrat (*engl. NEES = normalized estimation error squared*):

$$NEES_{v,stationär} = \begin{bmatrix} v_x & v_y \end{bmatrix} \begin{bmatrix} P_{v_x v_x} & P_{v_x v_y} \\ P_{v_x v_y} & P_{v_y v_y} \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix}. \tag{6.1}$$

Dieser Wert ist für stationäre Objekte bei einer konsistenten Filterung χ^2-verteilt mit 2 Freiheitsgraden (Abbildung 6.7, links). Die Merkmalsberechnung bestimmt nun aus dem aktuellen *NEES* den Funktionswert der kumulativen Verteilungsfunktion $F(NEES(k))$. Dieser Merkmalswert gibt an, welcher statistische Anteil der stationären Objekte in einem niedrigeren *NEES* resultiert (Abbildung 6.7, rechts). Ein hoher Merkmalswert spricht also eher für bewegte Objekte, niedrige Merkmalswerte sind ein Anzeichen für stationäre Objekte.

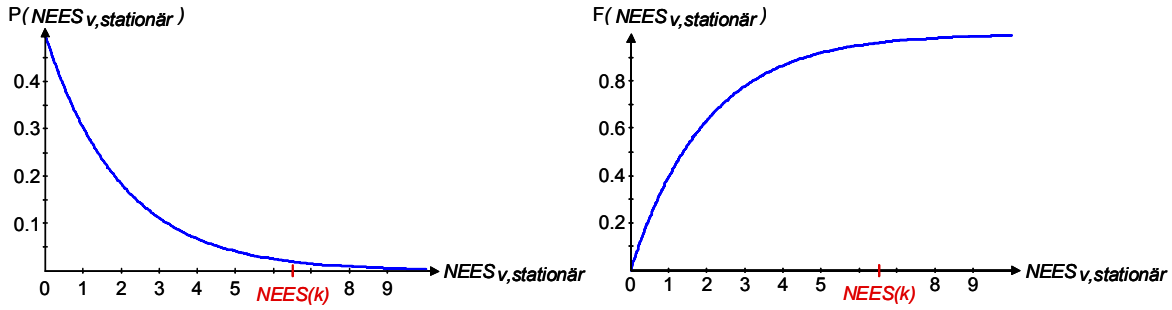

Abbildung 6.7: Der *NEES* der Geschwindigkeit stationärer Objekte ist χ^2-verteilt mit 2 Freiheitsgraden. Die Abbildung zeigt die links die Wahrscheinlichkeitsdichtefunktion und rechts die kumulative Verteilungsfunktion.

Ein weiteres Merkmal zur Klassifikation basiert auf der Konsistenzprüfung der Objektverfolgung nach [8] und [66]. Es beinhaltet die bei der Prüfung berechnete Bewertung der Objekthistorie.

Um Positionsrauschen erkennen zu können, wird weiterhin der zurückgelegte Weg von der ersten gemessenen Objektposition zur aktuell gemessenen Position des Objektes als Merkmal berechnet. In Abbildung 6.8 ist der Unterschied zwischen der zeitlichen Objekttrajektorie und dem direkten Weg illustriert.

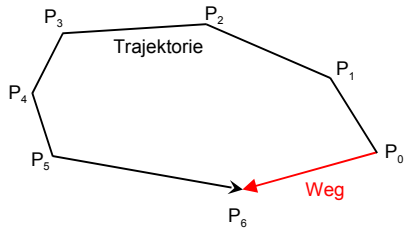

Abbildung 6.8: Zurückgelegter Weg: Der Abstand zwischen erster und aktueller Objektposition wird als Merkmal berechnet. Der Einfluss des zeitlichen Positionsrauschens auf diesen Weg ist deutlich geringer als auf die zurückgelegte Trajektorie.

Da der Laserscanner nur Positionen relativ zum Versuchsfahrzeug misst, werden die Geschwindigkeit und die Gierrate des Versuchsfahrzeugs in die Objektverfolgung integriert, um absolute Geschwindigkeiten messen zu können. Aus diesem Grund wirken sich Rauschprozesse bei der Bewegungsbestimmung des Versuchsfahrzeugs direkt auf die ermittelten

absoluten Objektgeschwindigkeiten aus. Die Stärke des Rauschens ist von der Bewegung des Versuchsfahrzeugs abhängig. So ist das Rauschen im Stand annähernd Null. Es wird stärker bei gerader Fahrt und ist am stärksten bei Kurvenfahrten. Damit der Klassifikator den Bewegungszustand des Versuchsträgers berücksichtigen kann, bilden dessen Gierrate und Geschwindigkeit weitere Merkmale. Zusätzlich wird noch die Objektgeschwindigkeit relativ zum Versuchsträger für die Klassifikation benutzt.

6.2 Künstliche neuronale Netze

Anhand der berechneten Merkmale soll nun mit Methoden der Mustererkennung die Objektklasse bestimmt werden. Künstliche neuronale Netze werden in vielen Bereichen zur Mustererkennung und Funktionsapproximation eingesetzt. Sie basieren auf kleinen Elementen, den Neuronen. Im neuronalen Netz werden viele Neuronen strukturiert miteinander verknüpft. Durch die dabei verschalteten Nichtlinearitäten lassen sich eine Vielzahl von Verteilungen mit dem Netz approximieren.

In den folgenden Abschnitten wird als Erstes das Neuron als Basiseinheit neuronaler Netze beschrieben. Danach werden verschiedene Netzstrukturen vorgestellt. Schließlich wird die Vorgehensweise zum Erlernen der zu approximierenden Verteilung kurz erläutert.

6.2.1 Neuronen

Neuronale Netze basieren auf kleinen, meist nichtlinearen Elementen, den Neuronen. Diese berechnen aus einer Reihe von Eingangswerten x_i eine Ausgabe y. Der prinzipielle Aufbau eines Neurons ist in Abbildung 6.9 dargestellt.

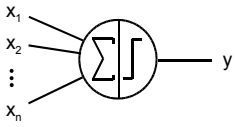

Abbildung 6.9: Aufbau eines Neurons: Die Eingänge werden durch einen gewichteten Mittelwert kombiniert. Aus dieser Aktivierung wird mit einer meist nichtlinearen Ausgabefunktion die Ausgabe berechnet.

Für die Berechnung der Ausgabe wird zunächst aus den Eingangswerten eine Aktivierung z als gewichteter Mittelwert berechnet, zu dem ein als Bias bezeichneter Offset addiert wird:

$$z = \sum_{j=1}^{n} w_j x_j + w_0. \tag{6.2}$$

Aus dieser Aktivierung wird mit der meist nichtlinearen Ausgabefunktion die Ausgabe des Neurons berechnet:

$$y = f(z). \tag{6.3}$$

Es existiert eine Vielzahl von möglichen Ausgabefunktionen. Abbildung 6.10 zeigt einige typische, häufig verwendete Ausgabefunktionen.

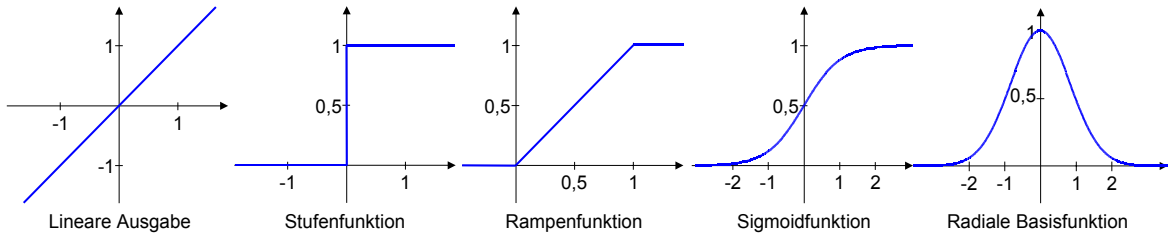

Abbildung 6.10: Häufig verwendete Ausgabefunktionen.

Während die Wahl der Ausgabefunktion ein Designparameter beim Entwurf neuronaler Netzwerke darstellt, werden die Gewichte zu den Eingangswerten und der Bias jedes Neurons beim Training eines Netzwerks automatisch adaptiert. Dieser Trainingsprozess bestimmt aufgrund vorgegebener Beispieldaten geeignete Werte zur Abbildung der zu approximierenden Datenverteilung.

6.2.2 Strukturen neuronaler Netze

Zum Training neuronaler Netze existieren zwei prinzipiell verschiedene Methoden. Liegen für die Beispieldaten keine Klasseninformationen vor, können sich Netze in ihrer Struktur lediglich an die zu approximierende Verteilung anpassen. In diesem Fall spricht man von unüberwachtem Lernen (*engl. unsupervised learning*). Beispiele für entsprechende Netze sind Kohonen-Netze sowie Strukturen wie Growing Neural Gas und Growing Cell Structures [42, 58].

Im hier vorliegenden Anwendungsfall der Objektklassifikation sollen bestimmten Bereichen im Merkmalsraum vorgegebene Klassen zugeordnet werden. Hierzu müssen Beispieldaten mit Klasseninformationen verwendet werden. In diesem Fall sollte das überwachte Lernen (*engl. supervised learning*) angewendet werden, das die lernbaren Netzwerkparameter basierend auf dem Klassifikationsfehler adaptiert. Hierfür existieren verschiedene Netzstrukturen [23].

Mehrlagige Perzeptrons (*engl. multi-layer perceptron*) sind eine relativ einfache, aber äußerst leistungsfähige Netzwerkstruktur. In dieser Struktur werden die Neuronen schichtenweise angeordnet. Die Eingänge des Netzes werden mit den Eingängen aller Neuronen der ersten Schicht verbunden. Die Ausgänge aller Neuronen einer Schicht werden mit den Eingängen aller Neuronen der Folgeschicht verbunden. Die Ausgänge der als Ausgabeschicht bezeichneten, letzten Schicht bilden die Ausgabe des Netzes (Abbildung 6.11).

Aufgrund der Struktur des mehrlagigen Perzeptrons, werden mehrere nichtlineare Ausgabefunktionen teilweise parallel und teilweise nacheinander durchlaufen. Dadurch erlangt das Perzeptron eine enorme Abbildungsfähigkeit.

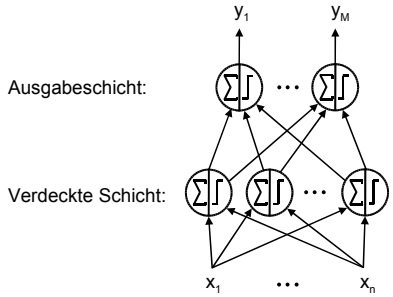

Abbildung 6.11: Mehrlagiges Perzeptron: Schichtenweise angeordnete Neuronen geben ihre Ausgabe jeweils an die folgende Schicht weiter. Durch diese Kombination der Nichtlinearitäten entsteht eine extrem leistungsfähige Netzstruktur.

Neben mehrlagigen Perzeptrons existiert noch eine Vielzahl weiterer, meist komplexerer Netzstrukturen [23]. Eine erste Gruppe von Netzen ist auf die Verarbeitung zeitlicher Sequenzen von Eingangssignalen spezialisiert. So führen partiell rekurrente Netze die Ausgabe einer bestimmten Netzschicht auf eine Schicht von Kontextneuronen zurück. Dabei werden die Ausgaben der Schicht von den Kontextneuronen nur kopiert und für den nächsten Zeitschritt gespeichert. Die korrespondierenden Gewichte sind also mit Eins festgelegt. Die Kontextneuronen geben nun im nächsten Zeitschritt die gespeicherte Ausgabe als Eingabe an eine frühere Ebene des Netzwerkes zurück. Beispiele für partiell rekurrente Netze sind Elman-Netze [25]. Bei diesen Netzen wird die Ausgabe der ersten verdeckten Schicht in den Kontextneuronen gespeichert und der gleichen Schicht im nächsten Zeitschritt als zusätzliche Eingabe präsentiert (Abbildung 6.12).

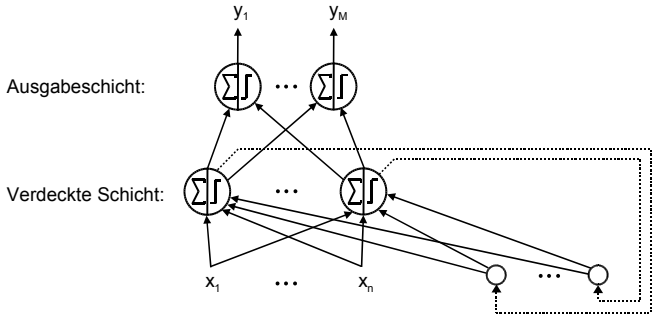

Abbildung 6.12: Elman-Netze verarbeiten zeitlich veränderliche Signale. Hierfür wird die Ausgabe der verdeckten Schicht in Kontextneuronen gespeichert und dieser Schicht zusammen mit der nächsten Eingabe wieder präsentiert.

Eine weitere Methode zur Verarbeitung zeitlich veränderlicher Signale ist die direkte Nutzung von Signalen aus vorherigen Zeitschritten. Hierbei werden dem Netz zusätzlich zu den aktuellen Signalen auch die Signale einer festgelegten Anzahl vorheriger Zeitschritte präsentiert. Dieser Ansatz wird in Time Delay Neural Networks verfolgt [159].

Weitere Netzstrukturen entstehen durch Ansätze zur Bestimmung einer optimalen Topologie. Während bei mehrlagigen Perzeptrons die Anzahl und Verteilung der Neuronen beim Entwurf vom Designer festgelegt werden muss, bestimmen diese Ansätze

selbstständig eine optimale Neuronenanzahl und -verteilung. Für die Bestimmung der Topologie existieren sowohl topologieminimierende Verfahren wie das Optimal Brain Damage als auch inkrementelle Strukturen wie Cascade Correlation Netze [23].

In dieser Arbeit wurde aus Gründen der Übersichtlichkeit und Analysierbarkeit des Klassifikators auf komplexere Strukturen verzichtet. Die verwendeten Klassifikatoren werden aus mehrlagigen Perzeptrons aufgebaut.

6.2.3 Lernverfahren

Das Lernen der Gewichte in mehrlagigen Perzeptrons basiert prinzipiell auf der Minimierung des quadratischen Fehlers der Ausgabe. In einem iterativen Prozess wird jeweils der Ausgabefehler E aus der aktuellen Ausgabe $\mathbf{y}$ und der vorgegebenen Ausgabe $\mathbf{t}$ berechnet:

$$E = \frac{1}{2}(\mathbf{t}-\mathbf{y})^T(\mathbf{t}-\mathbf{y}) = \frac{1}{2}\sum_i (t_i - y_i)^2 . \tag{6.4}$$

Die Gewichte der letzten Schicht werden durch ein Gradientenabstiegsverfahren jeweils in Richtung des niedrigeren Fehlers angepasst. Durch eine Propagierung des Fehlers in vorherige Schichten lassen sich auch deren Gewichte anpassen (Backpropagation). Das Prinzip der iterativen Parameteranpassung in Richtung des fallenden Fehlerverlaufs wird in Abbildung 6.13 dargestellt.

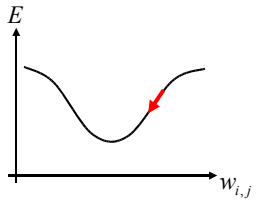

Abbildung 6.13: Die Gewichte des neuronalen Netzes werden iterativ jeweils durch ein Gradientenabstiegsverfahren in Richtung des niedrigeren Fehlers adaptiert.

Das Gewicht für den Eingang j des Neurons i der Schicht L wird mit $w^L_{i,j}$ bezeichnet. Die Anpassung des Gewichtes in jedem Trainingszyklus erfolgt nach folgender Vorschrift:

$$w^L_{i,j}(k+1) = w^L_{i,j}(k) + \Delta w^L_{i,j}(k) . \tag{6.5}$$

Das Update eines Gewichts $\Delta w^L_{i,j}$ wird aus zwei Komponenten berechnet. Die Ableitung des Fehlers E nach dem Gewicht $w^L_{i,j}$ gibt die Richtung der Anpassung vor, während die Lernrate η die Schrittweite der Anpassung beschreibt:

$$\Delta w^L_{i,j}(k) = -\eta\frac{\partial E}{\partial w^L_{i,j}} = -\eta\frac{\partial E}{\partial z^L_i}\frac{\partial z^L_i}{\partial w^L_{i,j}} = -\eta\frac{\partial E}{\partial z^L_i}x_j . \tag{6.6}$$

Für Neuronen der Ausgabeschicht lässt sich der Fehler E direkt nach der Aktivierung z_i^L ableiten:

$$\frac{\partial E}{\partial z_i^L} = \frac{\partial E}{\partial y_i^L}\frac{\partial y_i^L}{\partial z_i^L} = -\left(t_i - y_i^L\right) f\left(z_i^L\right)' . \tag{6.7}$$

Für ein Neuron einer verdeckten Schicht L wird der Fehler der Neuronen der folgenden Schicht $L+1$ zurückpropagiert (Abbildung 6.14). Für die Propagierung wird berücksichtigt, dass die Ausgabe des Neurons ein Eingang für jedes Folgeneuron ist.

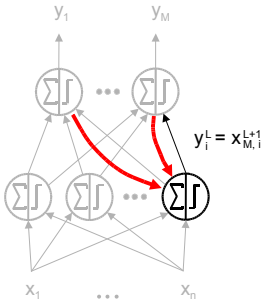

Abbildung 6.14: Backpropagation: Der Ausgang eines verdeckten Neurons ist gleichzeitig der Eingang für Neuronen der folgenden Schicht. Somit kann der Fehler dieser Neuronen auf das verdeckte Neuron zurückpropagiert werden.

Somit ist der Fehler eines verdeckten Neurons eine Funktion der Aktivierungen der $M\,(L+1)$ Neuronen in der folgenden Schicht $L+1$:

$$E = E\left(z_1^{L+1}, z_2^{L+1}, z_3^{L+1}, ..., z_{M(L+1)}^{L+1}\right). \tag{6.8}$$

Diese Fehlerfunktion muss zur Bildung des Fehlergradienten nach der aktuellen Aktivierung abgeleitet werden:

$$\frac{\partial E}{\partial z_i^L} = \sum_{m=1}^{M(L+1)} \frac{\partial E}{\partial z_m^{L+1}}\frac{\partial z_m^{L+1}}{\partial y_i^L}\frac{\partial y_i^L}{\partial z_i^L} = \sum_{m=1}^{M(L+1)} \frac{\partial E}{\partial z_m^{L+1}} w_{m,i}^{L+1} f\left(z_i^L\right)' . \tag{6.9}$$

Der Fehlergradient eines Neurons einer verdeckten Schicht lässt sich also aus den Gradienten der Neuronen der darüber liegenden Schichten ermitteln. Auf diese Weise können alle Fehlergradienten und somit alle Gewichtsadaptionen der Reihe nach berechnet werden, indem bei der Ausgabeschicht beginnend die Berechnung schichtenweise durchgeführt wird. Das Grundprinzip dieses Backpropagation-Algorithmus wird auch in vielen daraus abgeleiteten, verbesserten Trainingsverfahren beibehalten.

Der Backpropagation-Algorithmus beschreibt die Gewichtsanpassung für ein Paar von Beispieldaten und Ausgabe des Netzes. In der Praxis müssen dem Netz jedoch eine Vielzahl von solchen Datenpaaren präsentiert werden. Für die Berechnung der korrespondierenden Gewichtsanpassung existieren zwei prinzipielle Möglichkeiten. Die Gewichte der Neuronen können direkt nach jeder Präsentation eines Datenpaares angepasst (direktes Lernen) oder die Änderungen können über alle Daten akkumuliert und nach der Präsentation aller Datenpaare gemeinsam durchgeführt werden (*engl. batch learning*). Während der Erfolg des

direkten Lernens stark von der Reihenfolge der präsentierten Datenpaare abhängt und ein Oszillieren der Gewichte auftreten kann, werden diese Probleme beim zweiten Ansatz automatisch unterdrückt.

Aufgrund von Untersuchungen zu verfügbaren Trainingsverfahren als Teil einer Diplomarbeit [122] wurde das Verfahren skalierter konjugierter Gradienten [105] zum Training ausgewählt. Hierfür werden alle Gewichte eines Netzes im Gewichtsvektor $\mathbf{w}$ zusammengefasst. Das normale Gradientenabstiegsverfahren würde die Änderung des Fehlers nur mit einer Taylor-Reihe erster Ordnung approximieren:

$$E(\mathbf{w} + \Delta\mathbf{w}) = E(\mathbf{w}) + \mathbf{E}'(\mathbf{w})^T \Delta\mathbf{w}. \tag{6.10}$$

Hier bezeichnet $\mathbf{E}'(\mathbf{w})$ den Gradienten der Fehlerfunktion $E(\mathbf{w})$. Konjugierte Gradientenverfahren verwenden hier eine Approximation zweiter Ordnung:

$$E(\mathbf{w} + \Delta\mathbf{w}) = E(\mathbf{w}) + \mathbf{E}'(\mathbf{w})^T \Delta\mathbf{w} + \frac{1}{2}\Delta\mathbf{w}^T \mathbf{E}''(\mathbf{w}) \Delta\mathbf{w}. \tag{6.11}$$

Die Matrix $\mathbf{E}''(\mathbf{w})$ ist die Hesse-Matrix der Funktion $E(\mathbf{w})$. Zur Ermittlung einer optimalen Gewichtsänderung muss nun entsprechend 6.11 minimiert werden. Dies geschieht weiterhin iterativ durch die Auswahl einer Richtung $\mathbf{r}(k)$ der Gewichtsänderung und die Wahl einer geeigneten Schrittweite $\alpha(k)$ zur Minimierung entlang dieser Richtung:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \Delta\mathbf{w}(k) = \mathbf{w}(k) + \alpha(k) \cdot \mathbf{r}(k). \tag{6.12}$$

Als erste Richtung wird dabei wie gewohnt der negative Gradient verwendet. Bei der nächsten Iteration wird dann eine Richtung $\mathbf{r}(k+1)$ so gewählt, dass die erreichte Optimierung nicht wieder rückgängig gemacht wird. Die Richtungen werden dann als konjugiert bezeichnet. Eine neue Richtung kann dabei mit der Polak-Ribiere-Formel berechnet werden [23]:

$$\mathbf{r}(k+1) = -\mathbf{E}'(\mathbf{w}(k+1)) + \frac{\mathbf{E}'(\mathbf{w}(k+1))^2 + \mathbf{E}'(\mathbf{w}(k+1))\,\mathbf{E}'(\mathbf{w}(k))}{\mathbf{r}(k)^T\,\mathbf{E}'(\mathbf{w}(k))}\mathbf{r}(k). \tag{6.13}$$

Die optimale Schrittweite $\alpha(k)$ kann direkt berechnet werden [105]:

$$\alpha(k) = -\frac{\mathbf{r}(k)^T\,\mathbf{E}'(\mathbf{w}(k))}{\mathbf{r}(k)^T\,\mathbf{E}''(\mathbf{w}(k))\,\mathbf{r}(k)}. \tag{6.14}$$

Durch die Verwendung konjugierter Gradienten konvergiert das Optimierungsverfahren deutlich schneller als ein einfaches Gradientenabstiegsverfahren [23].

6.2.4 Mehrfachinitialisierung

Bei den verwendeten Lernverfahren kann in der Praxis nicht garantiert werden, dass die Fehlerminimierung im globalen Minimum konvergiert. Auch die Konvergenz in einem lokalen Minimum ist möglich. Das Ergebnis hängt stark von der Initialisierung der Gewichte ab, die zufällig erfolgt. Um lokale Minima zu vermeiden wird der Trainingsprozess immer mehrfach mit unterschiedlichen Initialisierungen aufgerufen. Als Ergebnis wird dann das Netzwerk verwendet, das den geringsten Fehler für die Validierungsdaten erreicht.

6.2.5 Automatische Code-Generierung

Im Rahmen dieser Arbeit wurde die „MATLAB Neural Network Toolbox“ für das Training der neuronalen Netze verwendet. Hier können auch Validierungsdaten eingebracht werden, um nach der in Abschnitt 2.6 beschriebenen Methode optimale Generalisierungseigenschaften zu erreichen.

Zur Verwendung der Mustererkennung war eine Konvertierung in die Programmiersprache C notwendig. Um die Rechenzeit der Mustererkennung zu beschleunigen, wurde ein MATLAB-Skript entwickelt, das für ein trainiertes Netzwerk automatisch entsprechenden, möglichst effizienten C-Code erzeugt. Hierbei werden die ermittelten Gewichte hart im Quelltext codiert. Die als Ausgabefunktion verwendeten Rampenfunktionen wurden ebenfalls direkt durch Fallunterscheidungen codiert. Auf diese Weise wird die aufwändigere Ausführung einer variablen Netzwerkstruktur inklusive der für die Verwendung der ermittelten Gewichte notwendigen Lookups eingespart.

Ein entsprechendes Skript wurde auch für das in dieser Arbeit entwickelte, komplexere Mustererkennungssystem erstellt, das in Abschnitt 6.4 vorgestellt wird.

6.3 Analyse neuronaler Netze

Im Automobilbereich gilt oft der Anspruch an Algorithmen, dass ein gewisses Verhalten garantiert werden muss. Dies ist insbesondere der Fall, wenn sicherheitsrelevante Algorithmen ausgeführt werden. Im Fall der Objektklassifikation wäre es zum Beispiel sinnvoll, Objekte ab einer gewissen horizontalen Ausdehnung nicht mehr als Zweirad zu klassifizieren. Die Realisierung einer solchen Vorgabe ist mit statistisch optimierten neuronalen Netzen nicht direkt möglich, kann jedoch mit einem zusätzlichen, regelbasierten Klassifikationsteil erreicht werden. Dieser Ansatz wird in Abschnitt 6.5 beschrieben.

In diesem Abschnitt werden Methoden zur Analyse und Extraktion der vom neuronalen Netz erlernten Abbildung untersucht.

6.3.1 Blackbox-Charakter neuronaler Netze

Prinzipiell erweist sich der Versuch aus einem trainierten neuronalen Netz die erlernte Abbildung exakt zu extrahieren als sehr schwierig. Dies ist durch die Verkopplung der Nichtlinearitäten in den vielen einzelnen Neuronen begründet. Es kann lediglich die Ausgabe

des Netzes für einzelne, präsentierte Eingabevektoren bestimmt werden. Dieses Vorgehen lässt jedoch keine Rückschlüsse auf die Ausgabe des Netzes in benachbarten Bereichen des Merkmalsraumes zu (Abbildung 6.15).

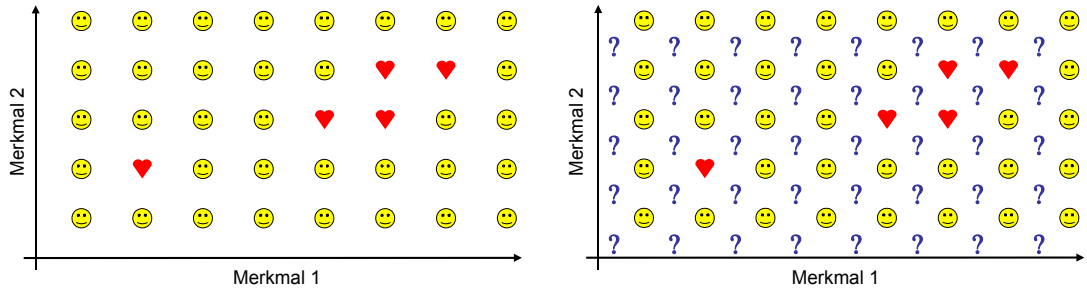

Abbildung 6.15: Blackbox-Charakter: Von trainierten neuronalen Netzen lässt sich stichprobenartig die Ausgabe für bestimmte Merkmalskombinationen bestimmen (links). Diese Ausgaben geben jedoch keinerlei Möglichkeiten für Rückschlüsse auf Ausgaben bei benachbarten Merkmalskombinationen (rechts).

Für die Analyse der erlernten Abbildung wäre die Extraktion einer Partition des Merkmalsraums deutlich hilfreicher. Unter einer Partition versteht man die Unterteilung des Merkmalsraums in nicht überlappende Unterräume. Die Ausgabe des Klassifikators sollte dann für jede Merkmalskombination des Unterraums die gleiche sein (Abbildung 6.16).

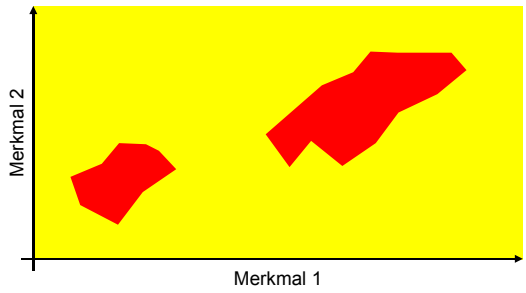

Abbildung 6.16: Partition des Merkmalsraumes: Für die Analyse der erlernten Abbildung wäre eine Unterteilung in nicht überlappende Unterräume mit jeweils konstanter Klassenausgabe hilfreich.

In der Literatur existieren verschiedene Ansätze zur Bestimmung nachvollziehbarer Regeln aus trainierten neuronalen Netzen. In [136] werden solche Regeln aus beliebigen Netzen extrahiert. Diese Regeln beschreiben jedoch nur einen Teil des Merkmalsraumes. Andere Arbeiten [95, 70] verwenden eine spezielle Netzwerkstruktur, die die Extraktion verständlicher Regeln erlaubt.

Im folgenden Abschnitt wird eine neue Methode zur Analyse normaler mehrlagiger Perzeptrons vorgestellt. Hierfür wird als Einschränkung lediglich gefordert, dass als Ausgabefunktionen nur stückweise lineare Funktionen verwendet werden.

6.3.2 Stückweise lineare Netze

In stückweise linearen Netzen sei die Ausgabefunktion auf die ersten drei Funktionen aus Abbildung 6.10 beschränkt. Die Analyse wird ohne Beschränkung der Allgemeinheit an dieser Stelle nur mit der Rampenfunktion durchgeführt.

Für ein Neuron der ersten verdeckten Schicht wird die Ausgabe dann folgendermaßen aus den Merkmalen $x_1, ..., x_n$ berechnet:

$$y = \begin{cases} 0, \text{ falls } w_1x_1 + w_2x_2 + ... + w_nx_n + w_0 < 0 \\ 1, \text{ falls } w_1x_1 + w_2x_2 + ... + w_nx_n + w_0 > 1 \\ w_1x_1 + w_2x_2 + ... + w_nx_n + w_0, \text{ sonst} \end{cases} . \tag{6.15}$$

Das bedeutet, alle Neuronen der ersten Schicht eines mehrlagigen Perzeptrons mit stückweise linearen Ausgabefunktionen unterteilen den Merkmalsraum jeweils in drei Teile. Die Ausgabe hat allgemein eine lineare Form mit bekannten Koeffizienten $v_0, ..., v_n$:

$$y = v_1x_1 + v_2x_2 + ... + v_nx_n + v_0. \tag{6.16}$$

Die einzelnen Teile des Merkmalsraumes sind durch endlich viele lineare Bedingungen mit bekannten Koeffizienten $u_{1,0}, .., u_{j,n}$ begrenzt:

$$\begin{aligned} u_{1,1}x_1 + u_{1,2}x_2 + ... + u_{1,n}x_n + u_{1,0} &< 0 \\ &\vdots \\ u_{j,1}x_1 + u_{j,2}x_2 + ... + u_{j,n}x_n + u_{j,0} &< 0 \end{aligned} \tag{6.17}$$

Für Neuronen mit stückweise linearer Ausgabe einer beliebigen Schicht kann gezeigt werden, dass diese den Merkmalsraum ebenfalls in eine Partition mit jeweils linearer Ausgabe zerlegen. Auch hier werden die Teile des Merkmalsraums durch eine begrenzte Anzahl linearer Bedingungen definiert. Um dies zu zeigen, wird davon ausgegangen, dass jeder Eingang eine Partition des Merkmalsraumes mit einer linearen Ausgabe und endlich vielen linearen Bedingungen durchführt. Ein Neuron einer höheren Schicht kombiniert also die einzelnen Partitionen der r Eingänge.

Kombiniert man zwei Partitionen mit l_1 und l_2 Teilen entsteht eine neue Partition. Dabei unterteilt die zweite Partition jeden Teil der ersten Partition in maximal l_2 neue Teile. Somit entsteht eine neue Partition mit maximal $l_1 \cdot l_2$ Teilen. Dieser Sachverhalt wird durch Abbildung 6.17 illustriert.

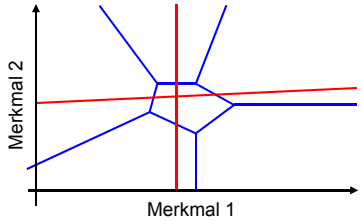

Abbildung 6.17: Kombination zweier Partitionen: Kombiniert man zwei Partitionen der Größen l_1 und l_2, dann entsteht eine neue Partition mit maximal $l_1 \cdot l_2$ Teilen.

Dementsprechend entsteht beim Kombinieren der Eingänge eine neue Partition mit maximal

$$l_{max,\ Aktivierung} = \prod_{i=1}^{r} l_i \tag{6.18}$$

Teilbereichen. Eine ähnliche Abschätzung lässt sich für die maximale Anzahl von Bedingungen pro Teilbereich definieren. Wenn $b_1 ... b_r$ die maximale Bedingungsanzahl pro Teilbereich für jede der Eingangspartitionen beschreibt, so ist die maximal mögliche Anzahl von Bedingungen pro Teilbereich für die kombinierte Partition:

$$b_{max,\ Aktivierung} = \sum_{i=1}^{r} b_i \ . \tag{6.19}$$

Jedes Neuronen mit einer Rampenfunktion als Ausgabefunktion partitioniert seinen Eingaberaum in drei Teilbereiche. Die Ausgabefunktion des Neurons unterteilt darum jeden Teilbereich noch einmal in maximal drei Teilbereiche. Es werden zu jedem entstehenden Teilbereich maximal zwei weitere Bedingungen hinzugefügt. Somit ist die maximale Anzahl von Teilbereichen für die Ausgabe beschränkt durch:

$$l_{max,\ Ausgabe} = 3 \cdot \prod_{i=1}^{r} l_i \ . \tag{6.20}$$

Die maximale Anzahl von Bedingungen pro Teilbereich ist dabei:

$$b_{max,\ Ausgabe} = 2 + \sum_{i=1}^{r} b_i \ . \tag{6.21}$$

Für Neuronen der zweiten Schicht mit r Eingängen entsteht dementsprechend eine Partition mit einer maximalen Teilbereichsanzahl von:

$$l_{max,\ Ausgabe} = 3 \cdot 3^r \ . \tag{6.22}$$

Die maximale Anzahl von Bedingungen jedes Teilbereiches ist dabei:

$$b_{max,\ Ausgabe} = 2 + 2r \ . \tag{6.23}$$

Weiterhin ist die Ausgabe der Neuronen einer höheren Schicht in den einzelnen Teilbereichen zu untersuchen. Hier ist speziell der mittlere Teil der Ausgabefunktion von Interesse, da in den anderen Teilen die Ausgabe konstant ist. Die zu untersuchende Ausgabe ist also gleich der Aktivierung des Neurons:

$$y = w_0 + w_1 \left(v_{1,1} x_1 + ... + v_{1,n} x_n + v_{1,0}\right) + ... + w_r \left(v_{r,1} x_1 + ... + v_{r,n} x_n + v_{r,0}\right) .$$

Da die einzelnen Werte der Faktoren $v_{1,0}, ..., v_{r,n}$ und $w_0, ..., w_r$ bekannt sind, lässt sich die Ausgabe unter Verwendung neuer Koeffizienten $v_0, ..., v_n$ in die folgende Form umschreiben:

$$y = z = v_1 x_1 + ... + v_n x_n + v_0. \tag{6.24}$$

Die Ausgabe entspricht somit ebenfalls der gewünschten linearen Form.

Demzufolge gilt für mehrlagige Perzeptrons unter Verwendung stückweise linearer Ausgabefunktionen: Jedes Ausgabeneuron unterteilt den Merkmalsraum in eine Partition mit endlich vielen Teilbereichen gemäß Abbildung 6.16. Die Ausgabe in jedem Teilbereich ist linear. Die einzelnen Teilbereiche sind durch lineare Bedingungen beschrieben.

Obwohl bei der Nutzung stückweise linearer Ausgabefunktionen eine gewünschte Partition entsteht, ist dieser Ansatz nur beschränkt nutzbar. Die Anzahl der Teile der Partition steigt theoretisch exponentiell mit der Anzahl der Neuronen. In der Praxis ist die auftretende Komplexität jedoch deutlich geringer. So entfallen einige Teilbereiche bei der Kombination unterschiedlicher Partitionen, da die kombinierten Teile häufig widersprüchliche Bedingungen enthalten. Weiterhin entfallen verschiedene Bedingungen, da bei der Kombination manche Bedingungen redundant sind. Trotzdem ist der Ansatz auf relativ kleine Netze beschränkt. Auch in der Praxis wird die Anzahl der Teilbereiche und Bedingungen für eine komplette manuelle Analyse sehr schnell zu komplex. Eine automatische Analyse ist jedoch auch noch bei mehreren tausend Teilbereichen möglich. Im Gegensatz zu Netzen mit beliebigen Ausgabefunktionen ist hier immer noch die Ausgabeanalyse für ganze Merkmalsbereiche möglich. Die praktische Nutzung der Analyse von neuronalen Netzen mit stückweise linearen Ausgabefunktionen wird im folgenden Abschnitt beschrieben.

6.4 Mustererkennung

Um eine spätere Anpassung der Ergebnisse entsprechend der Anforderungen der Assistenzsysteme durchführen zu können, bestimmt die Mustererkennung nicht nur die Objektklasse, sondern gibt für jede Objektklasse einen Zugehörigkeitswert aus. Bei diesem Zugehörigkeitswert entspricht Null einer unwahrscheinlichen Klasse, während Eins mit einer wahrscheinlichen Klasse korrespondiert.

6.4.1 One-VS-All-Methode

Um sinnvolle Zugehörigkeitswerte für die einzelnen Klassen bestimmen zu können, wurde das System nach der One-VS-All-Methode aufgebaut. Das bedeutet, dass für jede Klasse ein Mustererkennungssystem trainiert wird. Dieses System versucht die korrespondierende Klasse von den restlichen Objekten zu trennen (Abbildung 6.18). Dafür werden beim Training nur zwei Ausgabewerte präsentiert. Die gesuchte Klasse bekommt als Vorgabe für die Ausgabe den Wert Eins. Alle anderen Objekte werden zusammengefasst und werden mit der Ausgabe Null trainiert.

6.4.2 Optimierung der Merkmale

Beim Training künstlicher neuronaler Netze werden die Merkmale in der Regel auf vergleichbare Wertebereiche skaliert. Üblicherweise wird diese Skalierung manuell durchgeführt. In der Praxis ist jedoch häufig nicht bekannt, welcher Bereich des Merkmals für die Klassifikation von Bedeutung ist und welche Bereiche irrelevant sind. Gründe für dieses fehlende Vorwissen sind häufig in der Sensordatenverarbeitung auftretendes Rauschen und Messfehler. Es werden beispielsweise oft nur Teile der Objekte vermessen, wodurch die

geschätzte Objektgröße signifikant von der realen Größe abweichen kann. Darüber hinaus ist für komplexere Merkmale die genaue Verteilung auch nicht a priori bekannt. Diese müsste auch erst aufgrund einer Analyse der Beispieldaten bestimmt werden.

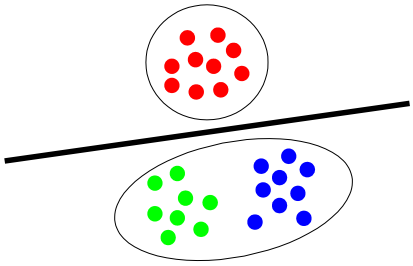

Abbildung 6.18: One-VS-All-Methode: Bei dieser Methode versucht das System zur Mustererkennung eine bestimmte Klasse von allen übrigen Objekten zu trennen.

Aus diesem Grund wurde hier ein neuer Ansatz zur automatischen, datengetriebenen Skalierung der Merkmale entwickelt. Dabei wird pro Objektklasse ein kleines Netz für jedes einzelne Merkmal trainiert (Abbildung 6.19). Die Anzahl der Neuronen in der verdeckten Schicht wurde hier manuell mit $r = 5$ festgelegt. Als Ausgabefunktion werden stückweise lineare Funktionen verwendet. Dies ermöglicht die nachträgliche Analyse der erlernten Abbildung gemäß Abschnitt 6.3.2.

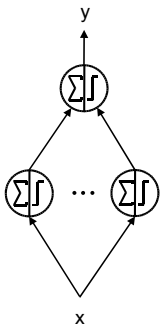

Abbildung 6.19: Datengetriebene Merkmalsoptimierung: Für jedes Merkmal wird ein kleines Netz trainiert, dass die Merkmale auf den Bereich zwischen Null und Eins skaliert.

Bei der Analyse reduziert sich die Zahl der Teilbereiche der Partition stark, da der Merkmalsraum eindimensional ist. So unterteilt das erste Neuron der verdeckten Schicht den Merkmalsbereich in drei Teile. Jeder Teil ist durch maximal zwei Merkmalswerte begrenzt. Jedes weitere verdeckte Neuron kann aufgrund der bei der Ausgabefunktion unterschiedenen drei linearen Bereiche nur zwei weitere Grenzen in den eindimensionalen Merkmalsraum einbringen. Dadurch können in diesem speziellen Fall nur jeweils zwei neue Teilbereiche im Merkmalsraum entstehen, die wiederum durch maximal je zwei Merkmalswerte begrenzt sind (Abbildung 6.20).

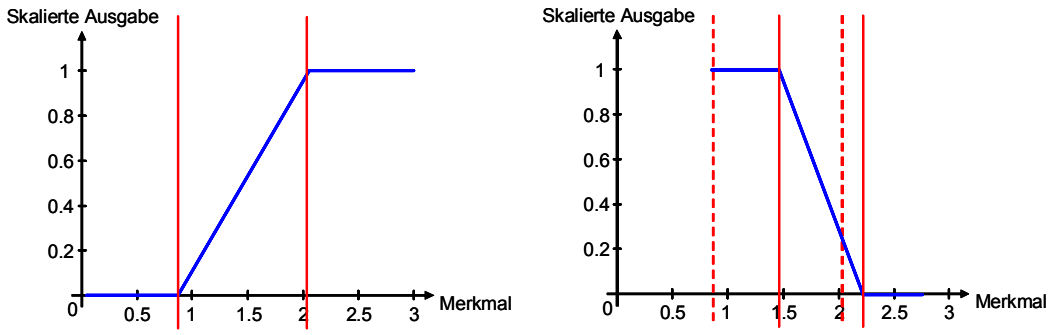

Abbildung 6.20: Unterteilung des eindimensionalen Merkmalsraums bei der datengetriebenen Optimierung: Das erste Neuron der verdeckten Schicht unterteilt den Merkmalsraum in drei Teile (links). Jedes weitere Neuron kann maximal zwei weitere Teile hinzufügen (rechts).

Das Ausgabeneuron unterteilt nun jeden Bereich noch einmal in maximal drei Teilbereiche. Somit ergibt sich für die Anzahl der entstehenden Teilbereiche:

$$l_{max,\,Skalierung} = 3 \cdot (1 + 2r)\,. \tag{6.25}$$

Als Begrenzungen jedes Teilbereiches fungieren dabei jeweils maximal zwei Werte des untersuchten Merkmals.

Das Netz lernt beim Training für die gesuchte Klasse eine Eins und für die restlichen Objekte eine Null. Auf diese Weise wird automatisch der Merkmalsbereich hervorgehoben, der die gesuchte Klasse am besten charakterisiert. In Abbildung 6.21 ist die erlernte Skalierung für die Objektbreite zur Klassifikation von PKWs dargestellt. Es ist zu erkennen, dass das Verfahren automatisch eine Breite von etwa 1,50 m für PKWs als typisch erlernt hat. Diese Breite entspricht auch in der Realität dem typischen Wert für PKWs, wenn man die Breite ohne Spiegel bestimmt.

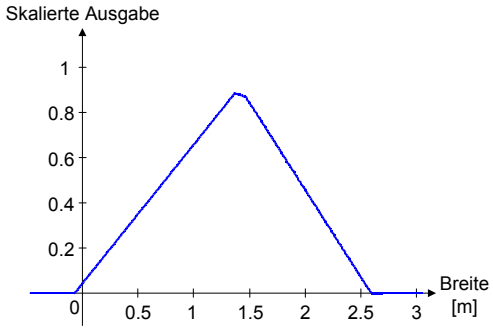

Abbildung 6.21: Merkmalsoptimierung für die Objektbreite zur Klassifikation von PKWs: Die typische Breite von etwa 1,50 m wird auf hohe Werte abgebildet. Je weiter die Breite abweicht, desto niedriger wird der Wert.

Die im obigen Beispiel erlernte Abbildung ähnelt einem aus der Fuzzy-Logik bekannten trapezförmigen Fuzzy-Set [161]. Fuzzy-Sets könnten ebenfalls datengetrieben erlernt

werden, um auf diese Weise eine automatische Merkmalsskalierung zu erreichen. Der vorgestellte Ansatz mit Hilfe eines neuronalen Netzes hat jedoch Vorteile bezüglich der Abbildungsfähigkeit der datengetriebenen Merkmalsoptimierung. Ein neuronales Netz mit 5 Neuronen in der verdeckten Schicht kann das Merkmal beispielsweise in bis zu 33 Bereiche mit linearer Ausgabe unterteilen. Die Ausgabe ist dadurch nicht auf ein Maximum beschränkt. So hat das optimierte Merkmal der Breite für die Klasse Hintergrund sinnvollerweise an den Stellen Maxima, die weder für Fußgänger oder Zweiräder noch für PKWs oder LKWs typisch sind. Darüber hinaus sind die entstehenden Minima und Maxima nicht auf 0 und 1 festgelegt, sondern können statistisch günstige Werte zwischen 0 und 1 annehmen. Schließlich ist der Anstieg von 0 auf 1 nicht auf einen lineraren Verlauf beschränkt. Stattdessen können mehrere aufeinander folgende Bereiche mit unterschiedlichem Anstieg erlernt werden. Von dieser Möglichkeit profitiert zum Beispiel die in Abbildung 6.22 dargestellte datengetriebene Merkmalsoptimierung der geschätzten Geschwindigkeit für PKWs. Bei sehr geringen Geschwindigkeiten steigt das optimierte Merkmal nur unwesentlich mit der Geschwindigkeit, da solche Geschwindigkeiten unter der im ersten Gang möglichen Schrittgeschwindigkeit liegen. Anschließend steigt die Ausgabe stark mit der Geschwindigkeit an. Ab einer Geschwindigkeit von etwa 8 $\frac{\text{m}}{\text{s}}$ steigt die Ausgabe deutlich weniger, da dann bereits Hintergrundobjekte, Fußgänger und Radfahrer über dieses Merkmal ausgeschlossen werden können. Ab einem Wert von etwa 32 $\frac{\text{m}}{\text{s}}$ erfolgt eine konstante Ausgabe, da auch LKWs diese Geschwindigkeit nicht mehr erreichen.

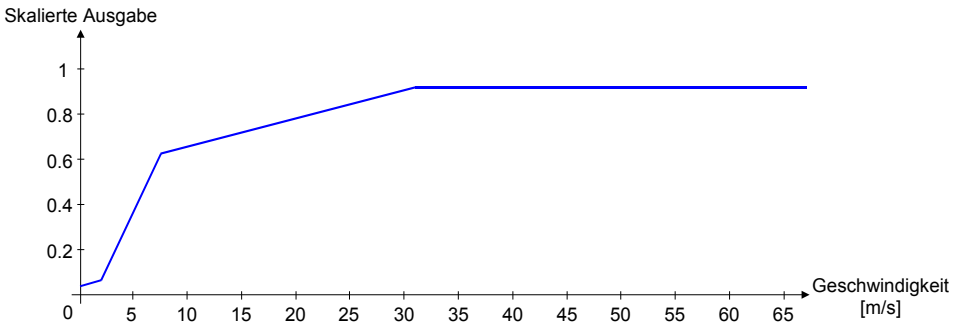

Abbildung 6.22: Automatische Skalierung der Geschwindigkeit für PKWs: Aufgrund der Skalierung mit einem neuronalen Netz kann die skalierte Ausgabe zwischen dem Minimum und dem Maximum in verschiedenen Bereichen unterschiedlich stark ansteigen. Der entsprechende Zusammenhang kann datengetrieben ermittelt werden.

Die Voraussetzung für eine datengetriebene Optimierung eines Merkmals ist die Eigenschaft, dass allein mit diesem Merkmal eine gewisse Charakterisierung der Objektklasse möglich ist. Dabei ist auch eine qualitativ eher schlechtere Charakterisierung ausreichend. Manche Merkmale werden jedoch nur als Zusatzinformation für die Auswertung anderer Merkmale in die Klassifikation eingebracht. Hierzu zählen zum Beispiel die Gierrate und die Geschwindigkeit des eigenen Versuchsfahrzeuges. Diese Werte dienen nur als Maß für die Verlässlichkeit der geschätzten absoluten Objektgeschwindigkeiten. Bei derartigen Merkmalen ist eine datengetriebene Optimierung nicht sinnvoll, da keinerlei Korrelation zwischen Merkmal und Objektklasse besteht.

Aus diesem Grund wird für solche Merkmale eine manuelle Skalierung durchgeführt. In Abbildung 6.23 ist der Verlauf der verwendeten manuellen Skalierung dargestellt. Für diese Skalierung müssen vier charakteristische Merkmalspunkte manuell festgelegt werden. Die Punkte $s_{1,0}$ und $s_{1,1}$ beschreiben den Verlauf der ersten Rampe, während $s_{2,0}$ und $s_{2,1}$ die zweite Rampe charakterisieren.

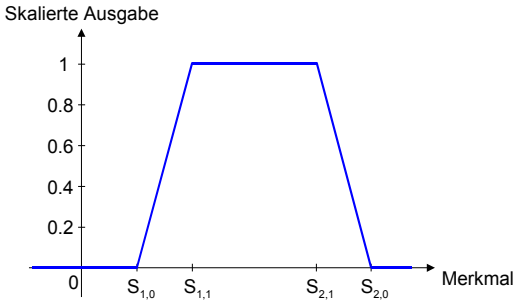

Abbildung 6.23: Verlauf der manuellen Skalierung: Der Verlauf wird durch vier Merkmalswerte festgelegt, die die beiden Rampen charakterisieren.

Die Skalierung wird in einem kleinen neuronalen Netz mit zwei verdeckten Neuronen codiert, um die formale Kompatibilität zu den anderen Merkmalen zu bewahren. Abbildung 6.24 zeigt die Struktur des Netzes und die zu berechnenden Gewichte und Biaswerte.

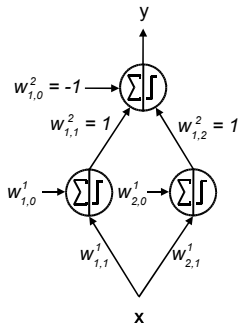

Abbildung 6.24: Codierung der Skalierung in einem neuronalen Netz: Das Ausgabeneuron übernimmt die Kombination der zwei Rampenfunktionen mit den Gewichten 1 und dem Bias -1. Die Gewichte und Biaswerte der verdeckten Neuronen müssen aus dem Verlauf der Skalierung berechnet werden.

Das Ausgabeneuron hat hier die Aufgabe, die beiden Rampenfunktionen der verdeckten Neuronen zu kombinieren. Dafür werden die Gewichte mit 1 und der Bias mit -1 festgelegt. Die Gewichte und Biaswerte der verdeckten Neuronen müssen aus den vier Merkmalspunkten aus Abbildung 6.23 berechnet werden.

Für das Gewicht und den Bias des ersten Neurons ergibt sich:

$$w_{1,1} = \frac{1}{s_{1,1} - s_{1,0}} \tag{6.26}$$

$$w_{1,0} = \frac{-s_{1,0}}{s_{1,1} - s_{1,0}}. \tag{6.27}$$

Analog ergibt sich für das zweite Neuron:

$$w_{2,1} = \frac{1}{s_{2,1} - s_{2,0}} \tag{6.28}$$

$$w_{2,0} = \frac{-s_{1,0}}{s_{2,1} - s_{2,0}}. \tag{6.29}$$

6.4.3 Gesamtklassifikator

Der Gesamtklassifikator jeder Objektklasse kombiniert nun die skalierten Merkmale. Dies wird durch ein neuronales Netz mit einer verdeckten Schicht durchgeführt. Auch hier wird die stückweise lineare Rampenfunktion als Ausgabefunktion verwendet. Die Anzahl der verdeckten Neuronen wird hier automatisch bestimmt. Für jede Anzahl zwischen 0 und 15 wird ein Trainingsprozess gestartet. Das Netz mit dem geringsten Abbildungsfehler für die Validierungsdaten wird schließlich ausgewählt.

Die Gesamtstruktur der Mustererkennung zur Berechnung des Zugehörigkeitswertes für eine Klasse hat die in Abbildung 6.25 dargestellte Struktur. Insgesamt fünf dieser Strukturen bestimmen die benötigten Zugehörigkeitswerte für die Klassen „Fußgänger“, „Zweirad“, „PKW“, „LKW“ und „Hintergrund“.

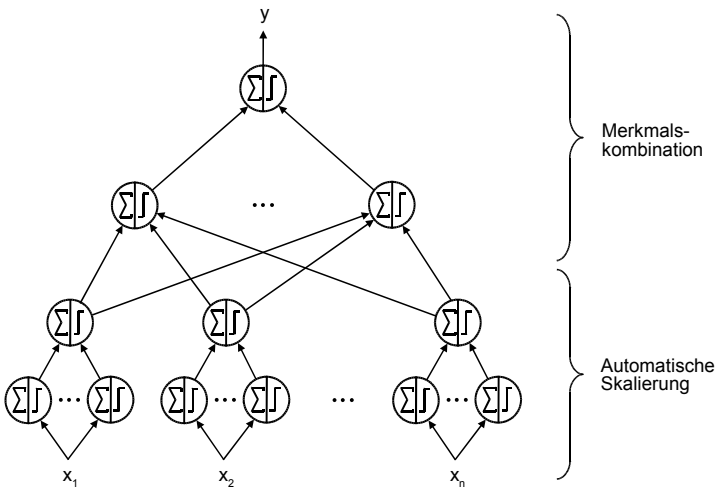

Abbildung 6.25: Die für jede Klasse unterschiedlich skalierten Merkmale werden pro Klasse mit einem weiteren Netz kombiniert. Als Ausgabefunktion kommt ebenfalls die Rampenfunktion zum Einsatz.

In der Regel wählt das Trainingsverfahren zur Kombination der Merkmale Netze mit 5 bis 10 verdeckten Neuronen aus. Die dadurch entstehenden Partitionen lassen sich nicht mehr manuell gemäß Abschnitt 6.3.2 analysieren. Automatische Analysen für bestimmte Merkmalsbereiche sind aber prinzipiell noch denkbar.

6.4.4 Signifikanz der Merkmale

Jedes in der Mustererkennung verwendete Merkmal erhöht die Anzahl der zu trainierenden Parameter in der Mustererkennung. Aus diesem Grund sollten prinzipiell nur Merkmale verwendet werden, die auch einen geeigneten Beitrag zur Klassifikation leisten.

Eine häufig verwendete Methode zur Auswahl geeigneter Merkmale für eine Klassifikationsaufgabe ist das Verfahren AdaBoost [117, 140, 141]. Dieses Verfahren wählt iterativ Merkmale aus einer Menge von verfügbaren Merkmalen aus. Dabei wird sukzessive jeweils das Merkmal mit dem besten Klassifikationsbeitrag zum Klassifikator hinzugefügt, bis eine gewünschte Klassifikationsleistung erreicht wurde. AdaBoost verwendet jedoch die Merkmale nur in einer Linearkombination. Somit werden auch lediglich Merkmale ausgewählt, die in einer Linearkombination die Klassifikationsleistung verbessern.

Die hier vorgestellte Mustererkennung kann jedoch deutlich komplexere Abbildungen des Merkmalsraums auf die einzelnen Objektklassen realisieren. Entsprechend komplexe Verknüpfungen der Merkmale können dementsprechend im trainierten System auftreten. Um nur Merkmale zu verwerfen, die auch in solchen Verknüpfungen keinen wesentlichen Beitrag zur Klassifikation leisten können, wurde hier eine andere Methode zur Signifikanzanalyse verwendet. Diese einfache Methode besteht darin, das System einmal mit allen Merkmalen und einmal unter Ausschluss des untersuchten Merkmals zu trainieren. Lässt die Klassifikationsleistung unter Ausschluss des untersuchten Merkmals nur unwesentlich nach, liefert das Merkmal nur einen sehr geringen Beitrag zur Gesamtklassifikation und kann entfernt werden.

Mit dem beschriebenen Verfahren lassen sich sukzessive einzelne Merkmale aus der Mustererkennung entfernen. Aufgrund der enormen Rechenzeit für das Training der neuronalen Netze konnte das Verfahren jedoch nur auf ausgewählte Merkmale angewendet werden. Im Rahmen der Untersuchungen der Klassifikation mit Laserscannermessdaten wurden verschiedene Merkmale ermittelt, deren Einfluss auf die Klassifikation aufgrund der beschriebenen Methode als gering bewertet wurde und die darum wieder aus dem System entfernt wurden. Diese Merkmale wurden in Abschnitt 6.1 jedoch nicht aufgelistet, so dass in dieser Arbeit nur die für die Mustererkennung eingesetzten Merkmale beschrieben werden.

6.5 Regelbasierte Klassifikation

Im regelbasierten Teil wird regelbasiertes a priori Wissen verwendet, um die Ergebnisse der Mustererkennung zusätzlich zu überprüfen und gegebenenfalls zu korrigieren. Hierfür gibt es verschiedene Gründe.

Die Automobilindustrie erwartet in der Regel nachvollziehbare Algorithmen, insbesondere wenn sicherheitsrelevante Anwendungen angesteuert werden. Nun ist der optimale Klassifikationsprozess aufgrund des hochdimensionalen Merkmalsraums in der Regel von außen nicht direkt nachvollziehbar. Dennoch lässt sich mit regelbasiertem Wissen nachträglich ein gewisses Verhalten erzwingen. Es werden zum Beispiel für bewegliche Objekte Maximalgrößen festgesetzt. Ist das verfolgte Objekt größer als die Maximalgröße einer Klasse, wird der Zugehörigkeitswert auf Null gesetzt. In vergleichbarer Weise wurde auch eine Maximalgeschwindigkeit für Fußgänger festgelegt. Ein den Regeln entsprechendes Verhalten sollte die Mustererkennung auch automatisch erlernen. Über Regeln lässt es sich jedoch garantieren.

Ein weiterer Grund für Manipulationen im regelbasierten Teil sind Zusatzinformationen, die nur bedingt verfügbar sind. So wird in Kapitel 8 ein Ansatz zur Nutzung von digitalen Karten vorgestellt. Dieser Typ Karten wird erwartungsgemäß nur an wenigen Kreuzungen verfügbar sein. Die korrespondierenden Informationen können regelbasiert im Online-Betrieb mit dem normalen System fusioniert werden.

Schließlich wird der regelbasierte Teil auch verwendet, um Objekte auszusortieren, die die Mustererkennung beim Training vorenthalten bekommt. So werden beim Training nur Objekte präsentiert, für die mindestens dreimal Messwerte vom Laserscanner vorlagen. Der Grund dafür ist die starke Abhängigkeit der Klassifikation von Bewegungsinformationen. Es werden vor der Klassifikation drei Messzyklen abgewartet, damit die Objektverfolgung eingeschwungen ist und sinnvolle Schätzwerte für die Geschwindigkeit liefert. Weiterhin wird die Messdistanz hier auf 150 m eingeschränkt, da in höheren Distanzen Objekte oft nicht mehr oder nur mit extrem wenigen Messwerten vermessen werden. Sinnvollerweise bräuchte für diese Objekte die Mustererkennung nicht ausgeführt werden. Die Mustererkennung benötigt jedoch mit weniger als 1 ms für alle Objekte eines Zeitschrittes vergleichsweise wenig Rechenzeit. Aus diesem Grund wird der Ausschluss von Objekten nicht explizit behandelt, sondern erst hier durchgeführt.

6.6 Anwendungsspezifische Ausgabe

Der letzte Schritt der Klassifikation besteht aus der Auswahl der zugeordneten Objektklasse. Diese Auswahl kann für jedes Assistenzsystem separat erfolgen. Als Grundlage der Klassenzuordnung dient ein Vektor mit Zugehörigkeitswerten für jede Klasse als Ergebnis der Klassifikation eines einzelnen Objektes:

$$\mathbf{c_{Klassifikation}} = \begin{bmatrix} c_{Hintergrund} \\ c_{Fußgänger} \\ c_{Zweirad} \\ c_{PKW} \\ c_{LKW} \end{bmatrix}. \tag{6.30}$$

Eine Klassenentscheidung kann nun aufgrund des maximalen Zugehörigkeitswertes getroffen werden.

In vielen Fällen werden von dem Fahrerassistenzsystem jedoch besondere Anforderungen an die Klassifikation gestellt. So muss beispielsweise ein Kompromiss zwischen einer geringen Anzahl von Falschalarmen und einer hohen Anzahl von korrekt klassifizierten Objekten gefunden werden. Dies geschieht über eine a priori Gewichtung der einzelnen Klassen. Eine höhere Gewichtung der Klasse „Hintergrund“ führt beispielsweise zu geringeren Falschalarmanzahlen. Die Gewichtung wird additiv mit einem a priori Vektor durchgeführt:

$$\mathbf{c}_{\mathbf{Ausgabe}} = \mathbf{c}_{\mathbf{a\ priori}} + \mathbf{c}_{\mathbf{Klassifikation}} \,. \tag{6.31}$$

Es wurden ebenfalls Experimente mit multiplikativer Gewichtung durchgeführt. Die additive Variante erwies sich in der Praxis jedoch als deutlich effektiver.

Da der sinnvollste a priori Vektor stark von den Anforderungen des Assistenzsystems abhängt, wird die Gewichtung im Modul der anwendungsspezifischen Ausgabe vorgenommen. Bei dieser Gewichtung können auch gleichzeitig bestimmte Klassen ausgeblendet werden, die für die jeweilige Anwendung nicht von Interesse sind. Die zugewiesene Klasse wird anschließend aufgrund des maximalen Zugehörigkeitswertes festgelegt.

6.7 Zusammenfassung

In diesem Kapitel wurde das entwickelte Klassifikationsframework beispielhaft für die laserscannerbasierte Klassifikation vorgestellt. Dafür wurden diverse Merkmale basierend auf den Laserscannermessdaten und der Objektverfolgung präsentiert. Die Klassifikation kombiniert Methoden der statistischen Mustererkennung und regelbasierte Ansätze. Weiterhin wurde eine spezielle Netzwerkstruktur präsentiert, die eine datengetriebene Merkmalsoptimierung durchführt und eine nachträgliche Analyse der durch das Netzwerk erlernten Abbildung ermöglicht. Die folgenden Kapitel beschreiben Erweiterungen der laserscannerbasierten Klassifikation durch Fusionsanätze mit Videomessdaten und präzisen digitalen Karteninformationen. Die Leistungsfähigkeit des Klassifikationsframeworks wird in Kapitel 10 quantitativ mit Hilfe von Testdaten untersucht.

7 Fusionsansätze mit Videomessdaten

Obwohl mit Hilfe des Laserscanners ein großer Anteil der Objekte korrekt klassifiziert werden kann, treten bei der ausschließlichen Nutzung des Laserscanners bei einigen Objekten Fehlklassifikationen auf. Hierbei handelt es sich meist um Objekte der Klassen „Fußgänger“, „Zweirad“, „PKW“ und „LKW“, die sich momentan nicht bewegen. Der Grund hierfür ist der starke Einfluss der Objektbewegung auf die korrekte Klassifikation. Eine Verbesserung der Klassifikationsleistung soll durch eine Fusion der Laserscannermessdaten mit weiteren Messdaten erreicht werden. Hierfür werden in diesem Kapitel Fusionsansätze mit Videomessdaten vorgestellt.

Die Fusion beruht auf einer zusätzlichen, klassenspezifischen Detektion des vermessenen Objekts im Videobild. In dieser Arbeit wurden Fusionsansätze für Fußgänger und PKWs untersucht, da diese die am häufigsten auftretenden beweglichen Objektklassen darstellen. Für Zweiräder sind Fusionsansätze ähnlich des Ansatzes für PKWs denkbar. Für LKWs sind entsprechende Fusionsansätze nur bedingt sinnvoll. LKWs sind nur in hohen Distanzen komplett im Sichtbereich der Kamera. Im städtischen Verkehr und insbesondere an Kreuzungen werden in der Regel nur Teile von LKWs mit der Kamera erfasst. Aus diesem Grund können hier nur Bildverarbeitungsalgorithmen angewendet werden, die sich auf Elemente von LKWs konzentrieren. Derartige Verfahren werden an dieser Stelle nicht untersucht.

Eine weitere, sensorbedingte Einschränkung betrifft den Bereich des Fahrzeugumfelds, in dem eine Fusion durchgeführt werden kann. Der von der Kamera beobachtete Winkelbereich ist auch bei dem verwendeten Weitwinkelobjektiv mit 71° deutlich geringer als der des Laserscanners. Auch die Reichweite der Kamera ist geringer, da Objekte ab einer gewissen Distanz im Kamerabild aufgrund der Auflösung nur noch aus sehr wenigen Pixeln bestehen und darum von den Bildverarbeitungsalgorithmen nicht mehr erfasst werden können. Abbildung 7.1 zeigt die Messbereiche der verwendeten Sensoren.

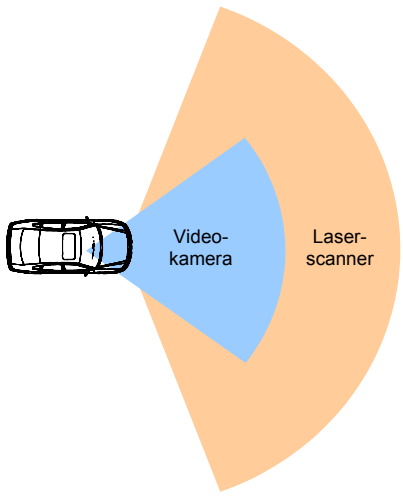

Abbildung 7.1: Unterschiedliche Sensormessbereiche: Die Videokamera deckt einen deutlich kleineren Winkelbereich als der Laserscanner ab. Auch die durch die Bildverarbeitung abgedeckte Reichweite ist geringer.

Obwohl der Messbereich, der für eine Fusion geeignet ist, deutlich kleiner als der des Laserscanners ist, können signifikante Verbesserungen erreicht werden. So deckt dieser Messbereich genau das Gebiet vor dem Fahrzeug ab, in das sich dieses bewegen wird. Somit ist dieser Bereich am kritischsten bezüglich möglicher Gefährdungen durch andere Objekte. Eine verbesserte Klassifikationsleistung in diesem Bereich ist demzufolge äußerst sinnvoll.

Die Verbesserung der Klassifikation basiert auf der Extraktion zusätzlicher Merkmale. Im Rahmen der Merkmalsextraktion werden Kaskadenklassifikatoren nach einem Verfahren von Viola und Jones [140, 141] eingesetzt. Aus diesem Grund werden im folgenden Abschnitt die Grundlagen dieses Verfahrens erläutert. Anschließend wird auf die Stellen im Gesamtsystem eingegangen, an denen eine Fusion durchgeführt wird. Nach einer Erläuterung der Aufmerksamkeitssteuerung wird auf die Detektion von Fußgängern und PKWs mit Kaskadenklassifikatoren eingegangen. Schließlich wird die Merkmalsextraktion aus den Detektionsergebnissen des Kaskadenklassifikators vorgestellt.

7.1 Kaskadenklassifikatoren

Die in dieser Arbeit verwendeten Kaskadenklassifikatoren beruhen auf einem Verfahren von Paul Viola und Michael Jones [140, 141]. Dieses Verfahren wurde ursprünglich zur Detektion von Gesichtern in Videobildern entwickelt. Aufgrund seiner Eignung zur echtzeitfähigen Anwendung und der ausgezeichneten Detektionsleistungen wurde es inzwischen auch im Automobilbereich für eine Vielzahl von Anwendungen eingesetzt [64, 72, 86, 91, 94, 108, 145, 164].

An dieser Stelle soll ein kurzer Überblick über die verwendeten Kaskadenklassifikatoren gegeben werden. Eine detailliertere Beschreibung und Analyse befindet sich in [154]. Das System kombiniert verschiedene Methoden aus dem Bereich der Bildverarbeitung und Klassifikation. Ein erster Ansatz, der erheblich zur Beschleunigung des Verfahrens beiträgt, sind kaskadierte Strukturen. In der Regel muss bei der Bildverarbeitung eine hohe Anzahl von in Frage kommenden Bildausschnitten auf Objekte überprüft werden. Durch kaskadierte Strukturen wird vermieden, dass ein rechenaufwändiges System jede Objekthypothese komplett überprüfen muss. Statt dessen werden nacheinander mehrere kleine Klassifikatoren durchlaufen, die jeweils einen Teil der Objekthypothesen verwerfen (Abbildung 7.2).

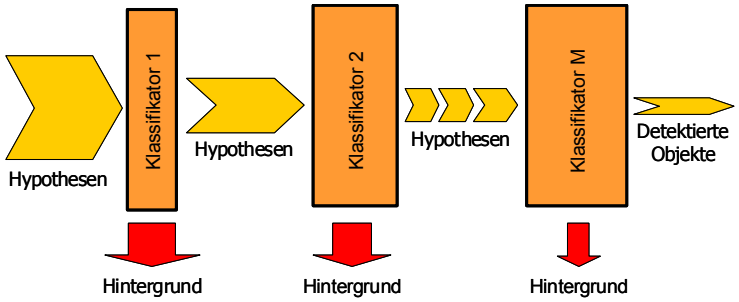

Abbildung 7.2: Kaskaden: Verschiedene Klassifikatoren reduzieren sukzessive die Objekthypothesen. Da viele Hypothesen nur von den ersten Stufen bewertet werden und diese Stufen nur einen Bruchteil der Gesamtrechenzeit beanspruchen, reduziert die Kaskade die Rechenzeit erheblich.

Innerhalb der einzelnen Kaskadenstufen werden sehr allgemeine, aber effizient berechenbare Merkmale verwendet. Hierfür werden als Erstes Integralbilder aus dem Originalbild berechnet. Dabei entspricht ein Pixel an der Position (x, y) im Integralbild der Summe der Pixelwerte des Originalbildbereiches links über dieser Position:

$$Integralbild\,(x, y) = \sum_{i=1}^{x} \sum_{j=1}^{y} Originalbild\,(i, j)\ . \tag{7.1}$$

Abbildung 7.3 zeigt die Fläche, deren korrespondierende Pixel im Originalbild zur Berechnung eines Pixels im Integralbild summiert werden.

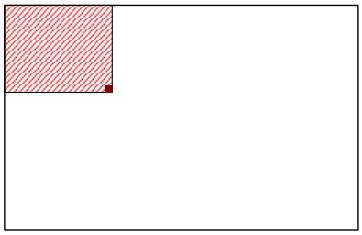

Abbildung 7.3: Integralbild: Jedes Pixel im Integralbild bekommt die Pixelsumme des links darüber liegenden Rechteckes des Originalbilds zugeordnet.

Mit Hilfe des Integralbildes können sehr effizient Pixelsummen beliebig großer, rechteckiger Bereiche des Originalbildes berechnet werden. Es werden lediglich vier Pixelwerte des Integralbildes benötigt (Abbildung 7.4).

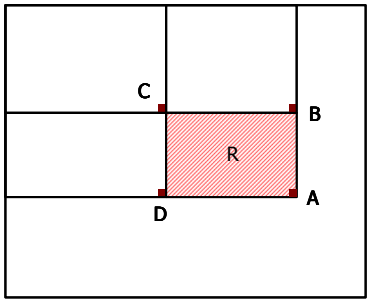

Abbildung 7.4: Effiziente Pixelsummen: Zur Berechnung der Pixelsumme eines beliebig großen Rechteckes werden lediglich vier Werte des Integralbildes benötigt.

Die Berechnung der Pixelsumme erfolgt durch eine geeignete Addition der vier Pixelwerte:

$$R = A - B + C - D\ . \tag{7.2}$$

Mit diesem Ansatz können nun verschiedene, einfache Merkmale sehr effizient berechnet werden. Abbildung 7.5 zeigt die prinzipiellen Merkmalsformen. Diese Merkmalsformen können an jeder beliebigen Position und in jeder passenden Größe auf die Objekthypothese angewendet werden (Abbildung 7.6).

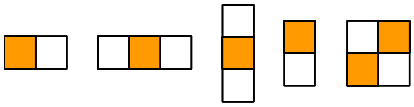

Abbildung 7.5: Merkmalsformen: Die abgebildeten Merkmalsformen können in jeder passenden Größe und an jeder Position auf der Objekthypothese verwendet werden. Zur Merkmalsberechnung werden jeweils die Pixelsummen der weißen Rechtecke von denen der blauen Rechtecke abgezogen.

Abbildung 7.6: Die verwendeten Merkmalstypen können in beliebiger Größe an jeder möglichen Position im überprüften Bildausschnitt angewendet werden.

Zur Merkmalsextraktion werden dann jeweils die Pixelsummen der weißen Rechtecke von denen der blauen Rechtecke abgezogen. Jede Kombination aus Merkmalstyp, Position und Größe führt zu einem sogenannten primitiven Klassifikator, der aus dem berechneten Merkmal mit einem anschließenden Schwellwert die Ausgabeklasse auf $y \in \{-1, 1\}$ festlegt. Der Vorteil der eher einfachen Merkmale liegt zum einen in der effizienten Berechnung durch Addition weniger Pixelwerte des Integralbildes. Zum anderen können die Merkmale bei der Suche nach größeren Objekten einfach skaliert werden, ohne den Rechenaufwand bei der Merkmalsextraktion zu erhöhen.

Aufgrund der zahlreichen verfügbaren Positionen und Größen existiert eine extrem große Anzahl von Merkmalen und korrespondierenden primitiven Klassifikatoren. Aus diesen verfügbaren Merkmalen müssen im Trainingsprozess geeignete Kombinationen zur Objektdetektion ausgewählt werden. Hierfür wird für jede Kaskadenstufe das Verfahren AdaBoost eingesetzt. Ähnlich wie andere Boosting-Verfahren handelt es sich hierbei um einen „Gierigen Algorithmus“ (*engl. greedy algorithm*). Bei dieser Klasse von Algorithmen wird iterativ jeweils der Folgezustand gewählt, der zum Zeitpunkt der Wahl den größten Gewinn verspricht.

Bei AdaBoost entspricht die iterative Bestimmung des Folgezustands der iterativen Auswahl der verwendeten primitiven Klassifikatoren. Als Grundlage zur Bewertung des Klassifikationsbeitrags einzelner Merkmale werden den Trainingsdaten Gewichte zugeteilt. Die Trainingsdaten bestehen dabei jeweils aus Paaren von Merkmalsvektor $\mathbf{x_i}$ mit Merkmalen $x_{i,1}, ..., x_{i,j}$ und Klasse $y_i \in \{-1, 1\}$. Diese enthalten m Beispiele des gesuchten Objekttyps und l Beispiele für Hintergrundobjekte. Die Initialisierung der Gewichte w_i erfolgt nach:

$$w_i\,(0) = \begin{cases} \frac{1}{2m}, \text{ falls der Datensatz } i \text{ ein gesuchtes Objekt enthält} \\ \frac{1}{2l}, \text{ falls der Datensatz } i \text{ ein Hintergrundobjekt enthält} \end{cases} . \tag{7.3}$$

In jeder Iteration k wird nun entsprechend der aktuellen Gewichte ein Merkmal ausgewählt.

Dafür werden als Erstes die Gewichte normiert:

$$w_i\,(k) = \frac{w_i\,(k)}{\sum\limits_n w_n\,(k)}\,. \tag{7.4}$$

Nun wird pro Merkmal j ein primitiver Klassifikator $h_j\left(k, x_j\right)$ trainiert, der mit einem geeigneten Schwellwert versucht, die Trainingsdaten unter Berücksichtigung der Gewichte bestmöglich zu klassifizieren. Die erreichbare Klassifikationsleistung ϵ_j des primitiven Klassifikators wird über die Summe der Gewichte falsch klassifizierter Daten bewertet:

$$\epsilon_j = \sum_i w_i\,(k) \left| \frac{h_j\left(k, x_{i,j}\right) - y_i}{2} \right|\,. \tag{7.5}$$

Der Klassifikator $h_{j(k)}\left(k, x_{j(k)}\right)$ mit der kleinsten Summe $\epsilon_{j(k)}$ leistet demzufolge den besten Beitrag und wird ausgewählt. Für die Verwendung des Klassifikators in der Kaskadenstufe wird ebenfalls ein Gewicht berechnet:

$$\beta_{j(k)} = \frac{1}{2} \ln\left(\frac{1 - \epsilon_{j(k)}}{\epsilon_{j(k)}}\right)\,. \tag{7.6}$$

Anschließend erfolgt eine Neuberechnung der Gewichte unter Berücksichtigung der Klassifikationsleistung des neuen Merkmals. Dabei werden die Gewichte falsch klassifizierter Daten erhöht und die der korrekt klassifizierten Daten verringert:

$$w_i\,(k+1) = \begin{cases} w_i\,(k) \cdot e^{-\beta_{j(k)}}, \textit{falls } h_{j(k)}\left(k, x_{i,j(k)}\right) = y_i \\ w_i\,(k) \cdot e^{+\beta_{j(k)}}, \textit{falls } h_{j(k)}\left(k, x_{i,j(k)}\right) \neq y_i \end{cases}\,. \tag{7.7}$$

Auf diese Weise wird erreicht, dass sich der Klassifikator in der folgenden Iteration auf die falsch klassifizierten Daten konzentriert. Das bedeutet, dass langfristig alle Daten durch geeignete primitive Klassifikatoren repräsentiert werden.

Die ausgewählten Merkmale werden schließlich in der Kaskadenstufe $H(\mathbf{x})$ als gewichtete Mehrheitsentscheidung kombiniert:

$$H(\mathbf{x}) = \begin{cases} +1, \text{falls } \sum\limits_k \beta_{j(k)} h_{j(k)}\left(k, x_{j(k)}\right) \geq T \\ -1, \text{falls } \sum\limits_k \beta_{j(k)} h_{j(k)}\left(k, x_{j(k)}\right) \leq T \end{cases}\,. \tag{7.8}$$

Hierbei kann mit der Schwelle T gesteuert werden, wieviel Daten von der Kaskadenstufe als Hintergrund verworfen werden. In der Regel wird die Schwelle so gewählt, dass annähernd alle Trainingsbeispiele gesuchter Objekte die Kaskadenstufe passieren.

Im Rahmen der hier durchgeführten Untersuchungen wurde eine Implementierung aus der Intel Open Computer Vision Library (OpenCV) [63] verwendet. Die Bibliothek enthält verschiedene Erweiterungen des ursprünglichen Algorithmus [81]. Dies beinhaltet unter anderem die Erstellung eines zusätzlichen Integralbildes mit einer Orientierung von 45° und die Verwendung entsprechender, um 45° gedrehter Merkmale. Weiterhin kann die Kaskadenstruktur in eine Baumstruktur erweitert werden [82].

7.2 Fusionsebene

Die Fusion findet an zwei Stellen statt. Für gemessene Objekte werden aus den aktuellen Laserscannermessdaten die Position, Größe und Orientierung bestimmt und in Form von Merkmalen an das Modul zur Bildverarbeitung geschickt. Dort erfolgt eine Fusion mit den Messdaten der Videokamera. Die Positions-, Größen- und Orientierungsinformationen werden mit einer Koordinatentransformation in die Bildebene transformiert. Mit diesen Informationen erfolgt bei der Bildverarbeitung eine Aufmerksamkeitssteuerung, die das Verfahren zur Objektdetektion auf relevante Bereiche einschränkt, dadurch erheblich beschleunigt und Falschalarme reduziert (Abbildung 7.7)

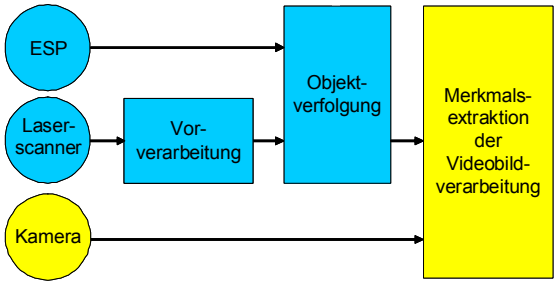

Abbildung 7.7: Aufmerksamkeitssteuerung: Vom Laserscanner werden gemessene Position, Größe und Orientierung für aktuelle Objekte an die Bildverarbeitung übergeben. Diese Informationen schränken die Bildverarbeitung auf die relevanten Bereiche ein.

Aus Kaskadenklassifikatoren werden zusätzliche Merkmale für bestehende Objekte extrahiert (Abschnitt 7.6). Diese werden mit den in Kapitel 6 beschriebenen Merkmalen durch die Mustererkennung fusioniert und tragen dadurch zu einer leistungsfähigeren Mustererkennung bei (Abbildung 7.8).

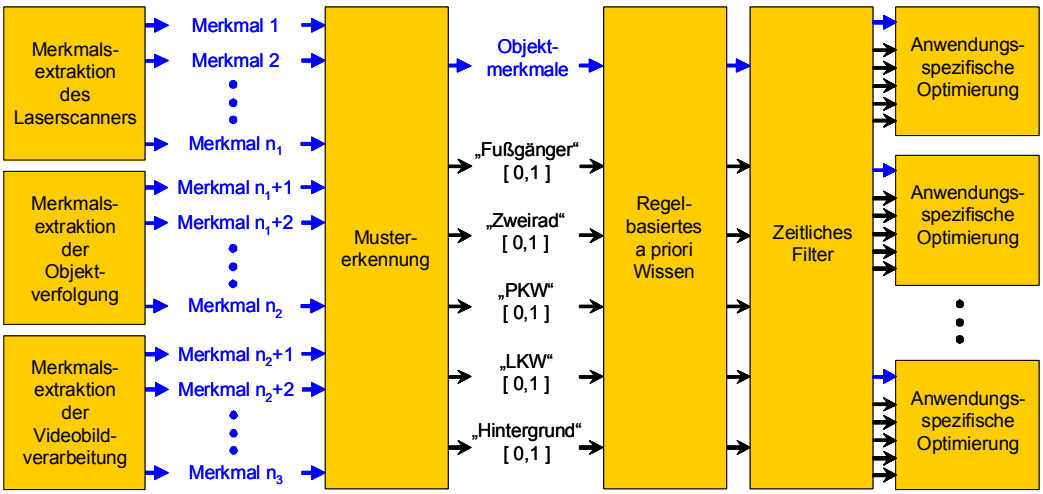

Abbildung 7.8: Verbessertes Klassifikationssystem: Durch zusätzliche Merkmale aus den Videomessdaten ist eine leistungsfähigere Mustererkennung realisierbar.

7.3 Aufmerksamkeitssteuerung

Verwendet man eine Kamera als einzigen Sensor, steht die Bildverarbeitung vor einer schwierigen Aufgabe. Die Detektion vorhandener Objekte im Videobild erweist sich als deutlich komplexer als eine Objektbildung aus den Messwerten eines entfernungsmessenden Sensors. Sowohl die Position als auch die Größe der Objekte ist in der Regel unbekannt (Abbildung 7.9, links). Aus diesem Grund muss in der Regel das gesamte Bild systematisch nach Objekten durchsucht werden. Hinzu kommt, dass Bildverarbeitungsalgorithmen meist nur nach Objekten einer bestimmten Größe suchen. Darum werden häufig Bildpyramiden verwendet. Das Bild wird mit einem entsprechenden Detektor an allen Positionen nach Objekten der kleinsten Objektgröße durchsucht. Anschließend wird das gesamte Bild auf eine kleinere Größe skaliert und erneut durchsucht. Auf diese Weise kann nach größeren Objekten gesucht werden. Die Skalierung wird dabei so oft durchgeführt, bis nach allen erwarteten Objektgrößen gesucht wurde (Abbildung 7.9, rechts).

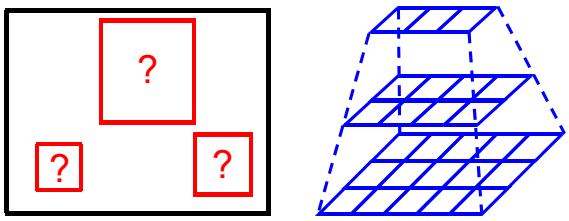

Abbildung 7.9: Bildpyramide: Bei der Bildverarbeitung ist oft unklar, an welcher Stelle und in welcher Größe Objekte auftreten (links). Aus diesem Grund muss das Bild systematisch durchsucht werden. Bei der Bildpyramide wird das Bild sukzessive durch Skalierung verkleinert. Ein Bildverarbeitungsalgorithmus durchsucht dann alle Skalierungen nach einer bestimmten Objektgröße (rechts).

Ansätze mit Bildpyramiden sind in der Regel sehr rechenaufwändig, da extrem viele Objekthypothesen mit der Bildverarbeitung überprüft werden müssen. Aus diesem Grund wird oft a priori Wissen verwendet, um die Anzahl der zu überprüfenden Hypothesen einzuschränken. Eine häufig eingesetzte Methode ist die Annahme einer ebenen Welt (*engl. flat world assumption*). Diese Methode wird auch in der Fahrzeugumfelderfassung oft eingesetzt [72, 108]. Bei dieser Annahme wird davon ausgegangen, dass sich alle Objekte auf der gleichen Ebene wie das Versuchsfahrzeug befinden. Unter dieser Voraussetzung korreliert die Objektposition mit der Objektgröße im Bild (Abbildung 7.10).

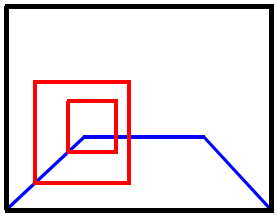

Abbildung 7.10: Annahme einer ebenen Welt: Unter der Annahme, dass alle Objekte auf einer Ebene stehen, korreliert die Objektposition mit der Objektgröße.

Leider stimmt die Annahme einer ebenen Welt in der Praxis nur selten. Bei Nickbewegungen des Versuchsträgers oder in bergigem Gelände treten starke Abweichungen von diesem Modell auf. Aus diesem Grund existieren Methoden, die leichte Abweichungen tolerieren (*engl. relaxed flat world assumption*). Zu diesem Zweck wird an einer durchsuchten Position die Bildverarbeitung auch für etwas größere und etwas kleinere Objektgrößen angewendet, als bei einer exakt ebenen Welt zu erwarten wäre.

Eine deutlich effektivere Methode zur Einschränkung der Objekthypothesen ist die Fusion mit einem entfernungsmessenden Sensor [69, 94]. Dieser misst die dreidimensionale Position des Objektes im Fahrzeugumfeld. Aufgrund einer Sensorkalibrierung können diese Informationen dann in die Bildebene der Kamera transformiert werden. Dadurch entsteht eine starke Einschränkung der Objektposition und -größe (Abbildung 7.11). Das Bild muss nun nur noch in einem sehr kleinen Bereich und für sehr wenige Objektgrößen durchsucht werden. Die durchgeführte Variation der Objektgröße muss nun nur noch reale Größenunterschiede von Objekten der gleichen Objektklasse ausgleichen.

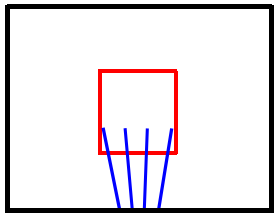

Abbildung 7.11: Einschränkung der Objekthypothesen durch Sensorfusion: Mit Entfernungssensoren erfasste Objekte können in die Bildebene der Videokamera transformiert werden. Dort müssen dann nur wenige Skalierungen und Positionen durchsucht werden.

Die Aufmerksamkeitssteuerung der Bildverarbeitung mit Hilfe eines entfernungsmessenden Sensors kann in erheblichem Maße die benötigte Rechenzeit reduzieren. Da im Bild aber auch nur noch an wenigen Stellen und in einem stark eingeschränkten Intervall möglicher Größen nach Objekten gesucht wird, erfolgt ebenfalls eine enorme Reduktion von Falschalarmen.

Eine Voraussetzung des Ansatzes ist jedoch die Notwendigkeit, dass der entfernungsmessende Sensor das Objekt erfassen muss. Ist dieser Sensor nicht in der Lage das Objekt zu erfassen, kann die Bildverarbeitung auch dessen Klassifikation nicht verbessern. Der verwendete Laserscanner liefert jedoch in der Regel von allen relevanten Objekten ausreichend Entfernungsmessungen, so dass dann lediglich Probleme bei der Klassifikation des Objektes auftreten können. In diesem Fall kann die Bildverarbeitung die Klassifikationsergebnisse mit Hilfe der im Bild enthaltenen Zusatzinformationen jedoch deutlich verbessern. In der Praxis können Objekte nur übersehen werden, wenn sich deren Entfernungsmessungen nicht von den Entfernungsmessungen anderer Objekte trennen lassen. Dies ist theoretisch möglich, wenn sich beispielsweise ein Fußgänger direkt an eine Wand lehnt. Aufgrund der Seltenheit entsprechender Fälle wird die Möglichkeit, auf diese Weise Objekte zu übersehen, in Kauf genommen, da eine enorme, für Fahrerassistenzsysteme deutlich relevantere Reduktion von Falschalarmen erreicht werden kann.

Der Ansatz der Aufmerksamkeitssteuerung wird in dieser Arbeit um zusätzliche Informationen erweitert. Aus den Entfernungsmessungen des Laserscanners lassen sich zusätzlich zur Position auch Informationen über horizontale und vertikale Ausdehnung sowie eine Orientierung bestimmen. Die Objekte werden dafür mit einer dreidimensionalen Box modelliert (Abbildung 7.12). Der Laserscanner liefert in der Regel Messwerte von einer oder zwei Seiten der Box. Das Videobild der Kamera enthält eine Mischung von bis zu drei Seiten der Box. Da in der Praxis jedoch der Einfluss der oberen Seite aufgrund der Kameraposition relativ gering ist, wird diese Seite ignoriert.

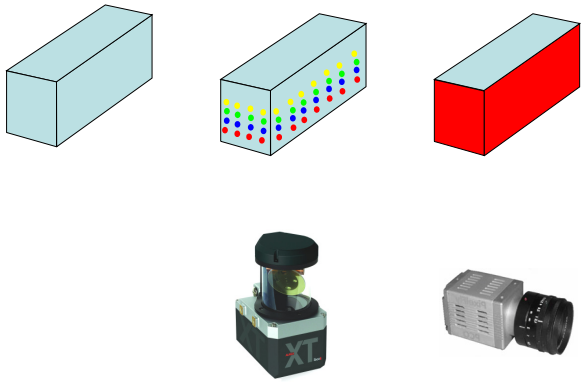

Abbildung 7.12: Dreidimensionales Objektmodell: Objekte werden durch eine dreidimensionale Box approximiert. Der Laserscanner liefert Messungen von einer oder zwei Seiten. Die Videobilder enthalten eine Mischung von bis zu drei Objektseiten. Die obere Seite kann in diesem Fall ignoriert werden, da ihr Einfluss aufgrund der Kameraposition gering ist.

Die Messwerte des Laserscanners beschreiben meist nicht das komplette Objekt. In horizontaler Richtung können Messwerte aufgrund von Verdeckungen oder ungünstiger Orientierung zum Sensor fehlen. Signifikantere Abweichungen sind in der ermittelten vertikalen Position und Größe enthalten. Da der Laserscanner nur in vier Ebenen Entfernungsmessungen durchführt, kann das Objekt meist nicht in der kompletten Höhe erfasst werden.

Zur Aufmerksamkeitssteuerung der Bildverarbeitung werden die aktuellen Messdaten des Laserscanners in einer dreidimensionalen, geometrischen Form zusammengefasst. Werden zwei Objektseiten gemessen, wird eine dreidimensionale Box berechnet. Bei einer vermessenen Objektseite kann nur ein Rechteck im dreidimensionalen Raum bestimmt werden (Abbildung 7.13). Diese geometrischen Formen beschreiben zwar nicht die komplette Objektgröße, sind jedoch mit Sicherheit komplett im Objekt enthalten. Da die Formen mit sehr wenigen Parametern beschrieben werden können, ist die Übertragung der Parameter an das Modul der Bildverarbeitung deutlich effektiver, als die Übertragung der kompletten Messdaten des Laserscanners. Im Bildverarbeitungsmodul wird für jedes Objekt pro gesuchter Objektklasse ein dreidimensionaler Suchbereich (*engl. ROI = region of interest*) festgelegt, der in die Bildebene transformiert wird und in dem dann die Bildverarbeitung zur Objektdetektion ausgeführt wird.

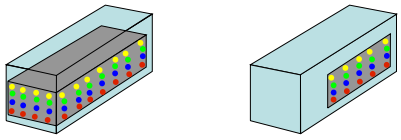

Abbildung 7.13: Die Entfernungsmessungen des Laserscanners werden in einer dreidimensionalen Form zusammengefasst. Wird nur eine Seite erfasst, so wird ein Rechteck berechnet. Werden zwei Seiten erfasst, ermittelt das System eine Box. Die ermittelte Form beschreibt nicht die komplette Ausdehnung des Objektes, ist aber mit Sicherheit darin enthalten.

7.4 Detektion von PKWs

Nach der allgemeinen Vorstellung der verwendeten Ansätze in den vorherigen Abschnitten wird in diesem Abschnitt die Anwendung zur Detektion von PKWs im Videobild vorgestellt. Während bisherige Arbeiten sich in diesem Bereich auf die Detektion von Fahrzeughecks konzentrierten, ist der hier vorgestellte Ansatz zusätzlich auch zur Detektion von Quer- und Gegenverkehr konzipiert, um darüber hinaus auch Kreuzungsszenarien abdecken zu können.

7.4.1 Erscheinungsbild von PKWs

Bei der Detektion von PKWs im Videobild ergibt sich neben der Lokalisierung und der Bestimmung der Objektgröße ein weiteres Problem. Das Aussehen eines PKWs ändert sich sehr stark mit der Orientierung zur Kamera (Abbildung 7.14). Dieser Sachverhalt stellt eine besondere Herausforderung an die Bildverarbeitung dar, da Detektionsalgorithmen Objekte üblicherweise aufgrund charakteristischer Konturen und Texturen erkennen.

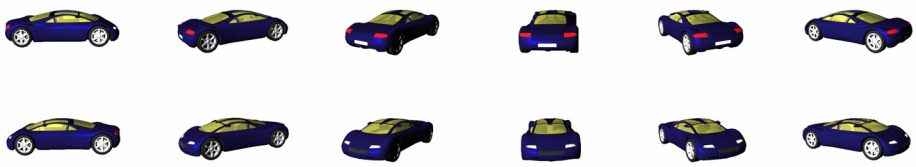

Abbildung 7.14: Das Erscheinungsbild eines PKWs ändert sich sehr stark mit der Orientierung zur Kamera. Dies ist eine besondere Herausforderung, da Detektionsalgorithmen meist Objete durch charakteristische Konturen und Texturen identifizieren.

Zur Handhabung des wechselnden Erscheinungsbildes existieren zwei bekannte Ansätze. In [121] wird ein komplexer Klassifikator trainiert, der alle möglichen Erscheinungsformen detektieren kann. Ein zweiter Ansatz besteht darin, mehrere Klassifikatoren zu trainieren. Dabei deckt jeder der Klassifikatoren einen bestimmten Bereich der Orientierung zur Kamera ab. Zumindest mit dem zweiten Ansatz ist ein erheblicher Rechenaufwand verbunden, da das Bild mehrfach durchsucht werden muss. Ein wesentlicher Nachteil beider Ansätze im

Fall von trainierten Bildverarbeitungsalgorithmen ist der enorme Bedarf an Trainingsdaten. Für jede mögliche Orientierung zur Kamera müssen in ausreichendem Maße Trainingsdaten vorbereitet und zum Training übergeben werden.

In dieser Arbeit wird deswegen ein neuer Ansatz verfolgt. Aufgrund der dreidimensionalen Objektinformationen des Laserscanners wird die Lage der Objektseiten im Videobild bestimmt. Die perspektivische Verzerrung dieser Seiten wird kompensiert. Im Anschluss wird mit jeweils einem Detektor nach Seitenansichten beziehungsweise nach Front- und Heckansichten gesucht. Auf diese Weise müssen lediglich Daten für diese Ansichten zum Training des Detektors bereitgestellt werden.

7.4.2 Objektauswahl

Die Anwendung der Bildverarbeitung ist nicht für jedes Objekt sinnvoll. Übersteigt die mit dem Laserscanner gemessene horizontale Ausdehnung die Maximalgröße eines PKWs, dann wird das entsprechende Objekt ignoriert. Weiterhin werden Objekte ignoriert, deren nach Abschnitt 6.1.1 berechnete maximale Größe die minimale Größe eines PKWs unterschreitet. Objekte mit einem Abstand von mehr als 120 m werden ebenfalls ignoriert, da die Kameraauflösung mit den verwendeten Objektiven dann so gering ist, dass PKWs nicht mehr detektiert werden können.

7.4.3 Berechnung der ROI

Aus der dreidimensionalen Form der Laserscannerdatenverarbeitung wird die Lage der Objektseiten ermittelt. In der Regel wird lediglich die Seite ausgewertet, die aufgrund ihrer Orientierung am besten von der Kamera erfasst werden kann. Bei einer Orientierung des Objekts zur Kamera 45° werden beide Seiten gleichermaßen erfasst. Die Seite von Fahrzeugen entspricht dem kubischen Objektmodell jedoch deutlich besser als die Front und das Heck, weil die Fahrzeugseite annähernd in einer vertikalen Ebene liegt. Aus diesem Grund wird die aufgrund des Winkels schlechter erfasste Seite auch ausgewertet, wenn aus den Laserscannermessdaten eine Länge von mehr als 3 m hervorgeht und darum von einer Seitenansicht ausgegangen werden kann. Zusätzlich muss die Breite der verzerrten Seitenansicht im Videobild größer als deren Höhe sein, um eine ausreichende horizontale Auflösung der Ansicht zu gewährleisten.

Für jede zu behandelnde Seite der dreidimensionalen Form muss eine geeignete dreidimensionale ROI definiert werden, die die korrespondierende Objektfläche mit Sicherheit enthält. Da die Seite der vom Laserscanner übermittelten Form nur einen Teil der realen Objektfläche darstellt, existieren dafür verschiedene Hypothesen. Die Extremfälle sind in Abbildung 7.15 blau gestrichelt dargestellt.

Wurde im Laserscanner nur eine Rechteckform bestimmt, kann diese in jeder Ecke der realen Objektfläche liegen. Im Fall einer Box ist eine vertikale Objektkante bekannt, demzufolge muss lediglich die vertikale Unsicherheit berücksichtigt werden. Die ROI wird so gewählt, dass sie alle in Frage kommenden Extremfälle enthält. Zusätzlich wird die ROI an allen Seiten um 15% vergrößert, damit für die Bildverarbeitung die komplette Kontur und ein Teil des Hintergrunds enthalten ist.

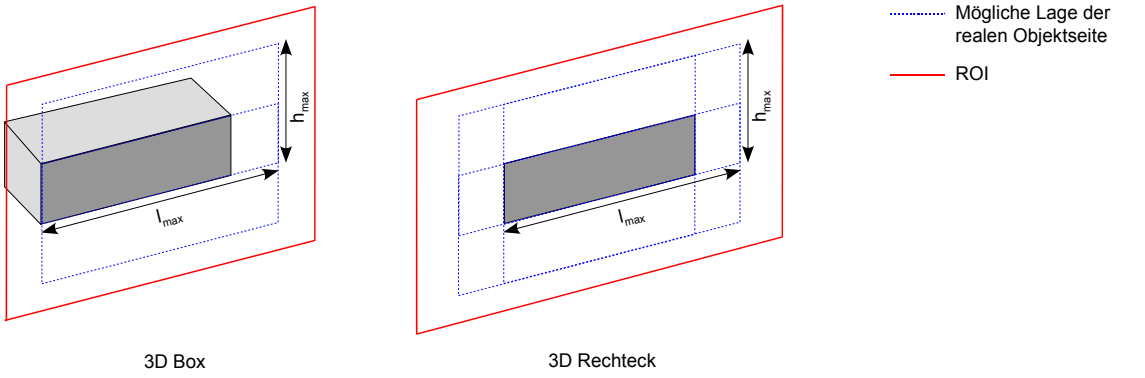

Abbildung 7.15: Bestimmung der ROI: Eine Seite der vom Laserscanner bestimmten dreidimensionalen Form stellt immer nur einen Teil der realen Objektseite dar. Diese ist deutlich größer und kann an verschiedenen Stellen lokalisiert sein. Die ROI wird so festgelegt, dass sie alle Extremfälle enthält und zusätzlich vergrößert, damit auch ein Teil des Hintergrunds enthalten ist.

Die zur Berechnung der ROI verwendete maximale Größe der realen Objektseite ist durch die maximale Länge und Höhe eines PKWs definiert, wenn nach Seitenansichten gesucht wird. Sollen Front- oder Heckansichten detektiert werden, ist die Größe der Objektseite durch die maximale Breite und Höhe eines PKWs bestimmt.

Die im dreidimensionalen Raum berechnete ROI kann im Anschluss in die Bildebene der Videokamera transformiert werden (Abbildung 7.16). Auf diese Weise wird ein direkter Bezug zu den Videodaten hergestellt. Befindet sich ein Großteil der transformierten ROI außerhalb des Bildes, kann davon ausgegangen werden, dass das Objekt nicht mit der Kamera erfasst wurde. In diesem Fall wird das Objekt ignoriert.

Abbildung 7.16: Die in das Videobild transformierte ROI enthält die perspektivisch verzerrte Ansicht einer Seite des vermessenen Objektes.

7.4.4 Perspektivische Rekonstruktion

Die in die Bildebene transformierte ROI enthält die perspektivisch verzerrte Seitenansicht des vermessenen Objektes (Abbildung 7.16). Um zu untersuchen, ob es sich dabei um eine Seiten-, Front- oder Heckansicht eines PKWs handelt, muss die Verzerrung kompensiert werden. Hierfür wird eine perspektivische Transformation des Bildinhaltes in der ROI durchgeführt (Abbildung 7.17). Zu diesem Zweck wurden bestehende Algorithmen der Intel Open Computer Vision Library [63] verwendet.

Bei der Transformation wird die Größe des Zielbildes so gewählt, dass ein Meter der in der Fahrzeugumgebung dreidimensional berechneten ROI auf eine feste Anzahl von Pixeln im Zielbild abgebildet wird. Auf diese Weise werden durch die Transformation automatisch Größenunterschiede kompensiert, die im Videobild durch unterschiedliche Abstände zwischen Objekt und Kamera entstehen.

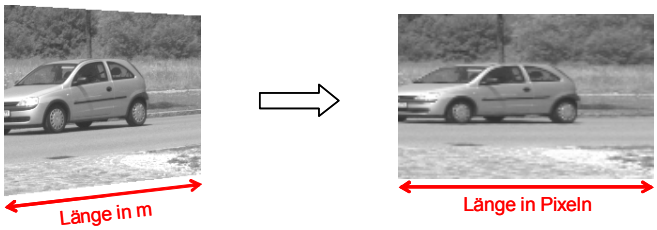

Abbildung 7.17: Perspektivische Rekonstruktion: Mit Hilfe einer perspektivischen Transformation des Bildinhaltes der ROI wird die korrespondierende Seite rekonstruiert. Dabei wird die Größe des Zielbildes so gewählt, dass automatisch abstandsbedingte Größenunterschiede verschiedender Objekte kompensiert werden.

Bei der perspektivischen Rekonstruktion der Objektseite entstehen an verschiedenen Stellen Deformationen. Dies ist dadurch begründet, dass die rekonstruierten Objekte in der Praxis nicht die angenommene Quaderform aufweisen. Bei PKWs fallen diese Deformationen relativ gering aus. Darum können die rekonstruierten Ansichten von PKWs trotz der auftretenden Fehler sehr gut von den verwendeten Kaskadenklassifikatoren detektiert werden.

Speziell Objekte, die keine PKWs darstellen, weichen oft enorm von der angenommenen Quaderform ab. Dadurch können bei der perspektivischen Rekonstruktion beliebig starke Deformationen entstehen. In der Praxis ist jedoch nicht zu erwarten, dass dies zu Ansichten führt, die einem PKW gleichen.

7.4.5 Anwendung des Kaskadenklassifikators

Für die Detektion der PKWs wurden zwei Kaskadenklassifikatoren trainiert, die auf die rekonstruierten Objektseiten angewendet werden. Untersuchungen haben gezeigt, dass ein Klassifikator die Detektion von Seitenansichten der linken und der rechten Fahrzeugseite übernehmen kann. Weiterhin genügt ein Klassifikator zur Detektion von Heck- und Frontansichten.

Bei der perspektivischen Rekonstruktion wird ein Meter der realen Objektgröße in eine feste Anzahl von Pixeln transformiert. Somit lässt sich aus dem entsprechenden Verhältnis und der erwarteten realen Objektgröße die zu erwartende Objektgröße im transformierten Bild berechnen. Es müssen mit dem Detektor also nur sehr wenige Skalierungen durchlaufen werden, um reale Größenunterschiede verschiedener PKWs zu kompensieren. Die deutlich signifikanteren Größenunterschiede im Bild, die aus unterschiedlichen Distanzen zum Sensor entstehen, werden automatisch kompensiert.

Aufgrund der stark eingeschränkten ROI müssen auch nur wenige Positionen durchsucht werden. Dadurch benötigt das verwendete System zur Detektion von PKWs relativ wenig Rechenzeit.

In Abbildung 7.18 sind Beispieldetektionen der angewendeten Kaskadenklassifikatoren dargestellt. Wenn innerhalb einer ROI mehrere Hypothesen bestätigt wurden, wurden diese vor der Ausgabe zusammengefasst.

Abbildung 7.18: Hier sind die Detektionsergebnisse der angewendeten Kaskaden für PKWs dargestellt. Auftretende Mehrfachdetektionen wurden zusammengefasst.

7.5 Detektion von Fußgängern

In diesem Abschnitt wird die Anwendung der vorgestellten Ansätze zur Detektion von Fußgängern beschrieben.

7.5.1 Erscheinungsbild von Fußgängern

Für Fußgänger kann aus den Messdaten des Laserscanners in der Regel keine ausreichend genaue Orientierung ermittelt werden. Darüber hinaus gleicht die Seitenansicht eines Fußgängers stärker der frontalen Ansicht. In der Vergangenheit wurden bereits erfolgreich Kaskadenklassifikatoren auf das Problem der Fußgängerdetektion angewendet [91, 142]. Aus diesen Gründen konnte auf eine perspektivische Transformation des Bildausschnittes für die Detektion von Fußgängern verzichtet werden.

7.5.2 Objektauswahl

Objekte, die die maximale horizontale Ausdehnung eines Fußgängers von 1,50 m in den Messdaten des Laserscanners überschreiten, werden in den weiteren Schritten der Fußgängerdetektion ignoriert. Weiterhin werden Objekte ignoriert, die weiter als 50 m entfernt sind. Diese sind bei der Benutzung des verwendeten Weitwinkelobjektivs für die Bildverarbeitung nicht mehr ausreichend aufgelöst.

7.5.3 Berechnung der ROI

Die ROI für die Detektion der Fußgänger wird ebenfalls im dreidimensionalen Raum berechnet, damit ihre Größe direkt aus der realen Maximalgröße eines Fußgängers abgeleitet werden kann. Reale unterschiedliche Personengrößen werden später durch die Abtastung in verschiedenen Skalierungen berücksichtigt.

Die ROI muss parallel zur Bildebene definiert werden, da in der Bildverarbeitung ohne Auswertung der Orientierung direkt im Originalbild nach Fußgängern gesucht werden soll. Dadurch wird die für die Detektion von PKWs notwendige zusätzliche Transformation des Bildausschnittes vermieden.

Für die Berechnung der gesuchten ROI wird die aus den Messdaten des Laserscanners bestimmte dreidimensionale Form in das dreidimensionale Koordinatensystem der Kamera transformiert. Hier ist die xy-Ebene parallel zur Bildebene der Kamera ausgerichtet. Nun wird ein zur xy-Ebene paralleles Rechteck aus den vom Laserscanner gemessenen Flächen bestimmt. Dieses wird jeweils durch die Minima und Maxima der x-Koordinaten und y-Koordinaten der Flächen definiert (Abbildung 7.19, links).

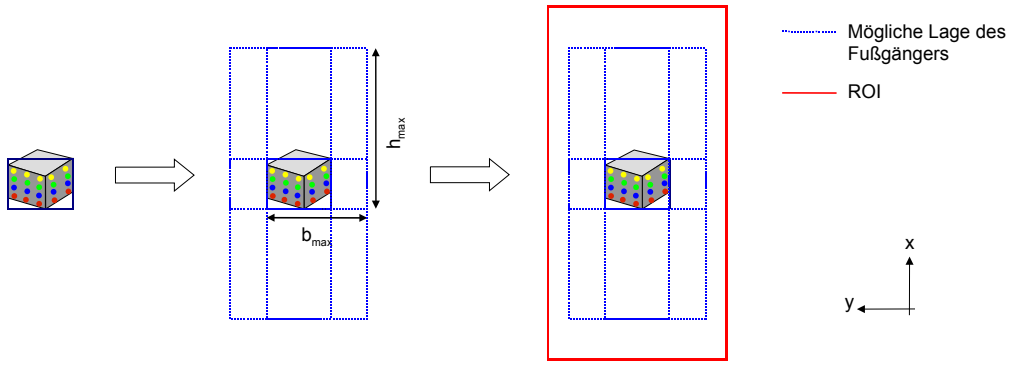

Abbildung 7.19: ROI zur Fußgängerdetektion: Die aus den Laserscannermessdaten bestimmte Form wird in das dreidimensionale Kamerakoordinatensystem transformiert. Hier wird ein zur Bildebene paralleles Rechteck bestimmt, dass die vom Laserscanner gemessenen Seiten enthält. Dieses wird so vergrößert, dass es mit Sicherheit einen korrespondierenden Fußgänger enthalten würde. Die ROI entsteht durch eine zusätzliche Vergrößerung um je 15% in alle Richtungen.

Dieses Rechteck ist mit Sicherheit Teil des beobachteten Objektes. Nun wird wie bei der Detektion der PKWs das Rechteck so vergrößert, dass ein Fußgänger mit Sicherheit darin liegt. Dafür wird die Maximalbreite und -höhe eines Fußgängers verwendet (Abbildung 7.19, Mitte). Schließlich wird das resultierende Rechteck wiederum in alle Richtungen um 15% vergrößert, um der Bildverarbeitung die komplette Kontur und einen Teil des Hintergrundes präsentieren zu können (Abbildung 7.19, rechts). Die auf diese Weise berechnete ROI wird nun in die Bildebene transformiert und erhält dort somit automatisch die passende Größe.

7.5.4 Anwendung des Kaskadenklassifikators

Zur Detektion von Fußgängern wurde lediglich ein Kaskadenklassifikator trainiert. Eine Verwendung von spezialisierten Kaskadenklassifikatoren, die zum Beispiel nur frontal aufgenommene Fußgänger oder nur Seitenansichten detektieren, war nicht notwendig. Der Kaskadenklassifikator wird direkt auf dem Originalbild angewendet. Die für PKWs notwendige Transformation der Bilddaten konnte hier vermieden werden. Da die Größe der Fußgänger im Originalbild von der Entfernung zu Kamera abhängt, muss die zu erwartende Größe der Fußgänger im Bild erst berechnet werden. Mit Hilfe des Lochkamera-Modells [34] kann für die zu erwartende Größe folgende Formel hergeleitet werden:

$$h_{Bild} = \frac{h_{Objekt}}{d_{Objekt}} f \ . \tag{7.9}$$

Hierbei entspricht f der Brennweite der Kamera, h_{Objekt} der realen Fußgängergröße und d_{Objekt} der Distanz des Fußgängers zur Kamera. Die Distanz des Fußgängers lässt sich aufgrund der Entfernungsmessungen des Laserscanners und des kalibrierten Sensoraufbaus bestimmen. Demzufolge fehlen lediglich genauere Angaben zur realen Größe des Fußgängers. Aus diesem Grund muss die Detektion in der ROI für reale Personengrößen von kleinen Kindern bis hin zu großen Erwachsenen in mehreren Skalierungen durchgeführt werden. Die im Videobild überprüfte Höhe wird dabei jeweils entsprechend Formel 7.9 berechnet. Die überprüfte Breite entspricht dann jeweils der halben Höhe, da dieses feste Verhältnis auch für die beim Training des Kaskadenklassifikators präsentierten Beispieldaten gewählt wurde.

Aufgrund der Berechnung der zu erwartenden Größe im Videobild und der Positionseinschränkungen durch die ROI müssen auch bei der Fußgängerdetektion nur wenige Skalierungen und Positionen überprüft werden. Abbildung 7.20 zeigt Beispiele für die Detektionsergebnisse der Klassifikatorkaskade. Auftretende Mehrfachdetektionen wurden vor der Ausgabe zusammengefasst.

Abbildung 7.20: Detektionsergebnisse für Fußgänger: Die Bilder enthalten Beispiele detektierter Fußgänger. Dabei wurden Mehrfachdetektionen des gleichen Objektes vor der Ausgabe zusammengefasst. Die Detektion im Bild ganz rechts zeigt, dass auch leichte Verdeckungen vom Kaskadenklassifikator toleriert werden.

7.6 Merkmalsextraktion

Die vorgestellten Kaskadenklassifikatoren erzielen bei der Detektion von Fußgängern und PKWs bereits hervorragende Ergebnisse. Darüber hinaus soll jedoch ein konsistentes Fahrzeugumfeldmodell erstellt werden, das alle Objektklassen enthält und nach Möglichkeit auch Objekte korrekt klassifiziert, die sich nicht im Messbereich der Kamera befinden. Aus diesem Grund müssen aus den angewendeten Kaskadenklassifikatoren Merkmale extrahiert werden, die dann im Rahmen der Mustererkennung mit den bisher bestimmten Merkmalen der Laserscannerdatenverarbeitung und der Objektverfolgung fusioniert werden. Die Merkmalsextraktion berücksichtigt dabei für jedes Objekt die Anzahl der durchlaufenen Kaskadenstufen und vorhandene Mehrfachdetektionen.

7.6.1 Mehrfachdetektionen

Da Kaskadenklassifikatoren meist verschiedene benachbarte Positionen und Skalierungen im Bild überprüfen, werden bei der Anwendung auf ein Bild der gesuchten Objektklasse meist mehrere Objekthypothesen von der Kaskade bestätigt (Abbildung 7.21). Eventuell auftretende Falschalarme bestehen in der Regel nur aus einer oder sehr wenigen bestätigten Objekthypothesen.

7.6.2 Merkmalsberechnung aus dem Kaskadenklassifikator

Aus jedem Kaskadenklassifikator, der auf die ROI eines Objektes angewendet wird, wird ein Merkmal für dieses Objekt berechnet. Die Merkmalsberechnung gliedert sich in zwei Stufen.

Bei der Anwendung eines Kaskadenklassifikators durchlaufen verschiedene Objekthypothesen nicht erfolgreich detektierter Objekte meist auch eine Vielzahl der Kaskadenstufen und werden erst in einer der letzten Stufen verworfen. Aus diesem Grund wird der primäre Merkmalswert über die letzte Kaskadenstufe bestimmt, die von mindestens einer

Objekthypothese passiert wird. Werden alle Stufen passiert, dann wird das Merkmal mit 10 auf den höchsten Wert gesetzt. Falls die letzte Stufe alle verbleibenden Hypothesen verwirft, also einzelne Hypothesen maximal die vorletzte Stufe passieren, beträgt der Merkmalswert 9. Mit jeder weiteren Stufe, die nicht passiert werden kann, wird das Merkmal um 1 reduziert. Als untere Grenze des Merkmals fungiert dabei der Wert 0 (Abbildung 7.22). Die genaue Wahl der Werte ist dabei willkürlich, da eine Skalierung der Merkmale automatisch vom Mustererkennungssystem vorgenommen wird.

Abbildung 7.21: Mehrfachdetektion: Bei der Anwendung von Kaskadenklassifkatoren auf ein Bild der gesuchten Objektklasse werden in der Regel mehrere Objekthypothesen bestätigt.

Zu dem über die höchste passierte Kaskadenstufe ermittelten Basiswert des Merkmals wird ein zusätzlicher Offset berechnet. Dieser Offset berechnet sich aus der Anzahl der Objekthypothesen, die diese Kaskadenstufe passieren. Jede Objekthypothese erhöht den Merkmalswert um jeweils 0,05. Bei der Offsetberechnung werden jedoch maximal 20 Objekthypotesen berücksichtigt, so dass der Wert des Offsets im Bereich (0,1] liegt.

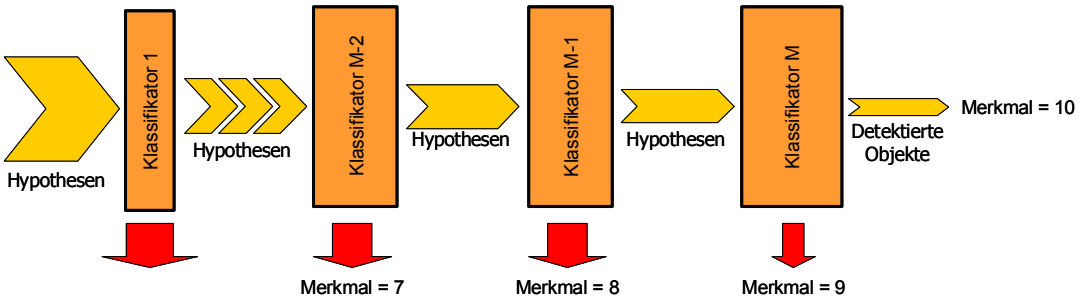

Abbildung 7.22: Merkmalsextraktion aus der Kaskade: Der primäre Merkmalswert wird aus der letzten Kaskadenstufe bestimmt, die von den Objekthypothesen einer ROI passiert wird. Wird die letzte Stufe passiert, ist der Wert 10. Für jede Stufe, die nicht passiert wird, reduziert sich der Wert um 1. Der minimale Merkmalswert ist jedoch 0.

7.6.3 Verwendete Merkmale

Für jedes Objekt werden insgesamt drei Kaskadenklassifikatoren ausgeführt, je einer für Fußgänger, Seitenansichten von PKWs sowie Front- beziehungsweise Heckansichten von PKWs. Für jede Kaskade werden zwei Merkmale bestimmt. Das erste Merkmal entspricht dabei jeweils dem im vergangenen Abschnitt beschriebenen Wert. Falls ein Klassifikator aufgrund der Objektauswahl nicht auf ein Objekt angewendet werden kann, wird das entsprechende Merkmal auf den Wert -1 gesetzt. Das zweite Merkmal ist das zeitliche Maximum des ersten Merkmals. Auf diese Weise können zum Beispiel Zeitschritte überbrückt werden, in denen starke Verdeckungen die Detektion im Videobild verhindern.

Die bisher beschriebenen Merkmale charakterisieren PKWs getrennt für Seitenansichten und Front- beziehungsweise Heckansichten. Weitere zwei Merkmale führen eine Kombination der Ergebnisse beider Kaskadenklassifikatoren durch, indem jeweils das Maximum der aktuellen Merkmalswerte beider Kaskadenklassifikatoren und das zeitliche Maximum dieser Werte ermittelt wird.

Schließlich geben zwei redundante Merkmale an, ob die Videobildverarbeitung für PKWs und Fußgänger stattgefunden hat. Diese Merkmale haben jeweils den Wert 0, wenn das Objekt in den bisherigen Zeitschritten nicht mit den entsprechenden Kaskadenklassifikatoren untersucht wurde. Sie werden mit dem Wert 1 belegt, wenn in vorherigen Zeitschritten ein Kaskadenklassifikator angewendet wurde, dies aber im aktuellen Zeitschritt nicht möglich war. Wurde im aktuellen Zeitschritt ein Kaskadenklassifikator angewendet, ergibt sich der Merkmalswert 2.

7.7 Zusammenfassung

In diesem Kapitel wurde die Berechnung neuer Merkmale mit Hilfe von Ansätzen zur Fusion von Laserscanner- und Videomessdaten vorgestellt. Die beschriebenen Methoden erlauben die Integration der Bildverarbeitung mit Kaskadenklassifikatoren in das Gesamtkonzept zur Objektklassifikation. Diese erfolgt durch den gemeinsamen Einsatz der bisherigen Merkmale und der neu berechneten Merkmale in dem in Kapitel 6 beschriebenen Klassifikationsframework. Die Methoden wurden für Fußgänger und PKWs umgesetzt. Eine besondere Neuerung ist dabei die dreidimensionale Modellierung von PKWs bei der Bildverarbeitung, die eine Detektion von Objekten mit beliebiger Ausrichtung zur Kamera erlaubt. Die erhebliche Verbesserung der Klassifikationsleistung durch die vorgestellten Fusionansätze mit Videomessdaten zeigt die quantitative Analyse in Kapitel 10.

8 Fusionsansätze mit präzisen digitalen Karten

In diesem Kapitel werden neue Methoden zur Verbesserung der Objektverfolgung und der Klassifikation unter Verwendung präziser digitaler Karten vorgestellt.

Zur Verbesserung der Objektverfolgung können in den digitalen Karten Gebiete markiert werden, in denen die Objektverfolgung nicht angewendet werden muss. Solche Gebiete sind zum Beispiel Häuser oder eingezäunte Bereiche. Alle Objekte in solchen Gebieten sind für Fahrerassistenzsysteme nicht von Interesse, da eine Kollision mit ihnen ausgeschlossen werden kann. Eine Einschränkung der Objektverfolgung auf die übrigen Gebiete ermöglicht eine Reduktion der benötigten Rechenzeit. Gleichzeitig kann durch die Reduktion der möglichen Zuordnungen in der Datenassoziation das Risiko einer Fehlassoziation reduziert werden.

Erhebliche Verbesserungen können bei der Klassifikation durch die Anwendung von a priori Wissen über mögliche Objektklassen in definierten Gebieten im vermessenen Fahrzeugumfeld erreicht werden. Ein Beispiel für die Nutzung von a priori Wissen ist in Abbildung 8.1 gegeben. Hier sind zwei Objekte abgebildet, deren von dem Laserscanner erfasste Objektzustand bis auf die Position identisch ist. Diese Objekte werden aufgrund ihrer Position auf beziehungsweise neben der Straße unterschiedlich klassifiziert. Das Objekt auf der Straße kann kein Hintergrundobjekt sein und wird aufgrund seiner Größe als LKW klassifiziert. Bei Objekten neben der Straße wird ausgeschlossen, dass es sich um PKWs oder LKWs handelt. Darum wird dieses Objekt als Hintergrund klassifiziert.

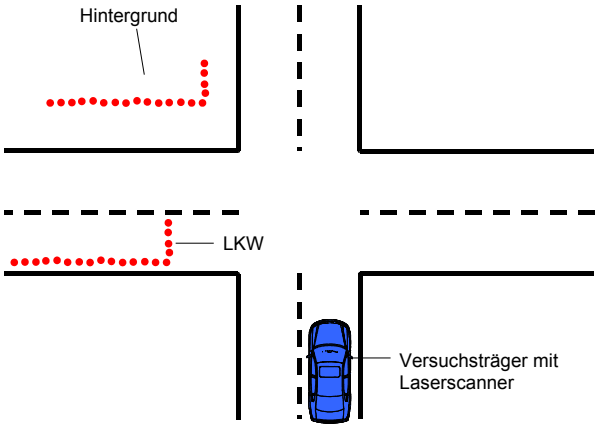

Abbildung 8.1: Anwendung von a priori Wissen aus Karten: Zwei in den Sensormessdaten bis auf die Position identisch erscheinende Objekte werden aufgrund ihrer Position unterschiedlich klassifiziert. Das Objekt auf der Straße muss aufgrund seiner Größe ein LKW sein. Das Objekt neben der Straße kann kein Fahrzeug sein und wird als Hintergrund klassifiziert.

8.1 Karteninhalt

Mit einem neuen Kartografierverfahren wurden verschiedene Kreuzungen vermessen und präzise digitale Karten erstellt [48, 147, 149, 152]. Alle Straßen und Parkplätze werden als

Polygone gespeichert. Das System erstellt aus diesen Polygonen automatisch ein großes, kombiniertes Polygon. Gebiete wie Verkehrsinseln, die in diesem großen Polygon liegen, werden in Form zusätzlicher Polygone gespeichert. Diese Gebiete erhalten den gleichen Status wie Gebiete, die nicht auf dem Straßenpolygon liegen. Zusätzlich werden Gebiete vermerkt, die von den Fahrerassistenzsystemen ignoriert werden können. Abbildung 8.2 zeigt symbolisch eine Beispielkarte für eine Kreuzung.

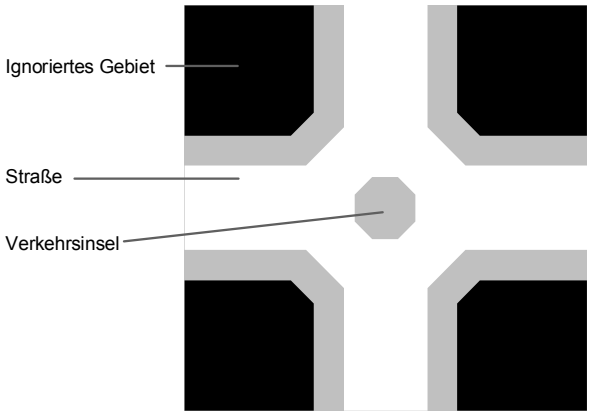

Abbildung 8.2: Symbolische präzise Karte: In den verwendeten Karten werden Straßen und Gebiete eingetragen, die von Algorithmen der Fahrzeugumfelderfassung ignoriert werden können. Zusätzlich werden Verkehrsinseln vermerkt, die den gleichen Status erhalten wie Gebiete, die sich nicht auf der Straße befinden.

Zusätzlich zu den unterschiedenen Gebieten werden in der digitalen Karte für die spätere Positionierung spezielle Landmarken mit einer geringen horizontalen Ausdehnung gespeichert. Hierzu eignen sich beispielsweise Bäume oder die Pfeiler von Schildern oder Ampeln (Abbildung 8.3) [48, 148].

Abbildung 8.3: Für eine spätere Positionierung werden die Positionen von Landmarken mit einer geringen horizontalen Ausdehnung in der Karte festgehalten. Sinnvolle Landmarken sind zum Beispiel Pfeiler von Schildern oder Ampeln.

Abbildung 8.4 zeigt die in der digitalen Karte gespeicherten Informationen. Momentan werden die Polygone und die Landmarken in einem eigenen Format gespeichert. Eine

Speicherung in den in Serienfahrzeugen verwendeten Navigationskarten ist prinzipiell ebenfalls denkbar. Diese Navigationsgeräte verwenden zur Zeit deutlich unpräzisere Karten im GDF-Format (*engl. GDF = geographic data file*). Die verwendeten Speicherstrukturen ermöglichen jedoch auch das Eintragen deutlich präziserer Informationen [26, 156]. Eine Umrechnung der Elemente im GDF-Format in eine Struktur von Polygonen ist ebenfalls möglich.

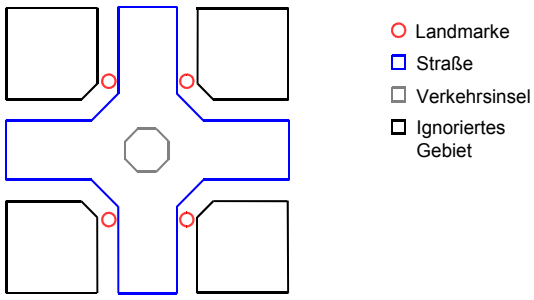

Abbildung 8.4: Kartenformat: Zusätzlich zu den für die Positionierung benötigten Landmarken werden bestimmte Bereiche in Form von Polygonen in die Karte eingetragen. Dazu zählen Straßenbereiche, darin enthaltene Verkehrsinseln und die zu ignorierenden Gebiete.

8.2 Positionierung

Um die Position gemessener Objekte in der Karte zu bestimmen, muss zunächst die Position und die Orientierung des eigenen Fahrzeugs relativ zur Karte bestimmt werden. Hierzu wird ein Algorithmus zur präzisen Positionierung basierend auf Landmarken verwendet, der in [146, 147, 148, 150, 151] vorgestellt und im Detail beschrieben wurde. Das Grundprinzip ist in Abbildung 8.5 dargestellt. Vom DGPS-Sensor werden grobe Positions- und Orientierungsinformationen ermittelt. Gleichzeitig werden mit dem Laserscanner verschiedene Landmarken gemessen. Der Positionierungsalgorithmus verschiebt nun die Position und dreht die Orientierung in der Karte derart, dass gemessene Landmarken möglichst gut auf korrespondierenden Landmarken in der Karte liegen.

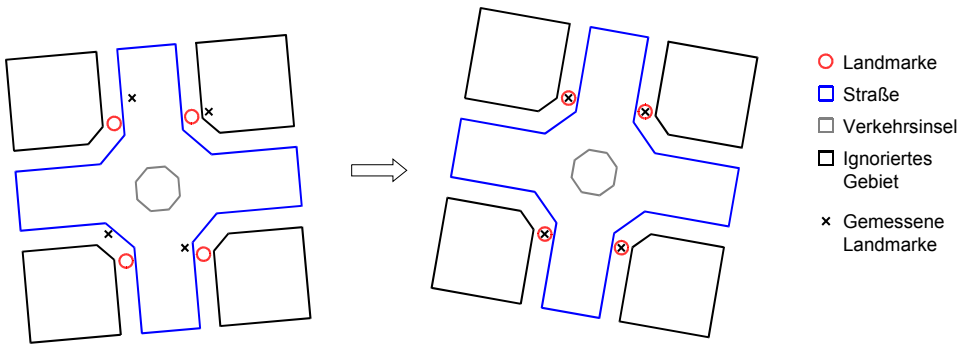

Abbildung 8.5: Positionierung: Grobe Positions- und Orientierungsinformationen werden vom DGPS-Sensor bestimmt. Diese werden so angepasst, dass vom Laserscanner gemessene Landmarken auf Landmarken in der Karte abgebildet werden.

8.3 Unsicherheiten

Die Positionsbestimmung ist mit Unsicherheiten behaftet. Diese werden in Form einer Positionsunsicherheit und einer Rotationsunsicherheit angegeben. Zur Rotationsunsicherheit wird ein entsprechender Rotationspunkt angegeben. Bei der Positionierung mit Hilfe der vom Laserscanner vermessenen Landmarken liegt der Rotationspunkt sinnvollerweise etwa mittig zwischen den Landmarken. In manchen Anwendungen sind jedoch auch andere Rotationspunkte sinnvoll. Bei Fusionsansätzen im Rahmen des EU-Projektes INTERSAFE [46, 62] erwies sich aufgrund der technischen Randbedingungen zum Beispiel die Position des Versuchsträgers als geeigneter Rotationspunkt.

8.3.1 Positionsunsicherheit

Die Positionsunsicherheit bedeutet, dass die korrekte Kartenposition eine gewisse Verschiebung zur berechneten Position aufweisen kann. Verschiebt man die Karte an alle potentiell möglichen Positionen, entstehen drei Gebiete bezüglich einer gespeicherten Struktur. Dieser Sachverhalt wird am Beispiel der Straße in Abbildung 8.6 dargestellt. Der mittlere Bereich, der an allen möglichen Positionen von der Struktur bedeckt wird, befindet sich mit Sicherheit auf der Straße. Der Bereich, der nie von der Struktur bedeckt werden kann, ist mit Sicherheit außerhalb der Straße. Der restliche Bereich kann auf oder neben der Straße liegen.

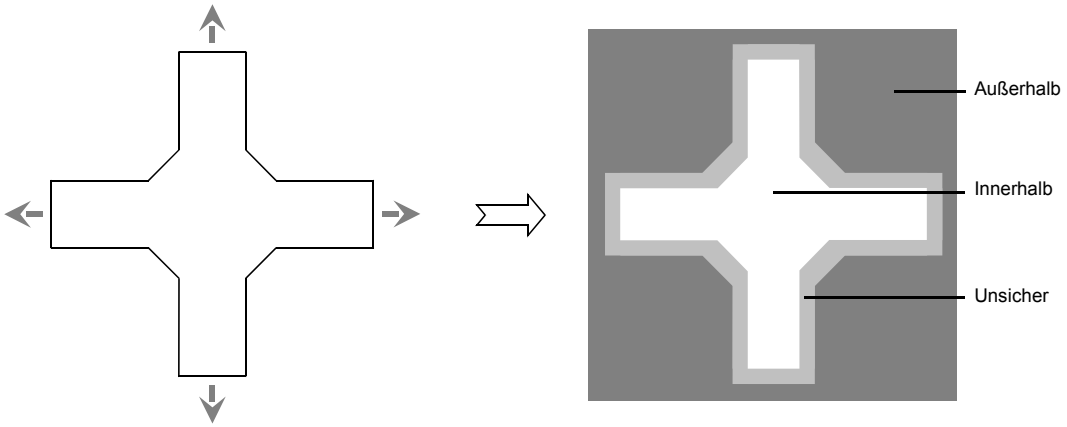

Abbildung 8.6: Positionsunsicherheit: Die reale Position der Straße kann bezüglich der berechneten Position leicht verschoben sein. Dadurch entstehen drei Bereiche bezüglich der Straße. Der innere Bereich liegt mit Sicherheit auf der Straße. Der äußere liegt mit Sicherheit daneben. Die Lage des mittleren Bereiches bezüglich der Straße ist unsicher.

8.3.2 Rotationsunsicherheit

Die Rotationsunsicherheit entspricht einer maximalen Rotation zwischen der berechneten Karte und der realen Lage der Strukturen. Hier liegt beim Beispiel der Straße der innere Bereich, der während der gesamten Rotation vom Polygon bedeckt wird, mit Sicherheit auf der Straße. Der äußere Bereich liegt mit Sicherheit nicht auf der Straße. Der Bereich dazwischen liegt teilweise auf und teilweise neben der Straße (Abbildung 8.7).

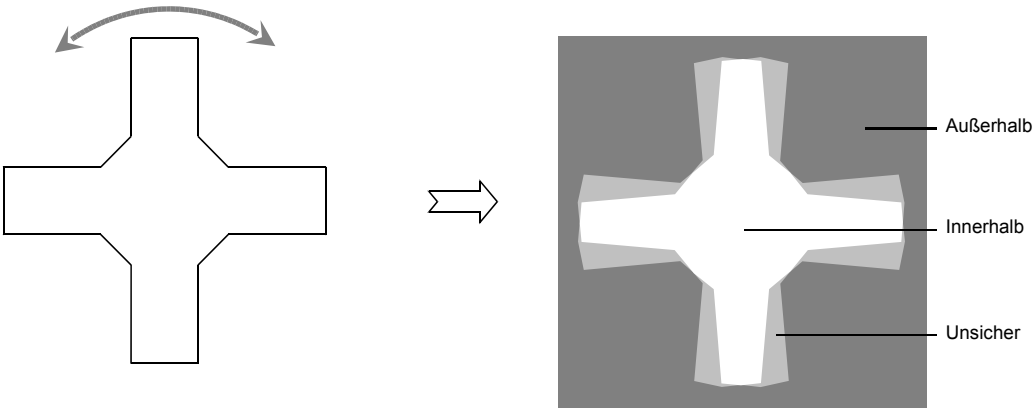

Abbildung 8.7: Rotationsunsicherheit: Auch die Unsicherheit in der Orientierung unterteilt das Fahrzeugumfeld für jede Kartenstruktur in drei Bereiche. Der mittlere liegt auf, der äußere neben der Struktur. Die Lage des mittleren Bereichs ist unsicher.

8.4 Kombination der Unsicherheiten

In der Praxis müssen beide Unsicherheiten kombiniert werden. Dabei werden die von den Positionierungsalgorithmen übermittelten Unsicherheiten als unkorreliert angenommen. Es wird ein 3σ-Bereich der übermittelten Standardabweichung verwendet. Die resultierende Fläche wird analytisch berechnet. Aus den in der Karte gespeicherten, polygonförmigen Strukturen werden verschiedene Flächen ermittelt. Diese Berechnungen werden in jedem Zeitschritt direkt nach der Positionierung durchgeführt. Der dazu entwickelte Algorithmus ist in Anhang B beschrieben. Das Ergebnis ist in Abbildung 8.8 dargestellt.

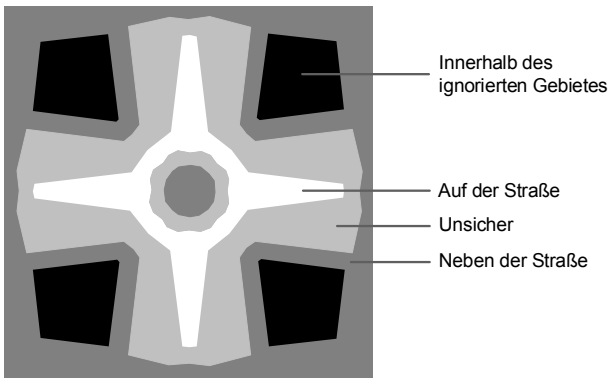

Abbildung 8.8: Ergebnisse der Anwendung der Unsicherheiten: Es wurden Bereiche ermittelt, die mit Sicherheit auf und neben der Straße liegen. Der dazwischen liegende Bereich ist unsicher. Zusätzlich wurden alle Bereiche berechnet, die mit Sicherheit ignoriert werden können.

8.5 Anwendung der Karteninformationen

Nach der Kombination der Unsicherheiten können einzelne Messdaten oder Objekte aufgrund der Position bestimmten Bereichen zugeordnet werden. Entsprechend erkannte Zuordnungen werden für verschiedene Modifikationen ausgenutzt (Abbildung 8.9).

Zunächst werden die berechneten Bereiche der Vorverarbeitung der Laserscannermessdaten übergeben. Dort werden vor allen Verarbeitungsschritten sämtliche Entfernungsmessungen entfernt, die mit Sicherheit in Gebieten liegen, die für Fahrerassistenzsysteme nicht von Interesse sind. Anschließend erfolgt die übliche Vorverarbeitung, Merkmalsextraktion und Datenassoziation.

Nach der Datenassoziation wird für alle Objekte in Form von Merkmalen vermerkt, auf welchen Bereichen sie sich befinden. Diese Merkmale werden im regelbasierten Teil der Klassifikation verwendet. Prinzipiell wäre es sinnvoll, auch diese neuen Merkmale der Mustererkennung zu präsentieren. Das neue Verfahren zur Positionierung benötigt allerdings immer eine ausreichende Menge von Landmarken in den kartografierten Gebieten. Aus diesem Grund wurden nur einige Kreuzungen kartografiert. Da die aus der Karte extrahierten Merkmale demzufolge nur in einem geringen Teil der Beispieldaten vorliegen, kann die Mustererkennung darauf nicht geeignet trainiert werden.

Der regelbasierte Teil der Klassifikation verändert die Zugehörigkeitswerte der Objekte zu den einzelnen Objektklassen basierend auf a priori Wissen. Alle Objekte, die sich mit Sicherheit auf der Straße aber nicht auf einer Verkehrsinsel befinden, können keine Hintergrundobjekte sein. Aus diesem Grund wird der entsprechende Zugehörigkeitswert auf Null gesetzt. Auch die gespeicherten Zugehörigkeitswerte der Klasse „Hintergrund“ für den zeitlichen Mittelwertfilter werden auf Null gesetzt. Alle Objekte die mit Sicherheit auf einer Verkehrsinsel oder nicht auf der Straße liegen, können keine PKWs oder LKWs sein. Die entsprechenden Zugehörigkeitswerte und auch die für den zeitlichen Mittelwertfilter gespeicherten Werte werden auf Null gesetzt.

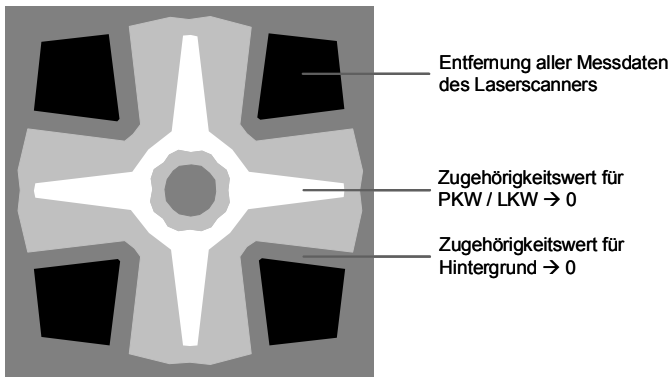

Abbildung 8.9: Nutzung der Karteninformationen: Alle Messdaten des Laserscanners, die mit Sicherheit in ignorierten Gebieten liegen, werden vor der Vorverarbeitung entfernt. Die Zugehörigkeitswerte der Klassen „PKW“ und „LKW“ werden für Objekte auf Null gesetzt, die sich mit Sicherheit auf einer Verkehrsinsel oder nicht auf der Straße befinden. Ähnliches gilt für den Zugehörigkeitswert der Klasse „Hintergrund“ für Objekte, die sich mit Sicherheit auf der Straße und nicht auf einer Verkehrsinsel befinden.

8.6 Zusammenfassung

In diesem Kapitel wurden Ansätze zur Fusion präziser digitaler Karteninformationen mit den Messdaten des Laserscanners vorgestellt. Die Ansätze profitieren von den verwendeten landmarkenbasierten Algorithmen zur präzisen Positions- und Orientierungsbestimmung in digitalen Karten. Somit sind die Fusionsansätze auf Gebiete beschränkt, die entsprechend präzise kartografiert wurden. Das deutliche Verbesserungspotenzial der Fusionsalgorithmen wird in Kapitel 10 durch quantitative Analysen belegt.

9 Nutzung drahtloser Datenübertragung

In diesem Kapitel werden neue Ansätze zur Erweiterung des Fahrzeugumfeldmodells durch drahtlose Datenübertragung präsentiert. Da sich bereits eine Vielzahl von Arbeiten mit möglichen Übertragungsmedien und -protokollen beschäftigen, wurden zu diesen Themen keine neuen Untersuchungen durchgeführt. Statt dessen konzentriert sich dieses Kapitel auf Methoden zur Integration der empfangenen Objektdaten in das sensorbasierte Fahrzeugumfeldmodell.

Bisherige Arbeiten verwenden drahtlose Datenübertragung zur Warnung vor Gefahrenquellen, wie zum Beispiel schlechtem Wetter oder Staus. Diese können mit Hilfe der globalen Position lokalisiert werden. Eine grobe Position relativ zum Fahrzeug kann dann mit Hilfe eines GPS-Sensors bestimmt werden. Die für derartige Ansätze verwendete grobe Position lässt sich ohne Komplikationen in das Fahrzeugumfeldmodell integrieren.

Das in dieser Arbeit vorgestellte Verfahren konzentriert sich auf die Übertragung von Informationen über bewegliche Objekte und deren Integration in das sensorbasierte Fahrzeugumfeldmodell. Für die Integration wird ein allgemeiner Ansatz vorgestellt, der für eine Vielzahl von Anwendungsfällen einsetzbar ist.

Objekte, die sich außerhalb des Sichtbereichs der Sensoren befinden, sollen als neue Objekte zum Fahrzeugumfeldmodell hinzugefügt werden. Interessanter ist die Fusion empfangener Objektinformationen mit Objektdaten, die von den Sensoren ermittelt wurden. Für diese Anwendung sind verschiedene Randbedingungen denkbar. Andere Verkehrsteilnehmer können ihre eigenen Daten drahtlos verschicken. Sie können aber auch die Daten von Objekten verschicken, die sie mit ihren Sensoren erfasst haben. Darüber hinaus ist auch die Erfassung anderer Objekte durch Sensoren in der Infrastruktur und die Übertragung der entsprechenden Daten möglich. Schließlich können Infrastrukturelemente wie Ampeln oder Schilder ihre eigene Position und ihren Status übermitteln. Ein weiterer Unterschied besteht in der Methode der Positionsbestimmung. Mit Hilfe von GPS- oder DGPS-Empfängern lässt sich die Position nur relativ ungenau bestimmen. Infrastrukturelemente kennen ihre Position und damit auch die der gemessenen Objekte deutlich präziser. Innovative Positionierungsalgorithmen, die GPS und Sensorinformationen fusionieren [148], können ebenfalls sehr präzise Positionsinformationen liefern.

Die Zuordnung der empfangenen Objektdaten zum sensorbasierten Objektmodell wird für die in dieser Arbeit verwendeten Sensoren untersucht. Die Ansätze sind jedoch prinzipiell auch auf andere Sensoren wie zum Beispiel RADAR übertragbar.

9.1 Motivation

Die Gründe für die Fusion der Sensormessdaten mit drahtlos empfangenen Objektdaten können anhand eines einfachen Beispiels für einen Kreuzungsassistenten erläutert werden. In Abbildung 9.1 befindet sich das Versuchsfahrzeug auf einer Nebenstraße an einer Kreuzung. Auf der abbiegenden Hauptstraße befindet sich ein weiteres Fahrzeug, das rechts abbiegen möchte und somit keine Gefahr für das Versuchsfahrzeug darstellt. Die Position, die Größe und die Geschwindigkeit des anderen Fahrzeugs können mit Hilfe der Sensoren des Versuchsfahrzeugs relativ präzise bestimmt werden. Verschiedene Informationen können mit

Sensoren allerdings nur sehr schwer, ungenau oder überhaupt nicht ermittelt werden. Hierzu zählt zum Beispiel die präzise Masse eines anderen Objekts, die für die Auslösung von Airbags im Fall einer Kollision von Bedeutung ist. Die Masse kann in der Regel nur mit Hilfe der Klassifikation grob abgeschätzt werden. Im aktuellen Beispiel kann die geplante Fahrtrichtung des anderen Fahrzeugs mit Hilfe der Sensoren nicht bestimmt werden. Aus diesem Grund müsste ein Kreuzungsassistent davon ausgehen, dass das andere Fahrzeug auf der Hauptstraße bleibt und die Vorfahrt gewährt werden muss. Das Ergebnis wäre eine überflüssige Warnung durch das Assistenzsystem.

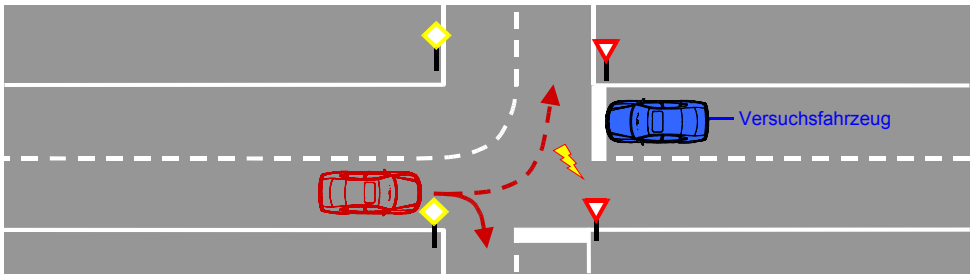

Abbildung 9.1: Sensoren können die Position, Größe und Geschwindigkeit anderer Fahrzeuge relativ präzise feststellen. Die geplante Fahrtrichtung lässt sich aber nur schwer ermitteln.

Betrachtet man das Szenario aus der Sicht der drahtlos übertragenen Daten, so kann das andere Fahrzeug einfach dem Versuchsfahrzeug mitteilen, dass es rechts abbiegen möchte, wenn der Fahrer zum Beispiel den Blinker gesetzt hat. Für Anwendungen, die nur auf drahtlosem Datenaustausch basieren, ist jedoch die Bestimmung der Position des anderen Fahrzeugs relativ zum Versuchsfahrzeug sehr schwierig. Die Positionen beider Fahrzeuge werden dann in der Regel mit DGPS-Sensoren gemessen. Der bei diesen Sensoren auftretende Positionsfehler von ein bis zwei Metern kann sich bei der Berechnung der relativen Position im Extremfall verdoppeln. Dadurch kann der Eindruck entstehen, dass das andere Fahrzeug sich auf der gleichen Spur wie das Versuchsfahrzeug befindet (Abbildung 9.2). Folglich entsteht aus der Sicht eines Assistenzsystems ebenfalls eine gefährliche Situation, da scheinbar eine Kollision bevorsteht.

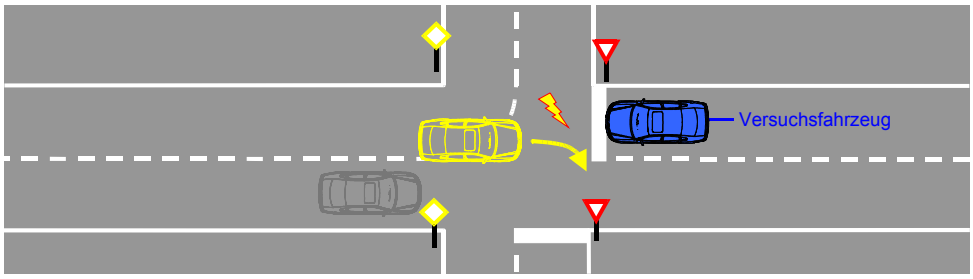

Abbildung 9.2: Über drahtlose Kommunikation können Informationen über gesetzte Blinker übertragen werden. Aufgrund der üblichen GPS-Messfehler lässt sich die Position aber nur auf einige Meter genau bestimmen. Dadurch können ungefährliche Situationen sehr gefährlich erscheinen.

Kombiniert man nun die sensorbasierten mit den drahtlos übertragenen Informationen in geeigneter Weise, sollte sich die im Beispiel dargestellte Situation korrekt erkennen lassen.

9.2 Übertragene Daten

Die zu übertragenden Objektdaten wurden unter der Zielstellung eines allgemein verwendbaren Ansatzes festgelegt. Objekte werden unabhängig voneinander übertragen.

Zur Identifikation des empfangenen Objekts muss eine eindeutige Objektnummer übertragen werden. Zu jedem Objekt wird pro Übertragung eine Position, eine Orientierung, eine Gierrate und eine Geschwindigkeit inklusive korrespondierender Unsicherheiten übermittelt. Die Position wird in WGS84-Koordinaten (*engl. World Geodetic System 1984*) angegeben. Dieses globale Koordinatensystem gibt die Position auf der Erde in Form von Längengrad λ, Breitengrad ϕ und Höhe h über dem Referenzellipsoid [106] an. Der Längengrad ist dabei in positiver Richtung nach Osten definiert. Der Breitengrad ist in positiver Richtung nach Norden ausgerichtet. Identische Koordinaten werden auch beim GPS verwendet. Die Orientierung ψ wird wie beim GPS bezüglich der Richtung Norden im Uhrzeigersinn angegeben. Die übertragene Gierrate ist dementsprechend auch im Uhrzeigersinn definiert. Im Gegensatz zur in globalen Koordinaten angegebenen Position wird die Geschwindigkeit in Metern pro Sekunde übertragen. Die Geschwindigkeitsrichtung ist über die Orientierung festgelegt. Zusätzlich wird jeweils eine UTC-Zeit (*engl. Coordinated Universal Time*) zur Synchronisierung übertragen. Diese globale Zeit kann ebenfalls vom GPS-Sensor ermittelt werden. Mit Hilfe der vorgeschlagenen Inhalte für die übertragenen Datenpakete lassen sich die verschiedenen Anwendungsfälle in einem gemeinsamen Ansatz realisieren. Unterschiedliche Datenquellen sowie unterschiedliche Positionierungsalgorithmen können durch die entsprechende Wahl der übertragenen Unsicherheiten berücksichtigt werden.

Die bisher aufgezählten Daten werden zur Lokalisierung der übertragenen Objekte benötigt. Darüber hinaus können diverse Zusatzinformationen für verschiedene Assistenzsysteme übertragen werden. Dazu zählen zum Beispiel die Objektklasse, die Größe, die Masse, Informationen über gesetzte Blinker oder aber auch die im Navigationssystem gewählte Route.

9.3 Ansatz zur Integration der übertragenen Daten

Die prinzipielle Vorgehensweise zur Integration der übertragenen Daten ist in Abbildung 9.3 dargestellt. Die Sensormessdaten werden, wie in den bisherigen Kapiteln beschrieben, verarbeitet. Damit eine Datenassoziation möglich ist, müssen die drahtlos empfangenen Daten auf den Messzeitpunkt der Sensordaten prädiziert werden. Weiterhin wird zur Berechnung der relativen Position empfangener Objekte auch die globale Position des eigenen Fahrzeugs benötigt. Da die Positionsbestimmung mit den verwendeten DGPS-Empfängern nur einmal pro Sekunde durchgeführt wird und zwischen Messzeitpunkt und Ausgabe eine signifikante Zeitspanne liegt, muss auch die eigene Position auf den Zeitpunkt der Sensormessdaten prädiziert werden. Nach einer Assoziation der empfangenen Objektdaten mit den durch Sensoren erfassten Objekten werden die Ergebnisse in das erweiterte Fahrzeugumfeld eingetragen.

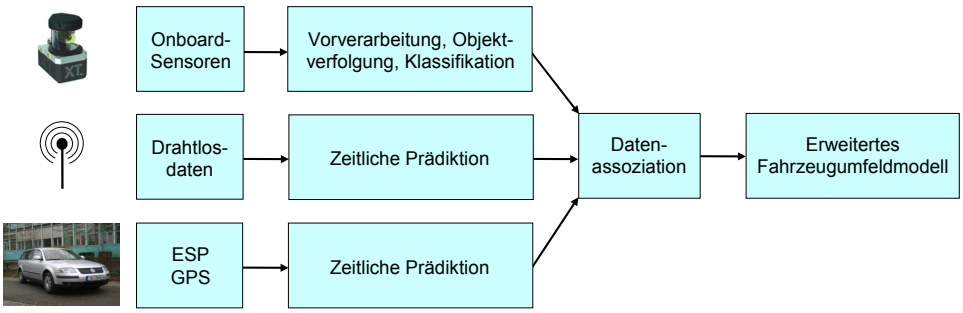

Abbildung 9.3: Ablauf der Integration der drahtlos übertragenen Daten: Die Objektdaten empfangener Objekte und die Position des Versuchsfahrzeugs werden auf den Zeitpunkt der Sensormessdaten prädiziert. Anschließend erfolgt die Assoziation mit den sensorbasierten Objektdaten.

9.4 Prädiktion des dynamischen Zustands

9.4.1 Motivation

Drahtlos empfangene Daten unterliegen einer signifikanten Latenz. Zum einen wird eine gewisse Zeitspanne für die Übertragung benötigt. Diese Zeitspanne kann stark variieren und ist von Faktoren wie Datenaufkommen, Distanz und Terrain abhängig. Für die hier durchgeführten Experimente wurde diese Zeitspanne mit 200 ms abgeschätzt. Zum anderen benötigt der DGPS-Empfänger eine Zeit von etwa 260 ms zur Berechnung und Ausgabe der Position. Darüber hinaus verzögert der verwendete Empfänger die Ausgabe um eine weitere Sekunde. Die Latenz der GPS-Messung tritt zusätzlich auch noch beim eigenen Fahrzeug auf.

Während dieser Zeit können sich PKWs um einige Meter von der gemessenen Position entfernen. Für die Datenassoziation wird eine relative Position berechnet, da das sensorbasierte Umfeldmodell Objekte relativ zum Versuchsfahrzeug speichert. Bei der Berechnung der relativen Position werden die latenzbedingten Positionsfehler kombiniert. Das Ergebnis ist ein signifikanter Unterschied zwischen der erwarteten Position empfangener Objektdaten und der korrespondierenden, mit Sensoren ermittelten Position (Abbildung 9.4).

Um eine Datenassoziation sinnvoll durchführen zu können, sollten diese Positionsunterschiede soweit wie möglich kompensiert werden. Aus diesem Grund wird der dynamische Zustand des übertragenen Objekts verwendet, um die erwartete Position und Orientierung auf den Zeitpunkt der Sensormessdaten zu prädizieren. Eine ähnliche Prädiktion wird für die Position und die Orientierung des Versuchsfahrzeugs durchgeführt.

Die zeitliche Prädiktion wird in dieser Arbeit mit einem Extended Kalman-Filter durchgeführt. Als Modell wird ein Einspurmodell in globalen WGS-84-Koordinaten verwendet.

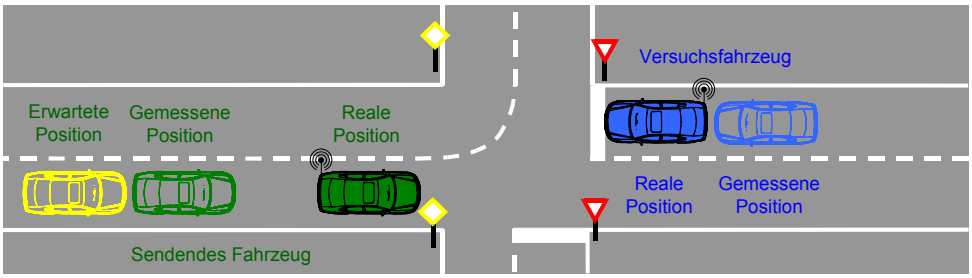

Abbildung 9.4: Einfluss verschiedener Latenzen auf die relative Positionsbestimmung für übertragene Objektdaten: Bei der Bestimmung des GPS-Messwerts und der Datenübertragung entstehen signifikante Latenzen. In dieser Zeit bewegen sich PKWs einige Meter. Aus den veralteten Messwerten würde eine erwartete, relative Position zum Versuchsfahrzeug berechnet, die sehr weit von der gemessenen Position des anderen Fahrzeugs entfernt wäre.

9.4.2 Einspurmodell in WGS84-Koordinaten

Zur Prädiktion der Objektdynamik wird ein Einspurmodell in globalen WGS84-Koordinaten verwendet. Dieses Modell muss die Geschwindigkeit in Metern pro Sekunde in eine globale Änderung der Länge λ und Breite φ umrechnen (Abbildung 9.5).

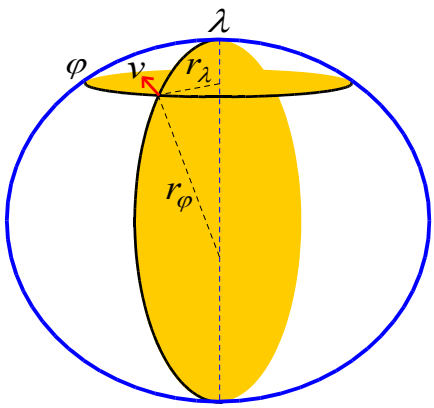

Abbildung 9.5: Zur Prädiktion der empfangenen Objektdaten und der eigenen Position auf den Messzeitpunkt der Sensordaten wird ein Einspurmodell in globalen WGS84-Koordinaten verwendet. Dieses Modell rechnet die Geschwindigkeit und die Orientierung in eine Änderung des Längen- und des Breitengrades um.

Die Berechnung der Positionsänderung erfolgt nach folgenden Gleichungen [106]:

$$\lambda\left(t+\Delta t\right) = \lambda\left(t\right) + \frac{\Delta t \cdot v \cdot \sin\psi}{r_\lambda} \tag{9.1}$$

$$\varphi\left(t+\Delta t\right) = \varphi\left(t\right) + \frac{\Delta t \cdot v \cdot \cos\psi}{r_\varphi}\,. \tag{9.2}$$

Der zugehörige Radius r_λ und der Ellipsenkrümmungsradius r_φ werden innerhalb eines Zeitschrittes als konstant angenommen und mit Hilfe der folgenden Formeln aus der aktuellen Zustandsschätzung berechnet:

$$r_\lambda = \left(\frac{a_{WGS84}}{\sqrt{1 - e^2_{WGS84} \sin^2 \varphi}} + h \right) \cos \varphi \ . \tag{9.3}$$

$$r_\varphi = \frac{a_{WGS84} \left(1 - e^2_{WGS84}\right)}{\left(1 - e^2_{WGS84} \sin^2 \varphi\right)^{\frac{3}{2}}} + h \tag{9.4}$$

$$e^2_{WGS84} = \frac{a^2_{WGS84} - b^2_{WGS84}}{a^2_{WGS84}} \tag{9.5}$$

Die Konstanten $a_{WGS84} \approx 6378\,\text{km}$ und $b_{WGS84} \approx 6357\,\text{km}$ beschreiben die Halbachsen des Erd-Ellipsoids im WGS84-System.

Für das Einspurmodell in globalen Koordinaten wird eine konstante Beschleunigung angenommen. Die Änderung der Beschleunigung wird als Rauschen modelliert. Die Schätzung der Beschleunigung ist hier im Vergleich zur Objektverfolgung sinnvoll, da in diesem Fall in der Regel Messwerte für die Objektgeschwindigkeit vorliegen. Die Beschleunigung muss somit nur als erste zeitliche Ableitung der Messwerte geschätzt werden.

Da sich die Höhe h in der Regel nicht wesentlich ändert, wird diese Größe nicht geschätzt. Es wird jeweils der letzte verfügbare Messwert verwendet. Der Zustandsvektor besteht demzufolge aus Länge λ, Breite φ, Geschwindigkeit v, Beschleunigung a, Orientierung ψ und Gierrate ω:

$$\mathbf{x} = \begin{bmatrix} \lambda & \varphi & v & a & \psi & \omega \end{bmatrix}^T . \tag{9.6}$$

Zur Herleitung des Prozessmodells wird die zeitliche Ableitung des Zustandsvektors verwendet:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \frac{v \cdot \sin \psi}{r_\lambda} \\ \frac{v \cdot \cos \psi}{r_\varphi} \\ a \\ 0 \\ \omega \\ 0 \end{bmatrix} . \tag{9.7}$$

Die zeitdiskrete Zustandsübergangsfunktion vom Zeitschritt k auf den Zeitschritt $k+1$ für eine Zeitspanne Δt wird durch die Integration der zeitlichen Ableitung berechnet[5]:

$$\mathbf{x}(k+1) = \mathbf{f_d}(\mathbf{x}(k), 0, \Delta t) = \mathbf{x}(k) + \int_0^{\Delta t} \dot{\mathbf{x}}(t)\, dt \ . \tag{9.8}$$

Somit ergibt sich als zeitdiskrete Zustandsübergangsfunktion:

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \begin{bmatrix} -\frac{v+a\Delta t}{r_\lambda \omega} C + \frac{a}{r_\lambda \omega^2} S + \frac{v}{r_\lambda \omega} \cos\psi - \frac{a}{r_\lambda \omega^2} \sin\psi \\ \frac{v+a\Delta t}{r_\varphi \omega} S + \frac{a}{r_\varphi \omega^2} C - \frac{v}{r_\varphi \omega} \sin\psi - \frac{a}{r_\varphi \omega^2} \cos\psi \\ a \cdot \Delta t \\ 0 \\ \omega \cdot \Delta t \\ 0 \end{bmatrix} \tag{9.9}$$

mit $C = \cos(\omega \Delta t + \psi)$ und $S = \sin(\omega \Delta t + \psi)$.

Da in dem ermittelten Modell durch die Gierrate ω dividiert wird, muss das System für kleine Gierraten auf ein Einspurmodell mit konstanter Orientierung umgeschaltet werden, um numerische Probleme zu vermeiden. Aus der entsprechenden Grenzwertbildung ergibt sich in diesem Fall die folgende zeitliche Ableitung des Zustandsvektors:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \frac{v \cdot \sin\psi}{r_\lambda} \\ \frac{v \cdot \cos\psi}{r_\varphi} \\ a \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{9.10}$$

Für die diskrete Zustandsübergangsfunktion ergibt sich dadurch:

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \begin{bmatrix} \frac{\left(v\Delta t + \frac{1}{2} a \Delta t^2\right) \cdot \sin\psi}{r_\lambda} \\ \frac{\left(v\Delta t + \frac{1}{2} a \Delta t^2\right) \cdot \cos\psi}{r_\varphi} \\ a \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{9.11}$$

9.4.3 Prozessrauschen

Die zeitliche Prädiktion der Objektdynamik wird durch Schätzung mit einem Extended Kalman-Filter durchgeführt. Für die Schätzung der Unsicherheiten muss zunächst das Prozessrauschen $\mathbf{Q}(k)$ für einen Zeitschritt bestimmt werden. Dafür wird zunächst das als zeitkontinuierlich auftretend angenommene Rauschen der Gierrate und der Beschleunigung als Rauschterm in folgender Matrix gespeichert [5]:

$$\mathbf{Q_c}(k) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q_a & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & q_\omega \end{bmatrix}. \tag{9.12}$$

Nun muss die Auswirkung des Rauschens während eines Zeitschrittes auf die anderen Zustandsgrößen berechnet werden. Hierfür wird für jeden Zeitpunkt $\Delta t - \tau$ die Auswirkung des aktuellen Rauschens auf die anderen Zustände zum Zeitpunkt Δt berechnet. Dazu wird die diskrete Zustandsübergangsfunktion für diesen Zeitraum benötigt:

$$\mathbf{f_d}\left(\mathbf{x}\left(k\right), \Delta t - \tau, \Delta t\right) = \mathbf{x}\left(k\right) + \int_{\Delta t - \tau}^{\Delta t} \dot{\mathbf{x}}\left(t\right) dt \ . \tag{9.13}$$

Mit Hilfe der Jacobi-Matrix der Funktion 9.13 lässt sich die Auswirkung des Rauschens am Zeitpunkt $\Delta t - \tau$ berechnen:

$$\mathbf{Q_d}\left(\Delta t - \tau, \Delta t\right) = \mathbf{J}_{\mathbf{f_d}(\mathbf{x}(k), \Delta t - \tau, \Delta t)} \cdot \mathbf{Q_c}\left(k\right) \cdot \mathbf{J}_{\mathbf{f_d}(\mathbf{x}(k), \Delta t - \tau, \Delta t)}{}^{T} \ . \tag{9.14}$$

Das komplette Rauschen über den gesamten Zeitschritt wird dann durch Integration ermittelt:

$$\mathbf{Q}\left(k\right) = \int_{0}^{\Delta t} \mathbf{J}_{\mathbf{f_d}(\mathbf{x}(k), \Delta t - \tau, \Delta t)} \cdot \mathbf{Q_c}\left(k\right) \cdot \mathbf{J}_{\mathbf{f_d}(\mathbf{x}(k), \Delta t - \tau, \Delta t)}{}^{T} \, d\tau \ . \tag{9.15}$$

Auf die Angabe der berechneten Kovarianzmatrix wird an dieser Stelle aus Gründen der Übersichtlichkeit verzichtet.

Die Werte der in der Berechnung verwendeten Rauschterme q_a und q_ω, können mit Hilfe einer Abschätzung des Prozessrauschens pro Zeitschritt für die Beschleunigung $\mathbf{Q}_{a,a}\left(k\right)$ und die Gierrate $\mathbf{Q}_{\omega,\omega}\left(k\right)$ festgelegt werden [5]:

$$\sqrt{\mathbf{Q}_{a,a}\left(k\right)} = \sqrt{q_a \cdot \Delta t} \tag{9.16}$$

$$\sqrt{\mathbf{Q}_{\omega,\omega}\left(k\right)} = \sqrt{q_\omega \cdot \Delta t} \tag{9.17}$$

Geeignete Werte für die Standardabweichung wurden für einen Zeitschritt von $\Delta t = \frac{1}{12}\,\mathrm{s}$ empirisch mit $\sqrt{\mathbf{Q}_{a,a}\left(k\right)} = 4\,\frac{\mathrm{m}}{\mathrm{s}^2}$ und $\sqrt{\mathbf{Q}_{\omega,\omega}\left(k\right)} = 0,087\,\frac{\mathrm{rad}}{\mathrm{s}}$ abgeschätzt. Für Zeitschritte einer anderen Länge muss die Standardabweichung skaliert werden:

$$\sqrt{\mathbf{Q}_{a,a}\left(k\right)} = 4 \cdot 12 \cdot \Delta t\,\frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.18}$$

$$\sqrt{\mathbf{Q}_{\omega,\omega}\left(k\right)} = 0,087 \cdot 12 \cdot \Delta t\,\frac{\mathrm{rad}}{\mathrm{s}^2} \tag{9.19}$$

9.4.4 Versuchsaufbau

Für die in dieser Arbeit durchgeführten Untersuchungen zur Assoziation drahtlos übermittelter Daten wurden zwei Versuchsfahrzeuge eingesetzt. Beide Fahrzeuge ermitteln ihre Position und Orientierung mit Hilfe eines DGPS-Empfängers. Die Geschwindigkeit und die Orientierung wird jeweils aus den ESP-Daten bestimmt. Darüber hinaus verwendete das eine Versuchsfahrzeug die eingebauten Sensoren zur Erfassung des Fahrzeugumfelds. Demzufolge stehen für die Prädiktion der eigenen Position und für die Prädiktion der empfangenen Objektdynamik vergleichbare Daten zur Verfügung.

Die Eigenbewegungsmessungen erfolgen relativ häufig und werden in der Sensormessfrequenz von 12 Hz ausgewertet. Aufgrund bestehender Untersuchungen wurde für die Geschwindigkeit eine Standardabweichung von $\sigma_v = 0{,}1\,\frac{\mathrm{m}}{\mathrm{s}}$ und für die Gierrate eine Standardabweichung von $\sigma_\omega = 0{,}23\,\frac{\mathrm{rad}}{\mathrm{s}}$ als Messrauschen verwendet.

Die Messungen des DGPS-Empfängers erfolgen im Sekundentakt. Aus dieser Messung wird die Position in Längen- und Breitengrad sowie die Orientierung verwendet. Das zugehörige Messrauschen wurde basierend auf verschiedenen Beispielmessdaten empirisch ermittelt. Die Orientierung wird im Empfänger über die zeitliche Änderung der Position bestimmt. Das Messrauschen der Orientierung wird also bei niedrigen Geschwindigkeiten deutlich höher. Aus diesem Grund wurde ein linearer Verlauf der Standardabweichung für Geschwindigkeiten zwischen $0\,\frac{\mathrm{m}}{\mathrm{s}}$ und $5\,\frac{\mathrm{m}}{\mathrm{s}}$ gewählt. Bei einer Geschwindigkeit von $0\,\frac{\mathrm{m}}{\mathrm{s}}$ ist $\sigma_\psi = 90°$. Für Geschwindigkeiten ab $5\,\frac{\mathrm{m}}{\mathrm{s}}$ wird $\sigma_\psi = 2°$ gewählt (Abbildung 9.6).

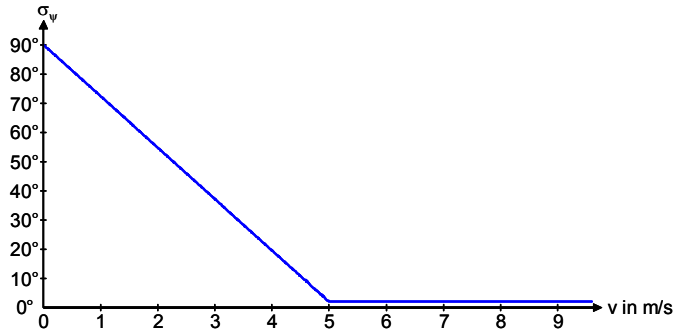

Abbildung 9.6: Bei niedrigen Geschwindigkeiten ist die Orientierungsmessung sehr ungenau, da diese aufgrund der Positionsschätzungen verschiedener Zeitschritte ermittelt wird. Dieser Sachverhalt ist im Messrauschen durch eine linear fallende, nach unten begrenzte Standardabweichung modelliert.

Aufgrund des Messrauschens des DGPS-Empfängers tritt in einigen Fällen im Stand ein Wandern der Position auf. Aus diesem Grund wird für die Position bei Geschwindigkeiten unter $5\,\frac{\mathrm{m}}{\mathrm{s}}$ eine Standardabweichung von 10 m gewählt, während bei höheren Geschwindigkeiten ein Wert von 1 m ausreicht. Diese Werte werden jeweils in Längen- und Breitengrade umgerechnet.

Zusätzlich weist der DGPS-Empfänger einen zeitlich annähernd konstanten Offset bei der Positionsmessung auf. Dieser Offset wirkt sich auf die interne Filterung des Empfängers aus, die nicht umgangen werden kann. Dadurch hängt der Offset der ausgegebenen Position auch vom individuellen Einschwingverhalten des internen Filters ab und variiert somit von Empfänger zu Empfänger. Auch Variationen zu unterschiedlichen Tageszeiten können durch veränderte Satellitenpositionen auftreten. Da die DGPS-Messungen dazu verwendet werden, eine Position der übertragenen Objektdaten relativ zum Versuchsfahrzeug zu bestimmen, äußern sich die unabhängigen Positionsoffsets der beiden DGPS-Empfänger hier in einem relativen Offset. Dieser relative Offset wird für jedes empfangene Objekt modelliert. Da der Offset unabhängig von der Orientierung der Objekte und damit unabhängig vom Fahrzeugkoordinatensystem des Versuchsfahrzeugs ist, muss er in globalen WGS84-Koordinaten modelliert werden. Ohne Zusatzinformationen kann der Offset nicht bestimmt werden, da

dieser über lange Zeit variiert. Aus diesem Grund wird im ersten Ansatz der Erwartungswert Null als bestmögliche Schätzung angenommen. Der Offset erhöht jedoch die Positionsunsicherheit um weitere 3,5 m. In einem zweiten Ansatz wird der Offset aufgrund sicherer Assoziationen der empfangenen Objektdaten zu gemessenen Objekten geschätzt. Auf diesen Ansatz wird in Abschnitt 9.6 eingegangen.

9.4.5 Zeitliche Realisierung des Filters

Beim Einbringen der Messwerte in das Kalman-Filter tritt ein weiteres Problem auf. Während die Eigenbewegungsmesswerte aus der ESP-Sensorik relativ zeitnah beim System eintreffen, benötigt der DGPS-Empfänger mehr als eine Sekunde für die Ausgabe der Position. Die Positionsmessung ist also zum des Empfangs im System schon veraltet (Abbildung 9.7).

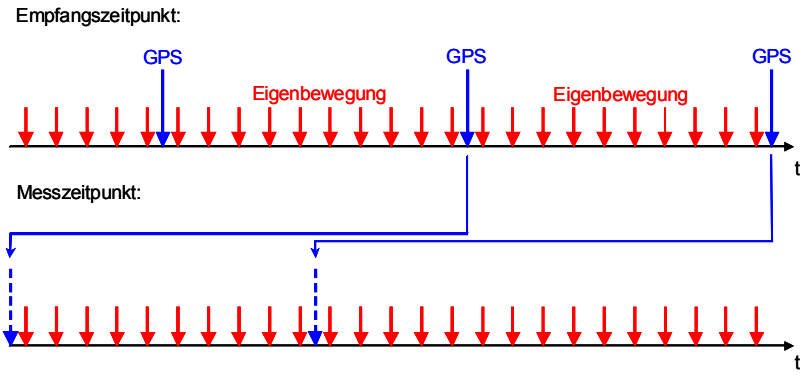

Abbildung 9.7: Bei der Positionsmessung des DGPS-Empfängers tritt eine signifikante zeitliche Verzögerung auf. Dadurch treffen vor dieser Messung mehrere Eigenbewegungsmessungen beim System ein, die deutlich später aufgenommen wurden.

Im Kalman-Filter wird erwartet, dass Messwerte zeitlich sortiert eingebracht werden. Es ist also nicht möglich, die Positionsmesswerte direkt einzubringen, wenn alle Eigenbewegungsmesswerte sofort eingebracht werden, sobald sie beim System eintreffen.

Aufgrund geeigneter Zeitstempel kann zumindest der genaue Messzeitpunkt der DGPS-Messwerte bestimmt werden. Ein Ansatz zur Bewältigung der Anforderung einer korrekten zeitlichen Reihenfolge wäre eine Zwischenspeicherung der neueren Eigenbewegungsmessungen bis zu dem Zeitpunkt, an dem die Positionsmessung eintrifft. Dadurch würden aber jeweils alle Eigenbewegungsmessungen der letzten Sekunde bei der Schätzung ignoriert. Um alle verfügbaren Messungen nutzen zu können, wurde ein aufwändigerer Ablauf der Filterung implementiert. Nach dem Einbringen einer DGPS-Messung wird eine Kopie des Zustandsvektors und der Kovarianzmatrix erzeugt. Alle darauffolgenden Eigenbewegungsmessungen werden in das Filter eingebracht und gleichzeitig in einem Puffer gespeichert.

Trifft eine neue DGPS-Messung ein, werden der dynamische Zustand und die Kovarianz auf die erzeugte Kopie zurückgesetzt. Nun werden alle vorhandenen Messungen wieder in kor-

rekter zeitlicher Reihenfolge eingebracht. Nach dem Einbringen der DGPS-Messung wird dabei wieder eine aktuelle Kopie des dynamischen Zustands und der Kovarianz erstellt.

9.4.6 Prädiktion auf den Zeitpunkt der Sensormessdaten

Nachdem alle vorhandenen DGPS- und Eigenbewegungsmessungen in das Filter eingebracht wurden, erfolgt eine Prädiktion des dynamischen Zustands und der Kovarianz auf den Zeitpunkt der Sensormessdaten. Diese Prädiktion überbrückt Latenzen, die bei der drahtlosen Übertragung der Daten entstehen. Nach dieser Prädiktion liegt die bestmögliche Zustandsschätzung für eine Assoziation mit sensorbasierten Objektdaten vor.

9.5 Assoziation

Die Assoziation empfangener Objektdaten zu gemessenen Objekten wird für jedes empfangene Objekt unabhängig von anderen empfangenen Objekten durchgeführt. In einem ersten Schritt wird die globale Position der empfangenen Daten in das lokale Fahrzeugkoordinatensystem des Versuchsfahrzeugs transformiert (Abbildung 9.8).

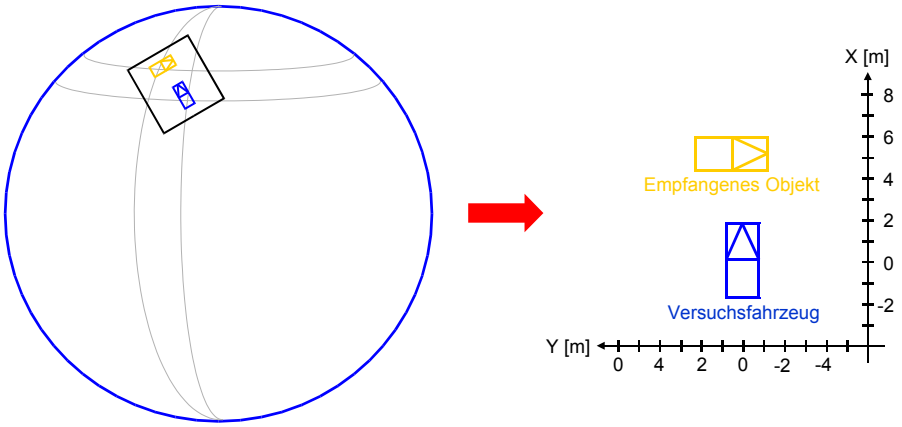

Abbildung 9.8: Vor der Assoziation der empfangenen Objektdaten wird die globale Position in das lokale Fahrzeugkoordinatensystem des Versuchsfahrzeugs transformiert.

Aufgrund der stark fehlerbehafteten Positionsmessungen des DGPS-Empfängers kann nach der Transformation eine signifikante Abweichung von der Objektposition auftreten, die mit den Sensoren des Versuchsfahrzeugs gemessen wurde (Abbildung 9.9). Aus diesem Grund wird bei der Assoziation neben der Position auch die Geschwindigkeit und die Orientierung berücksichtigt. Unter der Annahme einer normalverteilten Abweichung zwischen gemessenem und empfangenem Objektzustand wird mit Hilfe dieser Kriterien eine Assoziationswahrscheinlichkeit berechnet.

Durch eine zeitliche Filterung der Assoziationswahrscheinlichkeit können einige verbleibende Mehrdeutigkeiten aufgelöst werden. Die Entscheidung, ob eine Zuordnung wirklich erfolgt, wird erst in der Objektausgabe mit einer harten Schwellwertoperation getroffen.

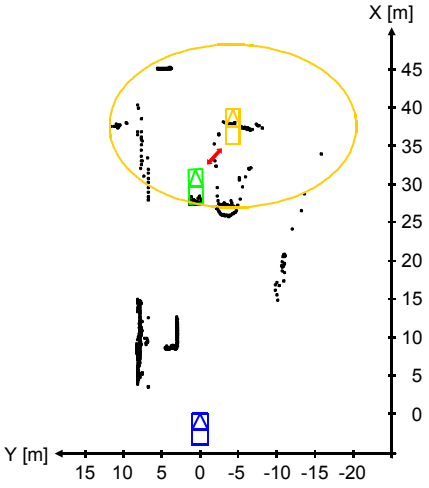

Abbildung 9.9: Nach der Transformation der empfangenen Objektposition (gelb) treten aufgrund des Messfehlers der DGPS-Empfänger signifikante Abweichungen von der mit den Sensoren gemessenen Objektposition (grün) auf. Dadurch können die Lidar-Messungen (schwarz) anderer Objekte deutlich dichter an der empfangenen Position liegen.

9.5.1 Transformation in das Fahrzeugkoordinatensystem

Für eine Integration der empfangenen Objektdaten in das lokale Fahrzeugumfeldmodell müssen diese zunächst in das Fahrzeugkoordinatensystem des Versuchsfahrzeugs transformiert werden. Dafür werden die aktuellen Schätzwerte der globalen Position und Orientierung des Versuchsfahrzeugs mit denen der globalen Position, Orientierung und Geschwindigkeit des empfangenen Objekts in einem kombinierten, globalen Zustandsvektor $\mathbf{x_g}$ gespeichert:

$$\mathbf{x_g} = \left[\lambda_{Ego}, \varphi_{Ego}, \psi_{Ego}, \lambda_{Obj}, \varphi_{Obj}, \psi_{Obj}, v_{Obj}\right]^T . \tag{9.20}$$

Die verwendeten Kalman-Filter liefern ebenfalls Prädiktionen für die Zustandskovarianzen des Versuchsfahrzeugs und des empfangenen Objekts. Aus diesen globalen Kovarianzmatrizen wird eine entsprechende kombinierte Kovarianzmatrix $\mathbf{P_g}$ erstellt.

Aus dem kombinierten, globalen Zustandsvektor kann die Position, die Orientierung und die Geschwindigkeit des empfangenen Objekts im kartesischen Fahrzeugkoordinatensystem des Versuchsfahrzeugs berechnet werden:

$$\begin{aligned}
\mathbf{x_r} &= \begin{bmatrix} x_r \\ y_r \\ v_r \\ \psi_r \end{bmatrix} = \mathbf{f_r}\left(\mathbf{x_g}\right) = \mathbf{f_r}\left(\lambda_{Ego}, \varphi_{Ego}, \psi_{Ego}, \lambda_{Obj}, \varphi_{Obj}, \psi_{Obj}, v_{Obj}\right) \\
\mathbf{x_r} &= \begin{bmatrix} r_\varphi \cos\left(\psi_{Ego}\right)\left(\varphi_{Obj} - \varphi_{Ego}\right) + r_\lambda \sin\left(\psi_{Ego}\right)\left(\lambda_{Obj} - \lambda_{Ego}\right) \\ r_\varphi \sin\left(\psi_{Ego}\right)\left(\varphi_{Obj} - \varphi_{Ego}\right) - r_\lambda \cos\left(\psi_{Ego}\right)\left(\lambda_{Obj} - \lambda_{Ego}\right) \\ v_{Obj} \\ \psi_{Obj} - \psi_{Ego} \end{bmatrix} .
\end{aligned} \tag{9.21}$$

Die zugehörige Kovarianz des Objektzustands im Fahrzeugkoordinatensystem kann dann mit Hilfe der Jacobi-Matrix der Funktion $\mathbf{f_r}$ aus der kombinierten, globalen Kovarianz $\mathbf{P_g}$ bestimmt werden:

$$\mathbf{P_r} = \mathbf{J_{f_r(x_g)}} \cdot \mathbf{P_g} \cdot \mathbf{J_{f_r(x_g)}}^T \ . \tag{9.22}$$

9.5.2 Gating

Um bei der Assoziation Rechenzeit zu sparen, werden die folgenden Assoziationsschritte auf sinnvolle Kandidaten gemessener Objekte eingeschränkt. Geeignete Objektkandidaten werden über den Abstand zur prädizierten Position des empfangenen Objekts ermittelt. Hierzu wird die Kovarianz der kartesischen Position des empfangenen Objekts berechnet:

$$\mathbf{P_{pos}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \cdot \mathbf{P_r} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \ . \tag{9.23}$$

Aufgrund der Kovarianz der Position $\mathbf{P_{pos}}$ wird eine Ellipse als Bereich gültiger Objekte definiert:

$$\left(\begin{bmatrix} x_m \\ y_m \end{bmatrix} - \begin{bmatrix} x_r \\ y_r \end{bmatrix} \right)^T \cdot \mathbf{H}\mathbf{P_{pos}}\mathbf{H}^T + \mathbf{R_m}^{-1} \cdot \left(\begin{bmatrix} x_m \\ y_m \end{bmatrix} - \begin{bmatrix} x_r \\ y_r \end{bmatrix} \right) \leq \gamma \ . \tag{9.24}$$

Als Schwelle γ wurde der Wert 9 gewählt, so dass 99,7 Prozent der Wahrscheinlichkeitsmasse innerhalb des Gültigkeitsbereichs liegen [6]. Weiterhin kann die Kovarianz der gemessenen Objekte $\mathbf{R_m}$ bei der Assoziation ignoriert werden, da diese deutlich geringer ist als die der empfangenen Objekte. Schließlich ist die Messmatrix $\mathbf{H}$ eine Einheitsmatrix. Somit kann die Bedingung für den Bereich gültiger Objekte folgendermaßen vereinfacht werden:

$$\left(\begin{bmatrix} x_m \\ y_m \end{bmatrix} - \begin{bmatrix} x_r \\ y_r \end{bmatrix} \right)^T \cdot \mathbf{P_{pos}}^{-1} \cdot \left(\begin{bmatrix} x_m \\ y_m \end{bmatrix} - \begin{bmatrix} x_r \\ y_r \end{bmatrix} \right) \leq 9 \ . \tag{9.25}$$

Alle gemessenen Objekte, deren gemessene Positionen $\left[x_m, y_m\right]^T$ der Bedingung 9.25 genügen, kommen für eine Assoziation in Frage.

9.5.3 Assoziationswahrscheinlichkeiten

Für jedes Objekt i, das innerhalb der berechneten Ellipse liegt, wird eine Assoziationswahrscheinlichkeit berechnet. Dabei wird von einer normalverteilten Abweichung zwischen empfangenem Objektzustand $\mathbf{x_r}$ und mit Sensoren gemessenem Objektzustand $\mathbf{x_i}$ ausgegangen. Die Assoziationswahrscheinlichkeit nimmt entsprechend dieser Verteilung mit zunehmender Abweichung des Objektzustands ab (Abbildung 9.10).

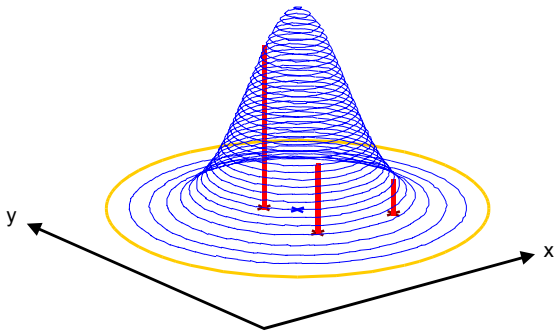

Abbildung 9.10: Für die Abweichung zwischen empfangenem und gemessenem Objektzustand wird eine Normalverteilung angenommen. Mit zunehmender Abweichung nimmt auch die Assoziationswahrscheinlichkeit entsprechend der Normalverteilung ab. Im Gegensatz zur dargestellten zweidimensionalen Normalverteilung wird der komplette Objektzustand bestehend aus Position, Orientierung und Geschwindigkeit verwendet.

Im ersten Schritt wird von der Annahme M ausgegangen, dass das empfangene Objekt auch mit den Sensoren des Versuchsfahrzeugs erfasst wurde. Für jedes geeignete Objekt i wird eine bedingte Assoziationswahrscheinlichkeit berechnet:

$$P\left(i|M\right) = \frac{\mathcal{N}\left(\mathbf{x_i}, \mathbf{x_r}, \mathbf{P_r}\right)}{\sum\limits_j \mathcal{N}\left(\mathbf{x_j}, \mathbf{x_r}, \mathbf{P_r}\right)} \,. \tag{9.26}$$

Hierbei bezeichnet $\mathcal{N}\left(\mathbf{x}, \boldsymbol{\mu}, \mathbf{P}\right)$ die multivariate Normalverteilung:

$$\mathcal{N}\left(\mathbf{x}, \boldsymbol{\mu}, \mathbf{P}\right) = \frac{1}{\left(2\pi\right)^{\frac{\dim(\mathbf{x})}{2}} \left|\mathbf{P}\right|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{P}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \,. \tag{9.27}$$

Im zweiten Schritt muss die Wahrscheinlichkeit $P\left(M\right)$, dass das empfangene Objekt mit den Sensoren gemessen wurde, bestimmt werden. Zu diesem Zweck wird das Ergebnis der in Kapitel 4 vorgestellten Verdeckungserkennung verwendet. Auch hier wird für den Positionsfehler der empfangenen Objekte eine Normalverteilung angenommen (Abbildung 9.11). Diese ist durch die Kovarianz $\mathbf{P_{pos}}$ der kartesischen Position des empfangenen Objekts charakterisiert.

Zur Bestimmung der Messwahrscheinlichkeit $P\left(M\right)$ wird die Verdeckungswahrscheinlichkeit gewichtet mit dem Funktionswert der Gaußglocke integriert. Die Integration wird numerisch approximiert, da die Verdeckungswahrscheinlichkeiten im Fahrzeugumfeld für diskrete Flächenelemente im Polarkoordinatensystem berechnet wurden. Im Polarkoordinatensystem werden alle Flächenelemente berücksichtigt, die sowohl zwischen einem minimalen Radius r_{min} und einem maximalen Radius r_{max} als auch zwischen einem minimalen Winkel α_{min} und einem maximalen Winkel α_{max} liegen. Die Werte r_{min}, r_{max}, α_{min} und α_{max} werden dabei so gewählt, dass sie bezüglich der Gaußglocke außerhalb der Ellipse liegen.

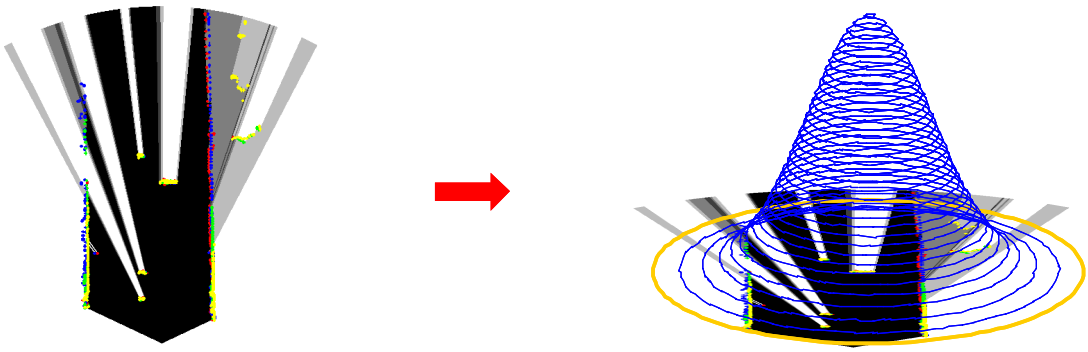

Abbildung 9.11: Für die Berechnung der Wahrscheinlichkeit, dass das empfangene Objekt verdeckt ist, wird ebenfalls ein normalverteilter Positionsfehler der empfangenen Objekte angenommen. Die Wahrscheinlichkeit kann dann durch numerische Integration der Verdeckungswahrscheinlichkeiten unter der Gaußglocke berechnet werden.

Die Verdeckungswahrscheinlichkeiten der betreffenden Flächenelemente werden mit ihrer korrespondierenden Wahrscheinlichkeitsmasse unter der Gaußglocke gewichtet und addiert:

$$P(M) = 1 - \sum_{r=r_{min}}^{r_{max}} \sum_{\alpha=\alpha_{min}}^{\alpha_{max}} P(r,\alpha)\ A(r,\alpha)\ g(r,\alpha)\,. \tag{9.28}$$

Hierbei bezeichnet $P(r,\alpha)$ die Verdeckungswahrscheinlichkeit eines Flächenelements, während $A(r,\alpha)$ dessen Flächeninhalt darstellt:

$$A(r,\alpha) = \frac{\Delta\alpha}{2}\left(\left(r+\frac{1}{2}\Delta r\right)^2 - \left(r-\frac{1}{2}\Delta r\right)^2\right). \tag{9.29}$$

Der Wert $g(r,\alpha)$ bezeichnet schließlich den Funktionswert der Gaußglocke für die Mitte des Flächenelements:

$$g(r,\alpha) = \frac{1}{2\pi\ \left|\mathbf{P_{pos}}\right|^{\frac{1}{2}}} e^{-\frac{1}{2}\Delta\mathbf{x}(r,\alpha)^T\ \mathbf{P_{pos}}^{-1}\Delta\mathbf{x}(r,\alpha)}. \tag{9.30}$$

Der Abstand $\Delta\mathbf{x}(r,\alpha)$ der Position des Flächenelements von der empfangenen Position ergibt sich durch:

$$\Delta\mathbf{x}(r,\alpha) = \begin{bmatrix} x_r \\ y_r \end{bmatrix} - \begin{bmatrix} r\cdot\cos\alpha \\ r\cdot\sin\alpha \end{bmatrix}. \tag{9.31}$$

Aus den Wahrscheinlichkeiten $P(M)$ und $P(i|M)$ kann nun die Assoziationswahrscheinlichkeit des Objekts $P(i)$ berechnet werden:

$$P(i) = \frac{P(i|M)\cdot P(M)}{P(M|i)}. \tag{9.32}$$

Unter der Bedingung, dass das gemessene Objekt i dem empfangenen Objekt entspricht, ist die Wahrscheinlichkeit, dass das empfangene Objekt gemessen wurde gleich Eins:

$$P\left(M|i\right) = 1 \;\rightarrow\; P\left(i\right) = P\left(i|M\right) \cdot P\left(M\right). \tag{9.33}$$

9.5.4 Zeitliche Filterung

Durch ähnliche dynamische Zustände der mit den Sensoren erfassten Objekte können trotz der Auswertung der Geschwindigkeit und der Orientierung in einigen Fällen Mehrdeutigkeiten bei der Assoziation entstehen. Diese Mehrdeutigkeiten lassen sich meist durch die Betrachtung der Objekthistorie auflösen. Zu diesem Zweck wird die berechnete Assoziationswahrscheinlichkeit für jedes Objekt zeitlich gefiltert. Der verwendete Filter mittelt die Wahrscheinlichkeiten mit einer stärkeren Wichtung der aktuellen Werte:

$$\begin{aligned} P_0\left(i\right) &= 0 \\ P_k\left(i\right) &= \frac{P\left(i\right) + \left(l_{Filter} - 1\right) \cdot P_{k-1}\left(i\right)}{l_{Filter}}. \end{aligned} \tag{9.34}$$

Als Filterlänge wurde $l_{Filter} = 20$ gewählt.

9.6 Offsetschätzung

Wie in Abschnitt 9.4.4 beschrieben, treten bei der Positionsbestimmung mit DGPS-Empfängern signifikante Offsets bezüglich der gemessenen Längen- und Breitengrade auf. Bei den verwendeten Empfängern verhalten sich diese Offsets über einen längeren Zeitraum konstant, unterscheiden sich jedoch von DGPS-Empfänger zu DGPS-Empfänger.

Die Offsets des empfangenen Objekts und des eigenen DGPS-Empfängers äußern sich bei der Transformation in das Fahrzeugkoordinatensystem des Versuchsfahrzeugs. Die kartesische Position des empfangenen Objekts weicht signifikant von der mit den Sensoren gemessenen Position ab (Abbildung 9.9).

Die kartesische Positionsabweichung kann gemessen werden, wenn von einer sicheren Assoziation der empfangenen Objektdaten zu gemessenen Objekten ausgegangen werden kann. Aus diesem Grund wird eine gemessene Abweichung nur verwendet, wenn die korrespondierende Assoziationswahrscheinlichkeit größer als 50 Prozent ist.

Da der Positionsoffset in Längen- und Breitengraden konstant ist und sich das kartesische Fahrzeugkoordinatensystem des Versuchsfahrzeugs mit dessen Orientierung dreht (Abbildung 9.8), kann eine Filterung des Offsets nur in globalen Koordinaten erfolgen. Diese Filterung ersetzt dann die in Abschnitt 9.4.4 beschriebene Methode der als Null angenommenen konstanten Offsets.

Die Positionsabweichung wird in kartesischen Koordinaten gemessen. Dazu muss als Erstes die kartesische Position der empfangenen Objektdaten ermittelt werden. Dies geschieht gemäß folgender Gleichung:

$$\begin{bmatrix} x_r \\ y_r \end{bmatrix} = \begin{bmatrix} r_\varphi \cos\left(\psi_{Ego}\right)\left(\varphi_{Obj} - \varphi_{Ego}\right) + r_\lambda \sin\left(\psi_{Ego}\right)\left(\lambda_{Obj} - \lambda_{Ego}\right) \\ r_\varphi \sin\left(\psi_{Ego}\right)\left(\varphi_{Obj} - \varphi_{Ego}\right) - r_\lambda \cos\left(\psi_{Ego}\right)\left(\lambda_{Obj} - \lambda_{Ego}\right) \end{bmatrix}. \tag{9.35}$$

Der gemessene Offset in kartesischen Koordinaten ergibt sich nun aus der Differenz zwischen der durch die Sensoren gemessenen Position und der kartesischen Position der empfangenen Objektdaten:

$$\mathbf{z_{offset}} = \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} x_m \\ y_m \end{bmatrix} - \begin{bmatrix} x_r \\ y_r \end{bmatrix}. \tag{9.36}$$

Der Offset ist somit eine Funktion von der Position und der Orientierung des eigenen Fahrzeugs, von der Position und der Orientierung des empfangenen Objekts sowie von der mit den Sensoren gemessenen Position des empfangenen Objekts. Diese Werte werden im Vektor $\mathbf{x_k}$ zusammengefasst:

$$\mathbf{z_{offset}} = \mathbf{f_{offset}}\left(\lambda_{Ego}, \varphi_{Ego}, \psi_{Ego}, \lambda_{Obj}, \varphi_{Obj}, x_m, y_m\right) = \mathbf{f_{offset}}\left(\mathbf{x_k}\right). \tag{9.37}$$

Prinzipiell sollten zur Ermittlung des kartesischen Offsets die Messgrößen der globalen Position für das Versuchsfahrzeug und das empfangene Objekt verwendet werden. Da die globalen Positionsmessungen zum Zeitpunkt ihres Eintreffens beim System schon stark veraltet sind, werden die gemäß Abschnitt 9.4 auf den Zeitpunkt der aktuellen Sensormessdaten prädizierten Positions- und Orientierungswerte verwendet. Darüber hinaus spricht auch die geringe Messfrequenz der DGPS-Empfänger für eine Prädiktion der globalen Positionen. Verwendet man ausschließlich die reinen DGPS-Messwerte, kann der Offset nur einmal pro Sekunde bestimmt werden. Ist zu diesem Zeitpunkt die Assoziation nicht ausreichend sicher, kann kein Offset bestimmt werden. Prädiziert man dagegen die globale Position, kann zu jedem Zeitpunkt, an dem eine sichere Assoziation mit Sensormessdaten aus dem Fahrzeugumfeld vorliegt, ein Offset ermittelt werden.

Der in kartesischen Koordinaten gemessene Offset ist eine Überlagerung verschiedener Effekte. Der primäre Einfluss stammt vom relativen Positionsoffset in globalen Koordinaten. Aber auch Schätzfehler der aktuellen globalen Position des empfangenen Objekts und der aktuellen globalen Position und Orientierung des eigenen Fahrzeugs wirken sich auf diese Messung aus. Weiterhin wird der gemessene Offset durch Fehler in der mit den Sensoren gemessenen kartesischen Position beeinflusst. Demzufolge sollte sich der gemessene kartesische Offset auf jede dieser Größen auswirken. Hierzu wäre jedoch eine Schätzung aller Größen in einem Kalman-Filter notwendig. Eine Schätzung mit einem gemeinsamen Filter für alle empfangenen und gemessenen Objekte ist praktisch extrem aufwändig, da die hohe Anzahl der Objekte zu sehr großen Zustandsvektoren führen würde. Der korrespondierende Aufwand für die notwendigen Matrixinvertierungen des Kalman-Filters würden einen sehr hohen Rechenaufwand erzeugen. Eine gemeinsame Schätzung des Offsets mit den globalen Positionen und Orientierungen des Versuchsfahrzeugs und des empfangenen Objekts ist prinzipiell möglich. Das vorgestellte System ist jedoch für die Integration mehrerer drahtlos empfangener Objekte konzipiert. Zusätzliche Objekte müssten bei einer gemeinsamen Schätzung zum gemeinsamen Filter hinzugefügt werden. Dies würde ebenfalls sehr schnell zu großen Zustandsvektoren und damit zu stark ansteigendem Rechenaufwand führen.

Um das System mit linear steigendem Rechenaufwand auf weitere empfangene Objekte erweitern zu können, wurde auf eine komplexe, gemeinsame Filterung verzichtet. Der in kartesischen Koordinaten gemessene Offset wird für eine von den anderen Zustandswerten

unabhängige Schätzung des Positionsoffsets in globalen Koordinaten $\mathbf{x_{offset}}\,(k) = [\Delta\lambda,\, \Delta\varphi]^T$ genutzt (Abbildung 9.12). Die Einflüsse der Positionsfehler und des Orientierungsfehlers werden im Rahmen des Messrauschens modelliert.

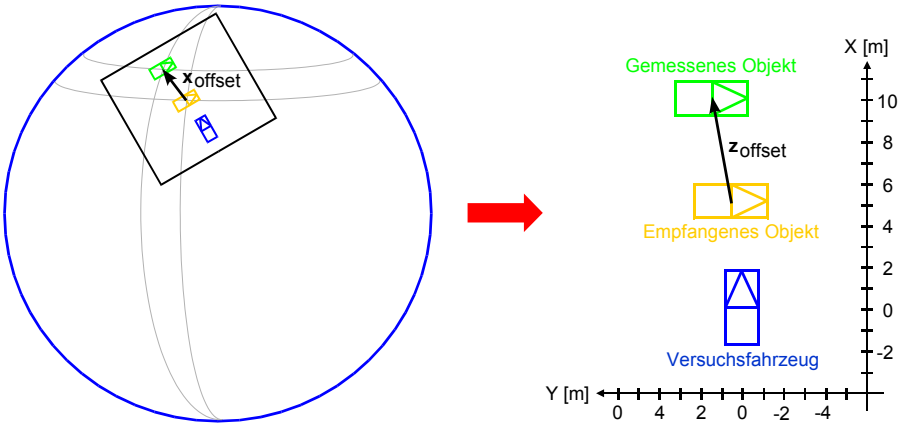

Abbildung 9.12: Offsetschätzung: Der in kartesischen Koordinaten gemessene Positionsoffset wird zur Schätzung des Positionsoffsets in globalen WGS84-Koordinaten verwendet.

Zur Berechnung des Messrauschens wird in einem ersten Schritt eine kombinierte Kovarianzmatrix erzeugt, die alle im Messrauschen modellierten Unsicherheiten enthält. Die ersten drei Zeilen und Spalten entsprechen der globalen Position und Orientierung des Versuchsfahrzeugs. Die entsprechenden Unsicherheiten werden aus der Kovarianzmatrix $\mathbf{P_e}$ des prädizierten globalen Zustands des Versuchsfahrzeugs entnommen. Die folgenden zwei Spalten und Zeilen entsprechen der globalen Position des empfangenen Objekts. Die entsprechenden Einträge stammen aus der prädizierten Kovarianzmatrix des empfangenen Objekts $\mathbf{P_o}$. Die letzten beiden Spalten und Zeilen korrespondieren mit der durch die Sensoren gemessenen Position. Die entsprechenden Einträge werden aus der Kovarianzmatrix $\mathbf{R_m}$ der mit den Sensoren gemessenen Objektposition kopiert. Die kombinierte Kovarianzmatrix $\mathbf{R_k}$ ergibt sich somit zu:

$$\mathbf{R_k} = \begin{bmatrix} \mathbf{P_e}(\lambda,\lambda) & \mathbf{P_e}(\varphi,\lambda) & \mathbf{P_e}(\psi,\lambda) & 0 & 0 & 0 & 0 \\ \mathbf{P_e}(\lambda,\varphi) & \mathbf{P_e}(\varphi,\varphi) & \mathbf{P_e}(\psi,\varphi) & 0 & 0 & 0 & 0 \\ \mathbf{P_e}(\lambda,\psi) & \mathbf{P_e}(\varphi,\psi) & \mathbf{P_e}(\psi,\psi) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{P_o}(\lambda,\lambda) & \mathbf{P_o}(\varphi,\lambda) & 0 & 0 \\ 0 & 0 & 0 & \mathbf{P_o}(\lambda,\varphi) & \mathbf{P_o}(\varphi,\varphi) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{R_m}(x,x) & \mathbf{R_m}(y,x) \\ 0 & 0 & 0 & 0 & 0 & \mathbf{R_m}(x,y) & \mathbf{R_m}(y,y) \end{bmatrix}.$$

Die Messmatrix für den kartesischen Offset wird aus dieser kombinierten Kovarianzmatrix mit Hilfe der Jacobi-Matrix $\mathbf{J_{f_{offset}(x_k)}}$ der Funktion $\mathbf{f_{offset}}$ aus Gleichung 9.37 berechnet. Darüber hinaus geht die quadratische Assoziationswahrscheinlichkeit $P_k\,(i)^2$ als Gewicht in das Messrauschen ein:

$$\mathbf{R_{offset}} = \frac{\mathbf{J_{f_{offset}(x_k)}} \cdot \mathbf{R_k} \cdot \mathbf{J_{f_{offset}(x_k)}}^T}{P_k\,(i)^2}. \tag{9.38}$$

Die in Gleichung 9.38 verwendete Skalierung weist Parallelen zu der in [5] vorgestellten Skalierung des Systemrauschens auf. Dort erfolgt bei einer hohen Abweichung der Messwerte eine Skalierung des Systemrauschens zur Korrektion der Modellannahme. Im Gegensatz dazu wird jedoch in Gleichung 9.38 bei niedrigen Assoziationswahrscheinlichkeiten von einem korrekten Modell und Abweichungen bei der Genauigkeit der Messwerte ausgegangen. Aus diesem Grund erfolgt eine Skalierung des Messrauschens über das Reziproke der quadratischen Assoziationswahrscheinlichkeit. Da das Filterupdate nur für Assoziationswahrscheinlichkeiten ab 50 Prozent durchgeführt wird, ist der Skalierungsfaktor auf $[1, 4]$ begrenzt. Bei einer unsicheren Assoziation wird somit die Standardabweichung des Messrauschens maximal verdoppelt.

Mit Hilfe der Messung $\mathbf{z}_{\mathbf{offset}}$ und dem zugehörigen Messrauschen $\mathbf{R}_{\mathbf{offset}}$ erfolgt eine Schätzung des globalen Positionsoffsets $\mathbf{x}_{\mathbf{offset}}(k) = [\Delta\lambda, \Delta\varphi]^T$ in einem Kalman-Filter. Das zugehörige Messmodell ist:

$$\mathbf{z}_{\mathbf{offset}}(k) = \begin{bmatrix} r_\lambda \cos\psi & r_\varphi \sin\psi \\ r_\lambda \sin\psi & -r_\varphi \cos\psi \end{bmatrix} \cdot \mathbf{x}_{\mathbf{offset}}(k) + \mathbf{w}(k)\,. \tag{9.39}$$

Als Prozessmodell wird ein konstanter Offset $\mathbf{x}_{\mathbf{offset}}(k)$ verwendet. Da bei unsicherer Assoziation möglicherweise lange Zeit keine Offsets gemessen werden können, wird die Schätzfehlerkovarianz $\mathbf{P}_{\mathbf{offset}}$ nach oben begrenzt. Als Maximalwert wird für Länge und Breite eine Standardabweichung von umgerechnet 3,5 m gewählt. Diese entspricht der additiven Positionsunsicherheit, die ohne Schätzung des Offsets angewendet wurde.

9.7 Anwendungsspezifische Ausgabe

Die Zuordnung der drahtlos empfangenen Daten zu Objekten, die mit den Sensoren erfasst wurden, kann anwendungsspezifisch erfolgen. Die im Folgenden beschriebenen Schwellwertoperationen werden dann entsprechend der Anforderungen des Assistenzsystems durchgeführt.

Eine Zuordnung ist nicht in jedem Fall sinnvoll. So sind niedrige Assoziationswahrscheinlichkeiten ein Indiz dafür, dass das Objekt möglicherweise aufgrund von Verdeckungen nicht von den Sensoren erfasst wurde. Aus diesem Grund wird ein Schwellwert $s_{Assoziation}$ für die minimal zulässige Assoziationswahrscheinlichkeit gewählt. Eine Zuordnung zum Objekt mit der größten Assoziationswahrscheinlichkeit erfolgt dann nur, wenn diese den vorgegebenen Schwellwert übersteigt.

In manchen Fällen ist eine sehr sichere Assoziation erwünscht. In diesem Fall kann auch die Differenz zwischen der größten und der zweitgrößten Assoziationswahrscheinlichkeit ausgewertet werden.

Weiterhin muss bei der Ausgabe entschieden werden, ob und welche der mit den Sensoren bestimmten Objektparameter von den drahtlos empfangenen Daten überschrieben werden. Bei einer ausreichend sicheren Assoziation ist es sinnvoll, die Objektgröße und die Klasse mit den empfangenen Daten zu ersetzen, da diese invariant sind und aufgrund der Sensormessung nur eine Schätzung vorliegt.

Für die verwendeten Sensoren war ein Ersetzen der Orientierung dann sinnvoll, wenn sich das empfangene Objekt nicht wesentlich bewegt und der Laserscanner nur wenige Entfernungsmessungen vom Objekt liefert.

Darüber hinaus hat sich gezeigt, dass für die untersuchten Szenarien auch ein Ersetzen der Geschwindigkeit sinnvoll ist. Dies ist dadurch begründet, dass die Objektverfolgung Geschwindigkeiten aus der Positionsänderung bestimmen muss. Da die empfangenen Geschwindigkeiten gemessen wurden und zusätzlich auch eine Beschleunigung modelliert wurde, kann das entsprechende Kalman-Filter trotz der durch die Übertragung entstehenden Latenz bessere Geschwindigkeitsschätzungen liefern. Diese Schätzung wird erst dann schlechter, wenn aufgrund von Übertragungsfehlern keine Geschwindigkeitswerte mehr empfangen werden. In diesem Fall sollte die durch die sensorbasierte Objektverfolgung ermittelte Geschwindigkeit bei der Ausgabe nicht ersetzt werden. Die Ergebnisse entsprechender Untersuchungen werden in Abschnitt 10.7.4 präsentiert.

9.8 Zusammenfassung

In diesem Kapitel wurde eine Methode zur konsistenten Integration drahtlos empfangener Daten in das sensorbasierte Fahrzeugumfeldmodell vorgestellt. Neben der Integration gebietsbezogener Gefahrensituationen wie beispielsweise Schlechtwettergebiete und Staus erlaubt dieser neue Ansatz insbesondere eine Fusion übertragener Detailinformationen über andere Verkehrsteilnehmer mit den korrespondierenden Sensormessdaten. Auf diese Weise lassen sich Objektinformationen im Fahrzeugumfeldmodell integrieren, die sensortechnisch nur schwer oder überhaupt nicht erfasst werden können. Die Performance der vorgestellten Algorithmen wird im folgenden Kapitel mit realen Sensormessdaten untersucht.

10 Experimentelle Untersuchungen

In diesem Kapitel werden die bisher vorgestellten Algorithmen bezüglich ihrer Leistungsfähigkeit ausgewertet. Zum Training der Klassifikatoren und zur Auswertung der Algorithmen wurden verschiedene Sequenzen realer Sensormessdaten aufgezeichnet. In diesen Sequenzen mussten manuell alle Vertreter der Klassen „Fußgänger“, „Zweirad“, „PKW“ und „LKW“ markiert werden. Im Abschnitt 10.1 werden diese Daten und Methoden zur vereinfachten Markierung der Objekte beschrieben. Im Anschluss wird in Abschnitt 10.2 die verwendete Methode zur Analyse der Rechenzeit vorgestellt. Danach werden die Algorithmen in der Reihenfolge ihrer Vorstellung in dieser Arbeit ausgewertet. Nach einer Auswertung verschiedener Aspekte der verbesserten Objektverfolgung wird die Klassifikation analysiert. Anschließend wird das Verbesserungspotential der Fusionsansätze mit Videodaten und der Nutzung präziser digitaler Karten untersucht. Schließlich folgen Analysen der Algorithmen zur Integration drahtlos empfangener Objektdaten.

10.1 Beispieldaten

10.1.1 Verwendete Datensätze

Für die Untersuchungen wurden verschiedene Datensätze verwendet. Die Trainingsdaten bestehen aus Sequenzen mit unterschiedlichen Szenarien. Diese beinhalten Kreuzungen, städtischen Verkehr, Landstraßen und Autobahnen. Insgesamt enthalten die Sequenzen Messdaten mit einer Laufzeit von etwa 64 min.

Die für das Training benötigten Validierungsdaten umfassen eine ähnliche Vielfalt von Szenarien. Sequenzen mit insgesamt etwa 50 min Laufzeit stehen zu diesem Zweck zur Verfügung.

Sämtliche Auswertungen der Objektverfolgung und der Klassifikation wurden mit Hilfe eines Datensatzes von Testdaten durchgeführt. Dieser Datensatz enthält verschiedene Sequenzen mit einer Gesamtlaufzeit von etwa 60 min. Auswertungen mit präzisen digitalen Karten konnten nur auf einem Teil dieses Datensatzes durchgeführt werden, da die präzise Positionierung nur für Kreuzungen und kleine städtische Gebiete mit einer ausreichenden Menge an Landmarken entwickelt wurde. Dieser Teil der Testdaten umfasst Sequenzen mit einer Laufzeit von etwa 22 min.

Für die Trainings-, Validierungs- und Testdaten wurden in den Laserscannermessdaten alle Vertreter der Klassen „Fußgänger“, „Zweirad“, „PKW“ und „LKW“ markiert. Diese Markierung stellt den aufwändigsten und aufgrund des enormen Zeitaufwands am stärksten limitierenden Schritt der Vorbereitung der einzelnen Datensätze dar. Aufgrund der Messfrequenz von 12,5 Hz und der hohen Anzahl gemessener Verkehrsteilnehmer pro Zeitschritt ergibt sich eine sehr hohe Anzahl markierter Beispielobjekte. Sämtliche Algorithmen zum Training der Mustererkennung und zur Auswertung nutzen für die entsprechenden Datensätze die den Laserscannermessdaten zugeordnete Klasse als Vorgabe. Die einzige Ausnahme bildet das Training der Kaskadenklassifikatoren zur Videobildverarbeitung. Für dieses Training wurden aus Teilen der Trainingsdaten verschiedene Videobeispieldaten ausgeschnitten.

Für die Algorithmen zur Integration drahtlos übermittelter Objektdaten musste ein gesonderter Satz von Sequenzen aufgenommen werden. Dabei wurde ein zweites Versuchsfahrzeug als sendendes Fahrzeug verwendet. Dieses wurde in allen Sequenzen in den Laserscannermessdaten des ersten Fahrzeugs markiert. Die insgesamt 15 Sequenzen enthalten Szenarien mit Gegen- und Querverkehr sowie Folgefahrten und Überholmanöver. Die Sequenzen wurden teilweise in der Stadt, teilweise auf Landstraßen aufgenommen. Da die Sequenzen mit Gegen- und Querverkehr in der Regel nur wenige Zeitschritte enthalten, umfasst der Datensatz zur Analyse der Integration drahtlos empfangener Objektdaten etwa 8 min.

10.1.2 Laserscannermessdaten

Zur Auswertung der Algorithmen mussten in den Laserscannermessdaten die Objektklassen markiert werden. Prinzipiell ist für eine Auswertung der Klassifikation eine Zuordnung von Klasseninformationen zu den in der Objektverfolgung erstellten Objekten ausreichend. Da jedoch an allen Teilen der Messdatenverarbeitung gearbeitet wurde, variiert auch das Ergebnis der Objektbildung aus den Sensormessdaten. Diese Variation wird primär durch unterschiedliche Einstellungen der Segmentierungsparameter und durch die Ansätze zur Segmentfusion beeinflusst.

Um die Objektklassen für alle Variationen der Objektbildung korrekt markieren zu können, wurden die Klassen direkt den Messdaten des Laserscanners zugeordnet. Dies erfolgt durch die Markierung der einzelnen Entfernungsmessungen des Laserscanners mit einer Box. Alle Messungen innerhalb der Box gehören zum gleichen Objekt aus dem Fahrzeugumfeld und bekommen die gleiche Klasse zugeordnet (Abbildung 10.1). Auf diese Weise muss nicht die exakte Position der Objekte markiert werden. Es muss lediglich für alle Messungen entschieden werden, ob sie zum Objekt gehören oder nicht.

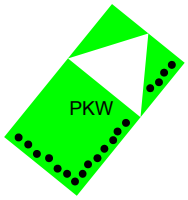

Abbildung 10.1: Die Entfernungsmessungen eines Objekts werden über Boxen markiert.

Innerhalb der Messdaten des Laserscanners ist die genaue Anzahl und die Größe von Hintergrundobjekten oft unklar. Als Gründe sind hier die hohe Objektdichte und die ungenaue Definition eines Hintergrundobjektes zu nennen. Entsprechende Beispiele finden sich in Bereichen mit Büschen, Wäldern oder Feldern. Hier kann in der Regel nicht entschieden werden, wie viele Objekte sich in dem entsprechenden Bereich befinden. Darüber hinaus existieren im Vergleich zu den anderen Objektklassen extrem viele Hintergrundobjekte. Aus diesen Gründen werden nur Objekte der Klassen „Fußgänger“, „Zweirad“, „PKW“ und „LKW“ markiert. Alle nicht markierten Objekte gehören folglich der Klasse Hintergrund an. Abbildung 10.2 zeigt die Markierungen für die Laserscannermessdaten eines Zeitschritts.

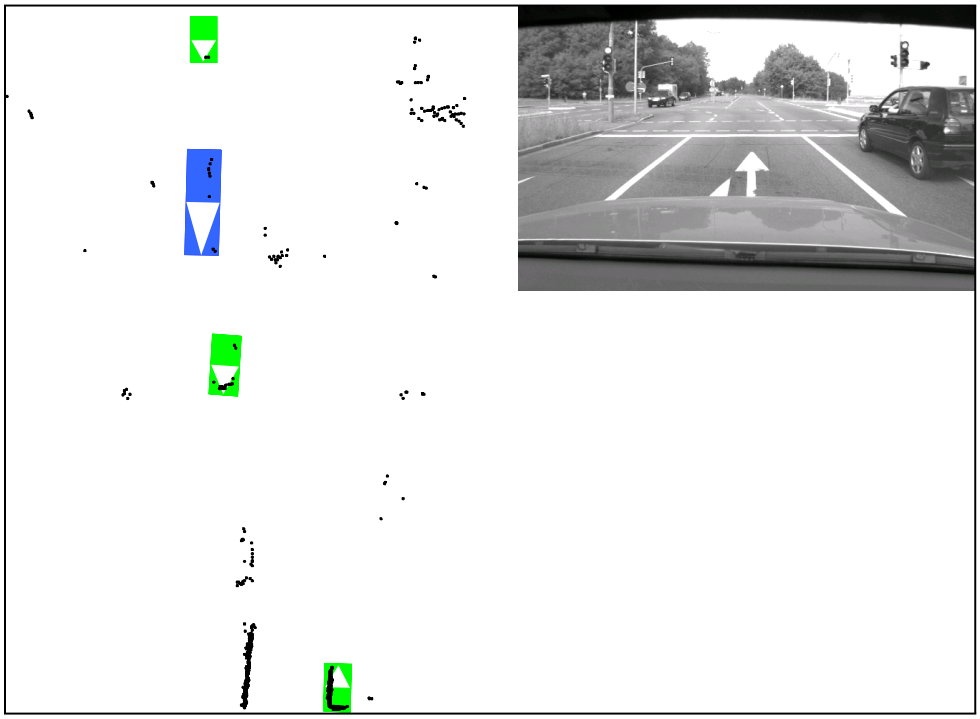

Abbildung 10.2: Beispielansicht für Objektmarkierungen eines Zeitschrittes der Laserscannermessdaten.

Zur Unterstützung des Markierungsvorgangs wurden verschiedene Hilfsmittel in das dafür entwickelte Programm integriert. Diese ermöglichen ein verbessertes Verständnis des Nutzers über die zu bearbeitenden Sensormessdaten und führen zu einer deutlichen Beschleunigung des Markierungsvorgangs. Zur visuellen Unterstützung wird ein Videobild der Szene angezeigt, in das sich auch die Messwerte des Laserscanners einblenden lassen. Die Sequenz aus Messdaten des Laserscanners und der Videokamera lässt sich vorwärts und rückwärts abspielen, um die Objekte aufgrund der Bewegung besser identifizieren zu können. Darüber hinaus wurde die Objektverfolgung auch in dieses Programm integriert. Der Nutzer bekommt nach der einmaligen Markierung eines Objekts im nächsten Zeitschritt einen Markierungsvorschlag an der neuen Objektposition. Dieser Vorschlag muss nun nur noch korrigiert werden, falls bei der Objektverfolgung ein Fehler auftritt.

Beim Markieren der Messdaten traten in der Praxis in hohen Distanzen oft Fehler auf. Verschiedene Objekte wurden übersehen, da sie nur gelegentlich vom Sensor erfasst wurden. Um diese Fehler nicht mit großem Aufwand manuell korrigieren zu müssen, werden alle Objekte in Distanzen von mehr als 150 m vor dem Versuchsfahrzeug oder in Distanzen von mehr als 100 m links oder rechts daneben ignoriert. Die so festgelegte ROI weist im Vergleich zu anderen Arbeiten dennoch eine beachtliche Größe auf.

Für das Training und die Auswertung der Klassifikation müssen nun aus den Markierungen in den entsprechenden Datensätzen Klassenvorgaben für die aktuellen Objekte extrahiert werden. Hierbei können in Einzelfällen verschiedene Mehrdeutigkeiten entstehen (Abbildung 10.3).

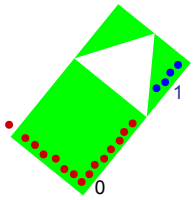

Abbildung 10.3: Bei der Zuordnung der markierten Klassen zu den aktuellen Objekten können verschiedene Mehrdeutigkeiten auftreten. Innerhalb einer Markierung können mehrere Objekte liegen. Manchmal können sich Objekte auch nur teilweise innerhalb der Box befinden (Objekt 0).

Falls ein Objekt aufgrund der Segmentierung in mehrere Segmente zerfällt und der Ansatz zur Segmentfusion nicht angewendet werden kann, können zwei Objekte in der Markierungsbox liegen. Da in der Realität jedoch nur ein Verkehrsteilnehmer existiert, dürfen nicht mehrere Objekte die entsprechende Klasse zugeordnet bekommen. Zusätzlich können Objekte aufgrund der Segmentierung mit Messungen anderer Objekte des Fahrzeugumfeldes verschmelzen. In diesem Fall liegt das Objekt nur teilweise in der Markierungsbox.

Für die Klassifikation wurden darum verschiedene Regeln zur Zuordnung der Klassenvorgabe festgelegt. Objekte, die nicht mindestens zu 80% innerhalb der Box liegen, werden ignoriert. Die markierte Klasse bekommt jeweils das größte Objekt in der Box zugeordnet. Die restlichen Objekte werden sowohl im Training, als auch in der Auswertung der Klassifikation ignoriert, da die Ursache der korrespondierenden Fehler in der Messdatenverarbeitung der Objektverfolgung und nicht in der Klassifikation zu suchen ist. Der Zerfall der Objekte in mehrere Segmente und die Methoden zur Segmentfusion werden in einem gesonderten Abschnitt analysiert.

10.1.3 Videomessdaten

Zum Training der Kaskadenklassifikatoren werden markierte Videodaten benötigt. Diese werden aus dem Trainingsdatensatz ausgeschnitten. Beispiele für Hintergrundobjekte werden in Form ganzer Bilder gesammelt, die keine der gesuchten Objektklassen enthalten. Der Trainingsalgorithmus schneidet aus diesen Bildern je nach Bedarf verschiedene Beispiele für Hintergrunddaten aus. Zum Training der verwendeten Kaskaden werden zusätzlich jeweils entsprechende Beispiele für die gesuchte Objektklasse benötigt. Damit der Klassifikator die Merkmalsverteilung der Objektklasse möglichst gut erlernen kann, müssen diese Beispiele so zugeschnitten werden, dass ein vergleichbares Aussehen beibehalten wird.

Zur Markierung dieser Beispieldaten wurden ebenfalls verschiedene Hilfsmittel in das entsprechende Programm integriert. Alle Algorithmen zum Markieren der Videodaten basieren auf den bereits markierten Messdaten des Laserscanners. Zur Markierung von Front- und Heckansichten von PKWs wird ein Ansatz zur Objektverfolgung durch Fusion von Video- und Laserscannermessdaten verwendet [66]. Die Markierungen werden so gelegt, dass sie genau das untere, das linke und das rechte Ende des PKWs kennzeichnen. Nach einer einmaligen Markierung eines PKWs wird im nächsten Zeitschritt für dieses Objekt automatisch ein Markierungsvorschlag erzeugt. Dafür wird aufgrund der mit dem Laserscanner ermittelten Abstandsänderung die erwartete Größenänderung im Vergleich zum

letzten Videobild berechnet. Diese Änderung wird auf den markierten Bildausschnitt des alten Videobildes angewendet. Anschließend wird mit Hilfe der Kreuzkorrelation der zugeschnittene Bildausschnitt im neuen Videobild gesucht. Der dadurch erzeugte Markierungsvorschlag muss vom Nutzer nur noch bestätigt und gegebenenfalls korrigiert werden.

Zum Training des Kaskadenklassifikators müssen die markierten Beispiele ausgeschnitten werden. Dafür wird jede Markierung nach oben erweitert, um eine quadratische Form zu erreichen. Dieses Quadrat wird in alle Richtungen um 12,5% erweitert und als Trainingsbeispiel ausgeschnitten. Anschließend werden alle Beispiele auf eine Größe von 32×32 Pixeln skaliert. In Abbildung 10.4 sind einige Beispiele für Heckansichten von PKWs dargestellt.

Abbildung 10.4: Ausgewählte Beispieldaten für Heckansichten von PKWs.

Ähnliche Methoden werden für Beispielansichten von Fußgängern verwendet. Aufgrund des variierenden Aussehens von Fußgängern sind die korrelationsbasierten Markierungsvorschläge nicht mehr in jedem Fall zweckmäßig und müssen häufiger korrigiert werden. Auch hier werden die Markierungen vergrößert, bevor die Trainingsbeispiele ausgeschnitten werden. Für das Training wurden die Beispielansichten auf eine einheitliche Größe von 20×40 Pixeln skaliert. In Abbildung 10.5 sind einige Beispiele für Fußgänger abgebildet.

Abbildung 10.5: Ausgewählte Beispieldaten für Fußgänger.

Zur Vorbereitung von Beispieldaten für Seitenansichten von PKWs wurde ein anderes Verfahren realisiert. Im Rahmen einer Studienarbeit [16] wurde ein Verfahren entwickelt, das Räder von PKWs im Videobild detektiert. Dabei werden PKWs aus den Markierungen

in den Laserscannerdaten erkannt. Die erwartete Radgröße wird aus der vom Laserscanner gemessenen Distanz bestimmt. Weiterhin wird die Orientierung der PKWs aus den Laserscannermessdaten bestimmt. Daraus kann die perspektivische Verzerrung der Räder im Videobild abgeschätzt werden. Entsprechende Ellipsen werden dann im Videobild mit Hilfe von Kantenbildern und einer Hough-Transformation für Ellipsen gesucht. Erkennen die Algorithmen zwei Räder für ein Objekt, wird aufgrund dieser Positionen eine Markierung berechnet, in der die Räder an einer festen Position liegen (Abbildung 10.6).

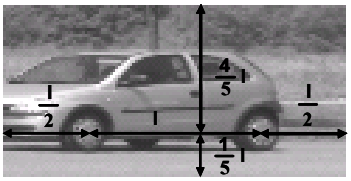

Abbildung 10.6: Beispiele für Seitenansichten von PKWs werden so zugeschnitten, dass sich die Räder jeweils an einer festen Position befinden.

Die Markierungen werden dann automatisch ausgeschnitten und auf eine Größe von 64 × 32 Pixeln skaliert. Der Nutzer muss nun lediglich die Beispieldaten überprüfen und fehlerhafte Beispiele entfernen. Alternativ können auch manuell Räder in Videobildern markiert werden. In Abbildung 10.7 sind einige Beispiele für Seitenansichten von PKWs dargestellt.

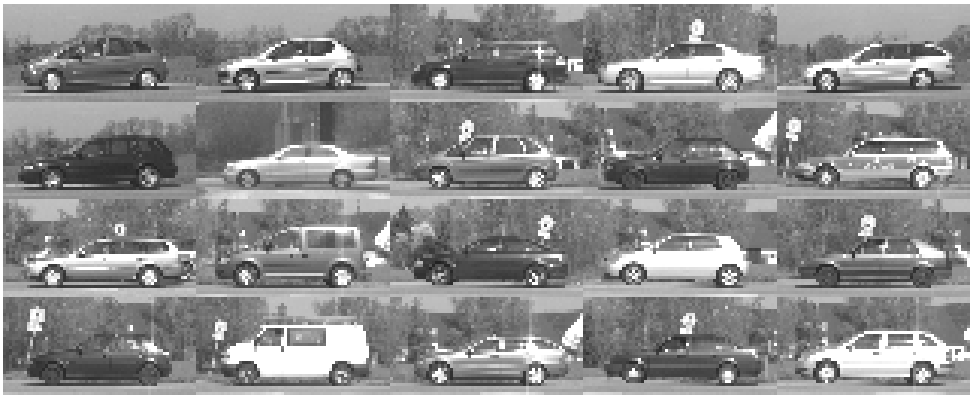

Abbildung 10.7: Ausgewählte Beispieldaten für Seitenansichten von PKWs.

10.2 Zeitmessungen

Für die Anwendung der Algorithmen in Echtzeit erfolgte im Rahmen der Auswertung auch eine Abschätzung der benötigten Rechenzeit. Üblicherweise wurde für die Analyse der Rechenzeit ein PC mit einem Intel Pentium D 3,4 GHz Prozessor verwendet. Der Prozessor enthält zwei Prozessorkerne, von denen jedoch nur einer für die Anwendung genutzt wurde. Eine Ausnahme bildet die Auswertung der durch Optimierungen erreichten Rechenzeitreduktion bei der Objektverfolgung. Diese ältere Analyse wurde mit einem Notebook mit einem Intel Pentium M 1,7 GHz Prozessor durchgeführt.

Während der Ausführung von Algorithmen unter einem normalen PC-Betriebssystem werden in der Regel auch Teile der Rechenzeit für andere Anwendungen und Prozesse verwendet. Dieser Einfluss ist insbesondere bei Einzelkernprozessoren sehr stark. Bei Systemen mit mehreren Prozessoren oder mehreren Prozessorkernen ist der Einfluss geringer, wenn die Algorithmen nicht alle verfügbaren Prozessorkerne ausnutzen. Hintergrundaktivitäten des Betriebssystems werden dann in der Regel mit den verbleibenden Prozessorkernen durchgeführt.

Die hier entwickelten Algorithmen werden auf einem normalen PC-Betriebssystem offline ausgewertet. Für eine Online-Anwendung werden jedoch in der Praxis Echtzeitbetriebssysteme oder zumindest angepasste Konfigurationen mit sehr geringer Hintergrundaktivität eingesetzt. Um den Einfluss anderer Prozesse bei der Ermittlung der Rechenzeit möglichst gering zu halten, wurde ein spezielles Verfahren zur Bestimmung der notwendigen Rechenzeit entwickelt. Das Verfahren beruht darauf, dass bei mehreren Offline-Durchläufen der deterministischen Algorithmen auf den selben aufgezeichneten Daten genau die selben Rechenoperationen durchgeführt werden müssen. Zur Ermittlung von Rechenzeiten werden darum sämtliche Testdaten mit den zu untersuchenden Algorithmen mehrfach durchlaufen. Anschließend wird für jeden Zeitschritt in den Testdaten das Minimum aller Durchläufe als notwendige Rechenzeit festgelegt. Höhere Rechenzeiten anderer Durchläufe sind dabei durch Hintergrundaktivitäten des Betriebssystems begründet. In Abbildung 10.8 werden die derart ermittelten Rechenzeiten der einzelnen Zeitschritte mit den Rechenzeiten eines einzelnen Durchlaufs für einen Einzelkernprozessor verglichen.

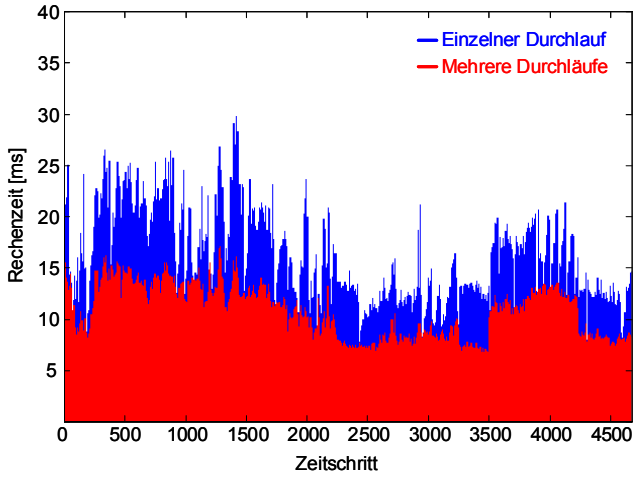

Abbildung 10.8: Bei Zeitmessungen unter Microsoft Windows variiert die benötigte Rechenzeit aufgrund von Hintergrundaktivitäten. Mit dem Minimum der Rechenzeiten aus mehreren Durchläufen des selben Rechenschrittes deterministischer Algorithmen lässt sich die Rechenzeit exakter bestimmen.

Mit Hilfe der ermittelten Rechenzeiten für die einzelnen Zeitschritte werden nun Aussagen über die benötigte Rechenzeit des Systems getroffen. Üblicherweise werden für Algorithmen die mittlere und die maximale Rechenzeit pro Zeitschritt angegeben. Diese Zeitangaben sind jedoch für Echtzeitanwendungen nur bedingt sinnvoll. Eine mittlere Rechenzeit in der Größe des Messintervalls kann nicht ausreichend sein, da in der Praxis höhere wie niedrigere

Rechenzeiten in Abhängigkeit vom Fahrzeugumfeld oft gehäuft auftreten. Treten höhere Rechenzeiten dabei über einen längeren Zeitraum auf, ist das System nicht in der Lage, sämtliche Messdaten zu verarbeiten. Eine komplette Verarbeitung aller Messdaten wird durch die Berücksichtigung der maximalen Rechenzeit des Systems sichergestellt. Diese Angabe ist jedoch meist zu hoch, da die maximale Rechenzeit sehr häufig nur als einzelner Peak auftritt. Ist das System in der Lage, Messungen zwischenzuspeichern, dann kann es auch kurzfristig Rechenzeiten tolerieren, die höher sind als das verwendete Messintervall. Die mittlere und die maximale Rechenzeit werden im Rahmen dieser Arbeit in folgender Form angegeben:

$$\textit{mittlere Rechenzeit}\ [\,\mathrm{ms}]\ \ (\textit{maximale Rechenzeit}\ [\,\mathrm{ms}])\ .$$

Um eine passendere Bewertung der Rechenzeit im Hinblick auf die echtzeitfähige Anwendung durchführen zu können, werden zusätzlich zu den hier angegebenen mittleren und maximalen Rechenzeiten noch zwei zusätzliche Maße vorgestellt. Diese Maße geben den Anteil der zur Verfügung stehenden Rechenzeit an, der für die Algorithmen benötigt wird. Die benötigte Rechenkapazität pro Messintervall entspricht dabei der maximalen Rechenzeit. Sie gibt an, welcher Anteil der 80 ms eines Messintervalls maximal von den Algorithmen benötigt wird.

In einem zweiten Maß wird die benötigte Rechenkapazität in einem Zeitintervall von einer Sekunde festgehalten. Dieses Maß entspricht einer Pufferung von bis zu 12 Messwerten. Dieses Maß ist eine fairere Bewertung, da ein System mit einer entsprechenden Pufferung immer noch echtzeitfähig ist, wenn die Verarbeitung eines einzelnen Messwertes mehr als 80 ms in Anspruch nimmt, in einer Sekunde aber alle 12 Messwerte verarbeitet werden können.

Im Rahmen dieser Arbeit wird die benötigte Rechenkapazität folgendermaßen angegeben:

$$\textit{Rechenkapazität pro Sekunde}\,[\%]\ \ (\textit{Rechenkapazität pro Messintervall}\,[\%])\ .$$

Nach der vorgestellten Bewertungsmethode kann also das für die Auswertung verwendete System mit der benötigten Rechenkapazität von 70% (105%) immer noch online alle Messwerte verarbeiten.

In dieser Arbeit wurden alle Algorithmen mit PC-Hardware ausgewertet. Auf diese Weise wird prinzipiell die Echtzeitfähigkeit der entwickelten Algorithmen gezeigt. Eine spätere Serienfertigung von Assistenzsystemen wird in der Regel auf speziellen Steuergeräten realisiert. Entsprechende Steuergeräte werden voraussichtlich eine geringere Rechenleistung zur Verfügung stellen. Dafür steht jedoch die komplette Rechenkapazität exklusiv den Algorithmen zur Verfügung. Zusätzlich bieten gezielte Optimierungen der Implementierung für spezielle Chips wie FPGAs, Methoden der parallelisierten Datenverarbeitung. Schließlich kann die benötigte Rechenzeit auch durch eine Begrenzung der verarbeiteten Objekte reduziert werden. In der vorliegenden Auswertung wurde keine derartige Begrenzung durchgeführt. Alle gemessenen Objekte wurden vom System verfolgt und klassifiziert.

10.3 Objektverfolgung

Eine detaillierte Auswertung der Objektverfolgung bezüglich der Genauigkeit der Positions- und Geschwindigkeitsschätzungen wurde bereits in [66] präsentiert. Da die Objektverfolgung grundsätzlich aus der entsprechenden Arbeit übernommen wurde, erfolgt hier lediglich eine Auswertung der veränderten Komponenten. Diese Auswertungen evaluieren die durch Optimierungen der Objektverfolgung erreichte Reduktion der Rechenzeit, die Orientierungsgenauigkeit der verwendeten Methode zur Formbestimmung und die Leistungsfähigkeit der entwickelten Ansätze zur Segmentfusion. Darüber hinaus wird die benötigte Rechenzeit der Objektverfolgung analysiert.

10.3.1 Rechenzeitreduktion durch die Optimierung der Objektverfolgung

Die bestehende Objektverfolgung wurde an verschiedenen Stellen zur Reduktion der benötigten Rechenzeit optimiert. Dies geschah insbesondere auch in Hinblick auf eine echtzeitfähige Anwendung auf weniger leistungsfähigen Rechnern. Die Analyse der benötigten Rechenzeit erfolgte mit einem 1,7 GHz Notebook. Zu Vergleichszwecken wurden dabei lediglich Mittelwerte für die benötigte Rechenzeit bestimmt.

In Abschnitt 5.5 wurde eine Optimierung der Messvektoren zur verbesserten Handhabung der zahlreichen Referenzpunkte präsentiert. Entsprechende Änderungen erfolgten in der Prädiktion und der Innovation. Zusätzlich wurde eine zweite Merkmalsprädiktion durchgeführt. Diese Veränderungen erzielten eine Reduktion der Rechenzeit von 5,5 ms auf 1,5 ms.

Eine neue Methode zur Formbestimmung ersetzt weiterhin den rechenintensiveren WICP-Algorithmus. Hier wurde eine Reduktion der Rechenzeit von 8,2 ms auf 1,3 ms erreicht.

Insgesamt konnte mit den durchgeführten Optimierungen also ein erheblicher Anteil der Rechenzeit eingespart werden.

10.3.2 Auswertung der Orientierung

Im Rahmen der Optimierung wurde für die Formbestimmung aus den einzelnen Entfernungsmessungen eine neue Methode entwickelt (Abschnitt 5.6). Dadurch verändert sich auch die Methode der Orientierungsbestimmung. Aus diesem Grund wurde das Messrauschen für die Orientierung neu ermittelt.

Die Bestimmung des Messrauschens wurde mit dem gleichen Versuchsaufbau, wie in [66] durchgeführt. Das Versuchsfahrzeug wurde parallel neben eine gerade Straße gestellt und zeichnete passierende Fahrzeuge in beiden Fahrtrichtungen auf (Abbildung 10.9). Die aufgezeichneten Fahrzeuge wurden mit den in Abschnitt 10.1 beschriebenen Methoden in den Laserscannermessdaten markiert. Die reale Orientierung der vermessenen PKWs ist somit 0° beziehungsweise 180°. Da aus den Laserscannermessdaten nur Orientierungen im Bereich $\pm 45^\circ$ geschätzt werden, wurde als Vorgabe immer 0° verwendet.

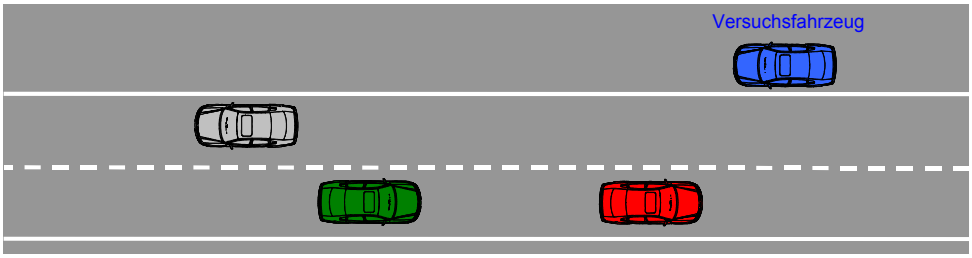

Abbildung 10.9: Zur Bestimmung des Messrauschens der Orientierung wurde das Versuchsfahrzeug parallel an einer geraden Straße aufgestellt. Die auf der Straße fahrenden Fahrzeuge haben somit eine reale Orientierung von 0°.

Auf diese Weise wurden insgesamt 5711 Orientierungsmessungen erzeugt. Auf dieser Basis wurde nun die Varianz des Messrauschens in Abhängigkeit von dem bei der Formbestimmung ermittelten Gütekriterium bestimmt. Das in Gleichung 5.24 beschriebene Gütekriterium s_{Form} setzt sich aus dem Abstand der Entfernungsmessungen von der Form und der Anzahl der Entfernungsmessungen zusammen.

Um die Varianz in Abhängigkeit von s_{Form} bestimmen zu können, wurden die Messungen nach dem Gütekriterium sortiert. Nun wurde jeweils für zusammenhängende Gruppen von je 200 Messungen eine Varianz bestimmt. Diese Varianzen wurden jeweils über dem mittleren Gütekriterium der Gruppe in Abbildung 10.10 dargestellt.

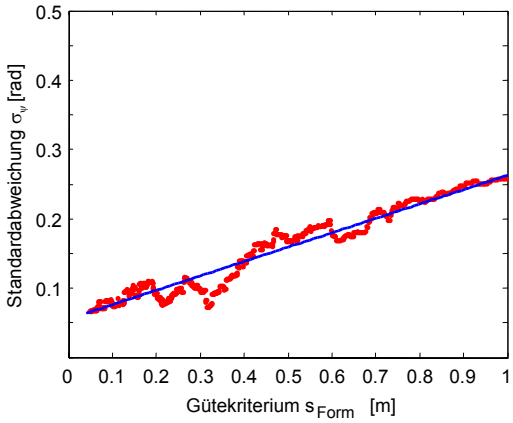

Abbildung 10.10: Bestimmung der Standardabweichung der gemessenen Orientierung in Abhängigkeit des Gütekriteriums der Formbestimmung: Für verschiedene Bereiche des Gütekriteriums wurde die Standardabweichung ermittelt. Die Verteilung der Standardabweichung wird durch eine Regressionsgerade approximiert.

Die Verteilung der Varianzen wurde mit einer Regressionsgerade approximiert. Damit ergibt sich für die Standardabweichung des Orientierungsmessrauschens:

$$\sigma_{\psi} = 0{,}21 \frac{\text{rad}}{\text{m}} \cdot s_{Form} + 0{,}05\,\text{rad}\,. \tag{10.1}$$

10.3.3 Segmentfusion

Die Auswertung der vorgestellten Ansätze zur Segmentfusion wurde mit dem in Abschnitt 10.1 vorgestellten Testdatensatz ausgewertet. Die Auswertung erfolgte nur mit markierten Objekten. Objekte der Klasse „Hintergrund“ wurden also nicht verwendet. Aufgrund der Markierung lässt sich feststellen, in wie viele Segmente ein markiertes Objekt zerfällt. Insgesamt 11,1% der markierten Verkehrsteilnehmer der Klassen „Fußgänger“, „Zweirad“, „PKW“ und „LKW“ bestanden aus mehr als einem Segment. Das bedeutet, dass in Ansätzen ohne Segmentfusion für diese Verkehrsteilnehmer mehr als ein Objekt vom System ausgegeben würde. Diese Mehrfachausgabe kann für das Fahrerassistenzsystem irreführend wirken und unter Umständen zu einem Fehlverhalten des Systems führen.

Durch die Anwendung der Segmentfusion wurden 66,4% der in mehrere Segmente zerfallenen Verkehrsteilnehmer vom System korrekt als ein Objekt ausgegeben. Das bedeutet eine Reduktion der durch Zerfall erzeugten Fehlausgaben auf 3,7% aller Verkehrsteilnehmer. Bei manchen Verkehrsteilnehmern konnte durch die Segmentfusion die Zahl der ausgegebenen Objekte reduziert werden, obwohl weiterhin mehr als ein Objekt ausgegeben wird. Insgesamt konnte die Ausgabe so bei 77,1% der zerfallenen Verkehrsteilnehmer verbessert werden.

Die verbleibenden Fehlausgaben sind durch die Kriterien der Segmentfusion begründet. Das Verfahren ist nur für sichere Geschwindigkeitsschätzungen anwendbar. Somit ist eine Anwendung während des Einschwingens des Kalman-Filters in den ersten Messzyklen nicht möglich. Weiterhin ist der berücksichtigte Sicherheitsabstand bei langsamen Objektgeschwindigkeiten manchmal für eine Fusion der zusammengehörenden Segmente nicht ausreichend.

Insgesamt wurden in allen Testdaten lediglich in 18 Fällen Segmente zusammengefügt, die nicht zum gleichen Objekt gehören. Dies entspricht etwa 0,02% der untersuchten Objekte. Hier wurden Segmente der Verkehrsteilnehmer mit Segmenten eines Hintergrundobjektes zusammengefügt. In diesen Fällen wurden Ungenauigkeiten in der geschätzten Geschwindigkeitsrichtung und sehr geringe Abstände zwischen Verkehrsteilnehmer und Hintergrundobjekt als Ursache ermittelt.

10.3.4 Rechenzeit

Die benötigte Rechenzeit von Vorverarbeitung und Objektverfolgung wurde mit den in Abschnitt 10.2 beschriebenen Methoden analysiert. Die Zeitmessungen wurden mit dem 3,4 GHz PC durchgeführt.

Die benötigte Rechenzeit pro Messintervall ist 6,3 ms (38,3 ms). Der Anteil der benötigten Rechenkapazität beträgt 32,3% (47,8%). Obwohl die benötigte Rechenzeit in der Regel sehr niedrig ist, treten in einigen Szenarien gehäuft Messintervalle auf, in denen deutlich mehr Rechenzeit benötigt wird. Ein markantes Beispiel für ein solches Szenario ist eine Sequenz an einer Einmündung im Wald. Der Laserscanner misst hier aufgrund des fehlenden Unterholzes sehr tief in den Wald hinein. Von der Objektverfolgung wird aufgrund des Abstandes für jeden gemessenen Baum ein Objekt aufgesetzt. Die auf diese Weise in der Objektverfolgung bearbeitete, sehr hohe Objektzahl begründet die höhere Rechenzeit. Eine

Reduktion der Rechenzeit in solchen Szenarien kann durch eine Reduktion der ausgewerteten Sensorreichweite oder eine Reduktion der maximalen Objektzahl erreicht werden. Für eine Reduktion der maximalen Objektzahl müssten jedoch Methoden entwickelt werden, mit denen entschieden werden kann, welche der gemessenen Objekte die Objektverfolgung bearbeiten soll. Als Basis für eine derartige Objektauswahl sind beispielsweise die Entfernung oder die Lage der Messdaten denkbar.

10.4 Klassifikation

10.4.1 Verwendete Bewertungsmethode

Die Ergebnisse der Klassifikation eines Testdatensatzes können in einer sogenannten Konfusionsmatrix zusammengefasst werden. Dabei werden die Ausgaben des Klassifikators den vorgegebenen Sollklassen gegenübergestellt. Nun wird die Anzahl aller korrekt klassifizierten Hintergrundobjekte als „richtig negative“ und die Anzahl der falsch klassifizierten Hintergrundobjekte als „falsch positive“ bezeichnet. Die gesuchten Objekte werden ähnlich unterschieden. Die Anzahl der korrekten Klassifikationen wird „richtig positive“ und die Anzahl der übersehenen Objekte „falsch negative“ genannt. In Abbildung 10.11 sind die entsprechenden Bezeichnungen für die Ergebnisse der Klassifikation des Zweiklassenproblems der Fußgängerdetektion als Konfusionsmatrix tabellarisch dargestellt.

	Hintergrund-objekt	Fußgänger
Klassifizierter Hintergrund	richtig negative	falsch negative
Klassifizierter Fußgänger	falsch positive	richtig positive

Abbildung 10.11: Mögliche Ergebnisse der Klassifikation für ein Zweiklassenproblem: Die Anzahl korrekt klassifizierter Hintergrundobjekte wird als „richtig negative“, die Anzahl der Fehldetektionen mit „falsch positive“ bezeichnet. Die Anzahl korrekt klassifizierter Exemplare der gesuchten Objekte werden „richtig positive“, die Anzahl übersehener Objekte der gesuchten Objektklasse „falsch negative“ genannt.

Zur Auswertung der Klassifikationsergebnisse muss die Anzahl der Falschalarme („falsch positive“) und die Anzahl der korrekt erkannten Objekte bewertet werden („richtig positive“). Üblicherweise werden für Zweiklassenprobleme die Detektionsrate und die Falschalarmrate berechnet. Die Detektionsrate wird auch als Sensitivität bezeichnet und gibt an, wie viel Prozent der gesuchten Objekte korrekt klassifiziert wurden:

$$Detektionsrate = \frac{„richtig\ positive“}{„richtig\ positive“ + „falsch\ negative“} . \tag{10.2}$$

Die Falschalarmrate bezeichnet das Verhältnis von falsch klassifizierten Hintergrundobjekten zu allen Hintergrundobjekten:

$$Falschalarmrate = \frac{„falsch\ positive“}{„richtig\ negative“ + „falsch\ positive“} . \tag{10.3}$$

Wie bereits in Abschnitt 10.1 angesprochen, ist die Anzahl der Hintergrundobjekte für das hier betrachtete Problem der Fahrzeugumfelderfassung nicht eindeutig bestimmbar. Darüber hinaus ist die Anzahl der Hintergrundobjekte im Vergleich zu den Vertretern der gesuchten Objektklassen extrem hoch. Da aus diesem Grund keine Markierung der Hintergrundobjekte erfolgt, ist eine Nutzung der Anzahl der korrekt klassifizierten Hintergrundobjekte (*„richtig negative“*) zur Auswertung nicht sinnvoll. Die Auswertung der Falschalarmanzahl (*„falsch positive“*) ist weiterhin möglich, da die entsprechenden Objekte in Form von Fehlklassifikationen gezählt werden können.

Um bei der Auswertung die Anzahl der korrekt klassifizierten Hintergrundobjekte umgehen zu können, existieren zwei mögliche Alternativen für die Falschalarmrate. Die erste Alternative stellt die Anzahl der Falschalarme der Anzahl der ausgewerteten Zeitschritte gegenüber und berechnet somit eine mittlere Anzahl von Falschalarmen pro Zeiteinheit. Im Rahmen dieser Arbeit wurde eine zweite Alternative verfolgt, die den Anteil von Falschalarmen an allen ausgegebenen, als Vordergrund klassifizierten Objekten angibt. Dieses Maß wird als Fehldetektionsrate bezeichnet:

$$Fehldetektionsrate = \frac{„falsch\ positive“}{„richtig\ positive“ + „falsch\ positive“} . \tag{10.4}$$

Die Anwendung der Klassifikation stellt in der Regel einen Kompromiss zwischen einer niedrigen Anzahl von Falschalarmen und einem hohen Anteil korrekt detektierter Objekte dar. Aus diesem Grund werden zur Bewertung eines Klassifikators üblicherweise mehrere mögliche Paare von Detektionsrate und Falschalarmrate in Form einer Kurve präsentiert. Dabei wird die Detektionsrate über der Falschalarmrate abgetragen. Diese Kurve wird ROC (*engl. ROC = receiver operating characteristic*) genannt.

Aufgrund der Nutzung der Fehldetektionsrate anstatt der Falschalarmrate sollen entsprechende Kurven hier als alternative ROC bezeichnet werden. Ebenso wie bei den üblichen ROCs würde ein fehlerfreier Klassifikator einen Punkt links oben im entsprechenden Koordinatensystem erreichen. An diesem Punkt werden alle Objekte korrekt detektiert, wobei keinerlei Falschalarme ausgegeben werden. Das Ziel bei der Entwicklung von Klassifikatoren ist also auch bei alternativen ROCs, mit der Kurve möglichst dicht an diesen Punkt heranzukommen.

Prinzipiell wird die ROC zur Auswertung von Zweiklassenproblemen verwendet. Eine Nutzung für die hier untersuchten Mehrklassenprobleme ist nur mit zusätzlichen Definitionen möglich. Während im Rahmen von [35] mehrdimensionale Formulierungen für ROCs untersucht wurden, soll hier weiterhin eine zweidimensionale Darstellung in Form von alternativen ROCs erfolgen.

Um dennoch mit den eingeführten Bezeichnungen für die Klassifikationsergebnisse arbeiten zu können, werden diese erweitert. Korrekt klassifizierte Hintergrundobjekte werden weiterhin als „richtig negativ“ bezeichnet. Die Bezeichnung „falsch positiv“ umfasst nun alle

Objekte, die fälschlicherweise als „Fußgänger“, „Zweirad“, „PKW“ oder „LKW“ klassifiziert wurden. Als „richtig positiv“ werden alle korrekt klassifizierten Objekte der Klassen „Fußgänger“, „Zweirad“, „PKW“ und „LKW“ bezeichnet. Schließlich wird die Bezeichnung „falsch negativ“ für alle Objekte der Klassen „Fußgänger“, „Zweirad“, „PKW“ und „LKW“ verwendet, die nicht korrekt erkannt werden konnten.

Diese Erweiterung bedeutet in der Praxis, dass manche Objekte doppelt als Fehler registriert werden, da sie sowohl den Kriterien für „falsch negativ“ als auch für „falsch positiv“ entsprechen. Ein Beispiel für ein solches Objekt ist ein als PKW klassifizierter LKW.

Zur Einschätzung der zu erwartenden Kurvenverläufe für alternative ROCs sollen hier entsprechende Kurven für zwei Beispiele zufälliger Klassifikation bestimmt werden. Die zufällige Klassifikation führt eine Klasseneinteilung ohne Berücksichtigung von Objektmerkmalen durch und kann somit als schlechtester Fall eines Klassifikationssystems angesehen werden. Klassifikatoren, die basierend auf berechneten Objektmerkmalen zusätzliche Informationen nutzen, sollten entsprechend bessere Leistungen erzielen.

Für die Beispielkurven wird hier eine a priori Verteilung der einzelnen Objektklassen in den Testdaten angenommen. Es gilt:

$$p_0^{soll} = p_{Hintergrund} = 0{,}60 \tag{10.5}$$
$$p_1^{soll} = p_{Fußgänger} = 0{,}15 \tag{10.6}$$
$$p_2^{soll} = p_{Zweirad} = 0{,}05 \tag{10.7}$$
$$p_3^{soll} = p_{PKW} = 0{,}15 \tag{10.8}$$
$$p_4^{soll} = p_{LKW} = 0{,}05\ . \tag{10.9}$$

Die zufällige Klassifikation führt nun unter Berücksichtigung der Klassifikationswahrscheinlichkeiten p_0^{ist}, ..., p_4^{ist} eine zufällige Klasseneinteilung für jedes der m Beispiele des Testdatensatzes durch.

Für das Zweiklassenproblem der Fußgängerdetektion ergibt sich dabei:

$$\text{„}richtig\ positive\text{“} = m \cdot p_1^{ist} \cdot p_1^{soll} \tag{10.10}$$
$$\text{„}falsch\ positive\text{“} = m \cdot p_1^{ist} \cdot p_0^{soll} = m \cdot p_1^{ist} \cdot \left(1 - p_1^{soll}\right) \tag{10.11}$$
$$\text{„}falsch\ negative\text{“} = m \cdot p_0^{ist} \cdot p_1^{soll} = m \cdot \left(1 - p_1^{ist}\right) \cdot p_1^{soll}. \tag{10.12}$$

Somit können Detektionsrate und Fehldetektionsrate berechnet werden:

$$Detektionsrate = \frac{m \cdot p_1^{ist} \cdot p_1^{soll}}{m \cdot p_1^{ist} \cdot p_1^{soll} + m \cdot \left(1 - p_1^{ist}\right) \cdot p_1^{soll}} = p_1^{ist} \tag{10.13}$$

$$Fehldetektionsrate = \frac{m \cdot p_1^{ist} \cdot \left(1 - p_1^{soll}\right)}{m \cdot p_1^{ist} \cdot p_1^{soll} + m \cdot p_1^{ist} \cdot \left(1 - p_1^{soll}\right)} = 1 - p_1^{soll}. \tag{10.14}$$

Da die a priori Verteilung der Klassen vorgegeben ist, kann die Kurve nur durch eine beliebige Wahl der Klassifikationswahrscheinlichkeit p_1^{ist} erzeugt werden. Die dadurch erzeugte alternative ROC ist in Abbildung 10.12 dargestellt.

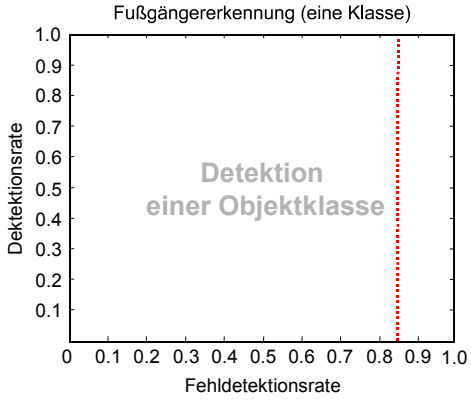

Abbildung 10.12: Alternative ROC für die zufällige Klassifikation des Beispiels der Fußgängerdetektion: Durch die Veränderung der Klassifikationswahrscheinlichkeit für Fußgänger wird die Detektionsrate verändert. Die Fehldetektionsrate bleibt konstant und entspricht der a priori angenommenen Auftrittswahrscheinlichkeit für Objekte, die keine Fußgänger darstellen.

Wird nun das Multiklassenproblem betrachtet, ergibt sich als Ergebnis der zufälligen Klassifikation:

$$\text{„}richtig\ positive\text{“} = m \cdot \sum_{i=1}^{4} \left(p_i^{ist} \cdot p_i^{soll} \right) \tag{10.15}$$

$$\text{„}richtig\ positive\text{“} + \text{„}falsch\ negative\text{“} = m \cdot \sum_{i=1}^{4} p_i^{soll} \tag{10.16}$$

$$\text{„}falsch\ positive\text{“} = m \cdot \sum_{i=1}^{4} \left(\left(1 - p_i^{soll} \right) \cdot p_i^{ist} \right) \tag{10.17}$$

$$\text{„}falsch\ positive\text{“} + \text{„}richtig\ positive\text{“} = m \cdot \sum_{i=1}^{4} p_i^{ist} \ . \tag{10.18}$$

Mit diesen Angaben lässt sich für das Mehrklassenproblem die maximal erreichbare Detektionsrate bestimmen:

$$Detektionsrate \ = \ \frac{m \cdot \sum_{i=1}^{4} \left(p_i^{ist} \cdot p_i^{soll} \right)}{m \cdot \sum_{i=1}^{4} p_i^{soll}} \tag{10.19}$$

$$Detektionsrate \ = \ \frac{\sum_{i=1}^{4} \left(p_i^{ist} \cdot p_i^{soll} \right)}{1 - p_0^{soll}}. \tag{10.20}$$

Nun wird die a priori Wahrscheinlichkeit der am häufigsten auftretenden Vordergrundklasse als p_{max}^{soll} bezeichnet. In diesem Fall ist $p_{max}^{soll} = p_1^{soll} = 0{,}15$. Somit lässt sich für die Detektionsrate eine obere Schranke finden:

$$Detektionsrate \leq \frac{\sum_{i=1}^{4} \left(p_i^{ist} \cdot p_{max}^{soll}\right)}{1 - p_0^{soll}} \leq \frac{p_{max}^{soll}}{1 - p_0^{soll}} \tag{10.21}$$

$$Detektionsrate \leq \frac{0{,}15}{1 - 0{,}60} = 0{,}375. \tag{10.22}$$

Analog lässt sich eine untere Schranke für die Fehldetektionsrate berechnen:

$$Fehldetektionsrate = \frac{m \cdot \sum_{i=1}^{4} \left(\left(1 - p_i^{soll}\right) \cdot p_i^{ist}\right)}{m \cdot \sum_{i=1}^{4} p_i^{ist}} \tag{10.23}$$

$$Fehldetektionsrate = \frac{\sum_{i=1}^{4} p_i^{ist} - \sum_{i=1}^{4} \left(p_i^{soll} \cdot p_i^{ist}\right)}{\sum_{i=1}^{4} p_i^{ist}} \geq 1 - p_{max}^{soll} \tag{10.24}$$

$$Fehldetektionsrate \geq 1 - 0{,}15 = 0{,}85. \tag{10.25}$$

Somit ist die alternative ROC der zufälligen Klassifikation für das in der Fahrzeugumfelderkennung verwendete Mehrklassenproblem auf das in Abbildung 10.13 dargestellte Rechteck beschränkt.

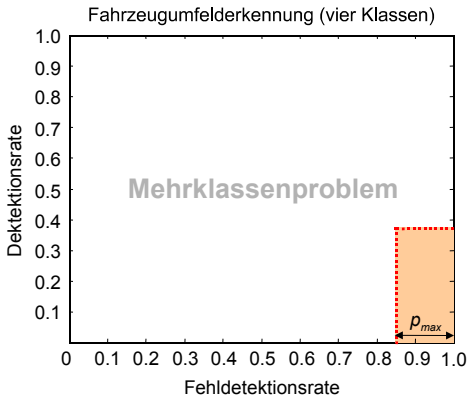

Abbildung 10.13: Alternative ROC für die zufällige Klassifikation des in der Fahrzeugumfelderkennung verwendeten Mehrklassenproblems: Im Vergleich zum Zweiklassenproblem lassen sich nur deutlich niedrigere Detektionsraten erreichen.

Beim Vergleich mit der alternativen ROC für das Zweiklassenproblem lässt sich die Komplexität des Mehrklassenproblems erkennen. Beim Mehrklassenproblem lassen sich mit der zufälligen Klassifikation nur deutlich niedrigere Detektionsraten erreichen. Dies ist dadurch begründet, dass Falschalarme nicht nur aus fehlerhaft klassifizierten Hintergrundobjekten bestehen, sondern hier auch Fehlklassifikationen innerhalb der gesuchten Objektklassen einfließen.

Die Bestimmung der alternativen ROC für trainierte Klassifikatoren erfolgt durch die Wahl geeigneter a priori Klassengewichte und einer Auswertung der Klassifikationsergebnisse für den Testdatensatz. Werden alle Testdaten mit einem trainierten System klassifiziert, ergibt sich lediglich eine Detektionsrate und eine Fehldetektionsrate. Durch die in Abschnitt 6.6 vorgestellte a priori Gewichtung einzelner Klassen können weitere Kombinationen dieser Bewertungsmaße ermittelt werden. Zur Erstellung der alternativen ROC werden mögliche a priori Gewichtsvektoren in einem vorgegebenen Raster durchlaufen. Das Ergebnis ist eine Punktwolke, deren Einhüllende als alternative ROC verwendet wird (Abbildung 10.14).

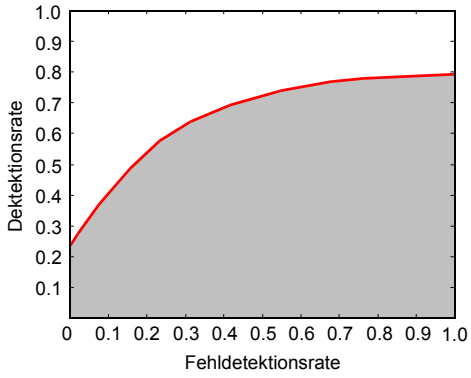

Abbildung 10.14: Alternative ROC: Für jede Veränderung der a priori Klassifikationswahrscheinlichkeiten der einzelnen Klassen durch einen additiven Gewichtsvektor ergibt sich eine Kombination aus Detektions- und Fehldetektionsrate. Iteriert man in einem vorgegebenen Raster durch alle Gewichtsvektoren ergibt sich eine Menge solcher Punkte. Die Einhüllende wird als alternative ROC verwendet.

10.4.2 Laserscannerbasierte Klassifikation

Zur Bewertung der Klassifikationsleistung der laserscannerbasierten Klassifikation wurde für den Testdatensatz eine alternative ROC ermittelt (Abbildung 10.15). Diese ROC zeigt die sehr gute Leistungsfähigkeit des Klassifikationssystems trotz der großen Herausforderung durch das zu bearbeitende Mehrklassenproblem. Zur Bewertung des in Abschnitt 6.4 vorgestellten Verfahrens zur Mustererkennung wurde dieses durch ein anderes Mustererkennungssystem aus fünf neuronalen Netzen ersetzt. In dieser Struktur trennt ebenfalls jedes Netz eine Klasse nach der One-VS-All-Methode von allen anderen. Jedes der neuronalen Netze besteht aus zwei verdeckten Schichten mit jeweils 30 Neuronen. Diese Schichten verwenden als Ausgabefunktion die Sigmoidfunktion. Das Ausgabeneuron verwendet als Ausgabefunktion eine Rampenfunktion. Diese verwendete Netzwerkstruktur wurde im Rahmen von Voruntersuchungen in einer Diplomarbeit [122] als am besten

geeignet ermittelt. Für dieses alternative Klassifikationssystem wurde eine zweite alternative ROC ermittelt und ebenfalls in Abbildung 10.15 dargestellt.

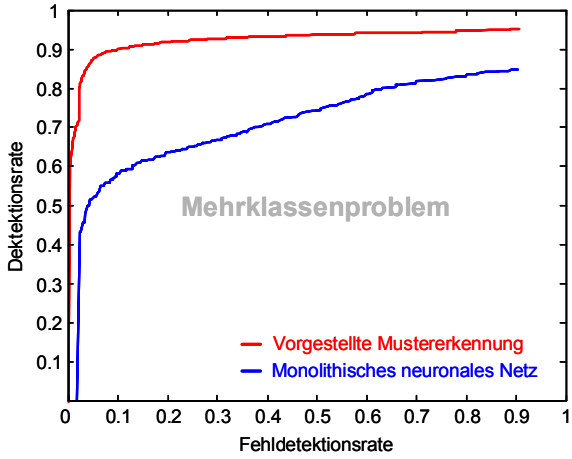

Abbildung 10.15: Vergleich der vorgestellten Mustererkennung mit normalen monolithischen neuronalen Netzen: Die alternative ROC für das laserscannerbasierte Klassifikationssystem unter Verwendung der vorgestellten Mustererkennung zeigt die deutlich höhere Leistungsfähigkeit im Vergleich zu monolithischen neuronalen Netzen.

Das vorgestellte Mustererkennungssystem konnte eine enorme Verbesserung der Klassifikationsleistung erzielen. Dafür sind zwei wesentliche Gründe zu nennen. Durch die automatische, datengetriebene Optimierung der Merkmale für jede Klasse findet bereits eine hilfreiche Vorverarbeitung der Merkmale statt. Dadurch sind weniger komplexe Netzwerke zur Kombination der Merkmale notwendig.

Bei der Verwendung monolithischer, neuronaler Netze werden zudem Netzwerkstrukturen trainiert, in denen deutlich mehr Gewichte zu bestimmen sind als bei dem vorgestellten Mustererkennungssystem. Dadurch besteht die Gefahr, dass die Menge der Trainingsdaten nicht ausreicht, um alle Parameter zu optimieren.

10.4.3 Anpassung an Anwendungsanforderungen

Das vorgestellte Klassifikationssystem berechnet pro Objekt für jede Klasse einen Zugehörigkeitswert. Aus den berechneten fünf Zugehörigkeitswerten muss vor der Ausgabe an das Assistenzsystem die ausgegebene Klasse festgelegt werden. Untersuchungen haben gezeigt, dass eine Anpassung dieser Selektion an die einzelnen Anforderungen der Assistenzsysteme deutliche Verbesserungen erreichen kann. Die Anpassung der Klassifikationsausgabe an das Assistenzsystem erfolgt durch die Wahl eines geeigneten a priori Vektors für die a priori Gewichtung der Klassen gemäß Gleichung 6.31 in Abschnitt 6.6. Darüber hinaus kann anwendungsspezifisches Vorwissen angewendet werden, um bestimmte Objektklassen oder Objekte auszuschließen. Bei der anschließenden Selektion der Klasse aufgrund des maximalen Zugehörigkeitswertes können dann für unterschiedliche

Assistenzsysteme unterschiedliche Klassen ausgegeben werden. Das Verbesserungspotential wird nun für zwei Anwendungen analysiert.

Als erste Anwendung wird der Fußgängerschutz untersucht. Diese Anwendung benötigt lediglich eine Klassifikation von Fußgängern. Fehlklassifikationen innerhalb der übrigen Klassen spielen für das Assistenzsystem keine Rolle. Somit kann das Klassifikationsproblem auf zwei Klassen reduziert werden. Als Falschalarme („falsch positive“) zählen lediglich fälschlicherweise als Fußgänger klassifizierte Objekte. Als übersehene Objekte („falsch negative“) werden hier nur nicht erkannte Fußgänger ausgewertet. Darüber hinaus kann der für den Fußgängerschutz signifikante Bereich über eine ROI eingegrenzt werden. Diese ROI wurde aus dem EU-Projekt SAFE-U [97] übernommen (Abbildung 10.16).

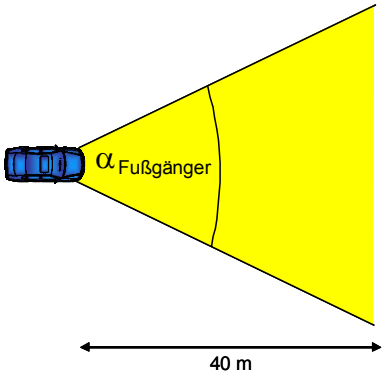

Abbildung 10.16: ROI für den Fußgängerschutz: Die ROI reicht bis 40 m vor das Versuchsfahrzeug. Die seitliche Begrenzung beginnt an den vorderen Ecken des Fahrzeugs und hat einen geschwindigkeitsabhängigen Öffnungswinkel $\alpha_{Fußgänger}$.

Die ROI reicht bis 40 m vor das Versuchsfahrzeug. Die seitliche Einschränkung beginnt an den vorderen Ecken des Versuchsfahrzeugs und begrenzt einen geschwindigkeitsabhängigen Winkelbereich der Größe $\alpha_{Fußgänger}$. Bis zu Geschwindigkeiten von 50 km/ h wird ein Winkel von $\alpha_{Fußgänger} = 63{,}7°$ verwendet. Zwischen 50 km/ h und 60 km/ h hat die ROI einen Öffnungswinkel von $\alpha_{Fußgänger} = 52{,}8°$. Bei Geschwindigkeiten über 60 km/ h beträgt $\alpha_{Fußgänger} = 44{,}9°$. Diese geschwindigkeitsabhängige Wahl des Öffnungswinkels ist dadurch begründet, dass Fußgänger in höheren Distanzen zur Straße bei hohen Geschwindigkeiten durch das Versuchsfahrzeug nicht gefährdet werden. Dieses würde die Fußgänger bereits passieren, bevor sie die Straße erreichen können. Objekte die außerhalb der ROI liegen, müssen demzufolge nicht an das Assistenzsystem übertragen werden.

Abbildung 10.17 zeigt, dass bei einer Anpassung der Ausgabe für die Anwendung Fußgängerschutz eine Verbesserung erreicht werden konnte. Die blaue Kurve zeigt die Leistung der Klassifikationsausgabe für Fußgänger, wenn bei der Wahl der a priori Vektoren alle Objekte und Objektklassen berücksichtigt werden. Diese Kurve ergibt sich aus der Anwendung des für fünf Klassen trainierten Systems, wenn für die Kurve nur die Fußgängerdetektion berücksichtigt wird. Die rote Kurve zeigt die verbesserte Klassifikationsleistung der angepassten Objektausgabe. Die Verbesserung entsteht durch die Einschränkung auf die Objekte innerhalb der ROI und die Wahl geeigneter a priori Vektoren für die Fußgänger-

detektion. In Abbildung 10.17 ist links die gesamte Kurve dargestellt, während der rechte Teil eine Vergrößerung des relevanten Bereichs darstellt.

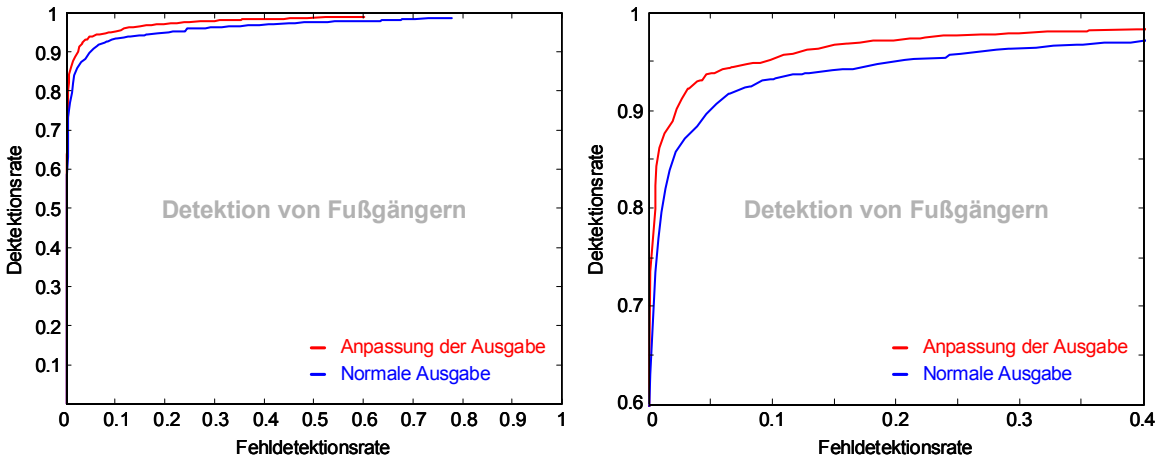

Abbildung 10.17: Verbesserung der Klassifikation durch Anpassung der Ausgabe für den Fußgängerschutz. Die blaue Kurve zeigt die Klassifikationsleistung der allgemeinen Klassifikation bei der Fußgängerdetektion. Die rote Kurve zeigt die verbesserte Klassifikationsleistung des optimierten Systems. Im linken Teil ist die komplette ROC und im rechten Teil ein vergrößerter Ausschnitt dargestellt.

Eine zweite Anwendung ist die Objektausgabe für abstandsgeregelte Tempomaten. Für diese Anwendung sollen Fahrzeuge identifiziert werden, die dem Assistenzsystem zur Abstandsregelung übergeben werden. In diesem Fall ist nur die korrekte Detektion von Fahrzeugen von Interesse. Eine Unterscheidung zwischen PKWs und LKWs ist nicht notwendig. Die Objektklassen „Fußgänger“ und „Zweirad“ werden ignoriert. Zusätzlich können die zu klassifizierenden Objekte mit Hilfe von anwendungsspezifischem Vorwissen eingeschränkt werden. Eine räumliche Einschränkung erfolgt mit einer ROI gemäß Abbildung 10.18. Diese beginnt ebenfalls an den vorderen Ecken des Versuchsfahrzeugs und hat einen Öffnungswinkel von 90°. Die betrachteten Objekte werden auch bezüglich ihrer Bewegung eingeschränkt. Hier sind nur Objekte von Interesse, denen das Versuchsfahrzeug folgen kann. Die Bewegungsrichtung der Objekte sollte darum nicht um mehr als 45° von der des Versuchsfahrzeugs abweichen. Die verbleibenden Objekte sind für das Assistenzsystem nicht von Interesse oder können als ortsfest angenommen werden. Eine Ausnahme bilden extrem langsame Fahrten in Stauszenarien. Diese erfolgen in der Regel jedoch mit extrem kurzen Abständen und sollten getrennt betrachtet werden.

Durch die gewählten Einschränkungen und die Reduktion auf eine einzelne Fahrzeugklasse wurde eine erhebliche Reduktion der Falschalarme erreicht. In Abbildung 10.19 ist die Verbesserung der Klassifikationsleistung dargestellt. Die blaue Kurve beschreibt die Klassifikationsleistung des allgemeinen Klassifikationssystems mit fünf Klassen für die Fahrzeugdetektion. Die rote Kurve beschreibt die Klassifikationsleistung bei der Optimierung der Ausgabe für die anwendungsrelevanten Objekte. Im linken Teil der Abbildung ist die gesamte ROC dargestellt. Der rechte Teil enthält eine vergrößerte Darstellung des wesentlichen Ausschnitts.

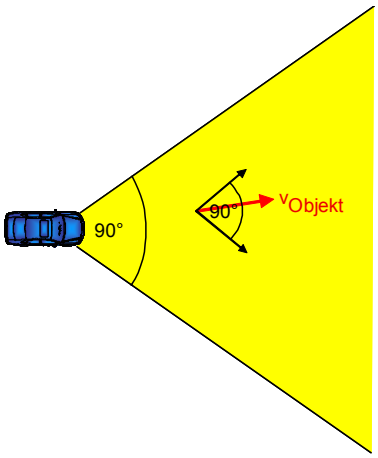

Abbildung 10.18: ROI für abstandsgeregelte Tempomaten: Die seitliche Begrenzung der ROI beginnt an den vorderen Ecken des Fahrzeugs und hat einen Öffnungswinkel von 90°. Zusätzlich werden nur Objekte betrachtet, die eine für die Verfolgung geeignete Bewegungsrichtung aufweisen.

Eine besondere Vereinfachung der Klassifikationsaufgabe für dieses Assistenzsystem stellt die Vereinigung der Klassen PKW und LKW dar. Analysen der Klassifikationsergebnisse haben gezeigt, dass LKWs in hohen Distanzen oft als PKWs klassifiziert werden. Die Ursache dafür besteht darin, dass in hohen Distanzen meist die Objektgröße nicht komplett vermessen wird. LKWs werden bei der Klassifikation mit Laserscannermessdaten in der Regel nur durch ihre Größe von PKWs unterschieden. In der Praxis ist eine Unterscheidung von LKWs und PKWs in hohen Distanzen jedoch nicht von höchster Priorität. Die Unterscheidung ist beispielsweise in kurzen Distanzen kurz vor einem Aufprall von Interesse, da der höhere Aufprallwiderstand eines LKWs beim Auslösen des Airbags berücksichtigt werden kann. In der Auswertung des Mehrklassenproblems wirken sich entsprechende Fehlklassifikationen in hohen Distanzen jedoch genau so negativ auf die Kurvenverläufe aus wie andere Fehlklassifikationen. Bei der Analyse der Ausgabe für abstandsgeregelte Tempomaten sind Fehlklassifikationen innerhalb der Klassen „PKW“ und „LKW“ hingegen nicht von Nachteil.

Die untersuchten Anwendungsbeispiele zeigen, dass die Anpassung der Ausgabe des Fahrzeugumfelderkennungssystems an die Anforderungen eines Assistenzsystems die Qualität dieser Ausgabe deutlich verbessern kann. Die Anpassung der Ausgabe benötigt extrem wenig Rechenaufwand, da pro Assistenzsystem lediglich eine Vektoraddition zur a priori Gewichtung der einzelnen Klassen durchgeführt wird. Darüber hinaus, kann noch eine Einschränkung der Ausgabe oder des klassifizierten Bereiches mit Hilfe einer ROI erfolgen. Diese Einschränkung lässt sich ebenfalls ohne nennenswerten Rechenaufwand erledigen. Somit können die im Fahrzeugumfeldmodell gespeicherten Informationen für diverse Assistenzsysteme angepasst werden, ohne dass der benötigte Rechenaufwand signifikant ansteigt.

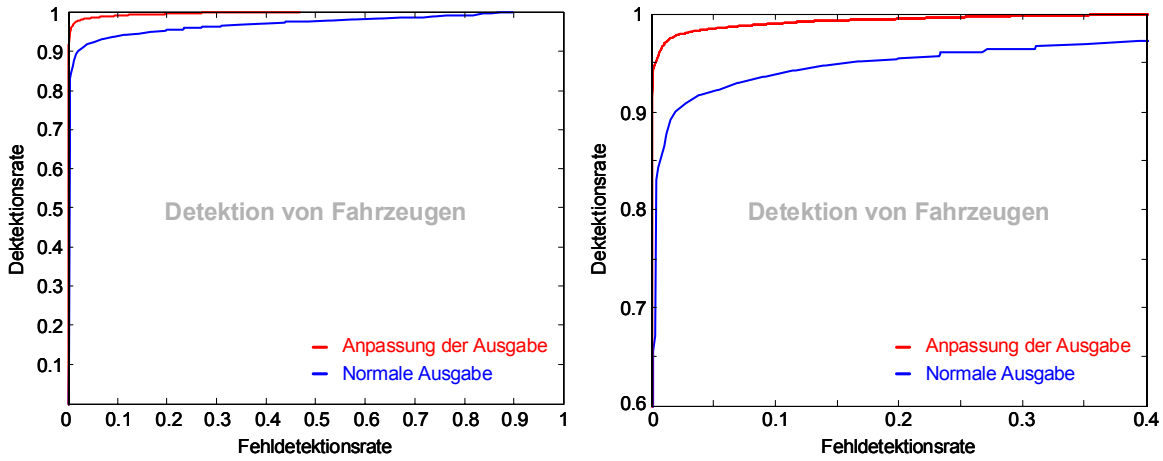

Abbildung 10.19: Verbesserung der Klassifikation für einen abstandgeregelten Tempomaten: Die Einschränkung der Objekte und Objektklassen durch anwendungsspezifisches Vorwissen führt zu einer enormen Verbesserung der Klassifikationsleistung. Der relevante Bereich der kompletten Kurve ist im rechten Teil vergrößert dargestellt.

10.4.4 Rechenzeit

Die vom Gesamtsystem für Vorverarbeitung, Objektverfolgung und Klassifikation benötigte Rechenzeit wurde mit den in Abschnitt 10.2 vorgestellten Methoden analysiert. Auch hier kam der beschriebene 3,4 GHz PC zum Einsatz.

Die benötigte Rechenzeit pro Messintervall ist 7,6 ms (42,1 ms). Der Anteil der benötigten Rechenkapazität nach dem in Abschnitt 10.2 beschriebenen Ansatz beträgt 36,4% (52,6%). Die Klassifikation benötigt demzufolge eine vergleichsweise geringe Rechenzeit von 1,3 ms (3,8 ms) und eine Rechenkapazität von 4,1% (4,8%).

10.5 Fusion von Laserscanner- und Videomessdaten

In diesem Abschnitt wird die Verbesserung der Klassifikation durch Fusionsansätze mit Videomessdaten untersucht. Dafür wurden zwei Mustererkennungssysteme verglichen. Dem ersten Mustererkennungssystem wurden beim Training nur Merkmale aus den Laserscannermessdaten und der Objektverfolgung präsentiert. Für das zweite System wurden zusätzlich die in der Bildverarbeitung aus den Kaskadendetektoren berechneten Merkmale verwendet. Beide Systeme wurden dann im beschriebenen Framework zur Fahrzeugumfelderkennung eingesetzt. Die erreichten Verbesserungen bei der Detektion von PKWs, der Detektion von Fußgängern und der Klassifikation mit allen verwendeten Objektklassen wurden untersucht. Die Untersuchungen beruhen auf den in Abschnitt 10.1 beschriebenen Testdaten.

10.5.1 Detektion von PKWs

Die Anwendung der Bildverarbeitung hat bei PKWs das Ziel, insbesondere die Klassifikation stehender Fahrzeuge zu verbessern. Ohne eine signifikante Objektgeschwindigkeit fehlt diesen Objekten ein wesentliches Merkmal zur Klassifikation, die dadurch erheblich erschwert wird. In Abbildung 10.20 ist die Klassifikationsleistung für PKWs mit alternativen ROCs dargestellt. Mit Hilfe der Bildverarbeitung konnte eine Verbesserung erreicht werden.

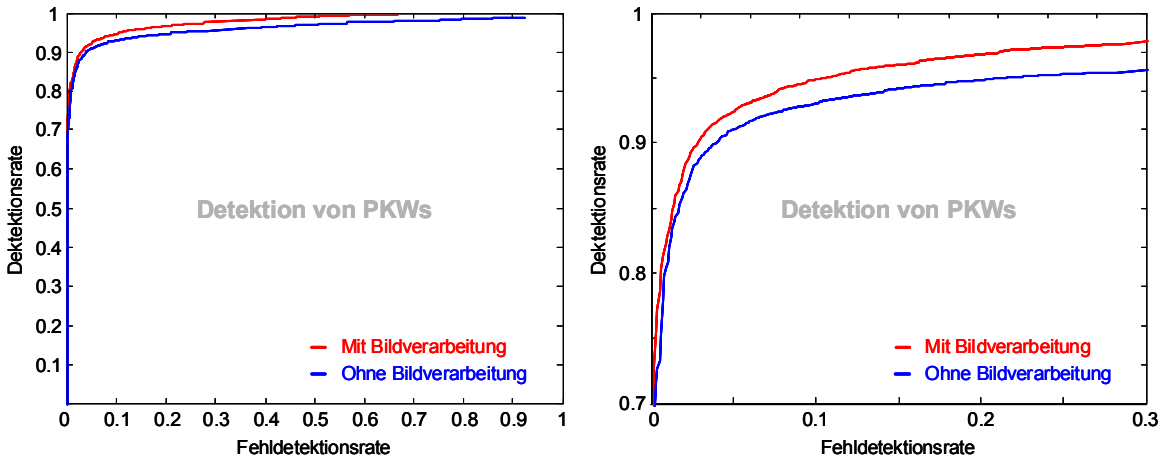

Abbildung 10.20: Verbesserung der Klassifikationsleistung durch die Bildverarbeitung bei der Detektion von PKWs: Links sind die alternativen ROCs mit und ohne Nutzung der Bildverarbeitung dargestellt. Im rechten Teil befindet sich eine Vergrößerung des interessanten Ausschnitts.

Die Bildverarbeitung konnte nur für einen gewissen Anteil der Objekte eine Verbesserung erreichen. Dies ist dadurch begründet, dass der Messbereich der Videokamera deutlich kleiner als der des Laserscanners ist. Darüber hinaus eignet sich die Videokamera nur zur Verbesserung der Klassifikation, wenn entsprechende Objekte nicht oder nur zu einem kleinen Teil verdeckt sind. Dieser Sachverhalt ist in Abbildung 10.21 am Beispiel einer Autobahnszene dargestellt. Die Bildverarbeitung kann im Videobild die zwei vor dem Versuchsträger fahrenden PKWs detektieren. Der Laserscanner detektiert zusätzlich noch drei weitere Fahrzeuge vor diesen PKWs. Diese lassen sich aufgrund von Verdeckungen im Videobild nicht mehr detektieren. Die Klassifikation erfolgt also lediglich mit den Merkmalen aus den Laserscannermessdaten und der Objektverfolgung.

Prinzipiell ist auch in den Laserscannermessdaten ersichtlich, dass die direkt vor dem Versuchsfahrzeug fahrenden PKWs um einiges besser erfasst werden, als die verdeckten und deutlich weiter entfernteren Fahrzeuge. Diese werden in manchen Messzyklen vom Laserscanner sogar überhaupt nicht erfasst. Aus diesem Grund ist die Klassifikation stark verdeckter Objekte deutlich schwieriger als die der komplett vermessbaren Objekte.

Betrachtet man nun die Bedeutung der einzelnen Objekte für Fahrerassistenzsysteme, wird klar, dass verdeckte Objekte in höheren Distanzen deutlich weniger von Interesse sind, als komplett sichtbare Objekte, deren Gefahrenpotential für das Versuchsfahrzeug aufgrund der direkten Nachbarschaft wesentlich höher ist.

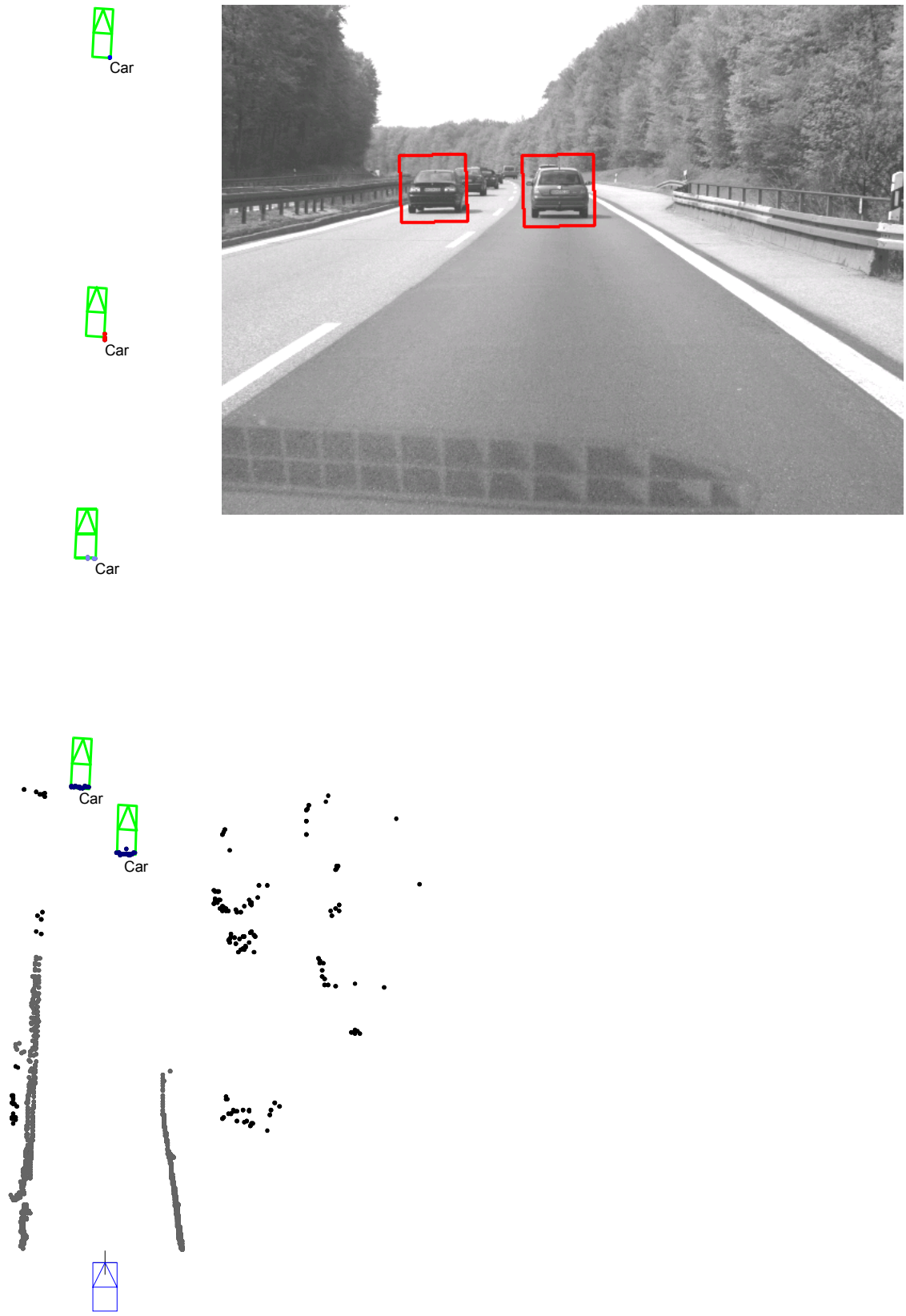

Abbildung 10.21: Beispielszene für unterschiedliche Sichtbereiche: Die Videokamera kann die Klassifikation nur bei sichtbaren Objekten unterstützen. Der Laserscanner detektiert jedoch auch verschiedene teilweise stark verdeckte Objekte, die hier aufgrund der Geschwindigkeit klassifiziert werden.

Zur Analyse des Verbesserungspotentials der Bildverarbeitung für die Klassifikationsleistung wurde diese gezielt für alle Objekte untersucht, die sowohl vom Laserscanner als auch von der Videokamera erfasst wurden. Die entsprechenden alternativen ROCs für die verglichenen Systeme sind in Abbildung 10.22 dargestellt. Es wird deutlich, dass bereits die laserscannerbasierte Klassifikation eine deutlich bessere Klassifikationsleistung erreicht. Dies ist durch die gute Messbarkeit der Objekte begründet. Die Fusion mit Videomessdaten erreicht auch hier eine Verbesserung. Diese Verbesserung äußert sich statistisch jedoch nur geringfügig. Das höchste Verbesserungspotential besteht für stehende PKWs, die nicht aufgrund ihrer Geschwindigkeit bereits mit Merkmalen aus der Objektverfolgung erkannt werden können. Diese sind statistisch allerdings eher selten, so dass die Unterstützung der Bildverarbeitung hier nur in speziellen Situationen greift. Trotzdem existieren verschiedene Situationen, in denen stehende PKWs erkannt werden müssen. Ein Beispiel sind Kreuzungsszenarien, in denen ein Fahrzeug bereits auf der Kreuzung steht. Solche Fahrzeuge stellen eine potentielle Gefahr dar, falls sie den Weg des Versuchsträgers nach dem Anfahren kreuzen werden. Eine Notwendigkeit der Erkennung besteht zum Beispiel für Kreuzungsassistenten, wenn diesen Fahrzeugen die Vorfahrt zu gewähren ist.

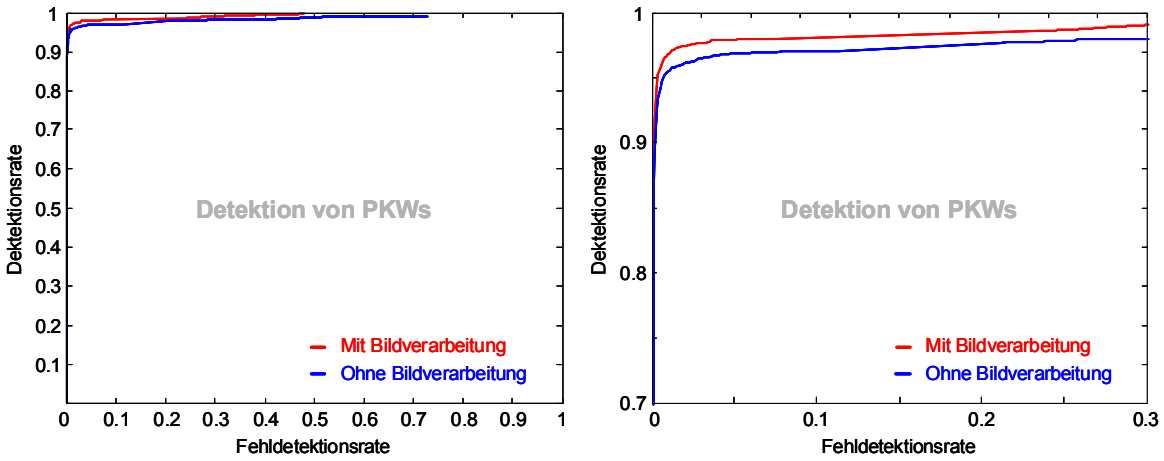

Abbildung 10.22: Verbesserte Detektion von PKWs mit Hilfe der Bildverarbeitung: Die dargestellten alternativen ROCs beziehen sich nur auf Objekte, die vom Laserscanner und der Kamera erfasst wurden. Rechts erfolgt wieder eine vergrößerte Darstellung des signifikanten Ausschnitts.

10.5.2 Detektion von Fußgängern

Die Verbesserungen der Detektion von Fußgängern mit Hilfe der Bildverarbeitung wurde mit den in Abschnitt 10.4.3 beschriebenen Anpassungen für das Assistenzsystem Fußgängerschutz durchgeführt. Für dieses Assistenzsystem bietet die verwendete Videokamera eine hervorragende Unterstützung, da sie mit dem Weitwinkelobjektiv die komplette ROI erfassen kann. Abbildung 10.23 zeigt, dass durch die Fusion mit den Videomessdaten eine enorme Verbesserung der Detektionsleistung für Fußgänger erreicht wurde. Die von der Kamera erfassten Kontur- und Texturinformationen sind hier eine besonders große Unterstützung, da Fußgänger nicht so signifikant hohe Geschwindigkeiten wie PKWs erreichen. Aufgrund des Messrauschens bei der Bestimmung der Bewegung des Versuchsfahrzeuges

werden für stehende Objekte mitunter fälschlicherweise Geschwindigkeiten geschätzt, die der Bewegung von Fußgängern gleichen. Somit ist die Bewegung als Merkmal für die Klassifikation von Fußgängern nicht so signifikant wie bei PKWs.

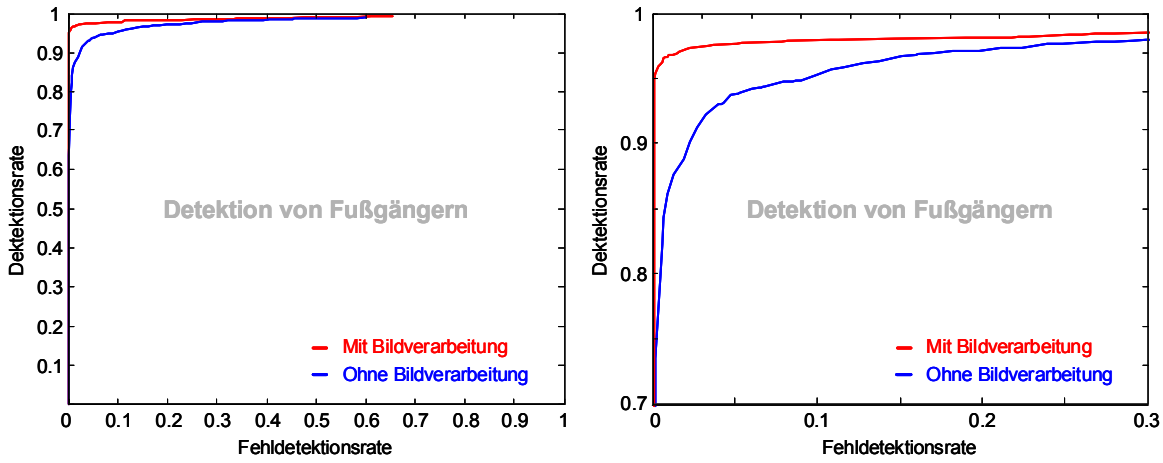

Abbildung 10.23: Verbesserte Detektion von Fußgängern durch Fusion mit Videomessdaten: Durch die zusätzlichen Kontur- und Texturinformationen des Videobilds kann eine erhebliche Steigerung der Detektionsleistung erreicht werden.

In Abbildung 10.24 wird die Detektionsleistung für Objekte untersucht, die vom Laserscanner und der Kamera erfasst wurden. Hier ist bei der laserscannerbasierten Klassifikation eine leichte Verbesserung gegenüber der Auswertung mit allen Objekten ersichtlich. Dies ist dadurch begründet, dass verschiedene verdeckte Objekte hier nicht mehr berücksichtigt werden. Eine deutliche Steigerung der Detektionsleistung ist bei der zusätzlichen Verwendung der Bildverarbeitung zu erkennen.

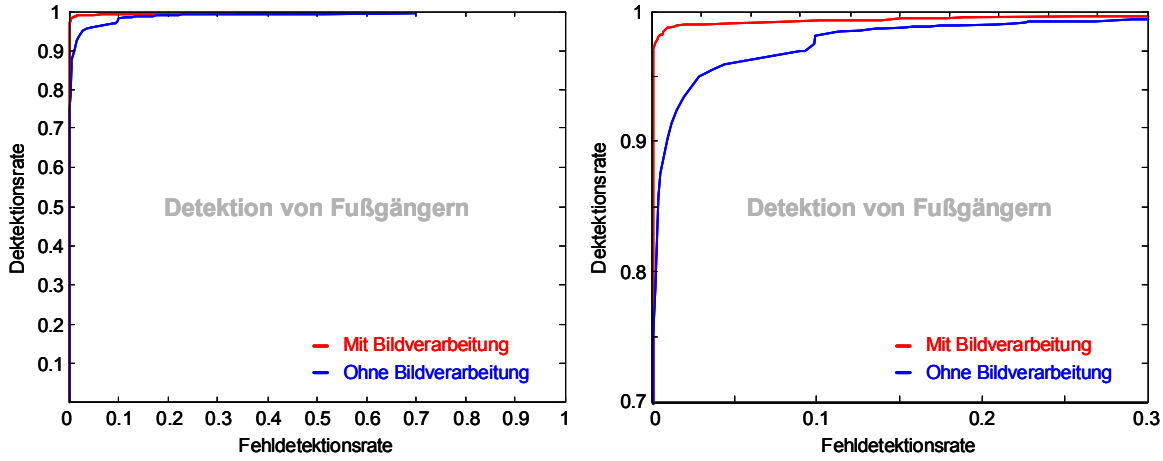

Abbildung 10.24: Verbesserte Fußgängerdetektion im gemeinsamen Sichtbereich von Laserscanner und Kamera: In dieser Darstellung werden nur die Objekte ausgewertet, für die die Videomessdaten auch einen Beitrag leisten können. Dadurch wird das große Verbesserungspotential deutlich.

10.5.3 Auswertung aller Objektklassen

In diesem Abschnitt wird der Einfluss der Fusion von Laserscanner- und Videomessdaten auf die komplette Klassifikation mit den fünf verwendeten Objektklassen untersucht. Im Wesentlichen stützen die Fusionsansätze lediglich die Detektion von PKWs und Fußgängern. Dabei können jedoch auch Verwechslungen mit anderen Objektklassen vermieden werden.

In Abbildung 10.25 wird die Gesamtklassifikationsleistung des Fusionssystems der rein laserscannerbasierten Klassifikation in Form von alternativen ROCs gegenübergestellt. Eine beachtliche Verbesserung kann insbesondere im für Fahrerassistenzsysteme wesentlichen Bereich der geringen Falschalarmanzahlen erreicht werden.

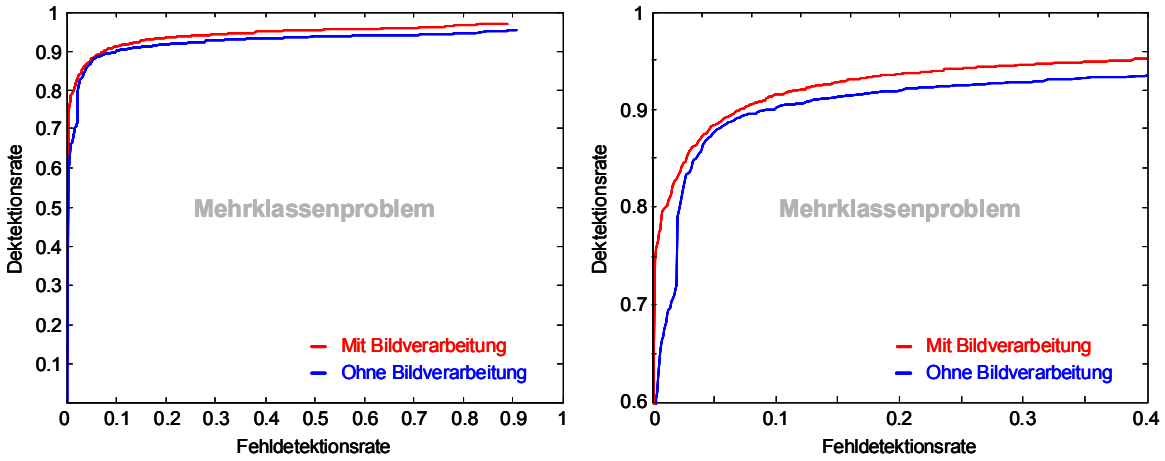

Abbildung 10.25: Verbesserte Objektklassifikation durch den Fusionsansatz mit Videomessdaten: Bei der zusätzlichen Verwendung von Videomessdaten lässt sich insbesondere im für Fahrerassistenzsysteme wesentlichen Bereich weniger Falschalarme eine beachtliche Verbesserung erzielen.

Da die Videokamera nur einen Teil aller ausgewerteten Objekte erfasst, lässt sich das Verbesserungspotential des Fusionsansatzes deutlicher zeigen, wenn die Auswertung auf diese Objekte eingeschränkt wird. Entsprechende alternative ROCs sind in Abbildung 10.26 dargestellt. Hier ist eine signifikante Verbesserung der Klassifikationsleistung zu erkennen.

10.5.4 Rechenzeit

Die vom Gesamtsystem benötigte Rechenzeit bei der Verwendung der Fusionsansätze mit Videomessdaten wurde entsprechend der in Abschnitt 10.2 vorgestellten Methoden analysiert. Insgesamt wird eine Rechenzeit von 22,5 ms (71,5 ms) benötigt. Dies entspricht einem Anteil der verfügbaren Rechenkapazität von 71,3% (89,3%). Somit benötigt die Bildverarbeitung und die Berechnung entsprechender Merkmale fast die Hälfte der verwendeten Rechenzeit. Trotzdem lassen sich selbst für Peakwerte der Rechenzeit alle Berechnungen innerhalb eines Messzyklus durchführen. Bei der Onlineanwendung bleibt also noch Rechenkapazität für zusätzliche Berechnungen.

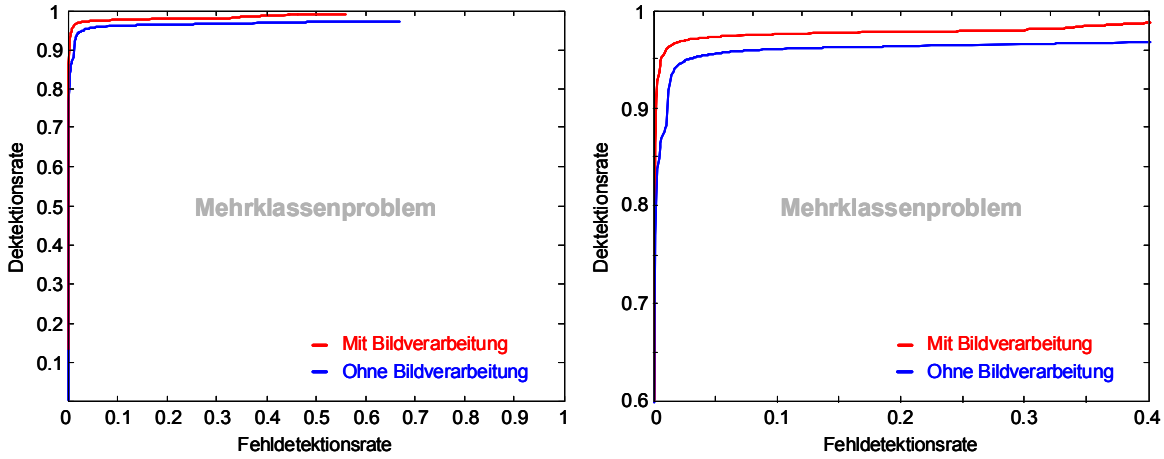

Abbildung 10.26: Auswertung der Klassifikationsleistung im gemeinsamen Messbereich von Kamera und Laserscanner: Die Kurvenverläufe zeigen eine signifikante Verbesserung bei der Fusion mit Videomessdaten.

10.6 Verwendung präziser Karten

Die Position detektierter Objekte innerhalb bestimmter Infrastrukturbereiche im Fahrzeugumfeld kann, wie in Kapitel 8 beschrieben, ebenfalls zur Klassifikation herangezogen werden. Darüber hinaus kann die Anzahl der zu verfolgenden Objekte durch gezielte Ausblendung von Messdaten in bestimmten Bereichen des Fahrzeugumfelds signifikant reduziert werden. Für die Untersuchung der entwickelten Algorithmen konnten nicht alle Sequenzen der vorhandenen Testdaten verwendet werden, da die Erstellung von präzisen Karten und die Positionierung in diesen Karten für Kreuzungen entwickelt wurde. Darum wurde nur der Teil der Testdaten, der Kreuzungsszenarien enthält, präzise kartografiert und zur Auswertung eingesetzt. Dieser Teil umfasst verschiedene Szenarien mit einer Laufzeit von insgesamt 22 min.

10.6.1 Verbesserung der Klassifikation

Zum Vergleich der Klassifikationsleistung wurden wieder zwei alternative ROCs erstellt (Abbildung 10.27). Die Leistung des Standardsystems inklusive Bildverarbeitung wird mit der Leistung bei zusätzlicher Nutzung der Karteninformationen verglichen. Die Klassifikationsleistung des Standardsystems für das eingeschränkte Testset ist im Vergleich zur Auswertung mit allen Testdaten etwas schlechter. Die Ursache dafür ist in der Komplexität der Kreuzungsszenarien zu sehen. Die im kompletten Testdatensatz enthaltenen Autobahnszenarien sind deutlich weniger komplex, da auf Autobahnen in der Regel die Verkehrsdichte niedriger ist, die Fahrzeuge sich mit signifikanten Geschwindigkeiten bewegen und die Straßenbegrenzung gut durch Leitplanken markiert ist. Aufgrund dieser geringeren Komplexität lassen sich Autobahnszenarien leichter klassifizieren und heben somit die Klassifikationsleistung statistisch leicht an.

Beim Einsatz präziser digitaler Karten ist eine deutliche Verbesserung der Klassifikationsleistung im Vergleich zum Standardsystem zu erkennen. Diese wird durch die regelbasierten

Eingriffe aufgrund der Zuordnung zu bestimmten Infrastrukturelementen erreicht. So wird beispielsweise die Einschwingzeit bei der Klassifikation von PKWs erheblich reduziert, da meist bei der ersten Detektion bereits klar ist, dass sich ein PKW auf der Straße befindet. Verbleibende Klassifikationsfehler bestehen zum einen aus Verwechslungen zwischen PKWs und LKWs. Zum anderen treten auch Fehler bei den Klassen „Fußgänger“ und „Zweirad“ auf, die durch den Einsatz digitaler Karten nicht behoben werden.

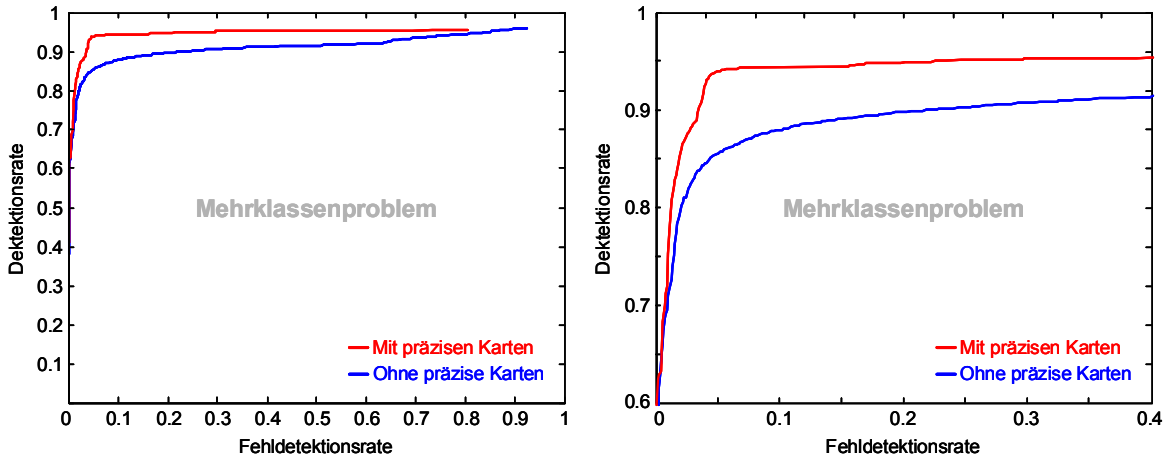

Abbildung 10.27: Durch die Nutzung von Karteninformationen im regelbasierten Teil der Klassifikation kann die Klassifikationsleistung signifikant verbessert werden.

10.6.2 Rechenzeit

Zur Bestimmung der Rechenzeit wurde das Gesamtsystem inklusive Bildverarbeitung einmal mit und einmal ohne die Anwendung präziser digitaler Karten mit der in Abschnitt 10.2 beschriebenen Methode analysiert. Für die Anwendung der Zusatzinformationen aus präzisen digitalen Karten müssen einige zusätzliche Algorithmen zur Positionierung und zur Berücksichtigung der Positionierungsunsicherheiten ausgeführt werden. Erst dann können erfasste Objekte bestimmten Bereichen der Infrastruktur zugeordnet werden. Diese Berechnungen benötigen eine zusätzliche Rechenzeit von 5,0 ms (9,9 ms). Dies entspricht einem Anteil von 11,1% (12,3%) der verfügbaren Rechenkapazität.

Mit Hilfe der präzisen digitalen Karten werden jedoch auch verschiedene Entfernungsmessungen des Laserscanners gezielt ausgeblendet, da die entsprechenden Bereiche für Fahrerassistenzsysteme nicht von Interesse sind. Dadurch wird die Anzahl der zu verfolgenden und zu klassifizierenden Objekte erheblich reduziert. Folglich wird bei der Objektverfolgung und bei der Klassifikation ein Teil der Rechenzeit gespart. Die gesparte Rechenzeit beträgt 6,7 ms (16,7 ms) und entspricht einem Anteil von 21,6% (21,9%) der verfügbaren Rechenkapazität. Somit sinkt die Rechenzeit für das Gesamtsystem, da die in der Objektverfolgung und bei der Klassifikation gesparte Rechenzeit den zusätzlichen Aufwand zur Positionierung und der Berücksichtigung der Unsicherheiten ausgleichen kann.

Das komplette System benötigt eine Rechenzeit von 18,9 ms (48,9 ms) und damit 46,3% (61,1%) der verfügbaren Rechenkapazität.

10.7 Drahtloskommunikation

In diesem Abschnitt wird die Assoziation drahtlos empfangener Objektdaten zu sensorbasierten Objektdaten untersucht. Diese Analyse wird am Beispiel eines Fahrzeugs durchgeführt, das eigene Objektdaten überträgt und gleichzeitig von den Sensoren des anderen Versuchsfahrzeugs erfasst wird.

10.7.1 Simulation der drahtlosen Übertragung

Verschiedene Hardware und Protokolle zur drahtlosen Übertragung von Daten im Fahrzeugumfeld wurden im Rahmen der Kooperation von der Firma IBEO Automobile Sensor GmbH untersucht. Da zum Zeitpunkt der Auswertung jedoch noch keine geeignete Übertragungskonfiguration zur Verfügung stand, wurde die drahtlose Übertragung für die Auswertung simuliert.

Um das Gesamtsystems trotzdem unter realistischen Randbedingungen bewerten zu können, wurden alle Untersuchungen basierend auf realen Messdaten zweier Versuchsfahrzeuge in realen Szenarien durchgeführt. Hierfür wurden mit zwei Versuchsfahrzeugen gleichzeitig Messdaten aufgezeichnet. Das eine Fahrzeug speicherte dabei die kompletten Messdaten mit einer korrespondierenden UTC-Zeit (*engl. UTC = coordinated universal time*), die vom DGPS-Empfänger gemessen wird und als globale Zeit zur Synchronisierung verwendet werden kann. Das zweite Fahrzeug speicherte lediglich die aus dem ESP bestimmten Messwerte der eigenen Bewegung mit den GPS-Positionen und der UTC-Zeit.

Bei der Offline-Auswertung werden beide Datenquellen über die UTC-Zeit synchronisiert. Die Daten des zweiten Fahrzeugs werden den Algorithmen dabei um die für die Übertragung notwendige Zeit verzögert zur Verfügung gestellt. Übertragungsfehler werden an dieser Stelle nicht simuliert. Da erwartungsgemäß die Reichweite einer drahtlosen Übertragung deutlich höher ist als die Reichweite der verwendeten Sensoren, wurde ebenfalls keine Begrenzung der Übertragungsreichweite implementiert.

10.7.2 Assoziation

In diesem Abschnitt wird die Assoziationsleistung des vorgestellten Systems analysiert. Als Übertragungszeit wurden 200 ms bei der Simulation der Übertragung angenommen. Dieser Wert wurde im Rahmen von Untersuchungen innerhalb eines Kooperationsprojektes mit der IBEO Automobile Sensor GmbH ermittelt.

Die Analyse wurde mit Hilfe der zu diesem Zweck vorbereiteten Testdaten durchgeführt. Das zweite Versuchsfahrzeug wurde in den Laserscannermessdaten des ersten Versuchsfahrzeugs markiert, wenn es vom Sensor erfasst wurde. Damit sind alle möglichen korrekten Assoziationen bekannt. Die Markierungen werden nun mit der Ausgabe des Systems verglichen.

Abbildung 10.28 vergleicht die Assoziationsleistung des Systems unter Verwendung der vorgestellten Offsetschätzung mit der Leistung ohne Offsetschätzung. Es wurde der erreichbare Anteil aller möglichen korrekten Assoziationen über der mittleren Anzahl von Fehlassoziationen pro Zeitschritt abgetragen. Die Kurvenverläufe zeigen, dass mit Hilfe der

Offsetschätzung ein deutlich höherer Anteil aller möglichen korrekten Assoziationen erreichen lässt. Insgesamt lassen sich so bis zu 98,5 Prozent aller möglichen korrekten Assoziationen erkennen.

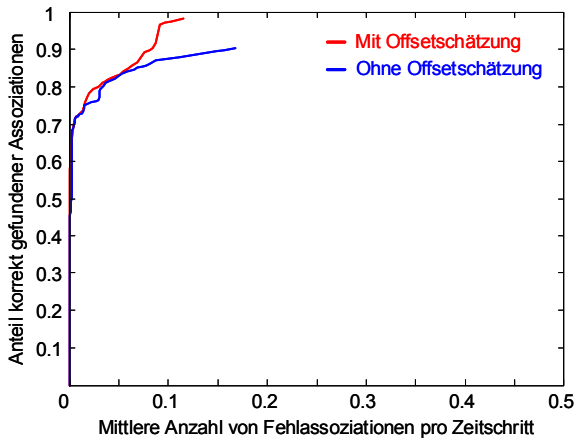

Abbildung 10.28: Verbesserung der Assoziationsleistung durch die Offsetschätzung: Mit Hilfe der Offsetschätzung lässt sich die Anzahl korrekter Assoziationen deutlich steigern. Akzeptiert man jedoch nur sehr wenige Fehlassoziationen, ist die Assoziationsleistung vergleichbar mit der eines Systems ohne Offsetschätzung.

Betrachtet man jedoch die erreichbare Assoziationsleistung bei möglichst wenigen Fehlassoziationen, verhalten sich beide Varianten vergleichbar. So lassen sich für sehr kleine Anzahlen von Fehlassoziationen zwischen 60 und 70 Prozent aller möglichen Assoziationen vom System bestimmen. Die nicht erkannten Assoziationen sind dadurch begründet, dass die Ausgabe von Fehlassoziationen nur durch eine hohe Schwelle der Assoziationswahrscheinlichkeit vermieden werden kann. Entsprechend hohe Assoziationswahrscheinlichkeiten können jedoch nur erreicht werden, wenn bei der Assoziation auch die passende Geschwindigkeit der laserscannerbasierten Objektverfolgung vorliegt. Für eine ausreichend genaue Geschwindigkeitsschätzung benötigt die Objektverfolgung allerdings nach der ersten Detektion des Objektes durch den Laserscanner zunächst einige Zeitschritte zum Einschwingen des Kalman-Filters.

Darüber hinaus wurde das zeitliche Mittelwertfilter der Assoziationswahrscheinlichkeiten ebenfalls so ausgelegt, dass ein Einschwingvorgang durchlaufen werden muss, bis hohe Assoziationswahrscheinlichkeiten ausgegeben werden. Durch diese Filterauslegung werden Fehlassoziationen unterdrückt, die durch einmalige hohe Assoziationswahrscheinlichkeiten zufällig passender Objekte entstehen würden.

Somit benötigt die Assoziation einige Zeit, bis hohe Assoziationswahrscheinlichkeiten ausgegeben werden. Bei hohen Schwellwerten zur Ausgabe der Assoziation kann das Verfahren in dieser Zeit keine Assoziation ausgeben, wodurch die erreichbare Zahl korrekter Assoziationen deutlich reduziert wird.

10.7.3 Einfluss der Übertragungszeit

Zur Spezifikation maximal vertretbarer Übertragungszeiten wurde das Verhalten der Assoziationsleistung für unterschiedliche Übertragungszeiten analysiert. In Abbildung 10.29 ist die erreichbare Assoziationsleistung für verschiedene Übertragungszeiten dargestellt. Es lässt sich erkennen, dass die Assoziationsleistung bei einer Übertragungszeit von 200 ms im Vergleich zur idealen Übertragungszeit von 0 s nur unwesentlich nachlässt. Auch bei Übertragungszeiten bis hin zu 1 s lässt die Leistung nur geringfügig nach. Bei einer Übertragungszeit von 2 s ist allerdings ein deutlicher Qualitätsverlust feststellbar. Der relativ geringe Einfluss von üblichen Übertragungszeiten unter 1 s ist durch die vom System durchgeführte Zustandsprädiktion der empfangenen Daten auf die aktuelle Zeit begründet. Erst wenn die dadurch überbrückte Zeit größer wird, fehlen entsprechende Messwerte für eine ausreichend genaue Zustandsprädiktion, die eine wesentliche Grundlage für die korrekte Assoziation mit aktuellen Sensormessdaten bildet. Dies ist insbesondere der Fall, wenn das sendende Fahrzeug seinen dynamischen Zustand zum Beispiel durch starke Beschleunigungsänderung signifikant variiert.

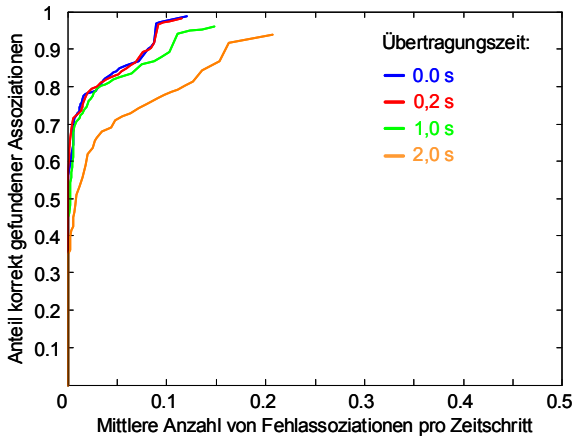

Abbildung 10.29: Auswirkung der Übertragungszeit auf die erreichte Assoziationsleistung: Eine realistische Übertragungszeit von 200 ms führt kaum zu einer Verschlechterung der Assoziationsleistung im Vergleich zu einer idealen Übertragungszeit von 0 ms. Erst bei Übertragungszeiten über 1 s lässt die Assoziationsleistung merklich nach.

10.7.4 Analyse der Geschwindigkeitsschätzung

Bei der drahtlosen Datenübertragung kann die Objektgeschwindigkeit direkt an das Versuchsfahrzeug übertragen werden, während die sensorbasierte Objektverfolgung die Geschwindigkeit indirekt über die gemessene Positionsänderung bestimmen muss. Beide Geschwindigkeitsschätzungen werden in einem Kalman-Filter ermittelt. Während bei der Drahtlosübertragung die Geschwindigkeitsänderung als konstante Beschleunigung modelliert wird, wird die Geschwindigkeitsänderung bei der Objektverfolgung nur über einen Rauschprozess modelliert. Die übertragene Geschwindigkeit sollte also genauer als die aus den Sensormessdaten ermittelte Geschwindigkeit sein. Ein Nachteil der Drahtlosübertragung

ist die Latenz. Die übertragenen Messwerte sind beim Eintreffen im Versuchsfahrzeug stark veraltet, so dass der aktuelle Zustand mit dem Kalman-Filter geschätzt werden muss.

An dieser Stelle sollen beide Methoden der Geschwindigkeitsschätzung verglichen werden, um zu analysieren, ob eine Umschaltung der ausgegebenen Geschwindigkeit auf die auf der Drahtlosübertragung basierende Geschwindigkeitsschätzung sinnvoll ist. Als Voraussetzung für eine Umschaltung der Geschwindigkeitsausgabe wird von einer korrekten Assoziation ausgegangen. Dies kann durch einen hohen Schwellwert für die bestimmte Assoziationswahrscheinlichkeit erreicht werden. In der durchgeführten Analyse wird von einer verlustfreien Übertragung ausgegangen. Eine Umschaltung der Geschwindigkeitsausgabe wird also nur in Betracht gezogen, wenn regelmäßig drahtlos übertragene Daten eintreffen.

Die Analyse der Methoden zur Geschwindigkeitsschätzung wurde beispielhaft mit einer 80 s langen Sequenz durchgeführt, in der das sendende Fahrzeug kontinuierlich vom Laserscanner erfasst wurde. Die korrekte Objektgeschwindigkeit wurde aus den Eigenbewegungssensoren des sendenden Fahrzeugs ermittelt. Zusätzlich erfolgen nun eine Geschwindigkeitsschätzung mit Hilfe der im empfangenden Fahrzeug durchgeführten laserscannerbasierten Objektverfolgung und eine Geschwindigkeitsschätzung basierend auf der verzögert empfangenen Objektgeschwindigkeit des sendenden Fahrzeugs. In Abbildung 10.30 wurde die korrekte Geschwindigkeit des sendenden Fahrzeugs und jeweils die zu diesem Zeitpunkt erfolgten Schätzungen beider Kalman-Filter über der Zeit abgetragen.

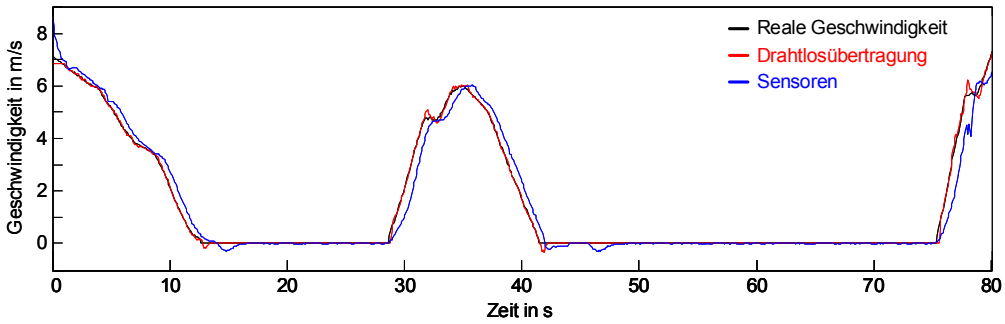

Abbildung 10.30: Vergleich der Geschwindigkeitsschätzung: Während beim laserscannerbasierten Kalman-Filter ein Nachlaufen der Schätzung zu beobachten ist, kann das Filter der drahtlos übertragenen Geschwindigkeitswerte trotz der Latenz gute Schätzwerte liefern.

In der Grafik ist ein Nachlaufen der aus den Sensoren ermittelten Geschwindigkeitsschätzung gegenüber der realen Geschwindigkeit erkennbar. Dies ist durch die indirekte Bestimmung der Geschwindigkeit im Filter erklärbar. Die Geschwindigkeitsschätzung mit Hilfe der drahtlos übertragenen Daten ist deutlich besser. In Phasen mit annähernd konstanter Beschleunigung kann durch die Zustandsprädiktion auf den aktuellen Zeitpunkt der Nachteil der Übertragungslatenz sehr gut ausgeglichen werden. Dies wird in Abbildung 10.31 durch eine Darstellung des Schätzfehlers deutlich.

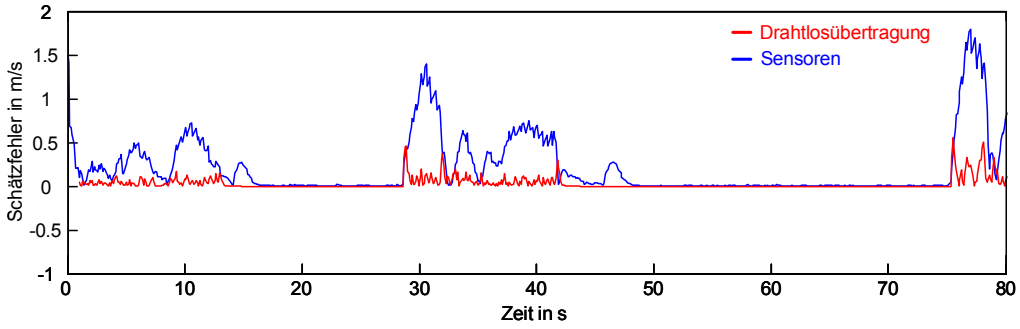

Abbildung 10.31: Schätzfehler beider Methoden zur Geschwindigkeitsschätzung: Der Schätzfehler der aus den Laserscannermessdaten bestimmten Geschwindigkeit ist deutlich höher als der Fehler der aus den drahtlos übertragenen Daten geschätzten Geschwindigkeit.

Zu wenigen, kurzen Zeiträumen ist die Geschwindigkeitsschätzung mit Hilfe der Laserscannermessdaten besser als die auf drahtlos übertragenen Daten basierende Schätzung (Abbildung 10.32).

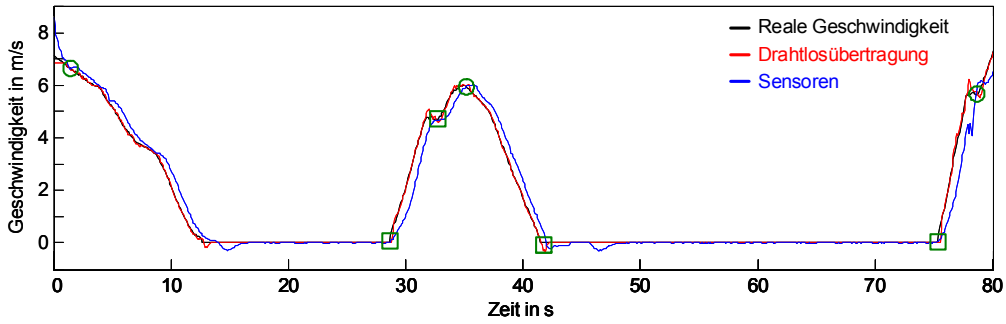

Abbildung 10.32: Analyse der Fälle, in denen die aus Laserscannermessdaten geschätzte Geschwindigkeit besser als die aufgrund von drahtloser Kommunikation bestimmte Geschwindigkeit ist: Durch die Übertragungslatenz verursachte Fälle sind mit Quadraten markiert. Fälle, die zufällig durch den Schnitt der Geschwindigkeitsverläufe entstehen, sind mit Kreisen markiert.

Im Wesentlichen existieren zwei mögliche Ursachen für schlechtere Schätzwerte des Kalman-Filters für drahtlos übertragene Geschwindigkeiten. Das Kalman-Filter prädiziert die Schätzungen immer mit der aus den bisherigen Messwerten ermittelten Beschleunigung. Da die Messwerte mit einer nicht zu vernachlässigenden Latenz beim System eintreffen, ist auch die geschätzte Beschleunigung veraltet. Dieser Sachverhalt wirkt sich in Phasen konstanter Beschleunigung nicht auf die Geschwindigkeitsschätzung aus. Ändert sich jedoch die Beschleunigung, passt sich das Filter erst nach der Übertragungszeit an die neue Beschleunigung an. Dadurch entstehen deutliche Abweichungen in der Geschwindigkeitsschätzung. Entsprechende Stellen sind in Abbildung 10.32 mit einem Quadrat markiert. Abbildung 10.33 zeigt einen vergrößerten Ausschnitt der Sequenz mit ähnlichen Markierungen.

Unabhängig von der Übertragungslatenz kann die Geschwindigkeitsschätzung mit Hilfe der Laserscannermessdaten zu einzelnen Zeitpunkten besser sein, wenn der Geschwindigkeitsverlauf gerade zum Zeitpunkt der Auswertung den realen Geschwindigkeitsverlauf schneidet. Derartige Schnittpunkte können durch ein Rauschen der geschätzten Geschwindigkeit auftreten. In jedem Fall treten Schnittpunkte beim Wechsel vom Beschleunigen zum Abbremsen auf. Aufgrund des Nachlaufens der Geschwindigkeitsschätzung ist die geschätzte Geschwindigkeit beim Beschleunigen kleiner und beim Abbremsen größer als die reale Geschwindigkeit. Dazwischen müssen sich beide Geschwindigkeitsverläufe zwangsweise schneiden. Entsprechende Stellen sind in den Abbildungen 10.32 und 10.33 mit einem Kreis markiert.

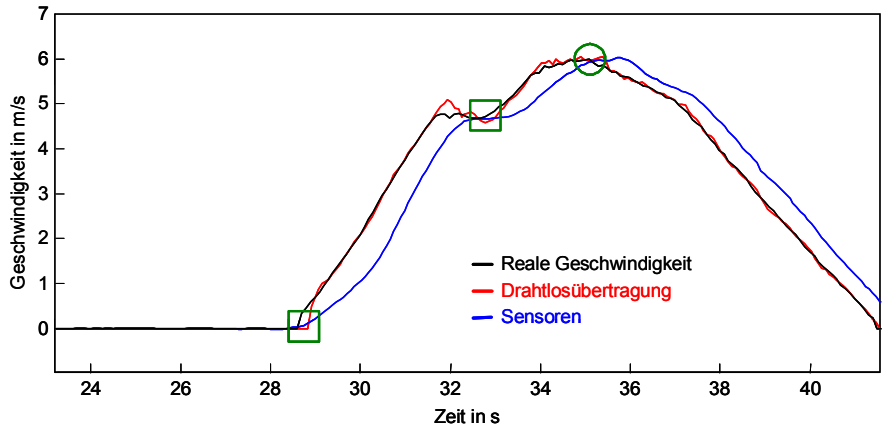

Abbildung 10.33: Analyse der Geschwindigkeitsschätzung: Bei dem Start des anderen Fahrzeugs und anderen Beschleunigungsänderungen liefert die Drahtloskommunikation schlechtere Schätzungen, bis erste Messwerte zum System übertragen wurden (quadratische Markierungen). Zusätzlich ist die laserscannerbasierte Schätzung an einzelnen Zeitpunkten besser, wenn sich die Geschwindigkeitsverläufe der Schätzung und der realen Werte schneiden (kreisförmige Markierung).

Im Rahmen der Untersuchungen wurde für eine bessere Vergleichbarkeit der Ergebnisse mit dem Kalman-Filter für drahtlos übertragene Daten auch für das auf Laserscannermessdaten basierende Kalman-Filter ein Modell mit konstanter Beschleunigung untersucht. Die Untersuchungen konnten jedoch keine Verbesserung der aus den Laserscannermessdaten geschätzten Geschwindigkeit feststellen. Die Schätzung zeigte ein ähnlich starkes Nachlaufen gegenüber der realen Geschwindigkeit. Zusätzlich trat ein stärkeres Rauschen der Schätzwerte auf. Die Ursache dafür liegt in der Tatsache, dass der Laserscanner nur Entfernungen messen kann. Die entsprechenden Positionsmessungen unterliegen zusätzlich einem gewissen Rauschen. Die aus der zeitlichen Positionsänderung geschätzte Geschwindigkeit und erst recht die durch Geschwindigkeitsänderung ermittelte Beschleunigungsschätzung passen sich aufgrund der Filterung nur langsam an reale Änderungen an. Bei der Schätzung der Geschwindigkeit aus drahtlos übertragenen Messwerten liegen nur gering verrauschte Geschwindigkeitsmesswerte vor. Dadurch passt sich auch die im Kalman-Filter geschätzte Beschleunigung sehr schnell realen Änderungen an.

Aufgrund der deutlich präziseren Schätzung der Geschwindigkeit unter Verwendung der drahtlos übertragenen Messwerte ist eine Umschaltung der an die Assistenzsysteme ausgegebenen Geschwindigkeit zu empfehlen. In der Praxis sollte diese Umschaltung allerdings nur bei ausreichend hohen Assoziationswahrscheinlichkeiten und einer regelmäßigen Übertragung der Geschwindigkeitsmesswerte durchgeführt werden. Diese beiden Bedingungen können vom System online überprüft werden.

10.7.5 Rechenzeit

In diesem Abschnitt wird die Rechenzeit analysiert, die für die Fusion drahtlos empfangener Daten mit dem sensorbasierten Fahrzeugumfeldmodell benötigt wird. Der in dieser Arbeit vorgestellte Ansatz bearbeitet jedes empfangene Objekt einzeln. Die einzige Berechnung, die nur einmal für alle empfangenen Objekte durchgeführt werden muss, ist die in Abschnitt 4.5 beschriebene Verdeckungserkennung. Der dafür benötigte Rechenaufwand ist vergleichsweise gering. Somit kann die Erhöhung des Rechenaufwands für jedes zusätzliche empfangene Objekt mit der Rechenzeit abgeschätzt werden, die in dieser Arbeit für ein einzelnes empfangenes Objekt ermittelt wurde.

Im Rahmen dieser Arbeit wurde die Integration der drahtlos empfangenen Objektdaten eines Fahrzeugs untersucht. Für die entsprechenden Berechnungen wurde eine Rechenzeit von 1,9 ms (6,5 ms) benötigt. Dies entspricht 6,9% (8,1%) der verfügbaren Rechenkapazität. Mit diesen Angaben lässt sich die Anzahl der empfangenen Objekte bestimmen, die das System mit der verfügbaren Rechenkapazität verarbeiten kann. Bei der Nutzung des Gesamtsystems inklusive Bildverarbeitung können mit der verbleibenden Rechenzeit vier drahtlos empfangene Objekte verarbeitet werden. Wird das Gesamtsystem ohne die Bildverarbeitung verwendet, können sogar neun drahtlos empfangene Objekte verarbeitet werden. Diese Kapazität ist für erste Assistenzsysteme ausreichend. Bei höheren Anzahlen empfangener Objekte kann beispielsweise mit Hilfe der Position und der Fahrtrichtung eine Auswahl der verarbeiteten Objekte erfolgen.

11 Zusammenfassung und Ausblick

Die stetig steigende Verkehrsdichte führt zu einer Zunahme der täglichen Belastung beim Fahren im Straßenverkehr. Neue Assistenzsysteme sollen die Verkehrssicherheit verbessern, indem sie dem Fahrer zum einen tägliche, automatisierbare Aufgaben abnehmen und ihn zum anderen in gefährlichen Situationen aktiv unterstützen. Aktuelle Systeme erfassen dazu das Fahrzeugumfeld in der Regel mit einem exklusiv genutzten, der Aufgabe angepassten Sensor. Aufgrund der zunehmenden Anzahl von Assistenzsystemen besteht jedoch ein deutlicher Bedarf für eine zentrale Fahrzeugumfelderfassung mit Hilfe einer gemeinsamen Sensorplattform. In dieser Arbeit wurde ein solches echtzeitfähiges System vorgestellt, das eine gemeinsame Sensordatenverarbeitung basierend auf einem ausgewählten Sensoraufbau realisiert. Als Schnittstelle zu den Assistenzsystemen wird eine zusätzliche Abstraktionsebene in Form eines Fahrzeugumfeldmodells eingeführt. Aus dieser detaillierten Fahrzeugumfeldbeschreibung können nun verschiedene Assistenzsysteme gleichzeitig mit den jeweils benötigten Informationen versorgt werden. Dafür wurden neue Methoden zur gezielten Anpassung der Ausgabe an die Anforderungen der einzelnen Assistenzsysteme entwickelt.

Obwohl sich erste Assistenzsysteme eher auf weniger komplexe Autobahn- und Landstraßenszenarien konzentrierten, werden zukünftige Systeme den Fahrer auch in hochkomplexen Szenarien in innerstädtischen Gebieten und auf Kreuzungen unterstützen. Aus diesem Grund wurde hier ein System präsentiert, das aufgrund der Sensorauswahl und der entwickelten Algorithmen sowohl bei herkömmlichen Szenarien als auch in komplexen Szenarien mit hohen Objektdichten sowie Quer- und Gegenverkehr eine enorme Leistungsfähigkeit erreicht.

Die vorgestellte Fahrzeugumfelderfassung basiert auf der Fusion verschiedener Informationsquellen in einem skalierbaren Gesamtsystem. Das Basissystem verwendet einen für den Automobilbereich entwickelten Laserscanner, der aufgrund seiner hohen Entfernungs- und Winkelauflösung präzise Positionsmessungen der erfassten Objekte liefert. Nach einer Vorverarbeitung der Sensordaten werden die mit dem Laserscanner vermessenen Objekte mit jeweils einem Kalman-Filter verfolgt. Die grundsätzlichen Methoden zur Objektverfolgung mit einem rechteckigen Objektmodell wurden hierfür aus bestehenden Arbeiten übernommen. Aufgrund verschiedener algorithmischer Verbesserungen konnte jedoch eine erhebliche Reduktion des Rechenaufwands erreicht werden.

Eine prinzipielle Herausforderung bei der laserscannerbasierten Objektverfolgung ist die Objektbildung. Aufgrund der abstandsabhängigen Segmentierung können andere Verkehrsteilnehmer im Fahrzeugumfeldmodell in mehrere Segmente zerfallen. Um eine Ausgabe in Form mehrerer Objekte zu vermeiden, wurde mit der Segmentfusion eine neue, geschwindigkeitsbasierte Methode präsentiert, die eine sinnvolle Zuordnung mehrerer Segmente zu einem Objekt ermöglicht.

Bisherige Arbeiten zur Objekterkennung im Fahrzeugumfeld beschäftigen sich in der Regel mit der Detektion einer ausgewählten Objektklasse. Da für verschiedene Assistenzsysteme auch verschiedene Verkehrsteilnehmer von Bedeutung sind, wurde für die gemeinsame Umfelderfassung ein neues Gesamtkonzept entwickelt, das verschiedene Verkehrsteilnehmer im Rahmen der Klassen „Fußgänger“, „Zweirad“, „PKW“, „LKW“ und „Hintergrund“ detektiert und unterscheidet. Dieses Klassifikationskonzept kombiniert statistische

Verfahren der Mustererkennung mit regelbasiertem a priori Wissen und einem zeitlichen Filter. Für die laserscannerbasierte Klassifikation wurden dazu verschiedene Merkmale aus den Laserscannermessungen und der Objektverfolgung ermittelt und zum Training sowie der Anwendung der neuronaler Netze eingesetzt.

Die Ausgabe der Klassifikationsalgorithmen besteht pro Objekt aus Zugehörigkeitswerten für jede Klasse. Basierend auf dieser Ausgabe kann für jedes Assistenzsystem in Abhängigkeit von den individuellen Anforderungen die zugeordnete Klasse ausgewählt werden. Dabei lassen sich die ausgegebenen Klassen zum einen für jedes Assistenzsystem auf die relevanten Objekttypen einschränken, zum anderen erfolgt der Kompromiss zwischen hohen Detektionsraten und niedrigen Falschalarmzahlen jeweils entsprechend der Anforderungen des Assistenzsystems. Durch diesen Mechanismus konnte eine deutliche Verbesserung der Klassifikationsergebnisse erzielt werden.

Durch analytische Untersuchungen wurde weiterhin gezeigt, dass die Unterscheidung verschiedener Objektklassen eine deutlich höhere Herausforderung für die Klassifikationsalgorithmen als die pure Detektion einer einzelnen Objektklasse darstellt. Trotzdem konnten im Rahmen dieser Arbeit sehr gute Klassifikationsergebnisse erzielt und mit Hilfe quantitativer Analysen basierend auf realen Messdaten belegt werden.

Im Automobilbereich und insbesondere bei sicherheitsrelevanten Anwendungen wird häufig eine Nachvollziehbarkeit der verwendeten Algorithmen gefordert. Mit Hilfe der Einschränkung der verwendeten mehrlagigen Perzeptrons auf stückweise lineare Ausgabefunktionen konnte eine Methode zur nachträglichen Analyse der trainierten Netze bezüglich der Ausgabe für ganze Bereiche des Merkmalsraumes vorgestellt werden. Weiterhin ermöglicht eine zweistufige Netzstruktur eine automatische, klassenabhängige Skalierung der Merkmale, die die sonst übliche Merkmalsskalierung ersetzt. Untersuchungen belegen die erreichte effiziente Approximation der vorliegenden Datenverteilung mit einer vergleichsweise geringer Anzahl zu erlernender Parameter.

Die guten Ergebnisse der laserscannerbasierten Fahrzeugumfelderfassung konnten durch verschiedene Fusionsansätze noch einmal signifikant verbessert werden. Ein wichtiger Schritt ist dabei die Fusion mit Messdaten einer Videokamera. Eine Aufmerksamkeitssteuerung mit Hilfe der Messdaten des Laserscanners schränkt die Bildverarbeitung auf wesentliche Bildausschnitte und Objektgrößen ein und führt somit zu einer erheblichen Reduktion der Rechenzeit. Während bisherige Arbeiten zur videobasierten Detektion von PKWs im Fahrzeugumfeld sich auf Heckansichten beschränkten, wurde in der vorliegenden Arbeit die Methode der Aufmerksamkeitssteuerung erweitert, so dass aus den Messdaten des Laserscanners nicht nur Positions- und Größeninformationen, sondern auch Orientierungsmessungen verwendet werden. Diese Zusatzinformationen erlauben die Rekonstruktion mindestens einer der im Videobild perspektivisch verzerrten Objektseiten. Auf diese Weise lassen sich PKWs beliebiger Orientierung wie beispielsweise im Gegen- und Querverkehr mit Hilfe der Bildverarbeitung detektieren.

Für die Detektion von PKWs und Fußgängern im Videobild wurden mehrere Kaskadenklassifikatoren nach dem Verfahren von Viola und Jones trainiert und jeweils in den durch die Aufmerksamkeitssteuerung ermittelten Bildbereichen angewendet. Zur Integration der Detektionsergebnisse in eine konsistente Klassifikation des Fahrzeugumfelds wurden aus den angewendeten Kaskadenklassifikatoren neue Merkmale extrahiert, die die Anzahl der durchlaufenen Kaskadenstufen und auftretende Mehrfachdetektionen berücksichtigen. Mit

Hilfe der so berechneten, zusätzlichen Merkmale konnte die Klassifikationsleistung des Gesamtsystems für jede der Objektklassen „Fußgänger“, „Zweirad“, „PKW“, „LKW“ und „Hintergrund“ erheblich verbessert werden.

Ein zweiter Fusionsansatz zur Verbesserung der Klassifikation wurde insbesondere für die komplexeren Szenarien auf Kreuzungen entwickelt. Mit Hilfe bestehender Verfahren lässt sich ein präziser Bezug zwischen hoch aufgelösten digitalen Karten und der aktuellen Position des Fahrzeugs bestimmen. Eine neu entwickelte Methode ordnet diese Informationen unter Berücksichtigung der Positions- und Orientierungsunsicherheit Objekten zu, die mit den Sensoren relativ zum Fahrzeug vermessen wurden. Diese Zuordnung ermöglicht den Einsatz regelbasierten a priori Wissens, mit dem sich eine deutliche Verbesserung der Klassifikationsergebnisse erzielen lässt.

Als zusätzliche Informationsquelle wurde ein Verfahren zur konsistenten Integration drahtlos empfangener Daten in das sensorbasierte Umfeldmodell vorgestellt. Im Gegensatz zu bisherigen Arbeiten über leicht integrierbare Informationen wie Schlechtwettergebiete oder Staus erfolgt eine neuartige, präzise Zuordnung von empfangenen Objektinformationen zu den mit den Sensoren bereits erfassten Objekten und somit eine Kombination der Vorteile von sensorbasierter und kommunikationsbasierter Fahrzeugumfelderfassung. Es können unterschiedliche Quellen für die empfangenen Objektdaten sowie unterschiedlich präzise Positionierungsverfahren durch die geeignete Wahl der korrespondierenden Unsicherheiten berücksichtigt werden. Auftretende Latenzen der GPS-Messungen und der drahtlosen Datenübertragung werden soweit möglich durch ein prädiktives Kalman-Filter ausgeglichen. Durch eine Auswertung der Position, der Orientierung und der Geschwindigkeit und eine Einbeziehung der Objekthistorie konnte trotz hoher Positionsungenauigkeiten der verwendeten DGPS-Empfänger ein Verfahren zur korrekten Assoziation der empfangenen Daten zu den sensorbasierten Objekten entwickelt werden. Die enorme Leistungsfähigkeit der Assoziationsalgorithmen wurde mit realen Messdaten zweier Versuchsfahrzeuge gezeigt.

Das Themengebiet zur Nutzung drahtlos empfangener Daten umfasst ein sehr breites Spektrum. Aus diesem Grund sind verschiedene, weiterführende Untersuchungen und Erweiterungen des vorgestellten Verfahrens sowie zusätzliche Anwendungen in zukünftigen Arbeiten denkbar. So sollten die entwickelten Algorithmen mit realen Übertragungsmedien analysiert werden. Ein Verbesserungspotential bietet der Einsatz von Verfahren zur präziseren Positionierung. Die Integration entsprechender Verfahren wurde über Unsicherheiten berücksichtigt, in dieser Arbeit aber nicht im Detail untersucht. Schließlich besteht die Möglichkeit, die Assoziation mit neuen Methoden zu verbessern, die bei mehreren empfangenen Objekten beispielsweise auch deren Konstellation auswerten könnten.

Alle in dieser Arbeit vorgestellten Algorithmen wurden mit Hilfe umfangreicher Sequenzen von Messdaten aus diversen Szenarien des realen Straßenverkehrs getestet. Zusätzlich wurden verschiedene Algorithmen im Rahmen von Drittmittelprojekten in Echtzeitanwendungen erfolgreich eingesetzt. Die sehr gute Leistungsfähigkeit und die für den Einsatz im Fahrzeug benötigte Echtzeitfähigkeit des Systems wurden darüber hinaus mit Hilfe detaillierter Auswertungen belegt.

A Liste der Merkmale für die Mustererkennung

Folgende Merkmale wurden im Rahmen dieser Arbeit für die Mustererkennung verwendet:

Aus den Laserscannermessdaten extrahierte Merkmale:

- zeitliches Maximum der gemessenen Breite [m]
- zeitliches Maximum der gemessenen Länge [m]
- zeitliches Maximum der gemessenen Höhen (Abstand der Scanlayer) [m]
- Mittlerer Abstand zwischen zwei benachbarten Scanpunkten (pro Segment) [m]
- Einfallswinkel der Laserstrahlen auf d. vordere/hintere Seite [rad]
- Einfallswinkel der Laserstrahlen auf d. rechte/linke Seite [rad]
- Maximalbreite unter Berücksichtigung von Verdeckung [m]
- Maximallänge unter Berücksichtigung von Verdeckung [m]
- zeitl. Min. der Maximalbreite unter Berücksichtigung begrenzender Messungen [m]
- zeitl. Min. der Maximallänge unter Berücksichtigung begrenzender Messungen [m]
- Winkel zwischen den Linien der L-Form [rad]
- Anzahl der Linien der bestimmten Form {0,1,2}
- Größe des Reflektors am vorderen Nummernschild [m]
- Abstand des Reflektors am vorderen Nummernschild von der erwarteten Position [m]
- Größe des Reflektors am hinteren Nummernschild [m]
- Abstand des Reflektors am hinteren Nummernschild von der erwarteten Position [m]
- Größe des Reflektors an der linken Rückleuchte [m]
- Abstand des Reflektors an der linken Rückleuchte von der erwarteten Position [m]
- Größe des Reflektors an der rechten Rückleuchte [m]
- Abstand des Reflektors an der rechten Rückleuchte von der erwarteten Position [m]
- Flag: Objekt befindet sich an linker Messbereichsgrenze des Laserscanners {0,1}
- Flag: Objekt befindet sich an rechter Messbereichsgrenze des Laserscanners {0,1}

Aus der Videobildverarbeitung extrahierte Merkmale:

- Merkmalswert aus der Kaskade für Front-/Heckansichten von PKWs
- Merkmalswert aus der Kaskade für Seitenansichten von PKWs
- Maximum der Merkmalswerte der Kaskaden für PKWs
- zeitl. Max. des Merkmalswerts aus der Kaskade für Front-/Heckansichten von PKWs
- zeitl. Max. des Merkmalswerts aus der Kaskade für Seitenansichten von PKWs
- zeitl. Max. des Maximums der Merkmalswerte der Kaskaden für PKWs
- Flag: Videobildverarbeitung für PKWs hat stattgefunden {0,1,2}
- Merkmalswert aus der Kaskade für Fußgänger
- zeitliches Maximum des Merkmalswerts aus der Kaskade für Fußgänger
- Flag: Videobildverarbeitung für Fußgänger hat stattgefunden {0,1,2}

Aus der Objektverfolgung extrahierte Merkmale:

- Gefilterte Breite [m]
- Gefilterte Länge [m]
- Unsicherheit der Breite [m]
- Unsicherheit der Länge [m]
- Länge des seit erster Messung zurückgelegten Weges [m]
- Aktuelle, geschätzte Geschwindigkeit $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$
- Mittlere Geschwindigkeit $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$
- Maximale Geschwindigkeit $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$
- Relative Geschwindigkeit $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$
- Median der geschätzten Geschwindigkeiten $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$
- Anzahl der Zyklen mit einer Geschwindigkeit von mehr als 4 $\frac{\mathrm{m}}{\mathrm{s}}$
- Wahrscheinlichkeit einer Objektbewegung
- Wahrscheinlichkeit eines korrekt verfolgten Objektes
- Geschwindigkeit des eigenen Fahrzeugs $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$
- Gierrate des eigenen Fahrzeugs $\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$

B Kombination der Positions- und Rotationsunsicherheit

Bei den verwendeten Positionierungsalgorithmen werden die Position und die Orientierung auf der Karte mit korrespondierenden Unsicherheiten ermittelt. An dieser Stelle werden die einzelnen Schritte zur Kombination der Unsicherheiten für verschiedene Bereiche der Karte erläutert. Die Ergebnisse sind in Abbildung B.3 dargestellt.

B.1 Bereich der Straße

Der erste berechnete Bereich entspricht dem Gebiet, das komplett auf der Straße liegt. Falls sich mehrere Straßenpolygone im Sichtbereich befinden, werden diese zu einem Polygon kombiniert. Liegen zwei getrennte Straßenverläufe vor, wird der Verlauf verwendet, auf dem sich das Versuchsfahrzeug befindet. Wahlweise könnten auch zwei verschiedene Bereiche nacheinander berechnet werden.

Als Erstes wird die Positionsunsicherheit berücksichtigt. Dies führt zu einer Verkleinerung des Polygons um diese Unsicherheit. Zur Berücksichtigung der Rotationsunsicherheit wird das resultierende Polygon um diese Unsicherheit um den Rotationspunkt nach links und rechts gedreht. Anschließend werden die korrespondierenden Punkte der beiden Polygone mit zusätzlichen Kanten verbunden. Diese Verbindungen approximieren den Weg der Punkte während der Rotation (Abbildung B.1, links).

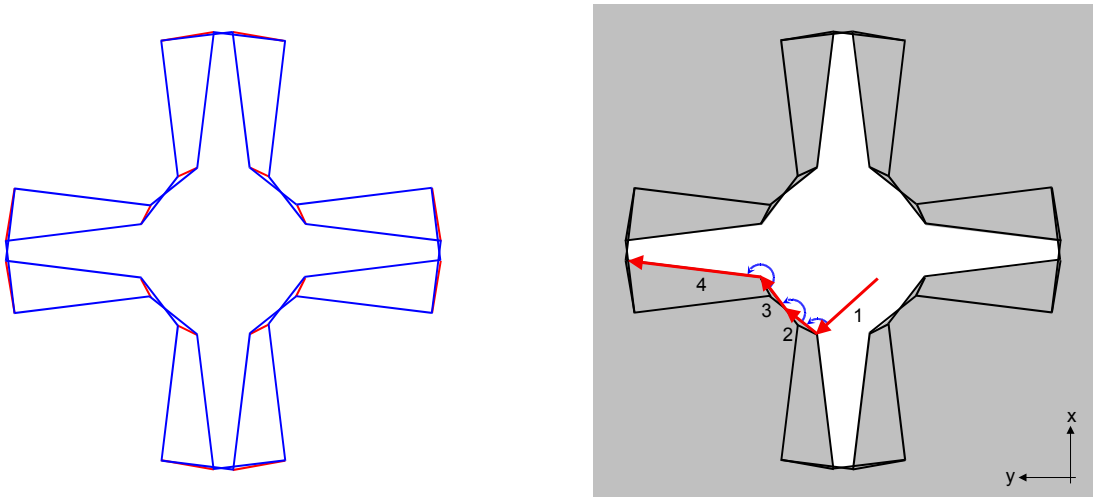

Abbildung B.1: Bereich der Straße: Das Polygon wird um die Positionsunsicherheit verkleinert und um die Rotationsunsicherheit nach links und rechts gedreht. Die korrespondierenden Punkte der zwei Polygone werden verbunden (links). Nun wird iterativ die Fläche ermittelt, die während der gesamten Rotation vom Polygon bedeckt ist.

Nun muss der Bereich extrahiert werden, der während der gesamten Rotation vom Polygon bedeckt ist. Dieser Prozess ist in Abbildung B.1 rechts dargestellt. Zuerst wird ein geeigneter Startpunkt gewählt. Ein solcher Eckpunkt des Polygons muss dementsprechend auch während der gesamten Rotation vom Straßenpolygon bedeckt werden. Der Bewegungsverlauf

eines solchen Eckpunkts entspricht der Verbindung zwischen zwei korrespondierenden Eckpunkten der beiden um die Rotationsunsicherheit nach links und rechts gedrehten Polygone. Darum eignet sich ein Eckpunkt des einen Polygons als Startpunkt, wenn dessen Verbindung zum korrespondierenden Eckpunkt des zweiten Polygons komplett im zweiten Polygon liegt. Zur Extraktion der Fläche wird eine Startrichtung von innerhalb des ersten Polygons gewählt (1). Nun wird iterativ jeweils eine Kante der gesuchten Fläche gewählt. Dabei wird von allen im aktuellen Punkt ausgehenden Kanten diejenige gewählt, die mit der letzten Richtung den kleinsten Winkel entgegen des Uhrzeigersinns einschließt. Diese Kante wird immer nur bis zum nächsten Schnittpunkt mit einer beliebigen anderen Kante verwendet (2).

B.2 Bereich außerhalb der Straße

Der zweite berechnete Bereich entspricht dem Gebiet, das mit Sicherheit außerhalb der Straße liegt. Hierfür muss das entsprechende Straßenpolygon um die Positionsunsicherheit vergrößert und wieder um die Rotationsunsicherheit gedreht werden. Die beiden Polygone werden wieder an den korrespondierenden Punkten verbunden (Abbildung: B.2, links). Nun wird der Bereich extrahiert, der während der Rotation nicht bedeckt wird.

Als Startpunkt wird hier der Punkt mit der kleinsten x-Koordinate gewählt. Die Startrichtung ist in positive x-Richtung definiert. Nun wird auch hier iterativ jeweils eine Kante des gesuchten Polygons gewählt. Als Kriterium gilt hier der minimale Winkel zwischen alter Orientierung und neuer Kante im Uhrzeigersinn (Abbildung: B.2, rechts).

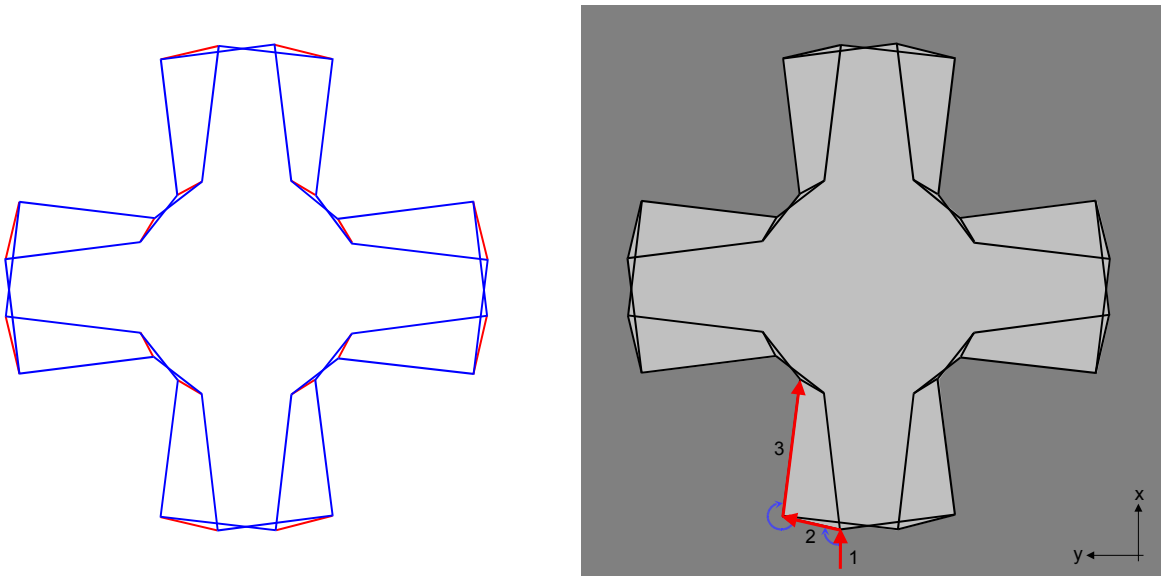

Abbildung B.2: Bereich außerhalb der Straße: Zur Berechnung wird das Straßenpolygon um die Positionsunsicherheit vergrößert und um die Rotationsunsicherheit nach links und rechts gedreht. Nun wird iterativ der Bereich berechnet, der während der Rotation nicht bedeckt wird.

B.3 Ignorierte Gebiete

Für Gebiete, die von der Objektverfolgung und der Klassifikation ignoriert werden können, wird entsprechend Abschnitt B.1 jeweils der Bereich berechnet, der mit Sicherheit im ignorierten Gebiet liegt (Abbildung B.3).

B.4 Verkehrsinseln

Für Verkehrsinseln werden analog zur Straße der Bereich auf und der außerhalb der Verkehrsinsel berechnet. Das Ergebnis ist in Abbildung B.3 dargestellt.

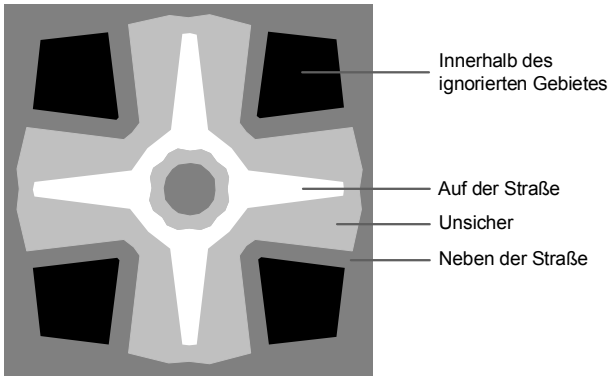

Abbildung B.3: Ergebnisse der Anwendung der Unsicherheiten: Auf alle Bereiche der Karte wurden die Positions- und die Rotationsunsicherheit angewendet.

Literatur

[1] O. Achler und M. Trivedi: “Camera Based Vehicle Detection, Tracking, and Wheel Baseline Estimation Approach”, in *Proceedings of IEEE 7th International Conference on Intelligent Transport Systems*, Washington, USA, 2004.

[2] A. C. Adrowitzer, H. Hlavacs und K. A. Hummel: “Throughput Measurements of IEEE 802.11g-based Ad-hoc Networks in Outdoor Scenarios”, in *Proceedings of WIT 2007: 4th International Workshop on Intelligent Transportation*, Hamburg, Deutschland, 2007.

[3] N. Apostoloff und A. Zelinsky: “Robust Vision based Lane Tracking using Multiple Cues and Particle Filtering”, in *Proceedings of the 2003 IEEE Intelligent Vehicles Symposium*, Columbus, USA, 2003.

[4] Autoliv, INC.: “Pedestrian Protection”, Online: http://www.autoliv.com/alv/connect/Home/What+We+Do/New%20Products/Pedestrian%20Protection (Stand: 20.08.2007).

[5] Y. Bar-Shalom und X.-R. Li: *Estimation and Tracking: Principles, Techniques, and Software*, Artech House, 1993.

[6] Y. Bar-Shalom und X.-R. Li: *Multitarget-Multisensor Tracking: Principles and Techniques*, 1995.

[7] C. Beder, B. Bartczak und R. Koch: “A Comparison of PMD-Cameras and Stereo-Vision for the Task of Surface Reconstruction using Patchlets”, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2007*, Minneapolis, USA, 2007.

[8] S. Blackman und R. Popoli: *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.

[9] V. Blervaque, K. Mezger, L. Beuk und J. Löwenau: “ADAS Horizon - How Digital Maps can contribute to Road Safety”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications 2006*, Berlin, Deutschland, 2006.

[10] K. Borgeest: “Car2x, a Platform for Wireless Communication Based Assistance Systems”, in *Proceedings of WIT 2007: 4th International Workshop on Intelligent Transportation*, Hamburg, Deutschland, 2007.

[11] Robert Bosch GmbH: “Sensor-Regler-Einheit”, Online: http://rb-k.bosch.de/de/einsatzgebiete/fahrspass/fahrerassistenzsysteme/adaptivecruisecontrolacc/sensor reglereinheit.html (Stand: 27.07.2007).

[12] J.-Y. Bouguet: “Camera Calibration Toolbox for Matlab”, Online: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html (Stand: 02.08.2007).

[13] A. Broggi, P. Cerri, P. Medici, P. P. Porta und G. Ghisio: “Real Time Road Signs Recognition”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[14] A. Broggi, P. Cerri, F. Oleari und M. Peterlini: “Guard rail detection using radar and vision data fusion for vehicle detection algorithm improvement and speed-up”, in *Proceedings of IEEE 8th International Conference on Intelligent Transport Systems*, Wien, Österreich, 2005.

[15] CAR 2 CAR Communication Consortium: “CAR 2 CAR Communication Consortium Manifesto -Overview of the C2C-CC System-”, Online: http://www.car-to-car.org/index.php?id=570 (Stand: 16.08.2007).

[16] S. Clemen: “Dreidimensionale Erkennung von Rädern durch Sensorfusion”, Studienarbeit, Universität Ulm, 2007.

[17] Continental Automotive Systems: “ACC mit Radar- oder Infrarot-Technologie”, Online: http://www.conti-online.com/generator/www/de/de/cas/cas/themen/produkte/elektr_brems_und_sicherheit/fahrerassistenzsysteme/radar_und_infrarottechnologie_de.html (Stand: 27.07.2007).

[18] DaimlerChrysler AG: *PREISE FÜR DIE CL-KLASSE*, 2007, Online: http://www.mercedes-benz.de (Stand: 02.08.2007).

[19] Defence Advanced Research Projects Agency: “Urban Challenge Website”, Online: http://www.darpa.mil/grandchallenge/ (Stand: 04.11.2007).

[20] Denso Corporation: “Laser Radar”, Online: http://www.globaldensoproducts.com/dcs/accs/lr.html (Stand: 20.07.2007).

[21] K. Dietmayer, N. Kämpchen, K. Fürstenberg und J. Kibbel: “Roadway Detection and Lane Detection using Multilayer Laserscanner”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications 2005*, Berlin, Deutschland, 2005.

[22] K. Dietmayer, J. Sparbert und D. Streller: “A System Architecture for the Detection and Modelling of Dynamic Traffic Scenes”, in *Proceedings of 8th World Congress on Intelligent Transportation Systems*, Sydney, Australien, 2001.

[23] R. O. Duda, P. E. Hart und D. G. Stork: *Pattern Classification*, John Wiley & Sons, 2001.

[24] D. Ehmanns und T. Kosch: “Short Range Wireless Communication for Intersection Assistance: The INTERSAFE System”, in *Proceedings of 13th World Congress on Intelligent Transportation Systems*, London, Großbritannien, 2006.

[25] J. L. Elman: “Finding Structure in Time”, *Cognitive Science*, vol. 14, S. 179–211, 1990.

[26] ERTICO: “GDF - Geographic Data Files”, Online: http://www.ertico.com/en/links/links/gdf_-_geographic_data_files.htm (Stand: 12.10.2007).

[27] ERTICO: “The Integrated Project PReVENT”, Online: http://www.prevent-ip.org/ (Stand: 07.10.2007).

[28] P. J. Escamilla-Ambrosio und N. Lieven: “Sensor Fusion Approaches to Guideway and Obstacle Detection in the Autotaxi System”, in *2005 8th International Conference on Information Fusion*, Philadelphia, USA, 2005.

[29] European Commission: “eSafety related Research Activities”, Online: http://ec.europa.eu/information_society/activities/esafety/research_activities/index_en.htm (Stand: 23.08.2007).

[30] Europe’s Information Society: “Automotive short-range radars (SRRs) at 24 GHz and 79 GHz”, Online: http://ec.europa.eu/information_society/policy/radio_spectrum/by_topics/srr/index_en.htm (Stand: 27.07.2007).

[31] Europe’s Information Society: “In-vehicle emergency call system eCall (Second eSafety Communication)”, Online: http://europa.eu/scadplus/leg/en/cha/c11328.htm (Stand: 20.08.2007).

[32] B. Fardi, U. Scheunert, H. Cramer und G. Wanielik: “Multi Modal Detection and Parameter-based Tracking of Road Borders with a Laser Scanner”, in *Proceedings of the 2003 IEEE Intelligent Vehicles Symposium*, Columbus, USA, 2003.

[33] B. Fardi und G. Wanielik: “Hough Transformation Based Approach for Road Border Detection in Infrared Images”, in *Proceedings of the 2004 IEEE Intelligent Vehicles Symposium*, Parma, Italien, 2004.

[34] O. Faugeras: *Three-Dimensional Computer Vision*, MIT Press, 2001.

[35] J. E. Fieldsend und R. M. Everson: “Formulation and comparison of multi-class ROC surfaces”, in *Proceedings of the 2nd ROC Analysis in Machine Learning Workshop, part of the 22nd International Conference on Machine Learning (ICML 2005)*, Bonn, Deutschland, 2005.

[36] N. Floudas: “Track-Level Fusion”, *PReVENT ProFusion e-Journal*, vol. 1, S. 8–9, 2006, Online: http://prevent.ertico.webhouse.net/download/News/eJournalionr_E-Journalr_v1.pdf (Stand: 01.08.2007).

[37] F. Fölster, H. Rohling und M.-M. Meinecke: “Pedestrian recognition based on automotive radar sensors”, in *Proceedings of 5th European Congress on Intelligent Transport Systems and Services in Europe*, Hannover, Deutschland, 2005.

[38] U. Franke, D. Gavrila, S. Görzig, F. Lindner, F. Paetzold und C. Wöhler: “Autonomous Driving Goes Downtown”, *IEEE Intelligent Systems*, vol. 13, no. 6, S. 40–48, 1998.

[39] U. Franke und C. Rabe: “Kalman Filter based Depth from Motion with Fast Convergence”, in *Proceedings of the 2005 IEEE Intelligent Vehicles Symposium*, San Francisco, USA, 2005.

[40] W. Franz: “Car-to-Car Communication - Anwendungen und aktuelle Forschungsprogramme in Europa, USA und Japan”, in *VDE Kongress 2004 Berlin - Innovationen für Menschen*, Berlin, Deutschland, 2004.

[41] R. Freymann: “Möglichkeiten und Grenzen von Fahrerassistenz- und Aktiven Sicherheitssystemen”, in *Aktive Sicherheit durch Fahrerassistenz*, Garching, Deutschland, 2004.

[42] B. Fritzke: “A growing neural gas network learns topologies”, in *Advances in Neural Information Processing Systems 7*, G. Tesauro, D. S. Touretzky und T. K. Leen, Eds., S. 625–632. MIT Press, 1995.

[43] K. Fürstenberg und K. Dietmayer: “Fahrzeugumfelderfassung mit mehrzeiligen Laserscannern”, *Journal Technisches Messen 71 (2004) 3*, 2004.

[44] K. Fürstenberg, D. Linzmeier und K. C. J. Dietmayer: “Pedestrian Recognition and Tracking of Vehicles using a vehicle based Multilayer Laserscanner”, in *Proceedings of 9th World Congress on Intelligent Transportation Systems*, Madrid, Spanien, 2003.

[45] K. Fürstenberg und R. Schulz: “Laserscanners for Driver Assistance”, in *Proceedings of WIT 2006: 3rd International Workshop on Intelligent Transportation*, Hamburg, Deutschland, 2006.

[46] K. C. Fürstenberg: “Intersection Safety - The EC-Project INTERSAFE”, San Francisco, USA, 2005.

[47] K. C. Fürstenberg und K. Dietmayer: “Object Tracking and Classification for Multiple Active Safety and Comfort Applications using a Multilayer Laserscanner”, in *Proceedings of the 2004 IEEE Intelligent Vehicles Symposium*, Parma, Italien, 2004.

[48] Kay Ch. Fürstenberg und Thorsten Weiss: “Feature-Level Map Building and Object Recognition for Intersection Safety Applications”, in *Proceedings of IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005, S. 490–495.

[49] D. M. Gavrila und S. Munder: “Multi-Cue Pedestrian Detection and Tracking from a Moving Vehicle”, *International Journal of Computer Vision*, vol. 73, no. 1, S. 41–59, 2007.

[50] S. K. Gehring, S. Wagner und U. Franke: “System Architecture for an Intersection Assistant Fusing Image, Map, and GPS Information”, in *Proceedings of the 2003 IEEE Intelligent Vehicles Symposium*, Columbus, USA, 2003.

[51] M. Gerlach, A. Festag, T. Leinmüller, G. Goldacker und C. Harsch: “Security Architecture for Vehicular Communication”, in *Proceedings of WIT 2007: 4th International Workshop on Intelligent Transportation*, Hamburg, Deutschland, 2007.

[52] J. Giebel: *Learning Dynamic Shape Models for Bayesian Tracking*, Dissertation, Universität Mannheim, Mannheim, Deutschland, 2005.

[53] T. Goering: “Sensors for Active and Passive Safety Systems”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications*, 2007.

[54] D. Göhring: “Digitalkameratechnologien - Eine vergleichende Betrachtung”, Seminararbeit, Humboldt Universität zu Berlin, 2002.

[55] J. Gunnarsson, L. Svensson, L. Danielsson und F. Bengtsson: “Tracking vehicles using radar detections”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[56] Y. Günter und H. P. Großmann: “Usage of Wireless LAN for Inter-Vehicle Communication”, in *Proceedings of IEEE 8th International Conference on Intelligent Transport Systems*, Wien, Österreich, 2005.

[57] B. Hagebeuker, R. Lange und T. Ringbeck: “Bildverarbeitung mit dritter Dimension”, *SENSOR report*, vol. 2/2004, 2004.

[58] F. Hamker und D. Heinke: “Implementation and Comparison of Growing Neural Gas, Growing Cell Structures and Fuzzy Artmap”, Tech. Rep., Technische Universität Ilmenau, 1997.

[59] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner und W.v. Seelen: “An Image Processing System for Driver Assistance”, *Image and Vision Computing*, vol. 18, no. 5, S. 367–376, 2000.

[60] Internetseite der IBEO Automobile Sensor GmbH: “Products”, Online: http://www.ibeo-as.com/english/products.asp (Stand: 26.07.2007).

[61] C. Idler, R. Schweiger, D. Paulus, M. Mählisch und W. Ritter: “Realtime Vision Based Multi-Target-Tracking with Particle Filters in Automotive Applications”, in *Proceedings of the 2006 IEEE Intelligent Vehicles Symposium*, Tokio, Japan, 2006.

[62] Integrated Project PReVENT: “Internetseite des Projektes INTERSAFE”, Online: http://prevent-ip.org/en/prevent_subprojects/intersection_safety/intersafe/ (Stand: 26.07.2007).

[63] Intel Research Labs: “Open Source Computer Vision Library”, Online: http://www.intel.com/technology/computing/opencv/ (Stand: 05.10.2007).

[64] I. Kallenbach, R. Schweiger, G. Palm und O. Löhlein: “Multi-class Object Detection in Vision Systems Using a Hierarchy of Cascaded Classifiers”, in *Proceedings of the 2006 IEEE Intelligent Vehicles Symposium*, Tokio, Japan, 2006.

[65] R. E. Kalman: “A New Approach to Linear Filtering and Prediction Problems”, *Transaction of the ASME - Journal of Basic Engineering*, vol. 82, S. 35–45, 1960.

[66] N. Kämpchen: *Feature-Level Fusion of Laser Scanner and Video Data for Advanced Driver Assistance Systems*, Dissertation, Universität Ulm, 2007.

[67] N. Kämpchen, M. Bühler und K. C. J. Dietmayer: “Feature-Level Fusion for Free-Form Object Tracking using Laserscanner and Video”, in *Proceedings of the 2005 IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, 2005.

[68] N. Kämpchen, M. Clauss, Y. Günter, R. Schreier, M. Stiegeler, K. Tischler, K. Dietmayer, H. P. Grossmann, H. Kabza, A. Rothermel, C. Stiller und H. Neumann: “Vernetzte Fahrzeug-Umfelderfassung fÃijr zukünftige Fahrerassistenzsysteme”, in *Tagungsband 3. Workshop Fahrerassistenzsysteme*, Walting, Deutschland, 2005.

[69] N. Kämpchen und K. C. J. Dietmayer: “Fusion of Laserscanner and Video for Advanced Driver Assistance Systems”, in *Proceedings of 11th World Congress on Intelligent Transportation Systems*, Nagoya, Japan, 2004.

[70] N. Kasabov: “Evolving Fuzzy Neural Networks - Algorithms, Applications and Biological Motivation”, *Methodologies for the Conception, Design and Application of Soft Computing*, S. 271–274, 1998.

[71] N. Kawasaki und U. Kiencke: “Standard Platform for Sensor Fusion on Advanced Driver Assistance System using Bayesian Network”, in *Proceedings of the 2004 IEEE Intelligent Vehicles Symposium*, Parma, Italien, 2004.

[72] A. Khammari, F. Nashashibi, Y. Abramson und C. Laurgeau: “Vehicle Detection Combining Gradient Analysis and AdaBoost Classification”, in *Proceedings of IEEE 8th International Conference on Intelligent Transport Systems*, Wien, Österreich, 2005.

[73] A. Kirchner: *Sensordatenverarbeitung eines Laserscanners für autonome Fahrfunktionen von Kraftfahrzeugen*, Fortschritt-Berichte VDI: Reihe 12, Verkehrstechnik, Fahrzeugtechnik, Nr. 438. VDI-Verlag, 2000.

[74] H. Kobari, H. Uno, A. Kimata, S. Ueda und A. Takashi: “Utility Studies of a Warning Alarm Signaling the Presence of Pedestriens”, in *Proceedings of 13th World Congress on Intelligent Transportation Systems*, London, Großbritannien, 2006.

[75] S. Kolski, D. Ferguson, M. Bellino und R. Siegwart: “Autonomous Driving in Structured and Unstructured Environments”, in *Proceedings of the 2006 IEEE Intelligent Vehicles Symposium*, Tokio, Japan, 2006.

[76] T. Kvisterøy, H. Jakobsen, C. Vieider, St. Wissmar, P. Ericsson, U. Halldin, F. Niklaus, Fr. Forsberg, G. Stemme, J. E. Källhammer, H. Pettersson, D. Eriksson, J. Franks, J. VanNylen, H. Vercammen und A. VanHusel: “Far Infrared Low-Cost Uncooled Bolometer for Automotive Use”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications 2007*, Berlin, Deutschland, 2007.

[77] R. Labayrade, S. S. Ieng und D. Aubert: “A Reliable Road Lane Detector Approach Combining Two Vision-Based Algorithms”, in *Proceedings of IEEE 7th International Conference on Intelligent Transport Systems*, Washington, D.C., USA, 2004.

[78] R. Labayrade, C. Royere und D. Aubert: “A Collision Mitigation System using Laser Scanner and Stereovision Fusion and its Assessment”, in *Proceedings of the 2005 IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, 2005.

[79] U. S. Lages: *Untersuchungen zur aktiven Unfallvermeidung von Kraftfahrzeugen*, Fortschritt-Berichte VDI: Reihe 12, Verkehrstechnik, Fahrzeugtechnik, Nr. 446. VDI-Verlag, 2001.

[80] J. Lee und C. Crane: “Road Following in an Unstructured Desert Environment Based on the EM(Expectation-Maximization) Algorithm”, in *Proceedings of the International Conference on Control, Automation, and Systems*, Pusan, Korea, 2006.

[81] R. Lienhart, A. Kuranov und V. Pisarevsky: “Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection”, Tech. Rep., Microprocessor Research Lab, Intel Labs, 2002.

[82] R. Lienhart, L. Liang und A. Kuranov: “A Detector Tree of Boosted Classifiers for Real-Time Object Detection and Tracking”, in *Proceedings of 2003 International Conference on Multimedia and Expo*, 2003.

[83] R. Lindholm: “ECALL End-To-End Testing In Germany, Austria & Italy”, in *Proceedings of 6th European Congress on Intelligent Transport Systems and Services in Europe*, Aalborg, Dänemark, 2007.

[84] R. Lindl und L. Walchshäusl: “Three-Level Early Fusion for Road User Detection”, *PReVENT ProFusion e-Journal*, vol. 1, S. 19–24, 2006, Online: http://prevent.ertico.webhouse.net/download/News/eJournalionr_E-Journalr_v1.pdf (Stand: 02.08.2007).

[85] R. Lindl, L. Walchshäusl und V. Paraschoudis: “Autonomous Braking for Collision Mitigation Purposes by Means of Multi Sensor Perception”, in *Proceedings of 6th European Congress on Intelligent Transport Systems and Services in Europe*, Aalborg, Dänemark, 2007.

[86] F. Lindner, U. Kressel und S. Kälberer: “Robust Recognition of Traffic Signals”, in *Proceedings of the 2004 IEEE Intelligent Vehicles Symposium*, Parma, Italien, 2004.

[87] D. Linzmeier: “Sensor- und Objektmodell gestützte Assoziation mit Kalman-Filtern”, Diplomarbeit, Universität Ulm, 2003.

[88] D. Linzmeier: *Real-Time Detection of Pedestrians from a Moving Vehicle Using Thermopile and Radar Sensor Fusion*, Fortschritt-Berichte VDI: Reihe 12, Verkehrstechnik, Fahrzeugtechnik, Nr. 628. VDI-Verlag, 2006.

[89] J. Löwenau, J. Angenvoort und K. Mezger: “ADASIS: Real Time Map Information with a Standardized Interface for Advanced In-Vehicle Applications”, in *Proceedings of 6th European Congress on Intelligent Transport Systems and Services in Europe*, Aalborg, Dänemark, 2007.

[90] J. Löwenau, K. Gresser, D. Wisselmann, W. Richter, D. Rabel und S. Durekovic: “Dynamic Pass Prediction - A New Driver Assistance System for Superior and Safe Overtaking”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications 2006*, Berlin, Deutschland, 2006.

[91] M. Mählisch, M. Oberländer, O. Löhlein, D. Gavrila und W. Ritter: “A Multiple Detector Approach to Low-resolution FIR Pedestrian Recognition”, in *Proceedings of the 2005 IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, 2005.

[92] M. Mählisch, W. Ritter und K. Dietmayer: “De-cluttering with Integrated Probabilistic Data Association for Multisensor Multitarget ACC Vehicle Tracking”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[93] M. Mählisch, W. Ritter und K. Dietmayer: “Feature level video and lidar sensor fusion for ACC stop&go using Joint Integrated Probabilistic Data Association”, in *Proceedings of 6th European Congress on Intelligent Transport Systems and Services in Europe*, Aaalborg, Dänemark, 2007.

[94] M. Mählisch, R. Schweiger, W. Ritter und K. Dietmayer: “Sensorfusion Using Spatio-Temporal Aligned Video and Lidar for Improved Vehicle Detection”, in *Proceedings of the 2006 IEEE Intelligent Vehicles Symposium*, Tokio, Japan, 2006.

[95] J. J. Mahoney und R. J. Mooney: “Combining symbolic and neural learning to revise probabilistic theories”, in *Proceedings 1992 Machine Learning Workshop on Integrated Learning in Real Domains*, Aberdeen, Schottland, 1992.

[96] G. Malnati, C. Barberis und C. M. Cuva: “Gossip: Estimating Actual Travelling Time Using Vehicle to Vehicle Communication”, in *Proceedings of WIT 2007: 4th International Workshop on Intelligent Transportation*, Hamburg, Deutschland, 2007.

[97] M.-M. Meinecke, A. Obojski, D. Gavrila, E. Marc, R. Morris, M. Töns und L. Letellier: “Deliverable D6 Annex: Strategies in Terms of Vulnerable Road User Protection”, 2001, Deliverable of the EC project SAVE-U, Online: http://www.save-u.org/file_html/library.htm (Stand: 06.10.2007).

[98] U. Meis, M. Oberländer und W. Ritter: “Reinforcing the Reliability of Pedestrian Detection in Far-infrared Sensing”, in *Proceedings of the 2004 IEEE Intelligent Vehicles Symposium*, Parma, Italien, 2004.

[99] K.-H. Meitinger, B. Heißing und D. Ehmanns: “Linksabbiegeassistenz – Beispiel für die Top-Down-Entwicklung eines Aktiven Sicherheitssystems”, in *Aktive Sicherheit*, München, Deutschland, 2006.

[100] P. Mengel, L. Listl, B. Koenig, M. Pellkofer, U. Wagner und R. Wertheimer: “Time-of-Flight Camera for Pedestrian Protection and Collision Mitigation”, in *Proceedings of 6th European Congress on Intelligent Transport Systems and Services in Europe*, Aalborg, Dänemark, 2007.

[101] J. Miura, T. Kanda und Y. Shirai: “An Active Vision System for Real-Time Traffic Sign Recognition”, in *Proceedings of IEEE 3th International Conference on Intelligent Transport*, Dearborn, USA, 2000.

[102] Y. Miyake, K. Natsume und K. Hoshino: “Road-Shape Recognition Using On-Vehicle Millimeter-Wave Radar”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[103] R. Möbus, A. Joos und U. Kolbe: “Multi-Target Multi-Object Radartracking”, in *Proceedings of the 2003 IEEE Intelligent Vehicles Symposium*, Columbus, USA, 2003.

[104] R. Möbus und U. Kolbe: “Multi-Target Multi-Object Tracking, Sensor Fusion of Radar and Infrared”, in *Proceedings of the 2004 IEEE Intelligent Vehicles Symposium*, Parma, Italien, 2004.

[105] M. F. Moller: “A scaled conjugate gradient algorithm for fast supervised learning”, *Neural Networks*, vol. 6, S. 525–533, 1993.

[106] National Imagery and Mapping Agency: *Department of Defense World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems. Third Edition*, 1984, Online: earth-info.nga.mil/GandG/publications/tr8350.2/wgs84fin.pdf (Stand: 01.12.2006).

[107] N. Pettersson und Lars Petersson: “Online Stereo Calibration using FPGAs”, in *Proceedings of the 2005 IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, 2005.

[108] D. Ponsa, A. Lopez, F. Lumbreras, J. Serrat und T. Graf: “3D Vehicle Sensor based on Monocular Vision”, in *Proceedings of IEEE 8th International Conference on Intelligent Transport Systems*, Wien, Österreich, 2005.

[109] D. Ponsa, A. Lopez, J. Serrat, F. Lumbreras und T. Graf: “Multiple Vehicle 3D Tracking Using an Uncented Kalman Filter”, in *Proceedings of IEEE 8th International Conference on Intelligent Transport Systems*, Wien, Österreich, 2005.

[110] Preiss Technology Systems: “Sensortechnik”, Online: http://www.fen-net.de/walter.preiss/d/slomoinf.html (Stand: 30.07.2007).

[111] ProFusion: “Deliverable D13.400: Sensor Data Sheets - State-of-the-art of Sensors and Sensor Data Fusion for Automotive Preventive Safety Applications”, Tech. Rep., 2004.

[112] C. Rehbach: “Fusion zweier Laserscanner auf Merkmalsebene”, Studienarbeit, Universität Ulm, 2007.

[113] D. Reichardt, M. Miglietta, L. Moretti, P. Morsink und W. Schulz: “CarTALK 2000: Safe and Comfortable Driving Based Upon Inter-Vehicle-Communication”, in *Proceedings of the 2002 IEEE Intelligent Vehicles Symposium*, Versailles, Frankreich, 2002.

[114] C. Ress und M. Wiecker: “The future of vehicle navigation - dynamic and learning navigation systems”, in *Proceedings of WIT 2007: 4th International Workshop on Intelligent Transportation*, Hamburg, Deutschland, 2007.

[115] H. Rohling, A. Höß, U. Lübbert und M Schiementz: “Multistatic Radar Principles for Automotive RadarNet Applications”, in *Proceedings of IRS 2002: International Radar Symposium*, 2002.

[116] Bernd Rössler, Kay Fürstenberg und Ulrich Lages: “Intersection Safety - the EC Project INTERSAFE”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications*, Berlin, Germany, April 2006.

[117] A. Rottmann, O. Martínez Mozos, C. Stachniss und W. Burgard: “Place Classification of Indoor Environments with Mobile Robots using Boosting”, in *Proceedings of the National Conference on Artificial Intelligence*, Pittsburgh, PA, USA, 2005.

[118] SAFESPOT Consortium: “SAFESPOT Internetseite”, Online: http://www.safespot-eu.org/ (Stand: 08.10.2007).

[119] U. Scheunert, B. Fardi, N. Mattern, G. Wanielik und N. Keppeler: “Free space determination for parking slots using a 3D PMD sensor”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[120] B. Schiele: “Bestimmung des Auslösezeitpunkts einer Automatischen Notbremse”, Studienarbeit, Universität Ulm, 2005.

[121] H. Schneiderman und T. Kanade: “A Statistical model for 3D Object Detection Applied to Faces and Cars”, in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2000.

[122] M. Schönherr: “Klassifikation von Objekten im Fahrzeugumfeldbereich”, Diplomarbeit, Universität Ulm, 2005.

[123] R. Schweiger, K. Dietmayer, M. Serfling und W. Ritter: “Fusion of digital map and computer vision data for automotive safety applications”, in *Proceedings of 13th World Congress on Intelligent Transportation Systems*, London, Großbritannien, 2006.

[124] M. Skutek, D. T. Linzmeier, N. Appenrodt und G. Wanielik: “A PreCrash System based on Sensor Data Fusion of Laser Scanner and Short Range Radars”, in *2005 8th International Conference on Information Fusion*, Philadelphia, USA, 2005.

[125] J. Sparbert, K. Dietmayer und D. Streller: “Lane Detection and Street Type Classification using Laser Range Images”, in *Proceedings of IEEE 4th International Conference on Intelligent Transport Systems*, Oakland, USA, 2001.

[126] R. Stahn, G. Heiserich und A. Stopp: “Laser Scanner-Based Navigation for Commercial Vehicles”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[127] Statistisches Bundesamt: *STATISTISCHES JAHRBUCH 2006 Für die Bundesrepublik Deutschland*, Statistisches Bundesamt, 2006.

[128] D. Streller: *Multi-Hypothesen-Ansatz zur Erkennung und Verfolgung von Objekten in Verkehrsszenen mit Laserscannern*, Fortschritt-Berichte VDI: Reihe 12, Verkehrstechnik, Fahrzeugtechnik, Nr. 614. VDI-Verlag, 2006.

[129] D. Stüker: *Heterogene Sensordatenfusion zur robusten Objektverfolgung im automobilen Straßenverkehr*, Dissertation, Carl von Ossietzky-Universität Oldenburg, 2004.

[130] Z. Sun, R. Miller, G. Bebis und D. DiMeo: “A Real-time Precrash Vehicle Detection System”, in *Proceedings of the 2000 IEEE Intelligent Vehicles Symposium*, Dearborn, USA, 2000.

[131] Mate Szarvas, Utsushi Sakai und Jun Ogata: “Real-time Pedestrian Detection Using LIDAR and Convolutional Neural Networks”, in *Proceedings of IEEE Intelligent Vehicles Symposium*, Tokio, Japan, 2006.

[132] H.-S. Tan und J. Huang: “DGPS-Based Vehicle-to-Vehicle Cooperative Collision Warning: Engineering Feasibility Viewpoints”, *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 4, S. 415–428, 2006.

[133] T. Tatschke: “Early Fusion approach”, *PReVENT ProFusion e-Journal*, vol. 1, S. 6–7, 2006, Online: http://prevent.ertico.webhouse.net/download/News/eJournalionr_E-Journalr_v1.pdf (Stand: 02.08.2007).

[134] T. Tatschke, S.-B. Park, A. Amditis, A. Polychronopoulos, U. Scheunert und O. Aycard: “ProFusion2 - towards a Modular, Robust and Reliable Fusion Architecture for Automotive Environment Perception”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications 2006*, Berlin, Deutschland, 2006.

[135] Jörg Thiem: “Datafusion of two Driver Assistance System Sensors”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications*, Berlin, Deutschland, 2005.

[136] S. B. Thrun: “Extracting Provably Correct Rules from Artificial Neural Networks”, Tech. Rep., Institut für Informatik III, Universität Bonn, 1993.

[137] K. Tischler, M. Clauss, Y. Günter, N. Kämpchen, R. M. Schreier und M. M. Stiegeler: “Networked Environment Description for Advanced Driver Assistance Systems”, in *Proceedings of IEEE 8th International Conference on Intelligent Transport Systems*, Wien, Österreich, 2005.

[138] K. Tischler und B. Hummel: “Enhanced Environmental Perception by Inter-Vehicle Data Exchange”, in *Proceedings of the 2005 IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, 2005.

[139] B. Ulmer: “Introduction to ADASE2”, *ADASE2 - First Concertation Meeting*, 2001, Online: http://www.adase2.net (Stand: 20.08.2007).

[140] P. Viola und M. Jones: “Robust real-time Object Detection”, Tech. Rep., Compaq Cambridge Research Laboratory, 2001.

[141] P. Viola und M. Jones: “Fast and Robust Classification using Asymmetric AdaBoost and a Detector Cascade”, in *Advances in Neural Information Processing Systems 14*, 2002.

[142] P. Viola, M. J. Jones und D. Snow: “Detecting Pedestrians Using Patterns of Motion and Appearance”, in *Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003)*, 2003.

[143] VOLKSWAGEN AG: “Internetseite Volkswagen”, Online: http://www.volkswagen.de (Stand: 02.08.2007).

[144] T.-D. Vu, O. Aycard und N. Appenrodt: “Online Localization and Mapping with Moving Object Tracking in Dynamic Outdoor Environments”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[145] L. Walchshäusl, R. Lindl und K. Vogel: “Detection of Road Users in Fused Sensor Data Streams for Collision Mitigation”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications 2006*, Berlin, Deutschland, 2006.

[146] T. Weiss und K. Dietmayer: “Robust Host Localization in Skidding Situations Using Laserscanner”, in *Proceedings of WIT 2007: 4th International Workshop on Intelligent Transportation*, Hamburg, Deutschland, 2007.

[147] T. Weiss, N. Kämpchen und K. Dietmayer: “High Accuracy Ego-Localization in Complex Dynamic Environments using Laserscanners”, San Francisco, USA, 2005.

[148] T. Weiss, N. Kämpchen und K. Dietmayer: “Precise Ego-Localization in Urban Areas using Laserscanner and High Accuracy Feature Maps”, in *Proceedings of the 2005 IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, 2005.

[149] T. Weiss, B. Schiele und K. Dietmayer: “Robust Road Corridor Detection in Urban and Highway Scenarios Using a Laser Scanner and Online Occupancy Grids”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[150] T. Weiss, J. Spruck und K. Dietmayer: “A Scalable Sensor Fusion Framework for the Localization of a Vehicle on Detailed Digital Maps Using Laserscanners”, in *Proceedings of 2006 IEEE Conference on Multisensor Fusion and Integration*, Heidelberg, Deutschland, 2006.

[151] T. Weiss, S. Wender und K. Dietmayer: “Präzise Positions- und Zustandsschätzung eines Fahrzeugs mit Hilfe eines Laserscanners und detaillierten digitalen Karten”, in *4. Workshop Fahrerassistenzsysteme FAS 2006*, Löwenstein, Deutschland, 2006.

[152] T. Weiss, S. Wender und K. Dietmayer: “Precise Host Localization in Urban Areas”, *PReVENT ProFusion e-Journal*, vol. 1, S. 25–30, 2006, Online: http://prevent.ertico.webhouse.net/download/News/eJournalionr_E-Journalr_v1.pdf (Stand: 01.08.2007).

[153] G. Welch und G. Bishop: “An Introduction to the Kalman Filter”, Tech. Rep., Department of Computer Science, University of North Carolina at Chapel Hill, 2001.

[154] S. Wender: “Insassendetektion mit kaskadierten Klassifikatoren”, Diplomarbeit, Technische Universität Ilmenau, 2003.

[155] S. Wender, K. Fürstenberg und K. Dietmayer: “Object Tracking and Classification for Intersection Scenarios Using A Multilayer Laserscanner”, in *Proceedings of 11th World Congress on Intelligent Transportation Systems*, Nagoya, Japan, 2004.

[156] S. Wender, T. Weiss, K. Fürstenberg und K. Dietmayer: “Object Classification exploiting High Level Maps of Intersections”, in *Proceedings of AMAA: Advanced Microsystems for Automotive Applications 2006*, Berlin, Deutschland, 2006.

[157] L. Wenzel: “Kalman-Filter”, *Elektronik*, vol. 6, 8, 11 and 13, 2000, Online: digital.ni.com/worldwide/germany.nsf/3528e4c55a5f22d4852568e100532f33/e10103f0159a7f2ec12568e90057f03b?OpenDocument (Stand: 06.08.2007).

[158] H. Winner: “Mit aktiven Sensoren das Kfz-Umfeld erfassen - Funktion und Leistungsfähigkeit von Radar & Co”, *it - Information Technology*, vol. 49, S. 17–24, 2007.

[159] C. Wöhler, J. K. Anlauf, T. Pörtner und U. Franke: “A Time Delay Neural Network Algorithm for Real-Time Pedestrian Recognition”, in *Proceedings of 1998 IEEE International Conference on Intelligent Vehicles*, Stuttgart, Deutschland, 1998.

[160] E. Zabler, J. Berger, A. Herforth, M. Harder, H. Arnold, H. Winner, J. Töpfer, R. Giesel und V. Denner: *Sensoren im Kraftfahrzeug*, Robert Bosch GmbH, 2001.

[161] L. Zadeh: “Fuzzy Sets”, *Information and Control*, vol. 8, 1965.

[162] H. Zhao, Y. Chen und R. Shibasaki: “An efficient extrinsic calibration of a multiple laser scanners and cameras sensor system on a mobile platform”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[163] X. Zhou, Q. Yu, J. Rong und Dai S.: “Research of Traffic Guidance System Based on Narrowband Data Communication”, in *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, 2007.

[164] Y. Zhu, D. Comaniciu, V. Ramesh, M. Pellkofer und T. Köhler: “An Integrated Framework of Vision-based Vehicle Detection with Knowledge Fusion”, in *Proceedings of the 2005 IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, 2005.

Betreute Studien- und Diplomarbeiten

Sebastian Wiehr: „Filterung drahtlos übertragener Bewegungsdaten“, Studienarbeit, September 2008

Christof Rehbach: „Hintergrundausblendung mit Online Maps und Daten aus Fahrzeug-Fahrzeug-Kommunikation zur Verbesserung von Trackingverfahren mit Laserscannern“, Diplomarbeit, Februar 2008

Tobias Bund: „Fusion von Laserscanner und Video zur Fußgängerdetektion”, Studienarbeit, Februar 2008

Carolin Christine Hank: „Verwendung von Gaussian Mixture Models zur Objektklassifikation im Fahrzeugumfeld“, Studienarbeit, September 2007

Christof Rehbach: „Fusion zweier Laserscanner auf Merkmalsebene“, Studienarbeit, April 2007

Simon Clemen: „Dreidimensionale Erkennung von Rädern durch Sensorfusion“, Studienarbeit, September 2007

Michael Schönherr: „Klassifikation von Objekten im Fahrzeugumfeldbereich“, Diplomarbeit, Mai 2005

Liste der Publikationen

S. Wender, K. Dietmayer: „3D Vehicle Detection Using a Laser Scanner and a Video Camera“, in *IET Intelligent Transport Systems*, Volume 2, Issue 2, p. 105-112, June 2008

S. Wender, K. Dietmayer: „3D Vehicle Detection Using a Laser Scanner and a Video Camera“, in *Proceedings of 6th European Congress on ITS in Europe*, Aalborg, Dänemark, Juni 2007

H. Berndt, S. Wender, K. C. J. Dietmayer: „Driver Braking Behavior during Intersection Approaches and Implications for Warning Strategies for Driver Assistant Systems“, in *Proceedings of IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, Juni 2007

S. Wender, K. C. J. Dietmayer: „Extending Onboard Sensor Information by Wireless Communication“, in *Proceedings of IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, Juni 2007

S. Wender, K. C. J. Dietmayer: „A Feature Level Fusion Approach for Object Classification“, in *Proceedings of IEEE Intelligent Vehicles Symposium*, Istanbul, Türkei, Juni 2007

T. Weiss, S. Wender, K. Dietmayer: „Landmark Navigation for Robust Object Tracking in Skidding Maneuvers Using Laser Scanners“, in *Proceedings of Advanced Microsystems for Automotive Applications 2007*, Berlin, Deutschland, Mai 2007

S. Wender, T. Weiss, K. Dietmayer: „Integration of Wireless Communication Data in a Sensor Based Vehicle Environment Model“, *4th International Workshop on Intelligent Transportation*, Hamburg, Deutschland, März 2007

T. Weiss, S. Wender, K. Dietmayer: „Präzise Positions- und Zustandsschätzung eines Fahrzeugs mit Hilfe eines Laserscanners und detaillierten digitalen Karten“, *4. Workshop Fahrerassistenzsysteme*, Löwenstein, Deutschland, Oktober 2006

S. Wender, T. Weiss, K. Fürstenberg, K. Dietmayer: „Feature Level Fusion for Object Classification“, *PReVENT ProFusion e-Journal*, Vol. 1, pp 31-36, September 2006

T. Weiss, S. Wender, K. Fürstenberg, K. Dietmayer: „Precise Host Localization in Urban Areas“, *PReVENT ProFusion e-Journal*, Vol. 1, pp 25-30, September 2006

S. Wender, S. Clemen, N. Kämpchen, K. C. J. Dietmayer: „Vehicle Detection with Three Dimensional Object Models“, in *Proceedings of 2006 IEEE Conference on Multisensor Fusion and Integration*, Heidelberg, Deutschland, September 2006

S. Wender, K. C. J. Dietmayer: „An Adaptable Object Classification Framework“, in *Proceedings of 2006 IEEE Intelligent Vehicles Symposium*, Tokio, Japan, Juni 2006

S. Wender, T. Weiss, K. Fürstenberg, K. C. J. Dietmayer: „Object Classification exploiting High Level Maps of Intersections“, in *Proceedings of Advanced Microsystems for Automotive Applications 2006*, Berlin, Deutschland, April 2006

S. Wender, K. Dietmayer: „Statistical Approaches for Vehicle Environment Classification at Intersections with a Laserscanner“, in *Proceedings of ITS 2005, 12th World Congress on Intelligent Transportation Systems*, San Francisco, USA, November 2005

S. Wender, T. Weiss, K. Dietmayer: „Improved Object Classification of Laserscanner Measurements at Intersections using Precise High Level Maps“, in *Proceedings of the 8th*

International IEEE Conference on Intelligent Transportation Systems, Wien, Österreich, September 2005

S. Wender, M. Schönherr, N. Kämpchen, K. C. J. Dietmayer: „Classification of Laserscanner Measurements at Intersection Scenarios with Automatic Parameter Optimization“, in *Proceedings of 2005 IEEE Intelligent Vehicles Symposium*, LasVegas, USA, Juni 2005

S. Wender, K. Fürstenberg, K. Dietmayer: „Object Tracking and Classification for Intersection Scenarios Using A Multilayer Laserscanner“, in *Proceedings of ITS 2004, 11th World Congress on Intelligent Transportation Systems*, Nagoya, Japan, Oktober 2004

S. Wender, O. Löhlein, H.-M. Groß: „Multiple Classifier Cascades for Vehicle Occupant Monitoring Using an Omni-directional Camera“, in *Proceedings of SOAVE 2004: 3rd Workshop on SelfOrganization of AdaptiVE Behavior*, Ilmenau, Deutschland, September 2004

S. Wender, O. Löhlein: „A Cascade Detector Approach Applied to Vehicle Occupant Monitoring with an Omni-Directional Camera“, in *Proceedings of Intelligent Vehicles Symposium*, Parma, Italien, Juni 2004