

ulm university universität
uulm

Feature-Level Fusion of Laser Scanner and Video Data for Advanced Driver Assistance Systems

DISSERTATION

zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

(DR.-ING.)

der Fakultät für Ingenieurwissenschaften

und Informatik der Universität Ulm

von

NICO KÄMPCHEN

aus Hannover

1. Gutachter:	Prof. Dr.-Ing. K. Dietmayer
2. Gutachter:	Prof. Dr.-Ing. J.-U. Varchmin
Amtierender Dekan:	Prof. Dr. rer. nat. Helmuth Partsch
Datum der mündlichen Prüfung:	29. Juni 2007

Acknowledgments

First, I would like to thank Professor Dr.-Ing. Klaus Dietmayer for his exceptional support throughout the years. In particular I am grateful for the scientific freedom and encouragement he gave me, enabling me to follow my own ideas.

My gratitude also goes to Professor Dr.-Ing. Jörn-Uwe Varchmin whose suggestions and critique proved invaluable to this thesis.

I am indebted to my colleague and good friend Kay Fürstenberg who has been generous with his advise, expertise and experience. Throughout the years we shared the same office, our ideas and also the ups and downs of a PhD life.

I would like to thank the members of the ARGOS group for the many fruitful discussions and excellent team work. In particular, I should like to mention Stefan Wender, Thorsten Weiß, Mirko Mählisch, Daniel Streller, Jan Sparbert and Holger Berndt. Special thanks also goes to the students involved in the ARGOS project for their priceless contributions. These are: Matthias Bühler, Tobias Bühler, Jörg Kibbel, Ulrich Plöckl, Michael Schäfer, Michael Schönherr, Bruno Schiele, Alexander Skibicki, Andreas Wimmer and Markus Zocholl.

I would like to acknowledge Franz Degenhard, Martin Nieß, Oliver Betz and Thomas Löffler for their excellent work on maintaining the test vehicles, constructing the experimental setups and helping with the collection of data.

My gratitude also goes to the members of the company IBEO Automobile Sensor GmbH for the intensive and fruitful cooperation.

Special thanks go to my parents and brother for their numerous ways of encouragement. With "Gute Laune Drops", post cards, poems, phone calls and by many other means, they lightend up especially the last month of intense writing up.

Finally, I would like to thank my wife Katherine for her constant support, encouragement and love. She and our son Frederik brought me happiness and were a refreshing source of life outside the PhD.

"The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery every day. Never lose a holy curiosity."

Albert Einstein

Contents

Notation **ix**

1 Introduction **1**

2 State of the Art and Motivation **5**
2.1 Advanced Driver Assistance Systems 5
2.2 Sensors 9
2.2.1 Laser Scanner Data based Environment Perception 10
2.2.2 Video Data based Environment Perception 12
2.3 Sensor Data Fusion 13
2.4 State Estimation 15
2.5 Summary 17

3 Sensor Calibration **19**
3.1 Sensor and Vehicle Coordinate Systems 20
3.2 Calibration of a Multi-Layer Laser Scanner 21
3.2.1 Estimation of Pitch and Roll Angle 21
3.2.2 Estimation of the Yaw Angle 24
3.3 Calibration of a Camera 26
3.4 Worst Case Error Analysis 29
3.4.1 Erroneous Measurements and Calibration Parameters 30
3.4.2 Worst Case Errors of the Estimated Angles 31
3.4.3 Worst Case Projection Errors 32
3.5 Experiments 34
3.6 Discussion 35

4 Sensor Data Fusion **37**
4.1 Low-Level Fusion 37
4.1.1 Tracking 38
4.1.2 Classification 39
4.2 High-Level Fusion 39
4.2.1 Tracking 39
4.2.2 Classification 40
4.3 Feature-Level Fusion 40
4.3.1 System Architecture 40
4.3.2 Cuboidal Object Model 42
4.3.3 Sensor Specific Modules 43

4.3.4 Kalman Filter based State Estimation and Data Fusion 43
4.3.5 Classification 52
4.3.6 Object Management 53
4.3.7 Conclusion 56
4.4 Alternative Fusion Architectures 57
4.5 System Latency 57
4.5.1 Object Detection Latency 57
4.5.2 State Estimation Latency 58
4.6 Discussion 59

5 Laser Scanner Data based Feature Extraction and Association 61
5.1 The Multi-Layer Laser Scanner 62
5.2 Preprocessing 63
5.2.1 Reflector Classification 63
5.2.2 Ground Detection 65
5.2.3 Segmentation 67
5.2.4 Occlusion Detection 68
5.3 Feature Model and Feature Measurement Equation 68
5.3.1 Contour Model Features 68
5.3.2 Displacement Feature 70
5.4 Feature Extraction 70
5.4.1 Matching of Contour Models 70
5.4.2 Exploiting the Echo Pulse Width Information 75
5.4.3 Estimation of Object Displacement in Consecutive Scans 81
5.5 Association 84
5.5.1 Gating 85
5.5.2 Segment to Track Association 85
5.5.3 Shape Hypothesis Selection 85
5.6 Object Generation 85
5.7 Discussion 87

6 Video Data based Feature Extraction 89
6.1 Feature Model and Measurement Equation 91
6.2 Classical Model based Feature Extraction Algorithms 92
6.2.1 Edge Features 93
6.2.2 Matching of a Deformable Synthetic Template 94
6.3 Kalman Filter based Template Tracking 96
6.3.1 Feature Extraction 96
6.3.2 Augmented State Representation 100
6.3.3 Feature Measurement Equation 101
6.4 Discussion 101

7 Parameter Identification 103
7.1 Experimental Platform 103
7.2 Identification of the Measurement Error Probability Distributions 103
7.2.1 Laser Scanner — Azimuth 103

7.2.2 Laser Scanner — Distance . . . 113
7.2.3 Laser Scanner — Position . . . 113
7.2.4 Laser Scanner — Width and Length . . . 114
7.2.5 Laser Scanner — Orientation . . . 117
7.2.6 Laser Scanner — Displacement . . . 121
7.2.7 Camera — Template Tracking . . . 121
7.3 Discussion . . . 124

8 Evaluation 125
8.1 Motion Model Parameters . . . 125
8.2 Center of Gravity Tracking . . . 126
8.3 Performance Measures for Environment Descriptions . . . 127
8.3.1 Error Measures . . . 129
8.3.2 Filter Consistency . . . 129
8.4 Evaluation of the Environment Description for Passenger Cars . . . 131
8.4.1 Experimental Setup . . . 131
8.4.2 Position Estimate . . . 132
8.4.3 Velocity Estimate . . . 138
8.4.4 Velocity Estimation Error for Stationary Roadside Objects . . . 146
8.4.5 Detection Statistics . . . 148
8.5 Evaluation of the Environment Description for Trucks . . . 152
8.5.1 Stationary Truck . . . 152
8.5.2 Moving Truck . . . 155
8.5.3 Detection Rate and Precision . . . 157
8.6 Discussion . . . 159

9 Interacting Multiple Model Filter 161
9.1 Dynamic Models . . . 161
9.2 The Interacting Multiple Model Algorithm . . . 165
9.2.1 Interaction . . . 166
9.2.2 Model Specific Filtering . . . 166
9.2.3 Model Probability Update . . . 167
9.2.4 Combination . . . 167
9.3 Model Set Design . . . 167
9.3.1 Choice of Dynamic Models . . . 169
9.3.2 Choice of Markov Matrix . . . 169
9.4 Evaluation of Traffic Jam Situations . . . 170
9.4.1 Model Set . . . 170
9.4.2 Experimental Setup . . . 171
9.4.3 Results for Strong Acceleration Changes . . . 172
9.4.4 Results for Traffic Jam Situations . . . 174
9.5 Evaluation of Intersection Scenarios . . . 176
9.5.1 Model Set . . . 176
9.5.2 Experimental Setup . . . 177
9.5.3 Results for Strong Acceleration Changes . . . 178
9.5.4 Results for Turning Situations . . . 181

9.5.5 Results for Rapid Steering Maneuvers . . . 183
9.6 Discussion . . . 183

10 Situation Assessment for an Emergency Brake 187
10.1 Kamm's Circle and Resulting Trajectories . . . 187
10.2 Collision Prediction by Kopischke . . . 189
10.3 Situation Assessment . . . 190
10.3.1 Collision Prediction Algorithm . . . 190
10.3.2 Collision Index . . . 191
10.3.3 Matrix of Collision Indices . . . 192
10.3.4 Example of a Rear End Collision . . . 193
10.3.5 Example of a Collision at an Intersection . . . 195
10.3.6 Search for the Minimal Collision Index . . . 195
10.3.7 Stationary Obstacles . . . 195
10.4 Evaluation . . . 197
10.4.1 Rear End Collision . . . 197
10.4.2 Negative Tests . . . 202
10.5 Discussion . . . 204

11 Conclusions 207

A Proofs and Derivations 211
A.1 Existence of Maximally One Solution . . . 211
A.2 Argument under the Root . . . 212
A.3 Calculation of the Sub-pixel Correction Vector . . . 212
A.4 Calculation of the Reference Position . . . 214

Bibliography 217

Supervised Student Research Projects and Diploma Theses 231

Publications and Patent Applications 233

Notation

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
ABS	Antilock Brake System
ACC	Adaptive/Active Cruise Control
ACC S&G	Adaptive/Active Cruise Control Stop and Go
ANIS	Average normalized innovation squared
CA	Collision Avoidance
CA	Constant acceleration
CAD	Computer-aided design
CCD	Charge coupled device
CMOS	Complementary metal oxide semiconductor
COG	Center of gravity
CPU	Central processing unit
CT	Cuboid tracking
CT+TT	Cuboid tracking with template tracking
CV	Constant velocity
DC	Driving corridor
DIN	Deutsche Industrie-Normenausschuss (*engl.* German Industrial Standards Authority)
DOF	Degrees of freedom
DSC	Dynamic Stability Control
DT	Displacement tracking
EB	Emergency Brake
EKF	Extended Kalman filter
EPW	Echo pulse width
ESP	Electronic Stability Program
FMCW	Frequency-modulated continuous wave
FOV	Field of view
FSK	Frequency shift keying
F1	Filter 1 of the IMM
F2	Filter 2 of the IMM
F3	Filter 3 of the IMM
GPS	Global Positioning System
HC	Heading Control
IBA	Intelligent Brake Assistant

ICP	Iterative closest point
IMM	Interacting multiple model
LCA	Lane Change Assistant
LCW	Lane Change Warning
LDW	Lane Departure Warning
Lidar	LIght Detection And Ranging
NEES	Normalized (state) estimation error squared
NIS	Normalized innovation squared
Npp	Normal probability plot
PA	Parking Assistant
Pdf	Probability density function
Pel	Picture element
Radar	RAdio Detection And Ranging
RMS	Root mean square
ROI	Region of interest
S	Stationary
SKF	Single Kalman filter
SPKF	Sigma-point Kalman filter
UKF	Unscented Kalman filter
WICP	Weighted iterative closest point

General Notations

a	Scalar
$\mathbf{a}$	Vector
$\mathbf{A}$	Matrix
$\mathbf{A}'$	Transpose of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	Inverse of matrix $\mathbf{A}$
$\mathbf{I}$	Identity matrix

Latin Symbols

d	Distance
$\bar{d}$	Residual distance of the ICP algorithm
d^2	Mahalanobis distance
d_u	Width of the object in the image
d_v	Height of the object in the image
d_x	Length of the object
d_y	Width of the object
d_z	Height of the object
$\mathbf{f}(.)$	State transition function
$\mathbf{F}$	State transition matrix

$\mathbf{G}$	Input gain
$h(.)$	Measurement function
$\mathbf{h}(.)$	Measurement function vector
$\mathbf{H}_{a \to b}$	Homogeneous transformation matrix for transformation from coordinate system a to b
$\mathbf{H}$	Measurement matrix
k	Time instance
$\mathbf{K}$	Kalman filter gain
L	Track score
L	Likelihood
m	Dimension of the measurement vector
m	Weight of a vehicle
N_s	Number of sensors
o_u	u-coordinate of the estimated offset in the image
o_v	v-coordinate of the estimated offset in the image
p	Probability density
P	Probability
$\mathbf{p}^a$	Point given in coordinate system a
$\mathbf{P}$	Covariance matrix of the state
q	Process noise intensity
$\mathbf{q}$	Correspondence point
Q_{χ^2}	Cumulative probability of a chi-square distribution
$\mathbf{Q}$	Covariance matrix of the process noise
$\mathbf{Q}_d$	Diagonal covariance matrix of the process noise
r	Radial distance
r	Correlation coefficient
$\mathbf{R}$	Rotation matrix
$\mathbf{R}$	Measurement covariance matrix
s	Sample standard deviation
s	Scale factor
$\mathbf{S}$	Covariance matrix of the residual
t	Time
T	Time interval
$\mathbf{t}$	Mounting position of a sensor
$\mathbf{t}$	Translation vector
u	u-coordinate in the image
Δu	u-coordinate of a displacement in the image
$\mathbf{u}$	Input vector
v	v-coordinate in the image
v	Velocity
Δv	v-coordinate of a displacement in the image
v_x	x-component of the velocity vector
v_y	y-component of the velocity vector
$\mathbf{v}$	Process noise

$\mathbf{V}$	Process noise intensity matrix
h_c	Height of the laser scanner calibration object
w	Weighting factor
w_c	Width of the laser scanner calibration object
$\mathbf{w}$	Measurement noise
x	x-coordinate of a position in a cartesian coordinate system
Δx	x-coordinate of the laser scanners displacement measurement
$\mathbf{x}$	True state
$\hat{\mathbf{x}}$	Estimated state
$\tilde{\mathbf{x}}$	State estimation error
y	y-coordinate of a position in a cartesian coordinate system
Δy	y-coordinate of the laser scanners displacement measurement
z	z-coordinate of a position in a cartesian coordinate system
$\mathbf{z}$	Measurement vector

Greek Symbols

α	Azimuth angle
γ	Angle in Kamm's circle defining a driving maneuver
$\mathbf{\Gamma}$	Measurement gain matrix
ϵ	Normalized estimation error squared
$\bar{\epsilon}$	Average normalized estimation error squared
ϵ_ν	Normalized innovation squared
$\bar{\epsilon}_\nu$	Time-averaged normalized innovation squared
θ	Pitch angle
θ	Threshold
μ	Mean value
μ^i	Probability of model i in the IMM filter
$\boldsymbol{\mu}$	Initial model probabilities
$\boldsymbol{\nu}$	Measurement residual
π	State transition probability
$\mathbf{\Pi}$	Markov matrix
σ	Standard deviation
σ^2	Variance
ϕ	Roll angle
ψ	Yaw angle
ω	Yaw rate

1 Introduction

Advanced driver assistance systems are among the most significant innovations of the present automobile industry. The motivation for the research, development and production activities results mainly from increased demand for improved traffic safety, enhanced comfort and driving pleasure.

One of the first deployed advanced driver assistance systems aiming at increased comfort is Adaptive Cruise Control (ACC). In addition to the standard cruise control functionality of keeping the vehicle at a set speed, it adapts the velocity to slower preceding vehicles. Brake assistance systems in contrast aspire to reduce the risk of collisions. These systems detect an imminent collision and react depending on the systems' characteristics. In some cases the brake pads are moved closer to the brake discs in preparation for possible emergency braking, in others the braking force is increased, if the driver reacts too tentatively. The most advanced systems trigger an emergency brake autonomously.

All advanced driver assistance systems base their action on the behavior of other traffic participants, which are detected by onboard environment perceiving sensors. The sensors estimate the position and motion of the detected pedestrians, cars, trucks and other objects using advanced signal processing algorithms. A frequently used technology for vehicle environment perception is radar. However, laser based time-of-flight sensors and cameras are becoming increasingly important due to their higher angular resolution.

Present adaptive cruise control or brake assistance systems are optimized for the detection and tracking of preceding vehicles in the same lane. The limitations of the systems become apparent when vehicles change lane at close range. The most challenging scenarios are crossing or turning vehicles at intersections. The system has to cope with vehicles observed from different perspectives in a wide field of view. The vehicles often rapidly change their speed and driving direction. The diversity and complexity of possible vehicle trajectories increases dramatically compared to the mainly longitudinal oriented traffic on freeways. These challenges impose strong requirements on the sensors and demand novel and complex signal processing algorithms for vehicle detection, tracking and situation assessment.

The first advanced driver assistance systems base their action on the environment perception generated by only one sensor. Future cars will be equipped with several sensors and several assistance systems. Safety critical systems demand an almost infallible environment perception which is difficult to accomplish with only one sensor. It is hypothesized that the combination of several sensors looking in the same direction can increase the reliability and accuracy sufficiently for the high standards demanded by safety critical systems. The best choice of sensors and appropriate methods for the sensor data fusion remains a challenge for actual and future research.

The aim of this work is to enable a robust and precise tracking of vehicles in challenging traffic scenarios. A laser scanner with a wide field of view is utilized for vehicle detection

and tracking. A camera is expected to increase the accuracy of the vehicles' position and velocity estimates, exploiting its high angular resolution. A central sensor data fusion aims at an efficient combination of the sensor information. The resulting environment description is supposed to serve as a common information source for the subsequent advanced driver assistance applications. The aim of this work is an optimization of the whole signal processing chain including the sensor signal processing, the data fusion and in particular the object motion estimation in challenging scenarios. These situations comprise lane change maneuvers of preceding vehicles, intersection scenarios and the tracking of trucks on neighboring lanes. The development of a novel emergency brake algorithm aims at a holistic solution, which is in particular designed for so far unresolved collision situations at intersections.

An up-to-date state of the art for advanced driver assistance systems, the utilized sensors, the sensor signal processing and advanced state estimation methods is presented in Chapter 2. The unresolved challenges which serve as a motivation for the present work are identified.

Chapter 3 presents a novel sensor calibration method for the joint calibration of a laser scanner and camera. The calibration determines the orientation of the sensors mounted on the vehicle. This is necessary in order to geometrically relate the sensor information in the sensor data fusion process.

Different sensor data fusion architectures are discussed in Chapter 4. A feature-level fusion is suggested as an efficient compromise between the two extremes of low-level and high-level data fusion. A Kalman filter based fusion of the sensor measurements is described. The fusion architecture serves as a general framework independent from the actual sensors.

In Chapter 5 a laser scanner module, which is integrated into the feature-level fusion, is described. Novel algorithms for the detection and tracking of cars and trucks are presented. The signal processing takes the dimensions and the orientation of the vehicles into account, which is necessary for an accurate motion estimation especially at intersections. Special care has been taken to handle partial occlusions and sensor specific limitations.

The camera module integrated into the feature-level fusion architecture is presented in Chapter 6. Classical vehicle detection algorithms are discussed and a novel Kalman filter based template tracking approach described. The template tracking is especially promising as part of the feature-level fusion with the laser scanner object tracking.

The position, dimension and orientation estimates provided by the sensors are prone to errors. The measurement error probability distributions are determined in Chapter 7. Novel experimental reference systems were designed to enable a quantitative characterization of the sensors' signal processing performance. The thus derived accuracy measures serve for a parameterization of the Kalman filter.

Chapter 8 evaluates the novel tracking approaches and compares them to standard algorithms described in the literature. The position and velocity estimation accuracy of cars, trucks and stationary roadside obstacles is determined in representative traffic scenarios. Novel experiments were developed which enable a quantitative evaluation based on ground truth data.

The feature-level fusion evaluated in Chapter 8 uses a single Kalman filter for the estimation of object motion. Chapter 9 introduces a multiple model algorithm which

aims at an optimization of the estimates utilizing several Kalman filters. Each Kalman filter is designed for a special driving phase. Especially for highly dynamic scenarios the multiple model approach is expected to deliver significantly improved motion estimates.

A situation independent general solution to the situation assessment of an emergency brake is detailed in Chapter 10. In contrast to previous work, the collision prediction algorithm is not restricted to special traffic scenarios and addresses in particular the accidents at intersections.

Chapter 11 concludes the presented work. It provides a summary of the developed methods and their characteristics. The potentials for further improvement are identified for future research activities.

2 State of the Art and Motivation

The European project PROMETHEUS introduced in the eighties and nineties many ideas and first prototype developments in the area of advanced driver assistance systems [16, 17]. These still influence current research. Since PROMETHEUS, driver assistance systems exploiting environment perception have become reality in most premium class cars.

This chapter provides an overview of advanced driver assistance systems and the sensors necessary for environment perception. It motivates the demand for fusion architectures which combine data from diverse sensors to a consistent environment representation.

2.1 Advanced Driver Assistance Systems

Systems such as the Anti-lock Braking System (ABS) or the Electronic Stability Program (ESP) are well established safety systems which reduce the risk of crashes significantly [100]. Up to 10,000 fatal crashes could be avoided in the U.S. annually if all vehicles were equipped with ESP [70]. The risk of all single-vehicle crashes could be reduced by more than 40 percent — fatal ones by 56 percent. ESP compares the driver's steering and braking inputs to the vehicle's response, using lateral acceleration, yaw rate and wheel speed sensors. ESP then brakes individually front or rear wheels or reduces the engine power to help correct understeer or oversteer. This system therefore utilizes only sensors which observe the host vehicle's state.

The hope of future advanced driver assistance systems lies in the exploitation of sensors which perceive the vehicle's environment. With these sensors the objects surrounding the host vehicle can be detected and tracked. Potentially harmful situations can thus be recognized and countermeasures activated. The aim is an improvement of the road safety.

Apart from the demand for safety systems, the car manufacturers aim at an improvement of comfort in the vehicle. An established comfort system is Adaptive Cruise Control (ACC). Advanced driver assistance systems can therefore be classified into safety and comfort enhancing systems. There is however a smooth transition between comfort and safety as shown exemplarily in Figure 2.1 for some advanced driver assistance systems [45, 112]. Comfort applications, such as ACC, usually imply a safety enhancement which can be more or less pronounced depending on the system's parameterization.

Another possibility for the classification of driver assistance systems is used (with varying terminology) in a number of publications [45, 150, 176]. It differentiates the driver assistance systems by the characteristics of their action (Figure 2.1):

1. Informative, navigation.
2. Low response, automatic operation, guidance.
3. High response, intervention.

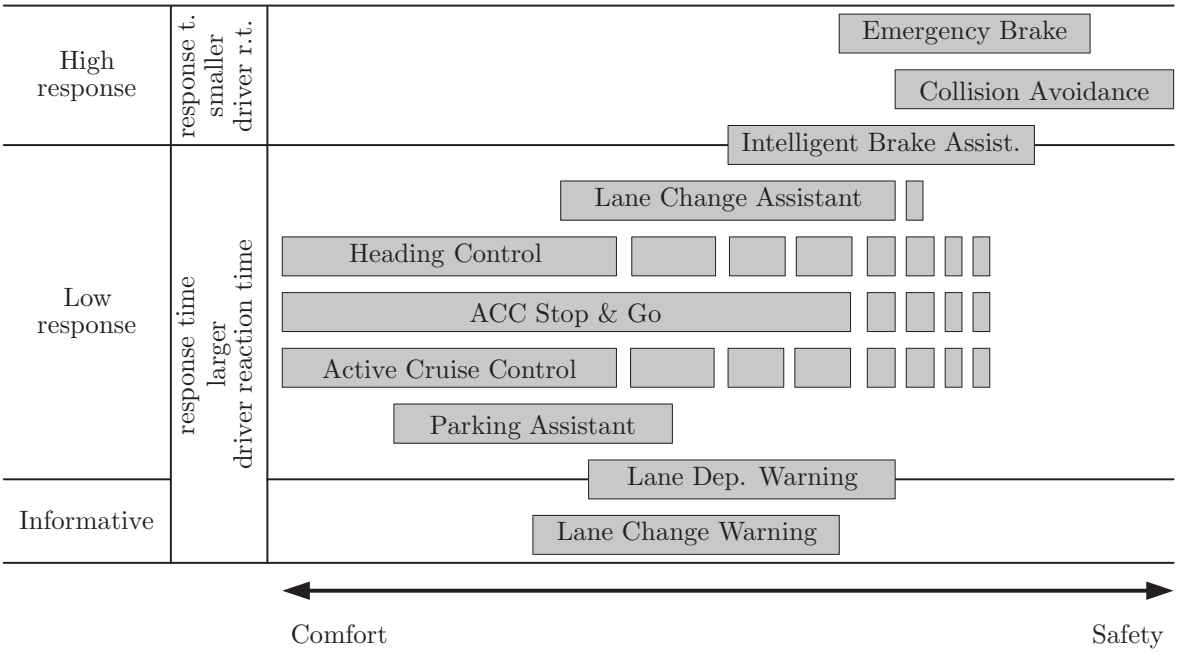


Figure 2.1: Exemplary classification of driver assistance systems [45, 112].

Informative systems, such as Lane Departure Warning (LDW) emit acoustic, haptic or visual signals to the driver in order to assist in the driving task. Low response systems, such as ACC, perform automatic operations and take over partially the guidance of the vehicle. However, these systems work only in certain situations and hand over to the driver, if the driving situation exceeds their limits. High response systems, such as an Emergency Brake (EB), intervene directly and rapidly by controlling the driving dynamics.

These categories of safety and comfort or informative, low and high response describe the type of action of the assistance system. For the development and deployment other aspects of the systems are of interest. A detailed list of criteria was compiled in the European project ADASE2 [40], which serves to assess different advanced driver assistance systems: complexity of system and system algorithms, complexity of sensors, requirements of infrastructure (including communication between vehicles and with infrastructure), requirement of communication between vehicles, complexity of the human machine interface, involvement of legal aspects, involvement of political and societal aspects, degree of driver assistance, degree of safety enhancement. These aspects aim at a holistic view, which serves to identify the needs for all disciplines involved in the development and deployment of advanced driver assistance systems.

For car manufacturers a central question is customer benefit [45, 122]. Advanced driver assistance systems which improve comfort and safety are assigned a high customer benefit as shown in Figure 2.2 [45]. However, market research has shown that there is a low acceptance of systems which intervene autonomously with the guidance of the car and therefore dominate the driver. Systems which are placed in the lower right corner of the diagram in Figure 2.2, such as the ESP (or DSC) system, are therefore more likely to perceive a high customer acceptance.

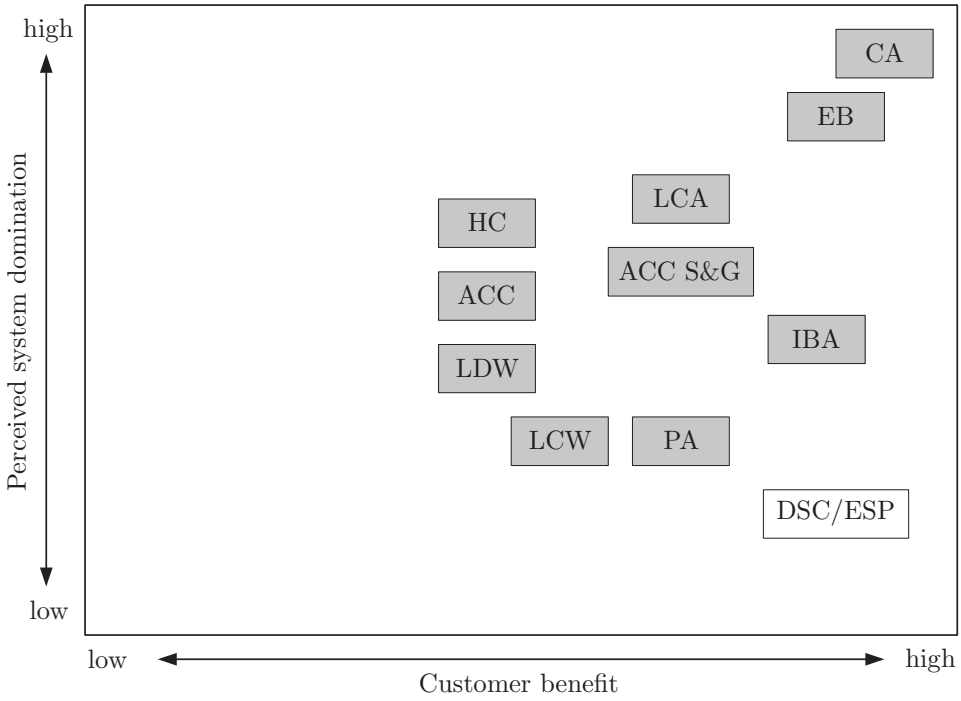


Figure 2.2: Customer benefit versus perceived system domination of advanced driver assistance systems and the ESP (or DSC) system [45].

In summary, there are many aspects characterizing advanced driver assistance systems which have to be considered for the development. However, for the motivation of research many of these aspects are of low relevance. A strong driving force for research comes from the aim of further reducing the number of accidents. In this context, the EU triggered an initiative for a reduction of fatalities on the road by 50 percent between 2000 and 2010 [41].

An application which aims at the reduction of fatalities by a mitigation of collisions is the emergency brake, which is exemplarily addressed in this thesis. An autonomous emergency brake performs a full brake, when an accident becomes unavoidable [96]. Even if this system can not avoid the accident in general, it reduces the energy of the crash impact. This system is therefore referred to as a collision mitigation system. Figure 2.3 depicts the different phases of a crash [45]. From the normal driving phase, a critical situation may arise. Collision avoidance measures are activated in the intervention phase. A collision mitigation system triggers in the pre-crash phase, when the accident is unavoidable and tries to reduce the consequences of the accident.

The 2005 accident statistics on German roads shown in Table 2.1 reveal that the dominating kind of accident is the collision between two vehicles (63 percent). An advanced driver assistance system which supports the driver in these kinds of accident scenarios has a high potential to reduce the number of accidents or at least the number of fatalities and seriously injured passengers. An emergency brake system which brakes if an accident becomes unavoidable is however only effective, if the impact of the crash is mainly caused by a longitudinal motion of the case vehicles. 59 percent of the accidents fall into

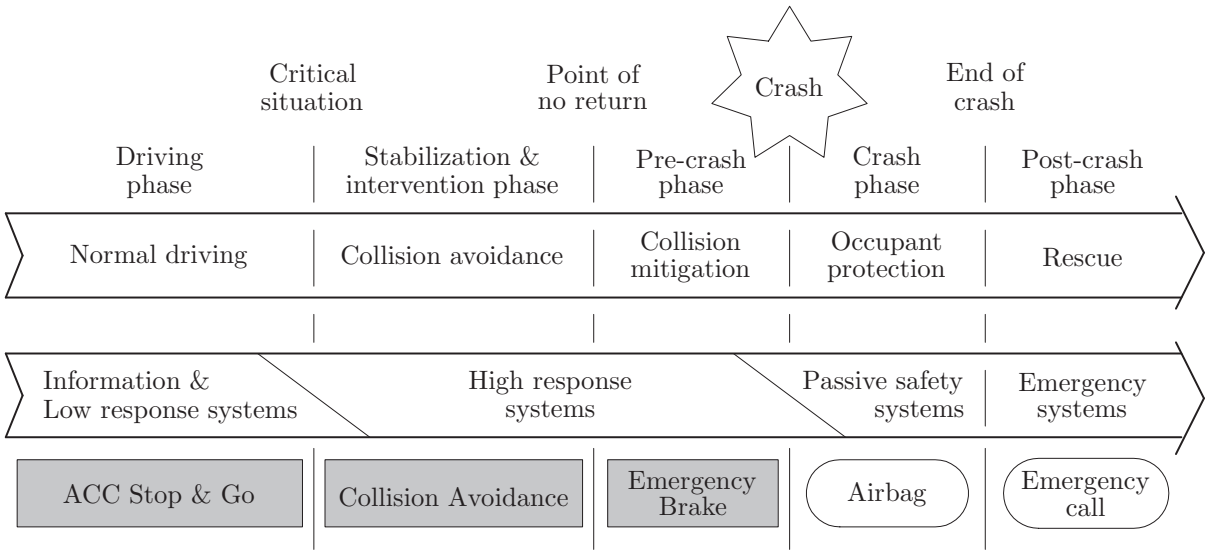


Figure 2.3: The crash sequence with exemplary advanced driver assistance systems [45].

this category. Additionally the system should react on average faster or more adequately than the driver. An analysis of frontal crashes with stationary obstacles exhibited that 21 percent of the drivers did not brake at all prior to the accident. 73 percent decelerate with less than 5 m/s^2 [96]. An emergency brake, which autonomously brakes prior to an accident, therefore has a good potential for mitigating the consequences of a frontal collision.

Kopischke [96] and Lages [99] implemented and tested emergency brake systems with a focus on rear-end collisions with stationary vehicles. They assume that the object vehicle remains stationary. For each recognized object, an analysis is performed. If the object is in the driving path and braking can not avoid the accident, the algorithm tests whether

Table 2.1: Distribution of accidents in Germany in 2005 [134].

Kind of accident	Percentage
Collision with vehicle which turns into or crosses a road	28.3
Collision with vehicle moving ahead or waiting	15.3
Collision with oncoming vehicle	8.3
Collision with vehicle which starts, stops or is stationary	7.1
Collision with vehicle moving laterally in the same direction	4.3
Collision with an obstacle in the carriageway	0.6
Collision between vehicle and pedestrian	9.4
Leaving the carriageway to the right or left	15.4
Accident of another kind	11.3

the collision may be avoided by a maximal steering maneuver. The system triggers an emergency brake if both tests come to the conclusion that the accident is inevitable. In both publications preliminary investigations are presented for crossing vehicles. However, the analyzed scenarios are limited to exactly perpendicular trajectories of the two colliding vehicles.

Jansson et al. calculate in [77] the probability of a collision for an emergency brake system. In this approach the joint probability density function of the host and the object vehicles' future positions is determined. If the probability that the vehicles physically overlap is one, the emergency brake is triggered. The probability density functions are obtained by a Kalman filter. This approach therefore takes both the sensor measurement uncertainty and the vehicle motion uncertainty into account. In [76] different risk metrics are discussed for the same system. In [91] the improvement by a particle filter approach is compared to the standard extended Kalman filter estimation.

Broadhurst et al. presented a prediction and planning framework for analyzing the safety and interaction of moving objects in complex road scenes [18, 19, 20]. All possible object motions and interactions are simulated and their danger assessed. Unlike previous approaches, the work considers simultaneously the path of several vehicles and does not restrict its analysis to critical object configurations. This framework has its advantages in situations in which the left and right collision avoidance paths are blocked by other vehicles or obstacles. However, the algorithm aims at long term prediction and planning over several seconds and not at short term emergency brake situations. The probability of driving maneuvers are modeled using Gaussian distributions. The collision risk is estimated using Monte Carlo simulations. The framework aims not explicitly at the recognition of certainly unavoidable accidents. This is however necessary for the emergency brake as defined in [96].

The above presented situation assessment algorithms are restricted to certain traffic scenarios or produce only collision risk probabilities. The problem of collision risk probabilities is their accuracy. Using Monte Carlo simulations, the accuracy depends on the number of samples. For the very high accuracy, necessary to prevent false engagements of emergency brake systems, it is hypothesized that the number of samples and therefore the computational burden becomes unrealistic.

A new approach for the calculation of the trigger time of an emergency brake will be presented in this work. The algorithm considers all physically possible motions of the object and host vehicle simultaneously. It can be applied to all different scenarios including rear-end collisions, collisions at intersections or with oncoming vehicles. Thus, 59 percent of possible accidents are addressed. The approach accounts for the object and host vehicles' dimensions. Unlike previous work, the orientation of the vehicles is incorporated into the collision estimation.

2.2 Sensors

Future advanced driver assistance systems require a detailed and accurate environment description in a wide field of view. The necessary environment perception poses high requirements on the sensors.

Established technologies for vehicle environment perception are radar (RAdio Detection

And Ranging), lidar (LIght Detection And Ranging), cameras, ultrasonic and digital maps together with GPS [171]. A detailed description and assessment of the different sensors would be beyond the scope of this thesis. A short summary of the common sensors is provided below. Further details are given in recently compiled reviews [142, 141, 171, 112].

Radar sensors measure the distance with the time of flight principle and the radial velocity with the Doppler principle. They usually deliver additionally an estimate of the angle to the object. The sensors can roughly be categorized by the frequency (24 GHz, 77 GHz, 79 GHz), the waveform (pulse, Frequency-Modulated Continuous Wave (FMCW), Frequency Shift Keying (FSK)) and the antenna (fixed (multi)-beam, mechanically scanning, electronically scanning).

Most lidar sensors measure the distance with the time of flight principle and the azimuth angle with several beams (multi-beam lidar) or a mechanically scanning device (laser scanner). There are lidar sensors which measure in only one elevation and multi-layer systems. The wavelength of the emitted laser pulse is either 900 nm (silicon) or 1500 nm (gallium-arsenic).

Cameras can be classified into visible light, near infrared light, far infrared light and 3D sensors. Visible light and near infrared sensitive cameras are based on CCD or CMOS technologies. Near infrared cameras are sensitive in the spectrum of 800 nm – 1000 nm. In contrast, far infrared sensors are sensitive to heat radiation in the spectrum of 8 µm – 12 µm. They can be differentiated into cooled systems exploiting photonic effects using narrow gap semiconductors and uncooled microbolometer based cameras, which exploit a change in resistance caused by thermal effects.

3D range cameras perform a time of flight range measurement simultaneously in different directions in azimuth and elevation. 3D cameras usually incorporate an additional visible light sensor.

Digital maps in combination with GPS provide information about the infrastructure in the environment of the vehicle. While the accuracy, the extent and the coverage of the information is rather limited today, they can be expected to improve in the near future.

For the selection of sensors or sensor combinations for specific advanced driver assistance systems, the sensors are usually contrasted with respect to different criteria [171, 112, 141]. There are purely technical aspects, such as the field of view, resolution, accuracy, update rate or measurement principle conditioned limits (sensitivity to bad weather, dependence on daylight). For serial development in a car, the dimensions, the price of the sensor and legal aspects are of additional importance.

In the present work a laser scanner and a visible light grayscale camera are used for vehicle environment perception. The laser scanner has a wide field of view which enables vehicle detection and tracking in many situations and in particular at intersections. A camera is expected to increase the accuracy of the vehicles' position and velocity estimates, exploiting its high angular resolution. The following sections provide a thorough overview of the research related to the signal processing of laser scanner and video data in the area of object tracking for driver assistance systems.

2.2.1 Laser Scanner Data based Environment Perception

The research into vehicle environment perception for driver assistance systems based on laser scanner data can be divided into:

- Road course and lane detection [94, 93, 133, 33, 151, 152, 38].
- Host vehicle localization and motion estimation [156, 158, 157, 162, 163, 161].
- Object classification [46, 169, 168, 167].
- Pedestrian recognition and safety [50, 49, 48, 47].
- Object tracking (detailed below).

As this work concentrates on the development of tracking methods, the advantages and disadvantages of laser scanner based object tracking already described in the literature will be presented in the following.

The first real-time object tracking approaches represent other vehicles as points [94, 93]. However, especially in the near range, there are several laser scanner measurements associated to a vehicle, as a vehicle is an extended object. The position of the object is therefore calculated from the mean position of all raw laser scanner measurements associated to the object. On freeways and rural roads this mean position usually represents the center of the rear of a preceding vehicle or the center of the front for oncoming cars. As a tracking applied to these positions delivers good estimates this approach is still utilized in recent publications [111].

However, researchers realized that in urban areas and particularly at intersections the point model assumption for a vehicle does not hold. This can be best explained for a vehicle crossing an intersection in front of the host vehicle. The mean position of all raw measurements associated to this vehicle, moves from the front over the side to the rear. A velocity estimate calculated from these positions delivers biased values. The motion of the position along the object's contour induces an additional virtual velocity.

Other approaches therefore account for the dimensions and the orientation of the object. A vehicle is here modeled by a rectangle in the birds eye view. This, however, introduces a new challenge. As the laser scanner observes either one side or two sides of a vehicle, the algorithm must determine the association between the measurements and the different sides. In [175] the association is solved by a Hough transformation, which generates line segments. Line segments which are close together and perpendicular are then grouped to build vehicle hypotheses.

Another widely exploited approach first builds segments from close raw measurements. It then determines the first and the last point in angular direction and the closest point to the laser scanner. If the laser scanner observes two sides of a vehicle, these three points represent the three visible corners of the vehicle [143, 46]. However, as this approach relies on only three points, it is sensitive to measurement outliers.

An extended method uses the closest point only for a separation of the point set associated to the vehicle. It then calculates a regression from the two resulting point sets [140, 139, 138]. However, vehicles often deviate significantly from the rectangular model, as the corners are often rounded. It becomes difficult to differentiate, if one or two sides of the vehicle were observed. Additionally, the determined sides are in general not perpendicular. As this method relies only on one point for the separation, it is very sensitive to measurement outliers.

Another possibility for the separation of the point set, which overcomes the previously stated problem, exploits the velocity vector of the observed vehicle [46]. As a vehicle

usually moves parallel to its longitudinal axis, the object sides parallel and perpendicular to the motion vector are determined. This approach however works only for significantly moving objects, which have been previously tracked and depends on the accuracy of the estimated motion vector.

In order to palliate for the remaining deficiencies of the above mentioned methods, a new approach is taken in this work. It uses predefined contour models, consisting of one, two or three line segments and matches the complete contour shape with the point set. One or two observed sides are represented either by an I-shape or L-shape. A U-shape has its advantages for the representation of the rear or front of vehicles with round corners.

Apart from the rectangular form, there are objects contours which are irregular. This prohibits the determination of a reference point with respect to a rigid generic model. The same problem arises for large vehicles, such as trucks, which are only partly in the field of view of the laser scanner. There are situations in which no corner of the truck can be observed. A position estimate is then impossible. Object tracking can be enabled by a displacement estimation of the measurement point set. This is similar to the calculation of optical flow in image processing. A new algorithm is devised, which addresses the challenging tracking of trucks which are passed by the host vehicle on freeways.

Another advance in laser scanner based object tracking described in this work exploits the reflectivity measurement provided by the laser scanner. As shown by Fürstenberg et al. in [48, 46] laser scanner measurements of reflecting surfaces can be identified. Measurements of the rear reflectors of vehicles or number plates can thus be detected. This was found to be useful in the present work for the classification of the vehicles' rears for an unambiguous orientation determination of stationary vehicles and additionally an improvement of the orientation estimation.

2.2.2 Video Data based Environment Perception

There is a long history of camera based applications in the research area of driver assistance systems. First real-time implementations date back to the late eighties and early nineties. Monocular image data sequences are used for vehicle, pedestrian and general obstacle detection, lane recognition, traffic sign recognition, ego-motion estimation and localization. In the context of this work vehicle detection and tracking algorithms are relevant.

For driver assistance applications, vehicle detection algorithms need to process the image content in real-time. The algorithms are therefore divided into several steps, which sequentially reduce the amount of information and (optimally) extract the most discriminating information in the first steps. The first steps use computationally cheap classification schemes, while the last step is computationally the most expensive. This coarse to fine strategy is in general referred to as cascade classification. The majority of older publications use a two step vehicle detection. In the hypothesis generation step potential vehicle locations are identified, while the subsequent hypothesis verification step tests the presence with more sophisticated methods. In [144] a recent review is given for different hypothesis generation and verification algorithms. Other recent publications are [136, 67, 55, 127, 65].

Hypothesis generation methods can be classified into appearance based and motion based methods. The utilized appearance based methods employ expert knowledge about

the appearance of vehicles in the image. The approaches focus mainly on symmetry features, color cues, shadows, corners, image edges, local entropy and bright spots. Significant early publications are [123, 36, 11, 35, 153, 21] and the references within.

Motion based hypothesis generation methods exploit optical flow [8, 135] for a detection of potential vehicle locations. Stationary and moving objects can be identified, as their corresponding optical flow fields diverge from the expected optical flow field induced by the host vehicle's motion [129, 3, 44, 81].

The hypothesis verification algorithms can be classified into knowledge based methods and statistical classification methods [153, 55, 54]. Knowledge based approaches determine the correlation between predefined structures of features with the feature occurrences extracted from the image. Synthetic template matching methods are a dominant representative of the knowledge based algorithm class [144]. In contrast, modern statistical methods learn the structure and parameters of the classifier from labeled data.

A holistic view regards the vehicle detection as a single optimization problem. In terms of coarse to fine classification, the most discriminating simple features must be found in the early steps and discriminating complex feature combinations in later steps. The optimization implies therefore the structure and the parameters of the classifier. Recently, efficient cascade classifiers and powerful training algorithms were developed which have been successfully applied to vehicle detection [166, 105].

Appearance based algorithms for vehicle localization were developed in this work. The algorithms are based on the above mentioned edge and shadow features and produce precise localization estimates for vehicles and trucks. However, they imply many rather restrictive assumptions about the object's appearance, position and orientation. So as to overcome these deficiencies of the utilized appearance based algorithms and to support the fusion system aimed at in this work, a model free template tracking approach was devised. It only assumes the object to be rigid. This approach however does not perform vehicle detection. It relies on the vehicle detection by the laser scanner and can only be used as part of a fusion system. In the fusion framework the template tracking serves to improve the accuracy of the lateral position and velocity estimates. The innovation of this approach lies in the feature-level fusion of template tracking in the image domain and laser scanner data based tracking.

The counterpart to the present work aims at an improvement of the classification of vehicles. In [166] laser scanner data based classification of vehicles is improved by the fusion of laser scanner data and image features as input to a common vehicle detector.

2.3 Sensor Data Fusion

It has been widely recognized, that several sensors are necessary to meet the requirements of advanced driver assistance systems [154, 9, 171, 113, 37, 34]. Therefore, sensor data fusion techniques must be applied for the generation of a consistent vehicle environment description. First systems employing sensor data fusion, such as the multiple-application system in the current S-Class Mercedes which is based on a long range radar and six short range radars [118], have become recently available on the market.

The aim of the sensor data fusion is:

- Improved state estimation accuracy.

- Increased object classification accuracy (Higher detection rate, fewer false alarms).
- Improved robustness for instance in bad weather conditions.
- Increased availability.
- Enlarged field of view.
- Enhanced level of detail of object description.

The short and long range radar fusion system of Mercedes described in [118] aims mainly at an enlargement of the field of view and an improvement of the object detection accuracy. The same applies for the long range radar and mid range laser scanner approach described in [111].

Other frequent fusion approaches combine data from a long range radar and a camera sensor [92, 137, 59, 55, 129] or a far range lidar and a camera sensor [128, 145, 146, 147, 105]. These systems usually aim at an improvement of the state estimation and object classification accuracy in the overlapping field of view. The wider field of view of the camera can be used for a detection of objects left and right of the host vehicle in the near range. This can be especially useful for an early detection of cut-in situations of preceding vehicles.

A classical sensor data fusion system can in general be decomposed into several modules: spatiotemporal alignment, data association, state estimation and identity classification. These modules are recursively executed when new measurements become available from the sensors.

The first module performs a spatiotemporal alignment of the sensor and object state estimates of the fusion system. The temporal alignment accounts for the different measurement times and synchronizes these with the fusion process' intrinsic time [6, 13]. In contrast, the spatial alignment accounts for the different sensor mounting positions and orientations. Its responsibility lies in the transformation of the data into a common coordinate system. As it is difficult to adjust the sensors with high precision, calibration procedures are utilized which estimate the parameters. A good overview of calibration methods in the automotive area is provided by [117]. However, joint laser scanner and video camera calibration methods are rare in the literature. The novel calibration approach, suggested in this work, estimates the orientation of a laser scanner and of a camera in a common calibration field. The related work will be discussed in detail at the beginning of Chapter 3.

The data association assigns the sensor data to already existing object tracks. Unassigned sensor data generates new objects. Different data association algorithms are described in [6, 13].

The state estimation fuses the information from the different sensors concerning position, velocity, dimensions, orientation and other object model parameters with the already accumulated information. A detailed overview of the research in different state estimation techniques will be given in Chapter 2.4.

The classification fuses the information from the different sensors concerning the objects' existence. A multi-class classification determines additionally the object class and differentiates for instances between cars, trucks, bikes and pedestrians. An overview of

possible classification algorithms are given in [61, 95]. Examples of a feature-level classification of fused video and lidar data for vehicle detection are given in [146, 166].

Recent research overcomes the separation of the data association, state estimation and identity classification. The algorithms are combined to a single optimization problem [62]. A first application of these approaches in the context of driver assistances system can be found in [104, 106].

A present research theme is the sensor data fusion level. Information theoretically a low-level fusion of the data is best. The full information content is then available for the central target tracking and classification algorithms. High-level fusion approaches in contrast combine target tracks and identity decisions of the different sensors. Thus the information content of the sensor data is already reduced at fusion level. The advantage of high-level approaches is that they are effortlessly scalable which promotes the reusability of system components. It is however an open question to which extent the information loss of high-level fusion systems is significant as compared to low-level fusion approaches. In [116] several recent publications can be found exploiting different levels of fusion in the context of advanced driver assistance systems.

In order to try to overcome the above stated deficiencies of low- and high-level fusion systems, a feature-level fusion approach is suggested here. The fusion is performed at an intermediate level, thus preserving a high amount of information for the central state estimation and classification. Additionally, the system architecture is designed to account for a good scalability and reusability of the system components.

2.4 State Estimation

The tracked objects relevant to driver assistance systems are traffic participants, such as cars, trucks, (motor-)bikes or pedestrians. These objects can be described by their position, velocity, dimensions and orientation. The aim of state estimation is the optimal estimation of the hidden variables describing an object recursively over time given noisy or incomplete observations. In the following an overview of different discrete-time estimation approaches and their underlying assumptions will be presented. Overviews of estimation methods can be found in [7, 4, 57, 165, 6, 13, 143].

The general optimal solution of the recursive estimation problem is given by the recursive Bayesian estimation. The exact probability density function representing all current probabilistic knowledge about the target state is propagated recursively over time. This approach is in general however not feasible for real-time systems, as there is no general analytical solution and the numerical solutions of the integral equations are computational too expensive.

The statistical estimation theory was therefore strongly advanced, when an optimal solution of the estimation problem was found with the development of the well known Kalman filter [89]. However, the Kalman filter applies only in a special case and the implied assumptions are rather restrictive. The optimal solution is only guaranteed for linear systems in which the noise perturbing the system and the observations is zero-mean, additive, white, mutually uncorrelated and has a Gaussian distribution.

As dynamic systems in real world applications are often nonlinear, approximate solutions were developed including the well known extended Kalman filter (EKF) [4, 57, 165,

13]. The EKF linearizes the nonlinear functions in each cycle and applies then the standard Kalman filter. The extended Kalman filter is the standard approach utilized in the automotive literature for target tracking with range sensors. The linearization implied in the EKF however leads to strong errors in highly nonlinear systems.

More recently a significant improvement was gained by the unscented Kalman filter (UKF) [80, 79]. The Gaussian probability density function is here approximated by sigma-points which are transformed by the nonlinear functions. The resulting state estimate and its covariance are finally calculated from the transformed sigma-points. This idea was further developed to a group of algorithms called sigma-point Kalman filters (SPKF) [155, 148, 149].

These recently invented Kalman filter based methods provide good estimates for many nonlinear problems. However, the measurement and process noise must again have a Gaussian distribution.

Sequential Monte Carlo methods, also known as particle filters, are powerful estimation techniques which apply for nonlinear systems and overcome the Gaussian assumption of Kalman filter based approaches [72, 73, 71]. The probability density function of the state estimate is here discretized and represented by random samples. The samples are propagated over time by Monte Carlo simulation. With sequential Monte Carlo methods it is possible to account for any arbitrary probability density distribution, including multi-modal or skewed functions. For an increasing number of samples, this approach converges against the optimal recursive Bayesian estimation. In the automotive area, sequential Monte Carlo methods are popular for target tracking in the image domain as the probability density function of the measurements is here often multi-modal [110, 127].

The Bayesian recursive estimation is a model based approach. A Markovian state space model is used for the propagation of the state probability density over time. Usually, there is one dynamic model describing the transition, such as a second order polynomial model, assuming the acceleration to be almost constant over time and only perturbed by random noise.

However, in real world applications an object's motion can often be better described by several models. The different models apply at different times, depending on the characteristics of the current motion. Multiple model approaches account for different system behaviors. They use a bank of filters based on a set of models that represent possible motion patterns [109, 102, 5]. The development of the cost-effective Interacting Multiple Model (IMM) [14] made the multiple model approaches practical for maneuvering target tracking. A first application of IMM for vehicle tracking in the context of driver assistance systems is described in [175]. Recently IMM algorithms have become rather popular as an alternative to single model Kalman filter approaches [28, 2, 114, 27, 26, 111, 84, 85, 159, 68, 66].

An often stated problem of IMM approaches is the parameterization of the model set. This includes model selection, parameterization of the model and the determination of a model transition probability matrix. In all publications in the automotive area known to the author, the model set design is performed rather arbitrarily, in particular that of the choice of the transition probability matrix.

In contrast to previous work, the parameterization of the IMM model set suggested in this work is performed statistically based on real data emanating from traffic jam scenarios. Evaluations were performed using ground truth data from challenging traffic

scenarios on freeways. In a second approach, novel motion model sets for intersection scenarios were devised. The aim is an improved estimation accuracy and consistency for highly dynamic scenarios at intersections which is of interest in particular for safety applications.

2.5 Summary

The state of the art in environment perception for advanced driver assistance systems has been reviewed and the so far unresolved challenges addressed in this work have been identified. The aim of this work is to enable a robust and precise tracking of vehicles in challenging traffic scenarios.

The sensors utilized in this work are a laser scanner and a grayscale camera. The laser scanner, with its wide angular field of view, detects the objects and the image processing of the camera aims at an improvement of the state estimates, by exploiting its higher angular resolution. The signal processing of the laser scanner estimates, apart from the position of the observed object, its width, length, height and orientation. The novel algorithms aim at an optimization of the object tracking at intersections and the tracking of trucks in challenging situations.

The optimal sensor data fusion architecture is still an open research issue and depends in general on the utilized sensors and desired applications. A feature-level fusion architecture will be suggested which tries to overcome the deficiencies of purely low- and high-level fusion systems.

The aim of the present work is to ensure an optimized object state estimation in challenging situations, such as intersection scenarios, lane change maneuvers of observed vehicles, traffic jams or emergency maneuvers. The optimization includes the laser scanner signal and image processing, the fusion architecture and finally the filter utilized for the state estimation process. Advanced multiple model approaches will be considered in this work as an alternative to the standard Kalman filter.

As an exemplary advanced driver assistance system an emergency brake application will be addressed. The situation assessment algorithms for an automatic emergency brake known from the literature are restricted to a defined set of scenarios. A novel situation assessment algorithm will be presented which overcomes this constraint and aims at a holistic solution applicable also to intersection scenarios.

3 Sensor Calibration

In order to geometrically relate the data from environment sensors it is important to know their relative position and orientation. This enables the transformation of laser scanner measurements into the image domain. Small image regions can therefore be associated with distance measurements, which is an advantage for subsequent algorithms over purely monocular image processing approaches. In automotive applications the sensors' pose must additionally be known with respect to a common vehicle coordinate system. Objects in the environment, which are detected by the sensors, can thus be related to the host vehicle.

Therefore a new method has been developed, which automatically estimates the orientation of the laser scanner and the camera with respect to the common vehicle coordinate system [22, 90]. The mounting position of the sensors is assumed to be known and can in general be measured by hand or extracted from the vehicle's CAD data provided by the car manufacturer. The design of the calibration method supports both a stand alone laser scanner calibration, which is necessary for the test-vehicles equipped with only a laser scanner, and the joint calibration of laser scanner and camera, where the precision of the relative orientation is of importance.

The calibration of the camera can be decomposed into internal and external calibration parameters. The external calibration parameters are the position and the orientation of the sensor, relative to a reference coordinate system. The internal parameters describe the projection of the 3D environment to the 2D image coordinate system, using the pinhole camera model. This work assumes the internal calibration to be known from a separate calibration process detailed in [15] and focuses on the estimation of the external parameters. The sensors are calibrated in a single calibration procedure using a joint calibration field. For the estimation of the yaw angle, which is the most relevant orientation parameter, a single calibration object viewed simultaneously by both sensors was designed.

A good overview of calibration methods in the automotive area provides [117]. Many patent applications address the calibration of only a laser scanner [98, 170, 164]. However, they focus mainly on the estimation of the yaw angle or aim at an automatic adjustment without estimating the magnitude of the sensor orientation [56]. A very general patent application addresses the joint calibration of several environment sensors without specifying the applied methods [43]. Some publications utilize a calibration setup which makes use of the visibility of the laser light to the camera [78, 74, 105]. The novel approach, suggested in this work, is independent of this necessity. Only recently, in October 2004, Zhang and Pless published a similar calibration algorithm for a joint laser scanner and camera calibration using a planar calibration pattern viewed simultaneously by both sensors [172]. Their method estimates additionally to the orientation, the position of both sensors with respect to a common reference coordinate system.

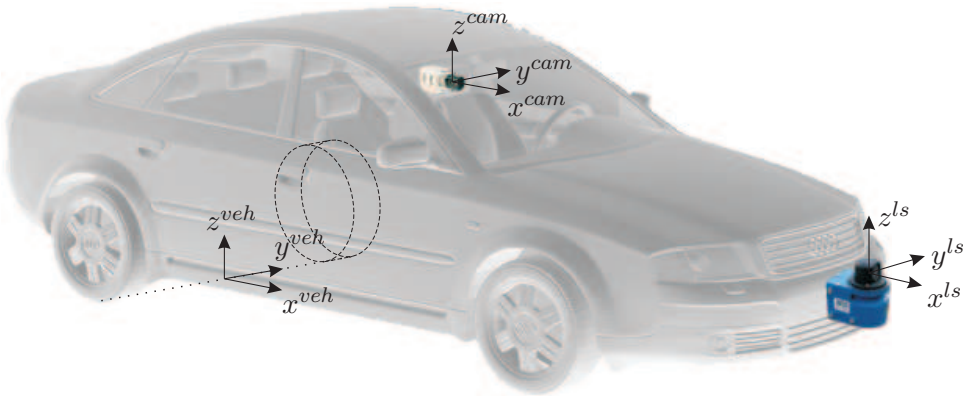


Figure 3.1: Sensor and vehicle coordinate systems.

3.1 Sensor and Vehicle Coordinate Systems

The sensor and vehicle coordinate systems shown in Figure 3.1 are defined according to the DIN 70,000 [39]. The origin of the vehicle coordinate system lies underneath the center of the rear axis on the road surface and is fixed to the car body[1].

A point given in a sensor coordinate system $\mathbf{p}^{sensor} = (x^{sensor}, y^{sensor}, z^{sensor}, 1)'$ is transformed into the vehicle coordinate system $\mathbf{p}^{veh} = \left(x^{veh}, y^{veh}, z^{veh}, 1\right)'$ using the homogeneous transformation matrix $\mathbf{H}_{sensor \to veh} \in \mathbb{R}^{4\times4}$

$$\mathbf{p}^{veh} = \mathbf{H}_{sensor \to veh} \mathbf{p}^{sensor} \tag{3.1}$$

The matrix $\mathbf{H}_{sensor \to veh}$ is defined as

$$\mathbf{H}_{sensor \to veh} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tag{3.2}$$

where $\mathbf{t} = (t_x, t_y, t_z)'$ is the mounting position of the sensor with respect to the vehicle coordinate system and $\mathbf{R} \in \mathbb{R}^{3\times3}$ is the rotation matrix defined as

$$\mathbf{R} = \mathbf{R}_\psi \mathbf{R}_\theta \mathbf{R}_\phi \tag{3.3}$$

The roll angle, which is the rotation around the x-axis of a sensor coordinate system, is denoted by ϕ and the corresponding rotation matrix $\mathbf{R}_\phi$ is given by

$$\mathbf{R}_\phi = \mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \tag{3.4}$$

The pitch angle θ defines the rotation around the y-axis

$$\mathbf{R}_\theta = \mathbf{R}_y = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \tag{3.5}$$

[1]For the calibration, the xy-plane of vehicle coordinate system is assumed to coincide with the road surface. However, in driving conditions the orientation and distance of the vehicle coordinate system with respect to the road surface varies and has to be estimated online if necessary.

whereas the yaw angle ψ denotes the rotation around the z-axis

$$\mathbf{R}_{\psi} = \mathbf{R}_z = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.6}$$

The transformation $\mathbf{H}_{sensor \rightarrow veh}$ of a point from a sensor coordinate system into the vehicle coordinate system given in Equation (3.2) therefore implies 6 degrees of freedom (DOF), which are the three angles and the three translation parameters. The transformation $\mathbf{H}_{ls \rightarrow cam}$ of a point given in the laser scanner coordinate system $\mathbf{p}^{ls}$ into the camera coordinate system $\mathbf{p}^{cam}$ which is given by

$$\mathbf{p}^{cam} = \mathbf{H}_{ls \rightarrow cam}\mathbf{p}^{ls} = \mathbf{H}^{-1}_{cam \rightarrow veh}\mathbf{H}_{ls \rightarrow veh}\mathbf{p}^{ls} \tag{3.7}$$

therefore implies 12 DOF, 6 DOF for each sensor. However, the mounting position of both sensors ($\mathbf{t}_{ls}$ and $\mathbf{t}_{cam}$) are assumed to be known in the following calibration and can in general be measured by hand or extracted from the vehicle's CAD data provided by the car manufacturer. The calibration is therefore reduced to the estimation of the 6 angles defining the orientation of both sensors with respect to the common vehicle coordinate system.

3.2 Calibration of a Multi-Layer Laser Scanner

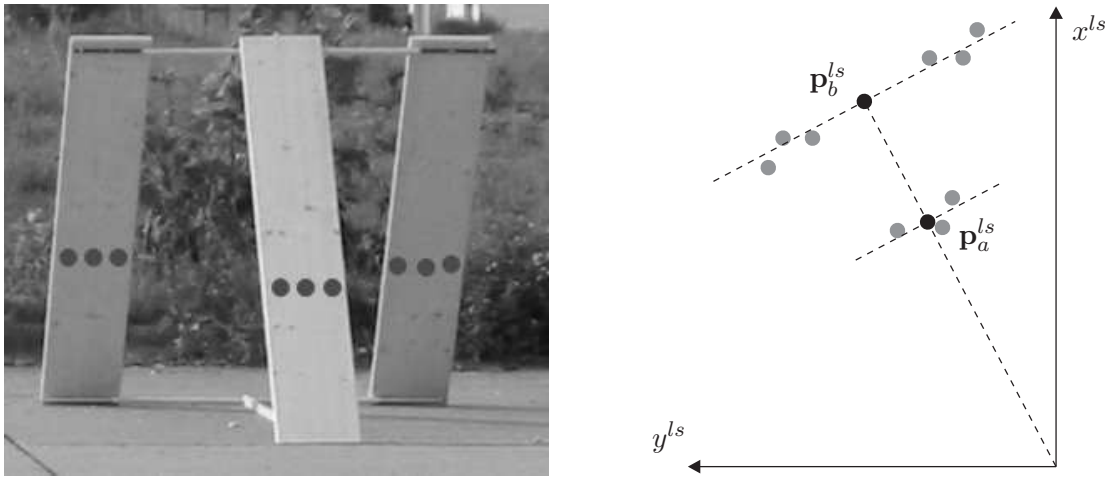
The calibration of the laser scanner comprises as a first step the estimation of the pitch θ and roll ϕ angle. In the second part the yaw angle ψ is estimated using the already determined pitch and roll angle. The calibration procedure is described for a multi-layer laser scanner which includes the special case of a single-layer laser scanner. The single-layer laser scanner measures in the xy-plane, whereas the multi-layer laser scanner produces data in 3d.

3.2.1 Estimation of Pitch and Roll Angle

The basic idea of the calibration is the exploitation of the accurate distance measurements of the laser scanner in terms of the repeat accuracy. Even if the error in the absolute distance measurement is in the order of 0.1 m the repeat accuracy is only prone to noise of a magnitude of $\sigma = \pm 0.05$ m [69]. Consider the calibration object shown in Figure 3.2. The inclination of the ray through $\mathbf{p}^{ls}_a$ and $\mathbf{p}^{ls}_b$ with respect to the ground plane depends directly on the distance between the two points $\mathbf{p}^{ls}_a$ and $\mathbf{p}^{ls}_b$. Using two of the proposed calibration objects as shown in Figure 3.3, the pitch and roll angle of the laser scanner can be deduced.

For the estimation of the pitch and roll angle a virtual coordinate system $\mathcal{C}_{ls^*}$ is introduced of which the xy-plane $\mathcal{L}^*_{xy}$ is parallel to the xy-plane $\mathcal{V}_{xy}$ of the vehicle coordinate system $\mathcal{C}_{veh}$ (Figure 3.4). The origin of the virtual coordinate system $\mathcal{C}_{ls^*}$ coincides with the laser scanner coordinate system $\mathcal{C}_{ls}$. A point given in the laser scanner coordinate system $\mathbf{p}^{ls}$ is transformed into the virtual coordinate system with the rotation matrix $\mathbf{R}_{ls \rightarrow ls^*}(\theta_{ls}, \phi_{ls}) \in \mathbb{R}^{3\times 3}$

$$\mathbf{p}^{ls^*} = \mathbf{R}_{ls \rightarrow ls^*}(\theta_{ls}, \phi_{ls})\,\mathbf{p}^{ls} \tag{3.8}$$



(a) (b)

Figure 3.2: (a) Calibration object. With a minimum of two such objects, the roll and pitch angle of the laser scanner can be estimated. The laser scanner measurements are painted into the image for visualization purposes. (b) shows the result of the pre-processing. The points $\mathbf{p}_a^{ls}$ and $\mathbf{p}_b^{ls}$ are deduced from constrained regressions. The distance between the two points is small, if the measurements are from the top of the calibration object and becomes larger for measurements from the bottom of the calibration object. The angle of the measurement plane depends therefore directly on the distance difference between the two points.

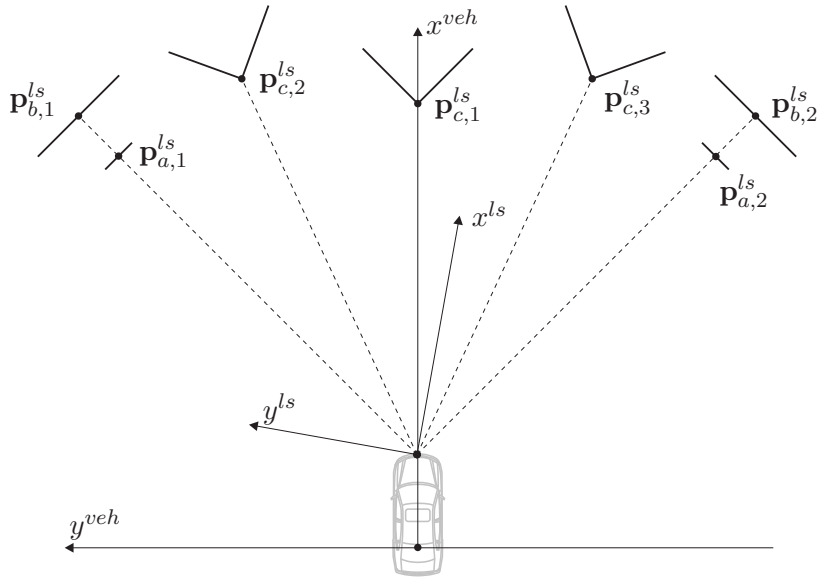


Figure 3.3: Calibration setup. The points $\mathbf{p}_{a,1}^{ls}$, $\mathbf{p}_{b,1}^{ls}$, $\mathbf{p}_{a,2}^{ls}$ and $\mathbf{p}_{b,2}^{ls}$ are used for the calibration of the laser scanner's pitch θ and roll ϕ. The yaw angle ψ of the laser scanner is calculated from point $\mathbf{p}_{c,1}^{ls}$. The camera is calibrated using the points $\mathbf{p}_{c,1}^{ls}$, $\mathbf{p}_{c,2}^{ls}$ and $\mathbf{p}_{c,3}^{ls}$.

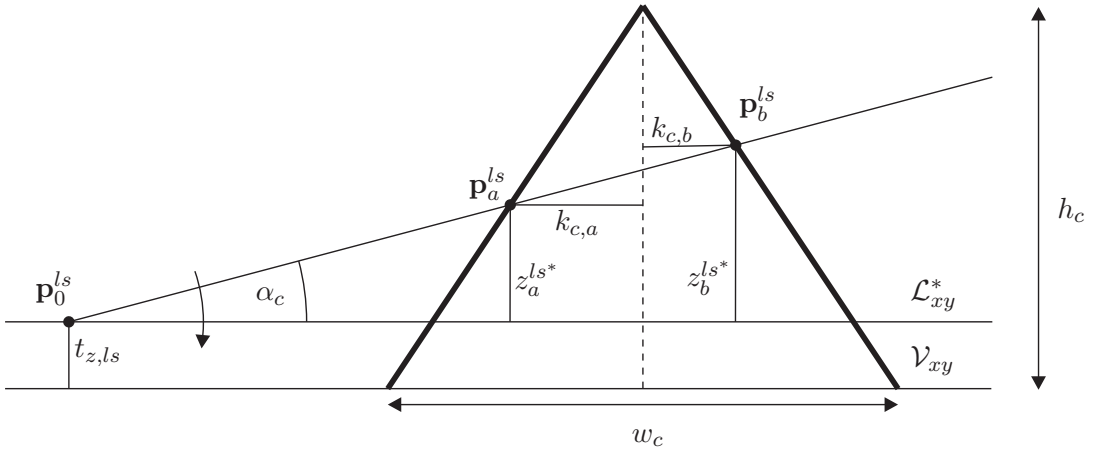


Figure 3.4: Calibration object for the laser scanner. The optical center of the laser scanner is denoted by $\mathbf{p}_0^{ls}$. The angle α_c depends mainly on the distance between the two points $\mathbf{p}_a^{ls}$ and $\mathbf{p}_b^{ls}$.

The rotation matrix $\mathbf{R}_{ls\rightarrow ls^*}(\theta_{ls}, \phi_{ls})$ depends on the pitch and roll angle.

As already mentioned the inclination α_c of the ray through the points $\mathbf{p}_a^{ls}$ and $\mathbf{p}_b^{ls}$ depends on the distance between the two points (Figure 3.4)

$$\cos\alpha_c = \frac{k_{c,a} + k_{c,b}}{\left\|\mathbf{p}_b^{ls}\right\| - \left\|\mathbf{p}_a^{ls}\right\|} \tag{3.9}$$

The parameters $k_{c,a}$ and $k_{c,b}$ can be deduced using the theorem on intersecting lines in Figure 3.4

$$\frac{2k_{c,a}}{w_c} = \frac{h_c - t_{z,ls} - z_a^{ls^*}}{h_c} \tag{3.10}$$

$$\frac{2k_{c,b}}{w_c} = \frac{h_c - t_{z,ls} - z_b^{ls^*}}{h_c} \tag{3.11}$$

where w_c is the width and h_c the height of the calibration object, $t_{z,ls}$ the z-coordinate of the laser scanner's mounting position and $z_a^{ls^*}$ and $z_b^{ls^*}$ are the z-coordinates of the points $\mathbf{p}_a^{ls}$ and $\mathbf{p}_b^{ls}$ in the virtual coordinate system $\mathcal{C}_{ls^*}$. These z-coordinates $z_a^{ls^*}$ and $z_b^{ls^*}$ are related to the inclination α_c, with

$$z_a^{ls^*} = \left\|\mathbf{p}_a^{ls}\right\| \sin(-\alpha_c) = -\left\|\mathbf{p}_a^{ls}\right\| \sin\alpha_c \tag{3.12}$$

$$z_b^{ls^*} = \left\|\mathbf{p}_b^{ls}\right\| \sin(-\alpha_c) = -\left\|\mathbf{p}_b^{ls}\right\| \sin\alpha_c \tag{3.13}$$

Substituting $z_a^{ls^*}$ and $z_b^{ls^*}$ in Equations (3.10) and (3.11) with (3.12) and (3.13) and then substituting $k_{c,a}$ and $k_{c,b}$ in (3.9) with the resulting expressions leads after a rearrangement to the following form

$$\cos\alpha_c + b\sin\alpha_c = c \tag{3.14}$$

where the coefficients b and c are given by

$$b = -\frac{w_c\left(\left\|\mathbf{p}_a^{ls}\right\| + \left\|\mathbf{p}_b^{ls}\right\|\right)}{2h_c\left(\left\|\mathbf{p}_b^{ls}\right\| - \left\|\mathbf{p}_a^{ls}\right\|\right)}, \quad c = \frac{w_c\left(h_c - t_{z,ls}\right)}{h_c\left(\left\|\mathbf{p}_b^{ls}\right\| - \left\|\mathbf{p}_a^{ls}\right\|\right)} \tag{3.15}$$

The solution[2] to the problem stated in Equation (3.14) is[3]

$$\alpha_c = \arctan\left(\frac{bc+\sqrt{1-c^2+b^2}}{1+b^2}, -\frac{b\left(bc+\sqrt{1-c^2+b^2}\right)}{1+b^2}+c\right) \tag{3.16}$$

Using Equations (3.12) and (3.16) the z-coordinate $z_a^{ls^*}$ of the point $\mathbf{p}_a^{ls^*}$ in the virtual coordinate system can be calculated. However, the x- and y-coordinate of the point $\mathbf{p}_a^{ls^*}$ are unknown. As all three coordinates of the point $\mathbf{p}_a^{ls}$ are known in the laser scanner coordinate system $\mathcal{C}_{ls}$, Equation (3.8) includes four unknowns

$$\hat{\mathbf{p}}_a^{ls}\left(\theta_{ls}, \phi_{ls}, x_a^{ls^*}, y_a^{ls^*}\right) = \mathbf{R}'_{ls\rightarrow ls^*}\left(\theta_{ls}, \phi_{ls}\right)\mathbf{p}_a^{ls^*}\left(x_a^{ls^*}, y_a^{ls^*}\right) \tag{3.17}$$

With $N \geq 2$ laser scanner calibration objects a least squares problem can be formulated

$$0 \stackrel{!}{=} \sum_{i=1}^{N} \left\| \hat{\mathbf{p}}_{a,i}^{ls}\left(\theta_{ls}, \phi_{ls}, x_{a,i}^{ls^*}, y_{a,i}^{ls^*}\right) - \mathbf{p}_{a,i}^{ls} \right\| \tag{3.18}$$

where $\hat{\mathbf{p}}_{a,i}^{ls}$ denotes the estimated position of the point $\mathbf{p}_{a,i}^{ls^*}$ of the laser scanner calibration object i given the estimated parameters θ_{ls}, ϕ_{ls}, $x_{a,i}^{ls^*}$ and $y_{a,i}^{ls^*}$. The measured position of the calibration object i in the laser scanner coordinate system $\mathcal{C}_{ls}$ is denoted by $\mathbf{p}_{a,i}^{ls}$. This nonlinear optimization problem is solved using the Levenberg-Marquard method [101, 107]. As an initial guess, the pitch and roll angle are set to zero ($\theta_{ls} = \phi_{ls} = 0$) and therefore $\mathbf{p}_{a,i}^{ls^*} = \hat{\mathbf{p}}_{a,i}^{ls} = \mathbf{p}_{a,i}^{ls}$ holds.

In order to reduce the uncertainty of the estimation, the points $\mathbf{p}_a^{ls}$ and $\mathbf{p}_b^{ls}$ in Figure 3.2 are constructed in several consecutive scan intervals. The mean of the individual estimates is used as an overall estimate in the optimization (3.18) assuming zero mean noise for the individual estimates.

So far the problem has been solved for one layer of a multi-layer laser scanner. Different strategies are possible in order to incorporate the data of several layers for the estimation of the sensor's orientation:

- The least-sqaures problem in Equation (3.18) can be solved using all available points $\hat{\mathbf{p}}_{a,i}^{ls}$ of the individual layers.
- Only one layer is chosen for the estimation.

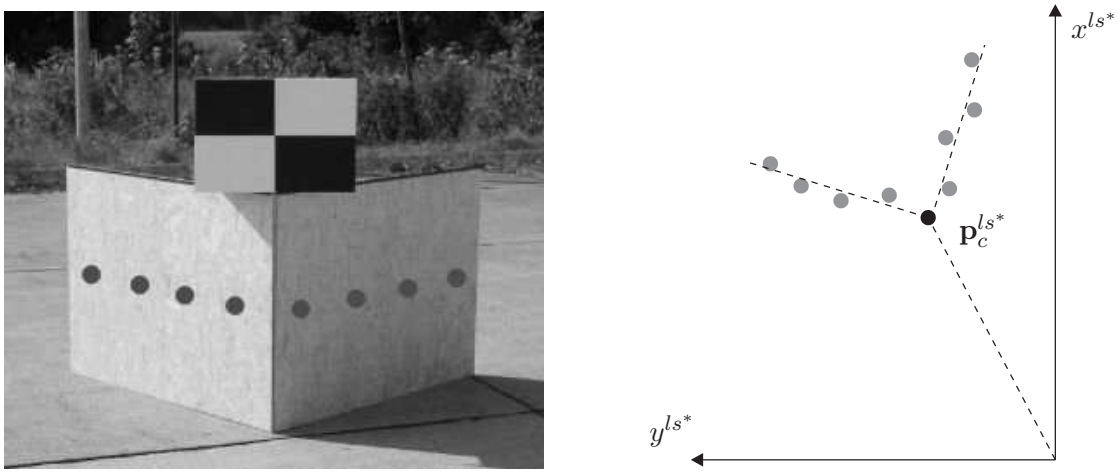
In practice it is advisable to use only one layer. Using the measurement layer which is almost parallel to the ground, the errors introduced by the inaccurate absolute distance measurement of the laser scanner can be minimized (Chapter 3.4).

3.2.2 Estimation of the Yaw Angle

Once pitch and roll of the laser scanner are estimated, the yaw angle is the only parameter left to be determined for a complete calibration of the laser scanner. The precision of

[2] There exists a second solution. This second solution however does not apply due to the constraints of the problem formulation. A proof can be found in Appendix A. Both solutions were found using Maple 9.5.

[3] The two-argument function $\arctan(y, x)$ computes the principal value of the argument of the complex number $x + iy$, with $-\pi < \arctan(y, x) < \pi$.



(a) (b)

Figure 3.5: (a) This central calibration object is used for the estimation of the yaw angle of the laser scanner. The laser scanner measurements are painted into the image for visualization purposes. (b) shows the result of the pre-processing. Two regression lines are matched to the laser scanner data of which the intersection defines the point $\mathbf{p}_c^{ls^*}$. The same calibration object is used for the camera calibration.

the yaw angle with respect to the vehicle coordinate system $\mathcal{C}_{veh}$ is most relevant, as it affects the precision of the position and velocity estimation of tracked objects. A calibration object was constructed (Figure 3.5(a)) which is placed on the x-axis of the vehicle coordinate system $\mathcal{C}_{veh}$. The calibration algorithm uses this constraint, in order to estimate the yaw angle ψ_{ls} of the laser scanner.

The calibration object shown in Figure 3.5 exploits the good repeat accuracy of the laser scanner. The point $\mathbf{p}_c^{ls^*}$ is constructed from the intersection of two regression lines. In order to further reduce the measurement noise for subsequent calculations the point is constructed in several consecutive scan intervals. The mean of the individual estimates is used as an overall estimate and will be subsequently referenced as $\mathbf{p}_{c,i}^{ls^*}$ for the calibration object i.

Figure 3.6 shows the given setup. The point $\mathbf{p}_{c,1}^{ls^*}$ of calibration object 1 is situated on the x-axis of the vehicle coordinate system $\mathcal{C}_{veh}$ and is viewed by the laser scanner. The point is transformed into the virtual laser scanner coordinate system $\mathcal{C}_{ls^*}$ using the already determined rotation matrix for the pitch and roll angle $\mathbf{R}_{ls\rightarrow ls^*}(\theta_{ls}, \phi_{ls})$. Using geometric constraints, the solution for the angle ψ_{ls} can be formulated differentiating two cases

$$\psi_{ls} = -\frac{\pi}{2} + \psi_1 + \psi_2 \qquad \text{with} \qquad y_{c,1}^{ls^*} \geq 0 \tag{3.19}$$

$$\psi_{ls} = \frac{\pi}{2} + \psi_1 + \psi_2 \qquad \text{with} \qquad y_{c,1}^{ls^*} < 0 \tag{3.20}$$

Figure 3.6(a) shows the first case ($y_{c,1}^{ls^*} \geq 0$) and the second case ($y_{c,1}^{ls^*} < 0$) is presented in Figure 3.6(b). The angles ψ_1 and ψ_2 are calculated with the y-component of the laser

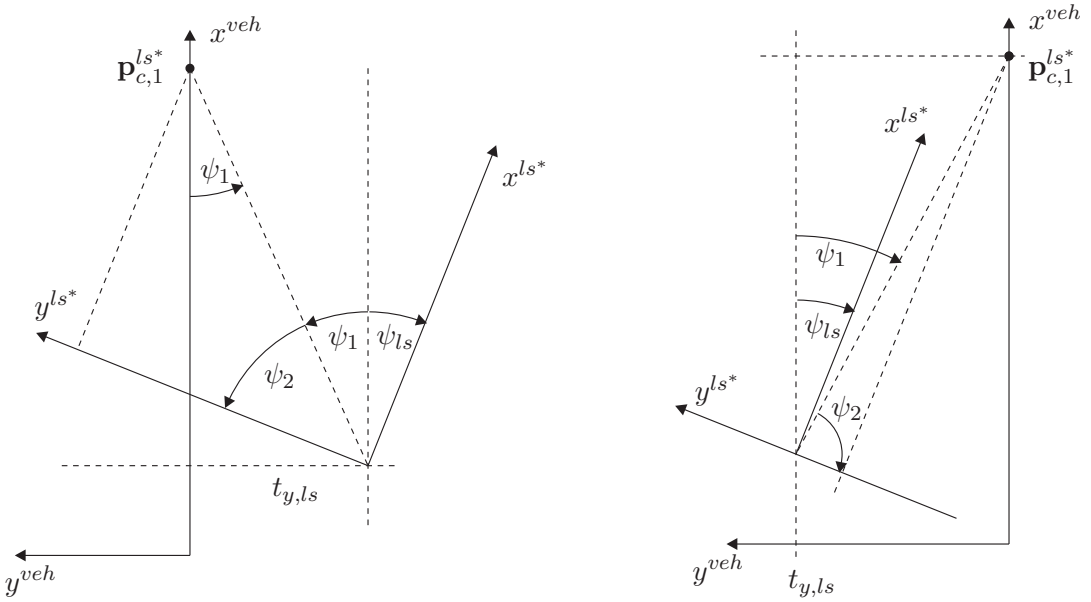


Figure 3.6: Two possible calibration setups for the estimation of the laser scanner's yaw angle ψ_{ls}. The figures show the vehicle coordinate system $\mathcal{C}_{veh}$ and the virtual laser scanner coordinate system $\mathcal{C}_{ls^*}$.

scanner's mounting position $t_{y,ls}$

$$\psi_1 = \arcsin\left(\frac{-t_{y,ls}}{\sqrt{\left(x_{c,1}^{ls^*}\right)^2 + \left(y_{c,1}^{ls^*}\right)^2}}\right) \tag{3.21}$$

$$\psi_2 = \arctan\left(\frac{x_{c,1}^{ls^*}}{y_{c,1}^{ls^*}}\right) \tag{3.22}$$

The overall rotation matrix is then given in accordance to the definition in Equation (3.3) by

$$\mathbf{R}_{ls\to veh} = \mathbf{R}_{ls^*\to veh}\left(\psi_{ls}\right)\mathbf{R}_{ls\to ls^*}\left(\theta_{ls}, \phi_{ls}\right) \tag{3.23}$$

The homogeneous matrix $\mathbf{H}_{ls\to veh}$ is constructed using Equation (3.2).

3.3 Calibration of a Camera

The calibration of the camera can be decomposed into an internal and an external calibration. The external calibration parameters are the position and the orientation of the sensor, relative to a reference coordinate system. The internal parameters describe the projection of the 3D environment to the 2D image coordinate system.

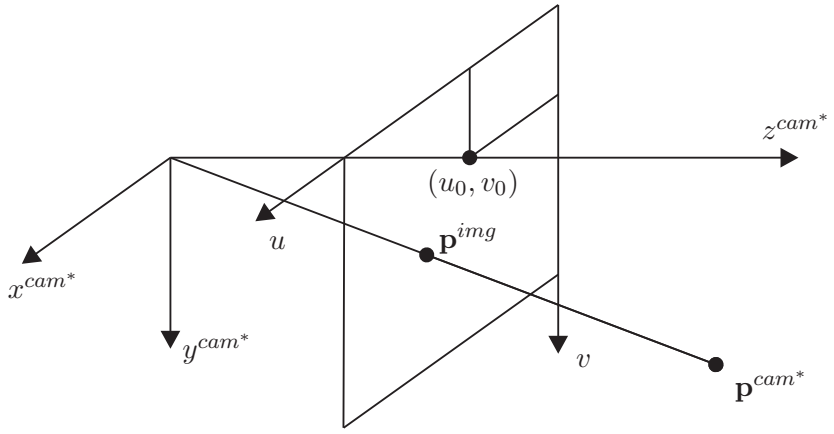


Figure 3.7: Pinhole camera model.

In order to keep the standard definition of the pinhole camera coordinate system $\mathcal{C}_{cam^*}$ shown in Figure 3.7 an additional transformation is necessary, which defines the relationship to the DIN 70,000 camera coordinate system $\mathcal{C}_{cam}$ (Figure 3.1)

$$\mathbf{H}_{cam^* \to cam} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.24}$$

The transformation from the 3d pinhole camera coordinate system $\mathcal{C}_{cam^*}$ into the vehicle coordinate system $\mathcal{C}_{veh}$ results in

$$\mathbf{p}^{veh} = \mathbf{H}_{cam \to veh} \mathbf{H}_{cam^* \to cam} \mathbf{p}^{cam^*} \tag{3.25}$$

The pinhole camera model defines the projection of a point in 3d homogeneous coordinates $\mathbf{p}^{cam^*}$ into a point $\tilde{\mathbf{p}}^{img}$ given in the image coordinate system $\mathcal{C}_{img}$ [42]

$$\tilde{\mathbf{p}}^{img} = \mathbf{H}_{cam^* \to img} \mathbf{p}^{cam^*} \tag{3.26}$$

$$\tilde{\mathbf{p}}^{img} = \begin{bmatrix} \tilde{p}_1^{img} \\ \tilde{p}_2^{img} \\ \tilde{p}_3^{img} \end{bmatrix} = \begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x^{cam^*} \\ y^{cam^*} \\ z^{cam^*} \\ 1 \end{bmatrix} \tag{3.27}$$

where the pixel coordinates (u, v) of the projected point are given by

$$\mathbf{p}^{img} = \begin{bmatrix} \frac{\tilde{p}_1^{img}}{\tilde{p}_3^{img}} \\ \frac{\tilde{p}_2^{img}}{\tilde{p}_3^{img}} \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} \tag{3.28}$$

The internal camera parameters, which are the focal lengths α_u and α_v as well as the optical center (u_0, v_0) are estimated once for each camera and lens configuration with the calibration toolbox described in [15].

For the external calibration the calibration object shown in Figure 3.5 is utilized. Using the already calibrated laser scanner, the calibration object is localized in the vehicle coordinate system $\mathcal{C}_{veh}$. However, the z-coordinate of the point $\mathbf{p}_c^{veh}$ is replaced by the known height of the calibration pattern. Therefore, the position of the calibration pattern is known in the 3D vehicle coordinate system. As a projection of a point $\mathbf{p}^{veh}$ into the image domain is with Equations (3.25) and (3.26) given by the homogeneous transformation matrix $\mathbf{H}_{veh\rightarrow img}$ with

$$\begin{aligned} \tilde{\mathbf{p}}^{img} &= \mathbf{H}_{cam^*\rightarrow img}\mathbf{H}_{cam^*\rightarrow cam}^{-1}\mathbf{H}_{cam\rightarrow veh}^{-1}\mathbf{p}^{veh} && (3.29)\\ &= \mathbf{H}_{veh\rightarrow img}\mathbf{p}^{veh} && (3.30) \end{aligned}$$

the projection of the point $\mathbf{p}_c^{veh}$ into the image domain follows as

$$\tilde{\mathbf{p}}_c^{img}\left(\psi_{cam},\theta_{cam},\phi_{cam}\right) = \mathbf{H}_{veh\rightarrow img}\left(\psi_{cam},\theta_{cam},\phi_{cam}\right)\mathbf{p}_c^{veh} \tag{3.31}$$

with

$$\hat{\mathbf{p}}_c^{img}\left(\psi_{cam},\theta_{cam},\phi_{cam}\right) = \begin{bmatrix} \frac{\tilde{p}_{1,c}^{img}}{\tilde{p}_{3,c}^{img}} \\ \frac{\tilde{p}_{2,c}^{img}}{\tilde{p}_{3,c}^{img}} \end{bmatrix} \tag{3.32}$$

The point $\hat{\mathbf{p}}_c^{img}\left(\psi_{cam},\theta_{cam},\phi_{cam}\right)$ denotes the current estimated image position of the calibration pattern, given the current estimates ψ_{cam}, θ_{cam} and ϕ_{cam}. Using image processing techniques, the calibration pattern is localized in the image with sub-pixel accuracy ($\mathbf{p}_c^{img}$).

The aim of the calibration is the estimation of the pitch θ, roll ϕ and yaw angle ψ of the camera, by minimizing the difference between the position of the calibration pattern projected into the image domain $\hat{\mathbf{p}}_c^{img}\left(\psi_{cam},\theta_{cam},\phi_{cam}\right)$ and the position of the pattern localized in the image $\mathbf{p}_c^{img}$. Using $N \geq 2$ calibration objects as shown in Figure 3.3 a least squares solution can be formulated

$$0 \stackrel{!}{=} \sum_{i=1}^{N} \left\| \hat{\mathbf{p}}_{c,i}^{img}\left(\psi_{cam},\theta_{cam},\phi_{cam}\right) - \mathbf{p}_{c,i}^{img} \right\| \tag{3.33}$$

The solution of this nonlinear problem is found with the Levenberg-Marquard method [101, 107].

In order to reduce the uncertainty of the estimation, the points $\mathbf{p}_c^{img}$ are localized in several consecutive images. In the optimization (3.33) the mean of the individual image positions is used as an overall estimate $\mathbf{p}_{c,i}^{img}$ for calibration object i.

In practice this optimization exhibited some imprecision in the relative yaw estimation of the sensors. In order to increase the precision, the same central calibration object ($\mathbf{p}_{c,1}^{ls}$) was chosen for the estimation of the yaw angle of both sensors. The optimization of the camera parameters now is divided into two steps

1. Least squares optimization as shown in Equation (3.33) with respect to the yaw ψ_{cam} and pitch angle θ_{cam} using the central calibration object in Figure 3.3.

2. Least squares optimization as shown in Equation (3.33) with respect to the roll angle ϕ_{cam} using the left and right camera calibration objects in Figure 3.3.

These two optimizations are performed iteratively until the change of the estimated parameters falls below a predefined threshold.

Additionally it would be possible to estimate the focal length. Assuming a quadratic size of the sensory receptors on the image chip, a single focal length $\alpha = \alpha_u = \alpha_v$ can be included into the optimization of the roll angle

$$0 \stackrel{!}{=} \sum_{i=2}^{3} \left\| \hat{\mathbf{p}}_{c,i}^{img}\left(\phi_{cam}, \alpha\right) - \mathbf{p}_{c,i}^{img} \right\| \tag{3.34}$$

However, it should be mentioned, that imprecise laser scanner measurements or an inaccurate laser scanner calibration distorts the estimation of the focal length. To which extent the refinement of the focal length improves the overall performance depends on the calibration setup, the laser scanner and the performance of the estimation procedure for the internal calibration [15].

3.4 Worst Case Error Analysis

In the previous sections the calibration procedure for the laser scanner and the camera have been introduced. While the laser scanner can be calibrated alone, the camera can only be calibrated in conjunction with an already calibrated laser scanner. This section identifies possible maximal errors of the sensors and the calibration setup and presents the resulting worst case errors of the estimated sensor rotation angles. Subsequently the worst case errors of the laser scanner measurements projected into the image domain are analyzed. The latter performance measure is more relevant to the application addressed in this thesis, as image regions are selected from projected laser scanner measurements for subsequent object localization algorithms in the image domain. The smaller the worst case projection error is, the smaller are the selected image regions and the more precise are the results of the object localization.

Alternatively to a worst case error analysis a stochastic error propagation was considered. However, as the probability density function of some parameters is not normally distributed a worst case error analysis seemed to be more appropriate.

The precision of the calibration depends on the sensor measurement errors and the calibration setup. The size of the calibration objects and the mounting position of the sensors are only known with a certain accuracy and the assumption of a perfect plane on which the vehicle and the calibration objects are standing is in general violated. The magnitude of the influence of the sensor measurement and calibration setup errors on the precision of the calibration process depends on the chosen calibration setup. For instance an erroneous sensor position distorts the calibrated sensor orientation more, if the corresponding calibration objects are close to the sensor. For an analysis of the error propagation a certain operating point must be chosen. Therefore a typical calibration setup used in practice is considered (Figure 3.3).

The calibration objects for the laser scanner shown in Figure 3.2 are situated in a distance of 20 m, ±45° off the x-axis

$$\mathbf{p}_{a,1}^{ls} = \begin{bmatrix} 20\text{ m} \cdot \cos(\pi/4) & 20\text{ m} \cdot \sin(\pi/4) & 0\text{ m} \end{bmatrix}' \tag{3.35}$$

$$\mathbf{p}_{a,2}^{ls} = \begin{bmatrix} 20\text{ m} \cdot \cos(-\pi/4) & 20\text{ m} \cdot \sin(-\pi/4) & 0\text{ m} \end{bmatrix}' \tag{3.36}$$

The width of the laser scanner calibration objects is $w_{c,i} = 1.844$ m and the height $h_{c,i} = 1.183$ m. The joint laser scanner and camera calibration object is at the position $\mathbf{p}_{c,1}^{veh} = \begin{bmatrix}20 & 0 & 1.22\end{bmatrix}'$ meters[4]. The second and third camera calibration objects are situated at the positions

$$\mathbf{p}_{c,2}^{veh} = \begin{bmatrix}20\text{ m} & 4\text{ m} & 1.22\text{ m}\end{bmatrix}' \tag{3.37}$$

$$\mathbf{p}_{c,3}^{veh} = \begin{bmatrix}20\text{ m} & -4\text{ m} & 1.22\text{ m}\end{bmatrix}' \tag{3.38}$$

The mounting position of the laser scanner $\mathbf{t}_{ls} = \begin{bmatrix}3.6\text{ m} & 0\text{ m} & 0.53\text{ m}\end{bmatrix}'$ is at the front bumper and the camera is situated close to the interior rearview mirror of the test vehicle (Volkswagen Golf), with $\mathbf{t}_{cam} = \begin{bmatrix}1.6\text{ m} & -0.15\text{ m} & 1.22\text{ m}\end{bmatrix}'$. The focal length of the camera used in the analysis (as well as in practice) is $\alpha = \alpha_u = \alpha_v = 1,275$ pixel. The resolution of the 1/2" chip is 640x480 pixel and the principle point is assumed to be in the center of the chip for this analysis. For this analysis the true orientation of the sensors is chosen such that all angles are zero.

In practice the exact position of the calibration objects is not of relevance. The only constraint is the position of the central calibration object for the joint laser scanner and camera calibration, which has to be placed on the vehicle's x-axis.

3.4.1 Erroneous Measurements and Calibration Parameters

The precision of the calibration depends on accurate calibration objects and on the sensor measurement errors. There are different erroneous parameters, which can be identified. The considered parameters of the calibration setup are:

- $w_{c,i}$: width of the laser scanner calibration object i for the laser scanner's pitch and roll angle estimation.
- $h_{c,i}$: height of the laser scanner calibration object i for the laser scanner's pitch and roll angle estimation.
- $t_{x,ls}, t_{y,ls}, t_{z,ls}$: mounting position of the laser scanner.
- $h_{g,ls,i}$: deviation of the real ground on which the calibration objects i for the laser scanner's pitch and roll angle estimation stand from a perfect plane. The error is due to unevenness of the ground.
- α_u, α_v: focal length of the camera.
- $t_{x,cam}, t_{y,cam}, t_{z,cam}$: mounting position of the camera.
- $h_{g,cam,i}$: deviation of the real ground on which the camera calibration objects i stand from a perfect plane.

[4]The z-coordinate of the calibration pattern should be approximately the same as the z-coordinate of the camera's mounting position in order to minimize the errors

Table 3.1: Worst case errors for the estimation of the orientation of the laser scanner. The different parameters ξ_i are shown together with the sensitivities of the different angles, the assumed worst case error $e_{max,i}$ of the parameter and the resulting worst case errors in the angles $e_{max,*,i}$. The parameter $\Delta p_{ab,i}$ denotes the distance $\left\|\mathbf{p}_b^{ls}\right\| - \left\|\mathbf{p}_a^{ls}\right\|$ for the laser scanner calibration object i.

ξ_i	$\frac{\partial\phi(\xi_i)}{\partial\xi_i}$	$\frac{\partial\theta(\xi_i)}{\partial\xi_i}$	$\frac{\partial\psi(\xi_i)}{\partial\xi_i}$	$e_{max,i}$	$e_{max,\phi,i}$	$e_{max,\theta,i}$	$e_{max,\psi,i}$
$w_{c,1}$	$0.696° \frac{1}{\text{m}}$	$-0.696° \frac{1}{\text{m}}$	$-0.022° \frac{1}{\text{m}}$	0.01 m	0.007°	0.007°	0.000°
$w_{c,2}$	$-0.696° \frac{1}{\text{m}}$	$-0.696° \frac{1}{\text{m}}$	$0.022° \frac{1}{\text{m}}$	0.01 m	0.007°	0.007°	0.000°
$h_{c,1}$	$0.885° \frac{1}{\text{m}}$	$-0.885° \frac{1}{\text{m}}$	$-0.028° \frac{1}{\text{m}}$	0.01 m	0.009°	0.009°	0.000°
$h_{c,2}$	$-0.885° \frac{1}{\text{m}}$	$-0.885° \frac{1}{\text{m}}$	$0.028° \frac{1}{\text{m}}$	0.01 m	0.009°	0.009°	0.000°
$t_{x,ls}$	$0.000° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	0.01 m	0.000°	0.000°	0.000°
$t_{y,ls}$	$0.000° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	$-3.494° \frac{1}{\text{m}}$	0.01 m	0.000°	0.000°	0.035°
$t_{z,ls}$	$0.000° \frac{1}{\text{m}}$	$3.951° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	0.01 m	0.000°	0.040°	0.000°
$h_{g,ls,1}$	$1.977° \frac{1}{\text{m}}$	$-1.977° \frac{1}{\text{m}}$	$-0.063° \frac{1}{\text{m}}$	0.02 m	0.040°	0.040°	0.001°
$h_{g,ls,2}$	$-1.977° \frac{1}{\text{m}}$	$-1.977° \frac{1}{\text{m}}$	$0.063° \frac{1}{\text{m}}$	0.02 m	0.040°	0.040°	0.001°
$\Delta p_{ab,1}$	$-1.267° \frac{1}{\text{m}}$	$1.267° \frac{1}{\text{m}}$	$0.041° \frac{1}{\text{m}}$	0.02 m	0.025°	0.025°	0.001°
$\Delta p_{ab,2}$	$1.267° \frac{1}{\text{m}}$	$1.267° \frac{1}{\text{m}}$	$-0.041° \frac{1}{\text{m}}$	0.02 m	0.025°	0.025°	0.001°
$\left\|\mathbf{p}_{a,1}^{ls}\right\|$	$0.000° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	0.1 m	0.000°	0.000°	0.000°
$\left\|\mathbf{p}_{a,2}^{ls}\right\|$	$0.000° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	0.1 m	0.000°	0.000°	0.000°
$\left\|\mathbf{p}_{c,1}^{ls*}\right\|$	$0.000° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	$0.000° \frac{1}{\text{m}}$	0.1 m	0.000°	0.000°	0.000°
β	0.000	0.000	-1.000	0.1°	0.000°	0.000°	0.100°
e_{max}					**0.161°**	**0.201°**	**0.140°**

The parameters of the calibration which depend on the precision of the sensor itself are listed below:

- $\Delta p_{ab,i} = \left\|\mathbf{p}_b^{ls}\right\| - \left\|\mathbf{p}_a^{ls}\right\|$: this distance difference can be imprecise due to a distance dependent bias in the range measurement of the laser scanner. However, the bias is locally small.
- $\left\|\mathbf{p}^{ls}\right\|$: the measured distance can be imprecise due to a distance dependent bias in the range measurement of the laser scanner.
- $\beta = \arctan\left(\frac{y_{c,1}^{ls*}}{x_{c,1}^{ls*}}\right)$: for the central calibration object an intersection of two regression lines is calculated from the laser scanner data (Figure 3.5(b)). This preprocessing is prone to inaccuracy.

3.4.2 Worst Case Errors of the Estimated Angles

The Tables 3.1 and 3.2 present the calibration parameters ξ_i which can be affected by a maximal error $e_{max,i}$. The maximal errors are either assumed or known from experiments. The assumed maximal errors determine the required precision of the calibration setup.

For instance, the width w_c and height h_c of the laser scanner calibration objects must known with an accuracy of 0.01 m.

The sensitivity of the roll $\frac{\partial\phi(\xi_i)}{\partial\xi_i}$, pitch $\frac{\partial\theta(\xi_i)}{\partial\xi_i}$ and yaw angle $\frac{\partial\psi(\xi_i)}{\partial\xi_i}$ with respect to the parameter ξ_i are calculated numerically by variation of the parameter ξ_i and observation of the change in the orientation estimate after the calibration. The calibration uses the least squares optimizations of (3.18) and (3.33). The sensitivities and the resulting worst case errors for the individual angles and parameters given by [64]

$$e_{max,\phi,i} = \left| e_{max,i} \frac{\partial\phi(\xi_i)}{\partial\xi_i} \right| \tag{3.39}$$

$$e_{max,\theta,i} = \left| e_{max,i} \frac{\partial\theta(\xi_i)}{\partial\xi_i} \right| \tag{3.40}$$

$$e_{max,\psi,i} = \left| e_{max,i} \frac{\partial\psi(\xi_i)}{\partial\xi_i} \right| \tag{3.41}$$

are shown in Tables 3.1 and 3.2.

The overall worst case error in the estimation of the angles

$$e_{max,\phi} = \sum_i \left| e_{max,i} \frac{\partial\phi(\xi_i)}{\partial\xi_i} \right| \tag{3.42}$$

$$e_{max,\theta} = \sum_i \left| e_{max,i} \frac{\partial\theta(\xi_i)}{\partial\xi_i} \right| \tag{3.43}$$

$$e_{max,\psi} = \sum_i \left| e_{max,i} \frac{\partial\psi(\xi_i)}{\partial\xi_i} \right| \tag{3.44}$$

is shown in the last line of the Tables 3.1 and 3.2.

Table 3.1 concludes the worst case error analysis for the laser scanner calibration and Table 3.2 presents the results for the camera calibration. The worst case error in the yaw angle of the laser scanner could be minimized to 0.140° and for the camera to 0.049° respectively. The pitch angle has a maximal error of 0.201° for the laser scanner and 0.074° for the camera. The less important roll angle can be estimated within the error bounds of 0.161° (laser scanner) and 0.334° (camera). The worst case errors in the laser scanner calibration are mainly affected by an imprecise mounting position, an uneven ground and the accuracy of the intersection point $\mathbf{p}_{c,1}^{ls^*}$ in Figure 3.5. The precision of the camera calibration is primarily influenced by an uneven ground and the mounting position of the camera.

The results presented in Tables 3.1 and 3.2 underline the good choice of the calibration setup. The relatively high errors in the absolute distance measurements of the laser scanner ($\left\|\mathbf{p}_{c,i}^{ls}\right\|$, $\left\|\mathbf{p}_{a,i}^{ls}\right\|$) have a negligible influence on the calibration errors. Pitch and roll angles are estimated exploiting the much smaller error of distance difference measurements $\Delta p_{ab,i}$.

3.4.3 Worst Case Projection Errors

Of special interest for the remainder of this work are the errors of laser scanner measurements projected into the image domain. For the association of image regions with laser

Table 3.2: Worst case errors for the estimation of the orientation of the camera. The different parameters ξ_i are shown together with the sensitivities of the different angles, the assumed worst case error $e_{max,i}$ of the parameter and the resulting worst case errors in the angles $e_{max,*,i}$. The parameter $\Delta p_{ab,i}$ denotes the distance $\left\|\mathbf{p}_b^{ls}\right\| - \left\|\mathbf{p}_a^{ls}\right\|$ for the laser scanner calibration object i. Some parameters from Table 3.1 were left out, as their influence on the result is insignificant.

ξ_i	$\frac{\partial\phi(\xi_i)}{\partial\xi_i}$	$\frac{\partial\theta(\xi_i)}{\partial\xi_i}$	$\frac{\partial\psi(\xi_i)}{\partial\xi_i}$	$e_{max,i}$	$e_{max,\phi,i}$	$e_{max,\theta,i}$	$e_{max,\psi,i}$
$h_{g,ls,1}$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	$-0.057°\frac{1}{\text{m}}$	0.02 m	0.000°	0.000°	0.001°
$h_{g,ls,2}$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	$0.057°\frac{1}{\text{m}}$	0.02 m	0.000°	0.000°	0.001°
$\Delta p_{ab,1}$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	$0.036°\frac{1}{\text{m}}$	0.02 m	0.000°	0.000°	0.001°
$\Delta p_{ab,2}$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	$-0.036°\frac{1}{\text{m}}$	0.02 m	0.000°	0.000°	0.001°
$\left\|\mathbf{p}_{c,2}^{ls}\right\|$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	0.1 m	0.000°	0.000°	0.000°
$\left\|\mathbf{p}_{c,3}^{ls}\right\|$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	0.1 m	0.000°	0.000°	0.000°
α_u, α_v	$0.000°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	10 pixel	0.000°	0.000°	0.004°
$t_{x,cam}$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	$0.025°\frac{1}{\text{m}}$	0.01 m	0.000°	0.000°	0.000°
$t_{y,cam}$	$0.000°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	$-3.114°\frac{1}{\text{m}}$	0.01 m	0.000°	0.000°	0.031°
$t_{z,cam}$	$0.000°\frac{1}{\text{m}}$	$3.114°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	0.01 m	0.000°	0.031°	0.000°
$h_{g,cam,1}$	$-0.536°\frac{1}{\text{m}}$	$-3.118°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	0.02 m	0.005°	0.031°	0.000°
$h_{g,cam,2}$	$7.420°\frac{1}{\text{m}}$	$0.060°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	0.02 m	0.148°	0.001°	0.000°
$h_{g,cam,3}$	$-6.884°\frac{1}{\text{m}}$	$-0.056°\frac{1}{\text{m}}$	$0.000°\frac{1}{\text{m}}$	0.02 m	0.138°	0.001°	0.000°
$p_{u,img,1}$	$0.000°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	$0.045°\frac{1}{\text{pixel}}$	0.2 pixel	0.000°	0.000°	0.009°
$p_{v,img,1}$	$-0.008°\frac{1}{\text{pixel}}$	$-0.045°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	0.2 pixel	0.002°	0.009°	0.000°
$p_{u,img,2}$	$0.000°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	0.2 pixel	0.000°	0.000°	0.000°
$p_{v,img,2}$	$0.107°\frac{1}{\text{pixel}}$	$0.001°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	0.2 pixel	0.021°	0.000°	0.000°
$p_{u,img,3}$	$0.000°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	0.2 pixel	0.000°	0.000°	0.000°
$p_{v,img,3}$	$-0.099°\frac{1}{\text{pixel}}$	$-0.001°\frac{1}{\text{pixel}}$	$0.000°\frac{1}{\text{pixel}}$	0.2 pixel	0.020°	0.000°	0.000°
e_{max}					**0.334°**	**0.074°**	**0.049°**

scanner distance measurements it is of great importance, that the projection errors, which stem from an imperfect calibration, are rather small.

Using Equations (3.1) and (3.25) – (3.28) the projection of a laser scanner point $\mathbf{p}^{ls} = (x^{ls}, y^{ls}, z^{ls}, 1)'$ given in homogeneous coordinates into the image domain $\mathbf{p}^{img} = (u, v)'$ can be calculated

$$\tilde{\mathbf{p}}^{img} = \mathbf{H}_{ls\rightarrow img}\mathbf{p}^{ls} = \mathbf{H}_{cam^*\rightarrow img}\mathbf{H}_{cam^*\rightarrow cam}^{-1}\mathbf{H}_{cam\rightarrow veh}^{-1}\mathbf{H}_{ls\rightarrow veh}\mathbf{p}^{ls} \tag{3.45}$$

and

$$\mathbf{p}^{img} = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{\tilde{p}_{1,img}}{\tilde{p}_{3,img}} \\ \frac{\tilde{p}_{2,img}}{\tilde{p}_{3,img}} \end{bmatrix} \tag{3.46}$$

The worst case error $e_{max,u}\left(x^{ls}, y^{ls}, z^{ls}\right)$ and $e_{max,v}\left(x^{ls}, y^{ls}, z^{ls}\right)$ in the u- and v-coordinate

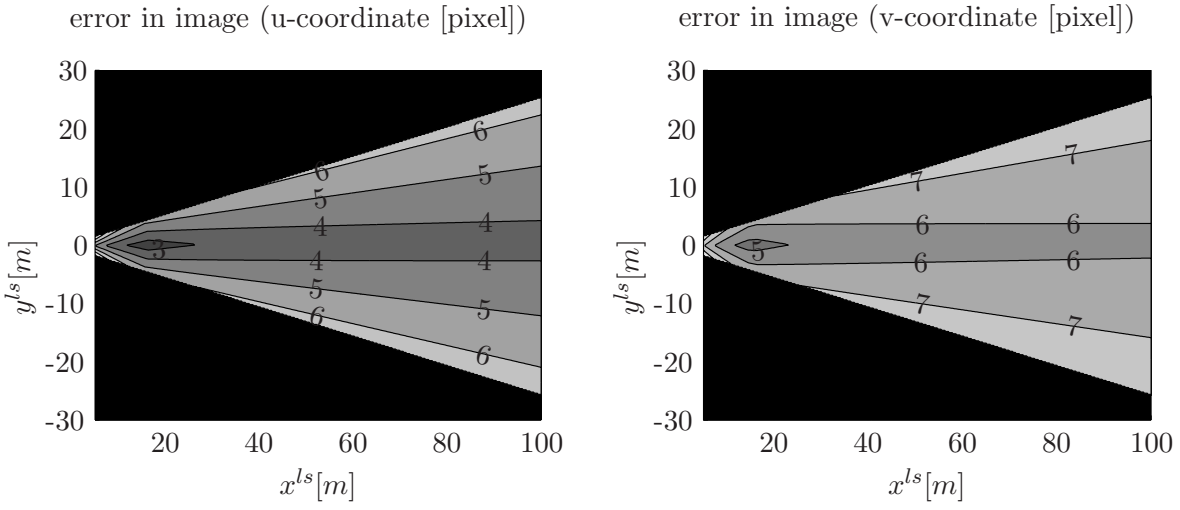


Figure 3.8: Worst case projection errors in image coordinates as a function of the x- and y-coordinates of the corresponding laser scanner measurements. The contour plot of the errors in the u-coordinate $e_{max,u}\left(x^{ls}, y^{ls}, z^{ls}=0\right)$ are shown in (a). (b) exhibits the errors in the v-coordinate $e_{max,v}\left(x^{ls}, y^{ls}, z^{ls}=0\right)$.

in the image is then determined as a function of the point $\mathbf{p}^{ls}$ according to

$$e_{max,u}\left(x^{ls}, y^{ls}, z^{ls}\right) = \sum_i \left| e_{max,i} \frac{\partial u(x^{ls}, y^{ls}, z^{ls}, \xi_i)}{\partial \xi_i} \right| \tag{3.47}$$

$$e_{max,v}\left(x^{ls}, y^{ls}, z^{ls}\right) = \sum_i \left| e_{max,i} \frac{\partial v(x^{ls}, y^{ls}, z^{ls}, \xi_i)}{\partial \xi_i} \right| \tag{3.48}$$

The resulting worst case projection errors are shown in Figure 3.8. The projection errors are visualized depending on the coordinate of the projected point in the laser scanner coordinate system $\mathbf{p}^{ls} = \begin{bmatrix} x^{ls} & y^{ls} & z^{ls}=0 & 1 \end{bmatrix}'$. For visualization purposes, the z-coordinate of the laser scanner measurement was always $z^{ls} = 0$.

3.5 Experiments

The calibration method has been used within the research group at the Department of Measurement, Control and Microtechnology, University of Ulm over several years. Many calibration cycles were performed with different cars and several different laser scanners and camera sensors in varying calibration setups. From visual inspection and experience with the object tracking applied to the fused sensor data, the calibration always exhibited a high reliability and a precision well in the presented worst case error bounds.

Figure 3.9 shows a street scenario on a highway, in which the laser scanner measurements are projected into the image after a successful calibration. The measurements on the truck driving ahead and on the cones delimiting the area of road works underline the precision of the presented calibration.

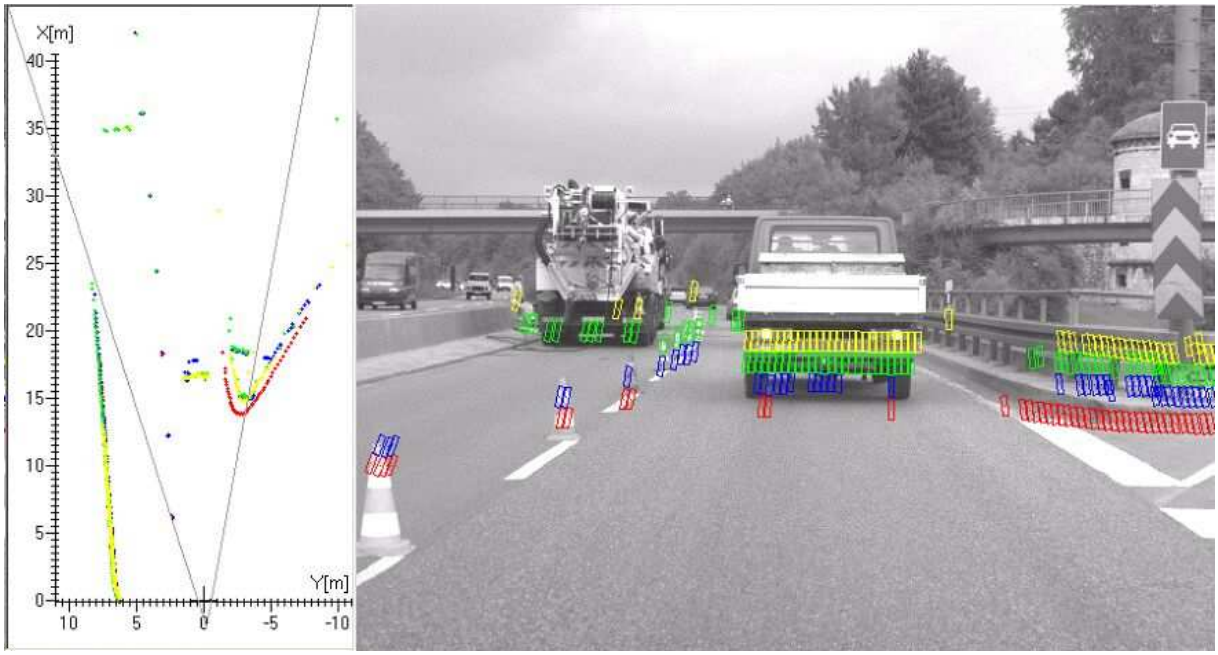


Figure 3.9: The laser scanner measurements are shown on the left side. The field of view of the camera is indicated. The laser scanner measurements are projected into the image, shown on the right side. The two sensors are synchronized and calibrated.

3.6 Discussion

A new method for a stand alone multi-layer laser scanner calibration and a subsequent camera calibration was developed. It can be used in a large variety of configurations:

- Calibration of multi-layer or single-layer laser scanner.
- Calibration of only the laser scanner's yaw angle. This is sufficient for most automotive applications using a laser scanner without a camera. Only one calibration object is necessary.
- Calibration of the laser scanner's yaw and pitch angle.
- Calibration of the yaw angle and adjustment of the pitch and roll angle. For most automotive applications the laser scanner should have a certain pitch and roll angle. Using the calibration procedure online, pitch and roll could be adjusted until the desired values are reached.
- Complete calibration of a laser scanner and a camera.
- Determination of the laser scanner's yaw and the camera's yaw and pitch. This configuration requires only one calibration object for both sensors. It applies, if the precision of the roll angles and the laser scanner's pitch angle are of minor importance.

The newly developed calibration objects exploit the sensors' characteristics in order to maximize the precision of the estimated parameters. Considering the calibration setup in Figure 3.3 the worst case error in the yaw angle of the laser scanner could be minimized to 0.140° and for the camera to 0.049° respectively. The pitch angle has a maximal error of 0.201° for the laser scanner and 0.074° for the camera. The less important roll angle can be estimated within the error bounds of 0.161° (laser scanner) and 0.334° (camera). An analysis of the joint calibration exhibits small errors of the laser scanner data projected into the image. The worst case error of the u-coordinate, which is used in the object tracking for the estimation of the lateral position and velocity, is smaller than 7 pixel in 99 % of the joint field of view of the camera and the laser scanner and decreases to 4 pixel in the most relevant regions. The worst case error of the less important v-coordinate is smaller than 8 pixel in 99 % of the joint field of view of the camera and the laser scanner.

In practice, however, the overall calibration error is due to the statistical nature of the parameter errors much below the stated worst case error.

4 Sensor Data Fusion

The sensor data fusion combines the data from several sensors to a unified description of the vehicle's environment. Components of this environment description are traffic participants, such as cars, trucks, bikes and pedestrians, the road with its lanes, traffic signs, other not further classified obstacles and the host vehicle. This environment description serves as an interface to one or several driver assistance applications.

The research in sensor data fusion for vehicle environment perception is motivated by the necessity of advanced driver assistance systems for several sensors. As soon as an application depends on the data of several sensors, sensor data fusion methods have to be applied. A focus of actual research lies on the choice of the fusion level. Low-level fusion approaches combine mainly raw sensor measurements, whereas in high-level systems tracked and classified objects are fused.

A low-level fusion is information theoretically optimal, as the full sensor information content is preserved at the fusion stage. Depending on the fusion architecture and the sensors a major drawback may be extensive communications between the sensors and the fusion processor [61].

In contrast, the high-level fusion combines preprocessed data. The main advantage of a fusion on an object level is a low required communication bandwidth and simple fusion algorithms [61]. Additionally the fusion system is widely sensor independent, as sensor specific preprocessing is performed in the sensor system. The advantage is a reusability of the sensor and fusion modules and a scalability of the system. A practical implication could arise considering the serial production of such a fusion system in a car. If the sensors are produced by different suppliers, the data interface will most likely be on a high abstraction level, as the different suppliers know best, how to interpret the raw sensor data. Therefore the fusion module must account for strongly preprocessed sensor data. However, the drawback of a high-level fusion is in general a suboptimal solution of the kinematic object state estimation and classification of the object type. A good overview of different fusion approaches is given in [6, 13, 95, 61].

The remainder of this chapter addresses the trade-off between different fusion levels and suggest a feature-level fusion architecture as a good compromise on an intermediate level for vehicle environment perception.

4.1 Low-Level Fusion

In the low-level fusion approach, raw data is transferred to the central fusion unit (Figure 4.1). The fusion unit performs a spatiotemporal alignment of the data, as the sensors are in general asynchronous and situated in different locations. However, in automotive applications the raw sensor signal in dynamic environments does usually not allow for a compensation of the object motion. Therefore synchronous sensors are required for a

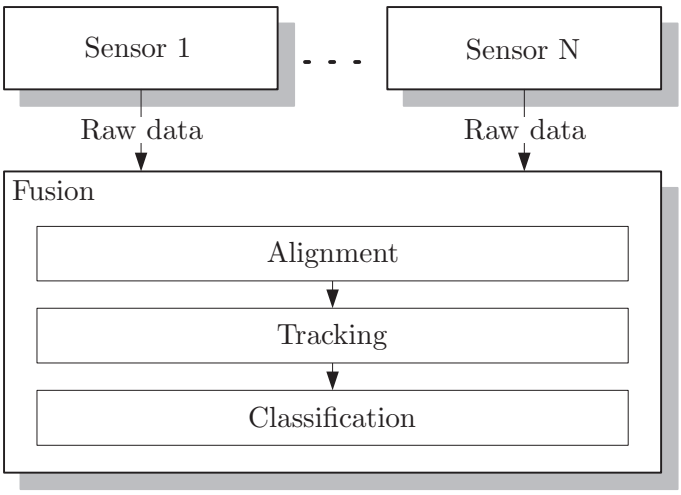


Figure 4.1: Low-level fusion architecture.

precise low-level fusion. The spatial alignment demands for a calibration or adjustment of the sensors. It might still not be possible to align the data by a simple transformation, if the transformation is ambiguous. Consider the images of two video cameras located at different positions in the vehicle. Each pixel is an angle-only measurement without a distance information and can therefore not be associated with a pixel in the other image[1].

If the association is possible the low-level data is fused. The tracking and classification are then performed on the fused raw data.

4.1.1 Tracking

In the multi-sensor multi-target tracking literature [6, 13] a different notation is used in order to differentiate multiple sensor system architectures. The categorization depends on the place in which the track processing takes place. A low-level fusion is called a *central-level tracking with centralized track file*. Theoretically an optimal tracking is possible, as all data is available. Klein summarizes the individual effects which lead to a high accuracy of the object tracking: [95]

- All data is processed in one place.
- The track is initialized using the data from all sensors simultaneously, thus producing an accurate initial state estimate and reducing the number of tracks established from partial data.
- Accurate association, as it can be done in the measurement space.
- Facilitates multiple hypothesis tracking on all data.

However, the disadvantages include high data transfer requirements and the potential problems involving the data alignment.

[1] Stereo algorithms calculate the most likely pixel associations and therefore estimate the alignment.

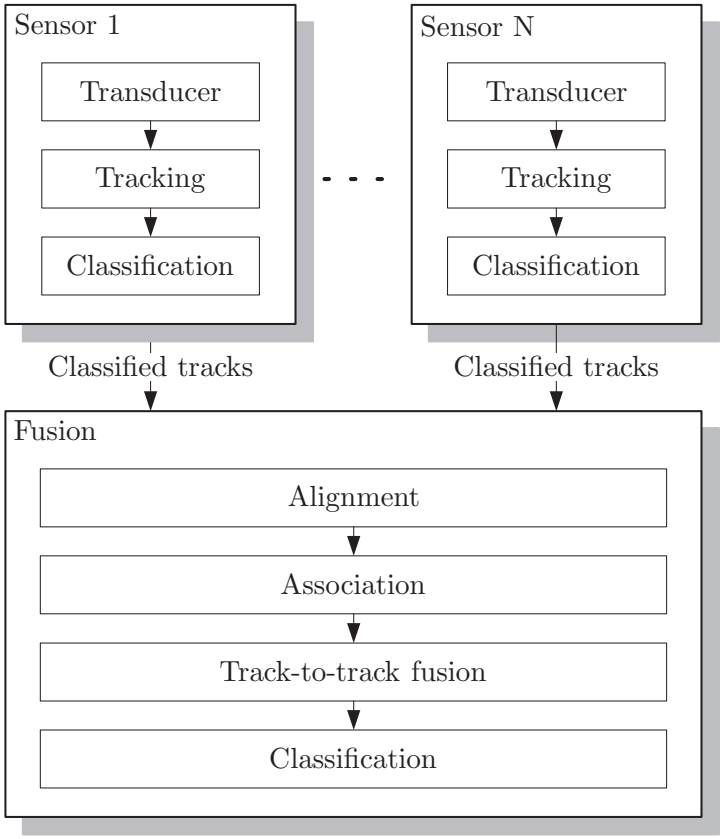


Figure 4.2: High-level fusion architecture.

4.1.2 Classification

In the context of classification or identity declaration the low-level fusion is often referred as pixel-level or data-level fusion. The feature extraction is performed on the fused data. The features are then input to a classifier as in a single sensor classification. For the classification, a low-level fusion is theoretically better than a fusion of processed data, as no information is lost during the processing.

4.2 High-Level Fusion

In a high-level fusion system the sensors perform the tracking and classification of the perceived vehicle's environment individually (Figure 4.2). The tracking and the classification results of the different sensors are then associated and combined in a track fusion module. In the multi-sensor multi-target tracking literature the tracking approach is called *sensor-level tracking with centralized track file*.

4.2.1 Tracking

A high-level track fusion approach in a central processor, which produces optimal global estimates, in terms of state estimation errors, is described in [6, Chapter 8.2]. However,

the optimality is only given under a number of conditions. The sensors have to be synchronized, must always both be able to track the objects if there are measurements available, and the association in the sensor must be globally optimal. These assumptions are in general not fulfilled, considering only partially overlapping sensor field of views or track losses in one sensor due to clutter, while the other sensor can still maintain the track.

In a second approach the track fusion unit sends the global estimates to the sensors after the fusion. Alternatively, without a central fusion processor, each sensor receives the tracks from the other sensors and fuses them to a global estimate. This approach is supposed to deliver optimal estimates, if the sensors are synchronized. It overcomes the above mentioned deficiencies, as all sensors work with the same global fused state estimates. Additionally the resulting redundancy of the track fusion enhances the reliability of the overall system [6]. However, the initial track verification period is in general longer than in low-level fusion schemes, as it is performed in each sensor separately. In the extreme, none of the sensors might even produce a confirmed track [13].

4.2.2 Classification

A high-level fusion of information used for classification is in the literature often referred as *identity fusion* or *decision-level fusion* [61, 95]. The results of the classification by the individual sensors are combined in a fusion module to a final decision about the object class or identity. Popular fusion algorithms include the probability theory, the possibility theory (or fuzzy logic) and the belief theory (or Dempster-Shafer theory) [25, 95]. In the automotive literature the Dempster-Shafer theory has recently found significant resonance [120, 58, 75, 103]. The advantage of a high-level combination of classification results is the sensor independence of the fusion algorithm. The classification algorithms can be parameterized for each sensor individually and there is no global adaption necessary.

However, the classification performance is suboptimal compared to fusion approaches using raw or slightly processed data, unless the sensor classification results are statistically independent. Even if the independence is often assumed in the literature, it is still recognized as one of the *"dirty secrets"* [62] in the current practice of data fusion systems.

4.3 Feature-Level Fusion

The aim of the proposed feature-level fusion is an optimal tracking and classification performance and at the same time a widely sensor unspecific central fusion module with a limited required bandwidth for the communication between the sensors and the fusion module. The advantages of the low- and the high-level approach are thus combined, while the disadvantages are kept within limits. Different aspects of this fusion architecture were motivated by the work described in [154, 31, 37, 143, 112, 34, 32].

4.3.1 System Architecture

Figure 4.3 depicts the proposed feature-level fusion architecture, which is composed of several sensors and a central fusion processor. In this approach the multi-object tracking

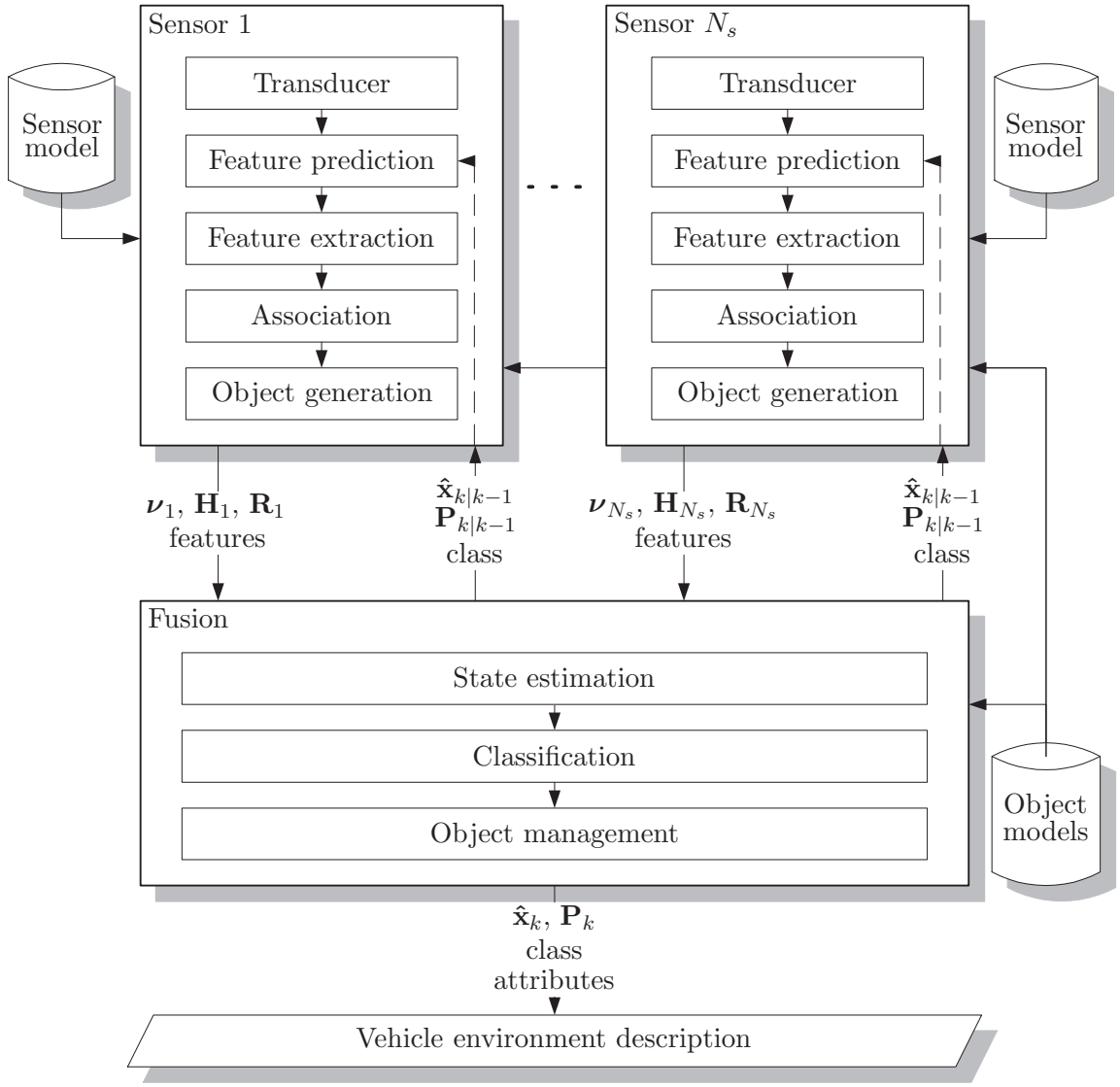


Figure 4.3: Feature-level fusion architecture.

is distributed over the fusion module and the sensors. The fusion processor manages centrally the object tracks and performs the filter update and the temporal prediction.

The sensors are responsible for the sensor specific parts of the tracking, which are the feature extraction, feature prediction, association of predicted and extracted features and the object generation from unassigned features. The fusion can thus be decoupled from the sensor models, which facilitates the reusability and scalability of the system.

Assuming a Kalman filter based state estimation approach, the data communicated from the sensors to the fusion processor are measurement residuals $\boldsymbol{\nu}_i$, the measurement matrices $\mathbf{H}_i$ and the measurement covariance matrices $\mathbf{R}_i$, together with selected features, which are used by the central classification. The predicted object states $\hat{\mathbf{x}}_{k|k-1}$ with their covariances $\mathbf{P}_{k|k-1}$ are send in return to the sensors. The required bandwidth of the communication channel is therefore significantly smaller than in pure low-level approaches.

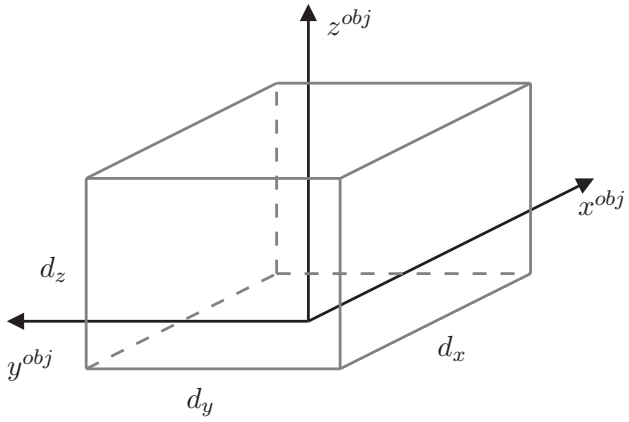


Figure 4.4: Cuboidal object model.

This system requires however that the fusion algorithms and the sensors must use the same object models. The object models describe mathematically the geometric and kinematic characteristics of the observed traffic participants. The parameters of these models form the object state representation. The sensors use the object model for a prediction of the feature. The fusion processor predicts the future state of the object utilizing the underlying kinematic behavior of the object.

4.3.2 Cuboidal Object Model

The shape of other traffic participants is within this work approximated by a cuboid of the length d_x, the width d_y and the height d_z as shown in Figure 4.4. As object velocities and yaw rate depend on the origin of the object coordinate system, it should be in the center of rotation of vehicles, which is in the middle of the rear axis[2]. The origin of the object coordinate system is therefore defined in accordance to the host vehicle coordinate system in Chapter 3.1. The x-axis is aligned with the longitudinal axis. As the exact position of the rear axis is not observable with the given sensors, it is a priori assumed to be $\frac{1}{4}d_x$ away from the rear. The object position $\begin{bmatrix} x & y & z \end{bmatrix}'$, orientation $\begin{bmatrix} \phi & \theta & \psi \end{bmatrix}'$ and motion are defined in the host vehicle coordinate system. As shown in Chapter 3.1 the position and the orientation deliver the parameters for a homogeneous transformation matrix $\mathbf{H}_{obj \to veh}$ (Equation (3.2)). This matrix $\mathbf{H}_{obj \to veh}$ enables the transformation from a point given in the object coordinate system $\mathbf{p}^{obj} = \begin{bmatrix} x^{obj} & y^{obj} & z^{obj} & 1 \end{bmatrix}'$ into the vehicle coordinate system $\mathbf{p}^{veh} = \begin{bmatrix} x^{veh} & y^{veh} & z^{veh} & 1 \end{bmatrix}'$

$$\mathbf{p}^{veh} = \mathbf{H}_{obj \to veh} \mathbf{p}^{obj} \tag{4.1}$$

[2]For stationary objects the position of the object coordinate system is not relevant.

As with the utilized sensors the roll and the pitch angle of the object are not observable, the transformation matrix is reduced to

$$\mathbf{H}_{obj\rightarrow veh} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 & x \\ \sin\psi & \cos\psi & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.2}$$

The pitch and roll angles are assumed to coincide with the host vehicle's. The motion of the object is for the following chapters given by the velocity vector $\begin{bmatrix} v_x & v_y \end{bmatrix}'$. In Chapter 9 more elaborate motion models and representations will be discussed. The complete state $\mathbf{x}$ of an object is thus defined by

$$\mathbf{x} = \begin{bmatrix} x & y & z & \psi & v_x & v_y & d_x & d_y & d_z \end{bmatrix}' \tag{4.3}$$

The orientation is in this model independent of the velocity vectors' orientation. Drift and skidding situations are therefore modeled by this approach.

4.3.3 Sensor Specific Modules

The sensors role is to provide the fusion module with measurements for the kinematic state estimation and with features for the classification. As the underlying algorithms are sensor specific, they will be explained in Chapter 5 and 6. In the following only a short description of the generic sensor tasks is provided.

The sensors preprocess the raw data and extract features which describe the object's position, orientation, motion and appearance characteristics. A feature might be the position of a corner of a vehicle in 3D coordinates or the position of the rear of a vehicle in a camera image. Other features which are subsequently used by the classifier describe the object's characteristics, such as the reflectivity of the surface, measured by a lidar system or the symmetry of the gray values in an image patch containing the object. These features are highly sensor specific and are measured in the sensor coordinate system. Therefore the association between measurements and predicted objects is performed by each sensor in the measurement space.

The feature prediction module transforms the object state estimates $\hat{\mathbf{x}}$ into the feature space of the sensor. This involves in general a transformation of the object state from the vehicle into the sensor coordinate system. Depending on the estimated object position and orientation, on the object class and on the sensor model, different features are predicted, which might be observed by the sensor.

The association then compares the predicted with the measured feature parameters in order to determine the object, of which the measured features originate from. New objects are generated from features which can not be associated with already existing objects. This involves the inverse sensor model.

4.3.4 Kalman Filter based State Estimation and Data Fusion

The function of the multi-target tracking mainly comprises the association of new measurements to already existing object tracks, the object state estimation and the object

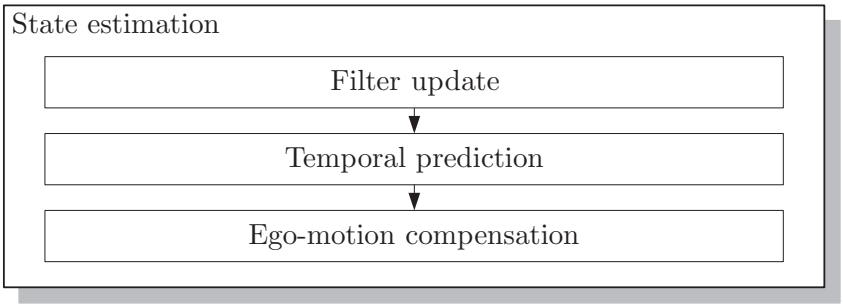


Figure 4.5: State estimation module.

management. A classical extended Kalman filter approach is used for the state estimation. In Chapter 9 more complex multiple model filter approaches will be discussed, however the structure of the tracking and the interfaces of the modules will be preserved. The Kalman filter based state estimation is composed of the filter update, the temporal prediction of the state estimates and their covariances and the ego-motion compensation (Figure 4.5). Each of the components will be detailed in the following.

There are many good introductions of the Kalman filter and of Kalman filter based multiple-object multiple-sensor tracking presented in the literature [89, 7, 4, 57, 165, 6, 13]. The approaches used in this work, including the central equations and notations, will be briefly summarized.

System Model

The state space model for a nonlinear discrete time stochastic system with additive noise is of the form

$$\mathbf{x}(k+1) \quad = \quad \mathbf{f}\left[k, \mathbf{x}(k), \mathbf{u}(k)\right] + \mathbf{v}(k) \tag{4.4}$$

where $\mathbf{x}$ is the state vector, $\mathbf{u}$ the input vector, $\mathbf{v}$ the process noise and $\mathbf{f}$ the time varying state transition function. The discrete-time measurement equation is

$$\mathbf{z}(k) \quad = \quad \mathbf{h}\left[k, \mathbf{x}(k)\right] + \mathbf{w}(k) \tag{4.5}$$

where $\mathbf{z}$ is the measurement at the measurement interval k, $\mathbf{h}$ the measurement function and $\mathbf{w}$ the measurement noise. The statistical assumptions of the Kalman Filter approach are summarized below. A detailed description can be found in [4].

- The initial state has known mean and covariance
- The process and measurement noise sequences are zero mean and white with known covariance matrices
- The process noise sequence, the measurement noise sequence and the initial state are mutually uncorrelated

As the objects are tracked from a moving platform the motion of the host vehicle is compensated. The motion of the host vehicle can be estimated by the sensors used for the stability control of the host vehicle or in conjunction with a road or lane estimation. The state transition can therefore be decomposed into an object state dependent linear process

$$\mathbf{x}_t(k+1) \;=\; \mathbf{F}(k)\mathbf{x}(k) + \mathbf{\Gamma}\mathbf{v}(k) \tag{4.6}$$

and an host vehicle motion induced nonlinear transformation

$$\mathbf{x}(k+1) \;=\; \mathbf{f}_{ego}\left[k, \mathbf{x}_t(k+1), \mathbf{u}_{ego}(k), \mathbf{v}_{ego}(k)\right] \tag{4.7}$$

where $\mathbf{F}$ is the state transition matrix, $\mathbf{\Gamma}$ the noise gain matrix, $\mathbf{u}_{ego}$ denotes the motion of the host vehicle and $\mathbf{v}_{ego}$ the noise of the ego-motion estimate.

Object Motion Model and State Prediction

A white noise acceleration model is utilized as a second order kinematic motion model with the state transition matrix $\mathbf{F}$ as described in [4]

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 & T & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & T & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.8}$$

where T is the, in general, time varying sampling period. The noise gain $\mathbf{\Gamma}$ and the covariance of the process noise $\mathbf{Q}_d$ are given by

$$\mathbf{\Gamma} = \begin{bmatrix} \frac{1}{2}T^2 & 0 & 0 & 0 \\ 0 & \frac{1}{2}T^2 & 0 & 0 \\ 0 & 0 & T & 0 \\ 0 & 0 & 0 & 1 \\ T & 0 & 0 & 0 \\ 0 & T & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{Q}_d = \begin{bmatrix} \sigma_{a_x}^2 & 0 & 0 & 0 \\ 0 & \sigma_{a_y}^2 & 0 & 0 \\ 0 & 0 & \sigma_z^2 & 0 \\ 0 & 0 & 0 & \sigma_\psi^2 \end{bmatrix} \tag{4.9}$$

with the covariance matrix $\mathbf{Q}$

$$\mathbf{Q} = \mathbf{\Gamma}\mathbf{Q}_d\mathbf{\Gamma}' \tag{4.10}$$

where $\sigma_{a_x}^2$ and $\sigma_{a_y}^2$ are the variances of the process noise modeling the acceleration, σ_z^2 is the variance of the process noise of the z-coordinate and σ_ψ^2 of the yaw angle. It should be

noted that the dimensions $\begin{bmatrix} d_x & d_y & d_z \end{bmatrix}'$ of the object are constant parameters of which the process noise is zero. However, the estimates of the dimensions are still time varying, as the initial covariance is non-zero. If the measurement interval T is varying over time an adaption of the variance has to be performed [6, Chapter 8.7.4]:

$$\sigma(k)^2 = [\sigma^* T(k)]^2 \tag{4.11}$$

The variance $(\sigma^*)^2$ can then be interpreted as the variance normalized to a measurement interval of one second.

The predicted state estimate $\hat{\mathbf{x}}_t(k|k-1)$ with its covariance $\mathbf{P}_t(k|k-1)$ is calculated from the estimates of the last measurement interval $\hat{\mathbf{x}}(k-1)$ and $\mathbf{P}(k-1)$

$$\hat{\mathbf{x}}_t(k|k-1) = \mathbf{F}\,\hat{\mathbf{x}}(k-1) \tag{4.12}$$

$$\mathbf{P}_t(k|k-1) = \mathbf{F}\,\mathbf{P}(k-1)\,\mathbf{F}' + \mathbf{\Gamma}\mathbf{Q}_d\mathbf{\Gamma}' \tag{4.13}$$

Ego-Motion Compensation

The motion of the host vehicle is given by the change in the position $\begin{bmatrix} \Delta x_{ego} & \Delta y_{ego} & \Delta z_{ego} \end{bmatrix}'$ and the change in the orientation $\begin{bmatrix} \Delta\phi_{ego} & \Delta\theta_{ego} & \Delta\psi_{ego} \end{bmatrix}'$, with

$$\mathbf{u}_{ego} = \begin{bmatrix} \Delta x_{ego} & \Delta y_{ego} & \Delta z_{ego} & \Delta\phi_{ego} & \Delta\theta_{ego} & \Delta\psi_{ego} \end{bmatrix}' \tag{4.14}$$

The most prominent are the longitudinal and lateral motions and the yaw rate. When using sensors without any angular resolution in the elevation, such as radar, ultrasonic or most lidar sensors, compensating for this motion is sufficient [143].

However, in the presented approach the multi-layer laser scanner and the video system deliver both measurements with an angular information in the elevation. Therefore, the ego-motion compensation must additionally account for the changes in the pitch and roll angles and in the z-coordinate.

According to Chapter 3.1 a rotation matrix $\mathbf{R}_{ego}$ can be calculated from the three angles $\begin{bmatrix} \Delta\phi_{ego} & \Delta\theta_{ego} & \Delta\psi_{ego} \end{bmatrix}'$. The ego-motion compensated, predicted state is then calculated from the predicted state $\hat{\mathbf{x}}_t(k|k-1)$ (Equation (4.12)) with

$$\hat{\mathbf{x}}(k|k-1) = \mathbf{f}_{ego}\left[\hat{\mathbf{x}}_t(k|k-1), \hat{\mathbf{u}}_{ego}(k)\right] = \begin{bmatrix} \mathbf{R}'_{ego} \cdot \left(\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} - \begin{bmatrix} \Delta x_{ego} \\ \Delta y_{ego} \\ \Delta z_{ego} \end{bmatrix} \right) \\ \psi - \Delta\psi \\ d_{x,t} \\ d_{y,t} \\ d_{z,t} \\ \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}' \begin{bmatrix} v_{x,t} \\ v_{y,t} \end{bmatrix} \end{bmatrix} \tag{4.15}$$

where r_{ij} are the entries of the rotation matrix $\mathbf{R}_{ego}$. It should be noted that apart from the three dimensional position transformation, the orientation and the velocity vector of the object are ego-motion compensated[3]. The dimensions of the object remain unaffected as it is defined in the object coordinate system.

[3] If estimated, the acceleration vector must also be ego-motion compensated.

The predicted, ego-motion compensated covariance $\mathbf{P}$ is then calculated with the predicted covariance $\mathbf{P}_t$ from Equation (4.13) and the covariance $\mathbf{P}_{ego}$ of the estimated ego-motion $\hat{\mathbf{u}}_{ego}$, with

$$\mathbf{P}(k|k-1) = \mathbf{F}_{ego} \begin{bmatrix} \mathbf{P}_t(k|k-1) & 0 \\ 0 & \mathbf{P}_{ego}(k) \end{bmatrix} \mathbf{F}'_{ego} \tag{4.16}$$

where the Jacobian $\mathbf{F}_{ego}$ of $\mathbf{f}_{ego}$ is given by

$$\mathbf{F}_{ego} = \left[\nabla_{\mathbf{x}} \mathbf{f}'_{ego}\right]' \Big|_{\mathbf{x}=\left[\hat{\mathbf{x}}'_t \;\; \hat{\mathbf{u}}'_{ego}\right]'} \tag{4.17}$$

In practise, only the ego-motion parameters Δx_{ego}, Δy_{ego} and $\Delta\psi_{ego}$ are estimated in the present system from the stability control sensor data. The change of the other parameters is however modeled by noise provided by the covariance matrix $\mathbf{P}_{ego}$, which is a diagonal matrix. The Jacobian of $\mathbf{f}_{ego}$ is then given by

$$\mathbf{F}_{ego} = \left[\nabla_{\mathbf{x}} \mathbf{f}'_{ego}\right]' \Big|_{\mathbf{x}=\left[\hat{\mathbf{x}}'_t \;\; \hat{\mathbf{u}}'_{ego}\right]', \Delta z_{ego}=0, \Delta\phi_{ego}=0, \Delta\theta_{ego}=0} \tag{4.18}$$

For the sake of simplicity some zero entries are omitted in the following Jacobian matrix

$$\mathbf{F}_{ego} = \begin{bmatrix} c & s & 0 & 0 & 0 & 0 & \cdots & -c & -s & 0 & 0 & -z_t & -s_1 + c_2 \\ -s & c & 0 & 0 & 0 & 0 & \cdots & s & -c & 0 & z_t & 0 & -c_1 - s_2 \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & s_1 - c_2 & c_1 + s_2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & c & s & \cdots & 0 & 0 & 0 & 0 & 0 & -s_1 + c_2 \\ 0 & 0 & 0 & 0 & -s & c & \cdots & 0 & 0 & 0 & 0 & 0 & -c_1 - s_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \mathbf{I}_{3\times 3} & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{4.19}$$

where $\mathbf{I}$ denotes an identity matrix. The elements are

$$\begin{aligned} s &= \sin(\Delta\psi_{ego}) \\ c &= \cos(\Delta\psi_{ego}) \\ s_1 &= \sin(\Delta\psi_{ego})\,(x_t - \Delta x_{ego}) \\ c_1 &= \cos(\Delta\psi_{ego})\,(x_t - \Delta x_{ego}) \\ s_2 &= \sin(\Delta\psi_{ego})\,(y_t - \Delta y_{ego}) \\ c_2 &= \cos(\Delta\psi_{ego})\,(y_t - \Delta y_{ego}) \end{aligned} \tag{4.20}$$

Of special interest are the elements in the third row, which exhibit the dependence of the objects' z-coordinate on the host vehicle's motion along the z-axis and the pitch and roll angle. As all three ego-motion parameters are not estimated and therefore set to zero, the z-coordinate is not affected by the ego-motion compensation. However, the uncertainty of the objects' z-coordinate increases due to the uncertainties in these motion parameters. This is a novel and important feature of this ego-motion compensation, which enables the calculation of accurate search and association gates from the predicted uncertainties in subsequent image processing algorithms.

State Update

The state update incorporates the measurements into the state estimate. For asynchronous sensors this is performed sequentially [6]. If the sensors are synchronized there are different possibilities [6, 13]. The first method is to update the measurements sequentially as in the asynchronous case. For nonlinear measurements care has to be taken, to update with the most precise measurement first. In a second method a composite measurement and covariance is computed from the individual measurements. The resulting measurement is then a weighted sum of the individual measurements in which the weights are calculated from the measurement covariances. This is however only applicable if the measurements have the same form.

A third method performs a parallel update, which is computationally more demanding than the sequential update. The sensors provide the measurement residual $\boldsymbol{\nu}_i$, the measurement covariance $\mathbf{R}_i$ and the Jacobian $\mathbf{H}_i$ of the measurement function, with

$$\mathbf{H}_i(k) = [\nabla_{\mathbf{x}} \mathbf{h}_i']' \big|_{\mathbf{x}=\hat{\mathbf{x}}(k|k-1)} \tag{4.21}$$

The measurement residual $\boldsymbol{\nu}_i$ is computed from the measurement $\mathbf{z}_i$ and the predicted measurement $\hat{\mathbf{z}}_i$

$$\boldsymbol{\nu}_i(k) = \mathbf{z}_i(k) - \hat{\mathbf{z}}_i(k|k-1) \tag{4.22}$$

The inverse of the updated covariance $\mathbf{P}(k)$ is computed from the measurement covariances $\mathbf{R}_i$ and the Jacobians $\mathbf{H}_i$ of the measurement function

$$\mathbf{P}(k)^{-1} = \mathbf{P}^{-1}(k|k-1) + \sum_{i=1}^{N_s} \mathbf{H}_i'(k) \mathbf{R}_i^{-1}(k) \mathbf{H}_i(k) \tag{4.23}$$

where N_s is the number of sensors and $\mathbf{P}(k|k-1)$ the predicted covariance. The individual contributions in the summation can be computed in parallel. The gains of the sensors are given by

$$\mathbf{K}_i(k) = \mathbf{P}(k) \mathbf{H}_i'(k) \mathbf{R}_i^{-1}(k) \tag{4.24}$$

The updated state estimate is computed according to

$$\hat{\mathbf{x}}(k) = \hat{\mathbf{x}}(k|k-1) + \sum_{i=1}^{N_s} \mathbf{K}_i(k) \boldsymbol{\nu}_i(k) \tag{4.25}$$

Synchronized sensors and a common state update are in many ways preferable, as will be discussed in Chapter 4.5.2. As the laser scanner and the video camera used in this work are synchronized, a block processing was chosen which is algebraically identical to the parallel update.

The residuals $\boldsymbol{\nu}_i$ and the linearized measurement matrices $\mathbf{H}_i$ are stacked together which are then used for the update

$$\boldsymbol{\nu}(k) = \begin{bmatrix} \boldsymbol{\nu}_1(k) \\ \vdots \\ \boldsymbol{\nu}_{N_s}(k) \end{bmatrix} \tag{4.26}$$

$$\mathbf{H}(k) = \begin{bmatrix} \mathbf{H}_1(k) \\ \vdots \\ \mathbf{H}_{N_s}(k) \end{bmatrix} \tag{4.27}$$

The block-diagonal covariance matrix $\mathbf{R}$ is constructed from the individual covariance matrices $\mathbf{R}_i$

$$\mathbf{R}(k) = \text{diag}\left[\mathbf{R}_i(k)\right] \tag{4.28}$$

The fusion module calculates from this data the residual covariance

$$\mathbf{S}(k) = \mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}'(k) + \mathbf{R}(k) \tag{4.29}$$

and the filter gain

$$\mathbf{K}(k) = \mathbf{P}(k|k-1)\mathbf{H}'(k)\mathbf{S}^{-1}(k) \tag{4.30}$$

The updated state $\hat{\mathbf{x}}$ is then given by

$$\hat{\mathbf{x}}(k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\boldsymbol{\nu}(k) \tag{4.31}$$

and its covariance is

$$\mathbf{P}(k) = \mathbf{P}(k|k-1) - \mathbf{K}(k)\mathbf{S}(k)\mathbf{K}'(k) \tag{4.32}$$

It should be noted that the covariance update in Equation (4.32) is susceptible to numerical errors. The covariance matrix can loose its positive definiteness due to the subtraction. Better alternatives are the Joseph form covariance update described in [4, Chapter 5.2] or the square-root filtering [4, Chapter 7.4] which is even less sensitive to round-off errors.

Observability

The observability of a system is important in order to guarantee convergence of the Kalman filter. Deterministic time-varying linear dynamic systems can be described in discrete time by the equations

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{F}(k)\mathbf{x}(k) + \mathbf{G}(k)\mathbf{u}(k) & (4.33)\\ \mathbf{z}(k) &= \mathbf{H}(k)\mathbf{x}(k) & (4.34) \end{aligned}$$

where k is the measurement cycle, $\mathbf{x} \in \mathbb{R}^{n_x}$ the state vector, $\mathbf{F} \in \mathbb{R}^{n_x \times n_x}$ the transition matrix, $\mathbf{u}$ the input vector and $\mathbf{G}$ the input gain, $\mathbf{z} \in \mathbb{R}^{n_z}$ the measurement vector and $\mathbf{H} \in \mathbb{R}^{n_z \times n_z}$ the measurement matrix. The state $\mathbf{x}$ of the system is completely observable, if the observability matrix

$$\mathbf{O}(k) = \begin{bmatrix} \mathbf{H}(k) \\ \mathbf{H}(k+1)\mathbf{F}(k) \\ \mathbf{H}(k+2)\mathbf{F}(k+1)\mathbf{F}(k) \\ \vdots \\ \mathbf{H}(k+n_x-1)\mathbf{F}(k+n_x-2)\cdots\mathbf{F}(k) \end{bmatrix} \tag{4.35}$$

has rank n_x for every k [30]. It is called locally observable for the time period $[k..k+n_x-1]$, if the corresponding matrix $\mathbf{O}(k)$ has rank n_x. Intuitively, observability is given if $\mathbf{H}$ is an identity matrix, thus if the measurements are directly mapped to the state. Then, one measurement cycle suffices for a complete observability, as the rank of $\mathbf{H}$ is n_x.

In the system described by Equation (4.8), the observability of the object state $\mathbf{x} = [x, y, z, \psi, v_x, v_y, d_x, d_y, d_z]'$ is not given in following cases:

1. The length or the width of a vehicle are not observed. This occurs, if only one side of the object is seen, the object is partially occluded by other objects or because of the limited FOV of the sensors.

2. The position of the object is not observable, if the laser scanner does not detect the object, as the video delivers only angular measurements and no distances.

3. In long distances the object orientation can not be determined by the laser scanner nor by the image processing algorithms.

Therefore, the convergence of the filter can not be guaranteed over the whole live cycle of an object under observation. Different measures ascertain the quality of the observability. One method investigates the diagonal elements of the covariance matrix at every cycle of the running system [30]. The higher these variances are with respect to a predefined reference, the poorer the corresponding state parameter is observable. This analysis accounts for the fact, that for some state parameters with very low process noise only few measurements, sparsely distributed over time, keep the estimated variance within reasonable limits. Another method presented in [63] determines the quality of the observability based on the eigenvectors and eigenvalues of a normalized covariance matrix.

Considering the three cases above in which the observability is not given, the following approach is suggested:

1. The length, width and height of vehicles are constant over time and the variance of the process noise is identical to zero. Therefore, their diagonal element in the covariance does not increase and the filter does not diverge[4].

2. The observability of the position estimates can be determined either by the diagonal elements in the covariance matrix or by the largest eigenvalue of the covariance submatrix of just the position entries. If the eigenvalue exceeds a predefined threshold, the position uncertainty of the object is too high and the object is deleted. Alternatively, a different object model could be used, which has less degrees of freedom.

3. Similarly the diagonal entry of the covariance matrix corresponding to the orientation is analyzed. There are different options, if this variance exceeds a threshold:

   – The object could be deleted. This would limit the detection range of the fusion system, as the orientation of cars can not be determined in the far field.
   – Reduction of the state. The orientation is set to zero or kept constant and used in the prediction as a constant parameter. Alternatively, the orientation of the velocity vector for $v_x^2 + v_y^2 \gg 0$ could be used for the prediction.

Other approaches try to prohibit a change of unobservable state parameters and their covariances, as soon as a single measurement does not provide a complete observability of all state parameters [143, Chapter 6.2]. This is attempted by setting the corresponding entries in the process noise covariance matrix $\mathbf{Q}$ and for the innovation in the measurement matrix $\mathbf{H}$ to zero. This approach is however not advisable for two reasons: First, it can be

[4]Unfortunately, due to the nonlinearity of the measurement function and numerical inaccuracies in the EKF approach, the covariance entries can slightly increase in certain situations.

shown, that the unobservable state parameters are in general still changed in the update, if the off-diagonal elements of the covariance matrix differ from zero. And secondly it is against the Kalman filter concept, in which covariances increase if there is missing information, independently from whether this is caused by missed detections or only by partial local observability.

In case the full state is not observable at the object initialization and it is clear from the system matrix that further measurements of the same form are not leading to an observability, the system state is reduced for the initialization and is augmented when the necessary information becomes available. The augmentation of the state also applies, if the state had to be reduced because of partial poor observability and then becomes again locally observable.

For the augmentation, the sensors must provide the new parameter vector $\hat{\mathbf{x}}_n$ with its covariance $\mathbf{P}_n$. After the usual filter update the state is augmented. The augmented state estimate $\hat{\mathbf{x}}_a$ is then calculated form the current state estimate $\hat{\mathbf{x}}$ and the additional state estimate $\hat{\mathbf{x}}_n$, with the state transformation function $\mathbf{g}_a$

$$\hat{\mathbf{x}}_a(k) = \mathbf{g}_a\left[\hat{\mathbf{x}}(k), \hat{\mathbf{x}}_n(k)\right] \tag{4.36}$$

The covariance of the augmented state $\mathbf{P}_a$ is derived from the Jacobian $\mathbf{G}_a$ of $\mathbf{g}_a$, the covariance of the current state $\mathbf{P}$ and the covariance of the additional state parameters $\mathbf{P}_n$, with

$$\mathbf{P}_a(k) = \mathbf{G}_a(k) \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_n \end{bmatrix} \mathbf{G}_a'(k) \tag{4.37}$$

$$\mathbf{G}_a(k) = \left[\nabla_{\mathbf{x}} \mathbf{g}_a'\right]'\Big|_{\mathbf{x}=[\hat{\mathbf{x}}_r' \;\; \hat{\mathbf{x}}_n']'} \tag{4.38}$$

This approach reflects the uncertainty of the additional state parameters in an increased covariance of the resulting augmented state. The approach however assumes the additional state parameters, which are in fact sensor measurements, to be uncorrelated with the updated state estimate.

Consider for instance, the tracking of a preceding vehicle. Only the rear of the vehicle is observable. The vehicle is passed by the host vehicle and the length becomes eventually observable. The current state estimate $\hat{\mathbf{x}}$ is then given by[5]

$$\hat{\mathbf{x}}(k) = \begin{bmatrix} x & y & \psi & d_y \end{bmatrix}' \tag{4.39}$$

where $[x, y]$ denotes the position, ψ the yaw angle and d_y the width of the tracked vehicle. The newly observed parameter is the length d_x of the vehicle. The additional state estimate is therefore given by $\hat{\mathbf{x}}_n = [d_x]$. The augmented state estimate $\hat{\mathbf{x}}_a$ from Equation (4.36) is a function of the current state estimate and the additional state (Figure 4.6)

$$\hat{\mathbf{x}}_a(k) = \begin{bmatrix} x_a \\ y_a \\ \psi_a \\ d_{x,a} \\ d_{y,a} \end{bmatrix} = \mathbf{g}_a\left[\hat{\mathbf{x}}(k), \hat{\mathbf{x}}_n(k)\right] = \begin{bmatrix} x + \frac{1}{4} d_x \cos(\psi) \\ y + \frac{1}{4} d_x \sin(\psi) \\ \psi \\ d_x \\ d_y \end{bmatrix} \tag{4.40}$$

[5]For simplicity reasons the state is given in a two dimensional coordinate system. The velocity components are omitted, as they are not affected by the augmentation process.

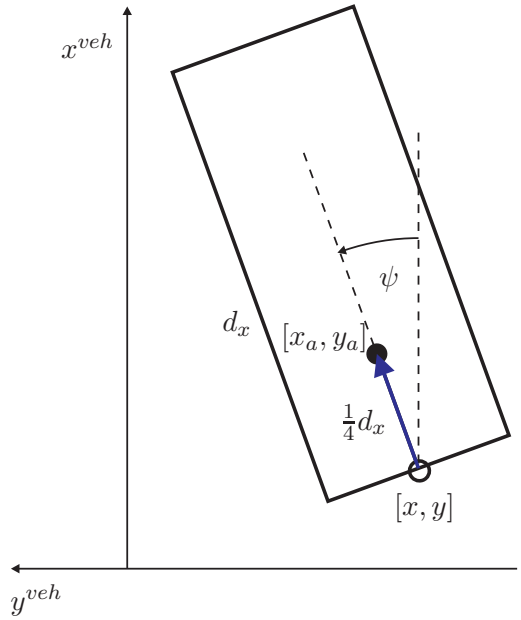


Figure 4.6: Object augmented by the length. The object coordinate system is shifted as indicated by the blue arrow.

The Jacobian $\mathbf{G}_a$ of $\mathbf{g}_a$ is provided by Equation (4.38)

$$\mathbf{G}_a = \begin{bmatrix} 1 & 0 & -\frac{1}{4}d_x \sin(\psi) & 0 & \frac{1}{4}\cos(\psi) \\ 0 & 1 & \frac{1}{4}d_x \cos(\psi) & 0 & \frac{1}{4}\sin(\psi) \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4.41}$$

Considering the calculation of the augmented covariance $\mathbf{P}_a$ given in Equation (4.37) it can be noted, that the uncertainty of the position increases depending on the uncertainty in the orientation and the length. This can be explained by the correlation of the new position estimate with the orientation and length.

In practise additional constraints must be provided by the sensors. For the expansion of a point to a line, a line to a rectangle or a rectangle to a cuboid, the direction of expansion must be known. And for the augmentation by the orientation the center of the rotation must be provided. The center of the rotation does in general not coincide with the origin of the object coordinate system, but with a point on the object's contour.

Additional to the sensors, the classification can provide a priori object dimensions and their uncertainties for an initialization or augmentation, based on the classified object type.

4.3.5 Classification

The classification as part of the vehicle environment perception for driver assistance systems either differentiates between relevant and irrelevant objects or is designed to be a multi-class problem as suggested in [48, 169].

The proposed feature-level classification is a central classification which uses discriminating features derived from the sensors. The advantage over high-level (or decision-level) fusion is in general a better performance [60, 61, 95]. A good overview of possible classification algorithms can be found in [61]. Examples of a feature-level classification of fused video and lidar data for vehicle detection is given in [146, 166].

For some sensors, which produce large amounts of data, it might be necessary to perform a preliminary classification in the sensor in order to reduce the number of hypothesized new objects. However, signatures should be provided by the sensors for the central classifier.

4.3.6 Object Management

In the tracking literature the object management is responsible for the fusion, splitting, confirmation and deletion of objects [13].

Object Fusion and Splitting

The object fusion combines several objects, if their state, type and the associated observations suggest that they are representations of the same object in reality. In contrary, splitting algorithms initialize several objects from one, if the evidence is high that there must be several objects in reality.

In this work track splitting was not found to be necessary as in most traffic scenarios the resolution of the laser scanner is high enough even for closely spaced objects. However, roadside objects are often partially fused, which is only problematic, if moving objects, such as pedestrians can not be discriminated from close stationary object. Multi hypothesis tracking approaches can perform better with additional computational cost [139].

Object Confirmation and Deletion

After the first detection of an object, it is usually not directly forwarded to the application. In order to reduce the false alarms, the object confirmation is delayed for several measurement cycles. Therefore temporal evidence is accumulated, that the object is not a ghost object produced by randomly associated clutter. An object is deleted, if the object is not supported by the sensorial observations over several cycles.

There are several different methods for the confirmation and deletion of objects, which base their decision on a track score $L(k)$. Depending on the thresholds θ_1 and θ_2 the object is confirmed, deleted or the test is continued, with

$$\begin{array}{rl} L(k) \geq \theta_2 & \text{declare object confirmation} \\ \theta_1 < L(k) < \theta_2 & \text{continue test} \\ L(k) \leq \theta_1 & \text{delete object} \end{array} \tag{4.42}$$

One method is to increase the track score, if there are measurements from at least one sensor and to decrease the score, if there are no detections in the cycle k.

$$L(k) = \begin{cases} L(k-1) + 1 & \text{at least one detection} \\ L(k-1) - 1 & \text{no detection} \end{cases} \tag{4.43}$$

An extended approach considers the reason, why the sensors do not provide any detection. It differentiates between a missing detection, where there should have been a detection and missing detections, which can be explained — for instance if the object is out of the sensors FOV or occluded by other objects.

$$L(k) = \begin{cases} L(k-1)+1 & \text{at least one detection} \\ L(k-1)-1 & \text{no detection} \\ L(k-1)-1 & \text{at least one unexplainable missed detection} \end{cases} \tag{4.44}$$

In order to realize a voting of several sensors, the increments could be chosen depending on the number of sensors supporting the object hypothesis against the number of the contradicting sensors.

An alternative method observes the estimated state covariance $\mathbf{P}$. Only if the diagonal elements P_{ii} are below a threshold, the object is confirmed. The object confirmation is therefore delayed initially, until enough information is accumulated. Additionally, the confirmation might be withdrawn from an already confirmed object, if the covariance $\mathbf{P}$ increases due to uncertain or missing sensor data.

A more sophisticated approach for the definition of the track score was developed by Sittler [130] and discussed in [12] and [13]. It defines a likelihood ratio LR to be

$$LR(k) = \frac{P_T}{P_F} = \frac{p(D|H_1)P_0(H_1)}{p(D|H_0)P_0(H_0)} \tag{4.45}$$

where H_1 is the true object hypothesis and H_0 the false alarm hypothesis. P_T is the probability associated to H_1 and P_F is the probability associated to H_0. The a priori probability of H_i is denoted by $P_0(H_i)$ and $p(D|H_i)$ represents the probability density function evaluated with the data D under the assumption that H_i is correct. With the data of K cycles, the likelihood ratio can be expressed by the product of $LR_M(k)$

$$LR(K) = \frac{P_T}{P_F} = LR_0 \prod_{k=1}^{K} LR_M(k) = LR_0 \prod_{k=1}^{K} \frac{p(D(k)|H_1)}{p(D(k)|H_0)} \tag{4.46}$$

where $LR_0 = P_0(H_1)/P_0(H_0)$ is the initial likelihood ratio. The track score $L(K)$ is then defined as the log likelihood ratio

$$L(K) = \ln\left[LR(K)\right] = \ln\left[LR_0\right] + \sum_{k=1}^{K} \ln\left[LR_M(k)\right] \tag{4.47}$$

For false alarms a uniform distribution over the measurement volume V_m is assumed

$$p(D(k)|H_0) = \frac{1}{V_m} \tag{4.48}$$

and for true observations a Gaussian distribution is assumed, with

$$p(D(k)|H_1) = \frac{1}{(2\pi)^{\frac{m}{2}}\sqrt{|\mathbf{S}(k)|}} e^{-\frac{1}{2}\boldsymbol{\nu}'(k)\mathbf{S}^{-1}(k)\boldsymbol{\nu}(k)} \tag{4.49}$$

where m denotes the dimension of the measurement vector, $\boldsymbol{\nu}$ the measurement residual vector from Equation (4.26) and $\mathbf{S}$ the measurement residual covariance from Equation (4.29). With the false track confirmation probability α and the true track deletion probability β the thresholds θ_1 and θ_2 in Equation (4.42) can be calculated

$$\theta_1 = \ln\left[\frac{\beta}{1-\alpha}\right], \qquad \theta_2 = \ln\left[\frac{1-\beta}{\alpha}\right] \tag{4.50}$$

The parameters α and β can then be defined from the systems requirements [13]. Even if the approach is theoretically sound, the underlying assumptions are unrealistic for the considered sensorial setup and its application to vehicle environment perception. False alarms are not uniformly distributed over the measurement space and depend highly on the scenario, which is in general not known. When substituting LR_M in Equation (4.47) using Equations (4.46), (4.48) and (4.49) and neglecting the initial likelihood ratio LR_0, the track score L becomes

$$\begin{aligned} L(K) &= \sum_{k=1}^{K} \ln\left[LR_M(k)\right] = \sum_{k=1}^{K} \ln \frac{p(D(k)|H_1)}{p(D(k)|H_0)} \\ &= \sum_{k=1}^{K} \left[\ln\left(\frac{V_m}{\sqrt{|\mathbf{S}(k)|}}\right) - \frac{m\ln(2\pi)}{2} - \frac{\boldsymbol{\nu}'(k)\mathbf{S}^{-1}(k)\boldsymbol{\nu}(k)}{2}\right] \end{aligned} \tag{4.51}$$

The last term produces the mahalanobis distance between the predicted and the current measurement, which is a measure for the consistency of the filter.

Therefore as an alternative track score, the time-averaged normalized innovation error squared (ANIS) defined in [4] is utilized

$$\bar{\epsilon}_\nu = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{\nu}'(k)\mathbf{S}^{-1}(k)\boldsymbol{\nu}(k) \tag{4.52}$$

If the innovations are white and with zero mean, then $K\bar{\epsilon}_\nu$ has a chi-square distribution with Km degrees of freedom, where m denotes the dimension of the measurement residual vector $\boldsymbol{\nu}$ and K the number of measurement cycles used for the averaging. In general, the dimension of the measurement vector varies over time. Therefore a different notation is chosen

$$\bar{\epsilon}_\nu = \sum_{k=1}^{K} \boldsymbol{\nu}'(k)\mathbf{S}^{-1}(k)\boldsymbol{\nu}(k), \qquad m_\Sigma = \sum_{k=1}^{K} m(k) \tag{4.53}$$

where m_Σ denotes the degrees of freedom of the chi-square distribution of $\bar{\epsilon}_\nu$. Calculating the cumulative probability Q_{χ^2} of the chi-square distribution for $\bar{\epsilon}_\nu$ with m_Σ degrees of freedom

$$Q_{\chi^2}(\bar{\epsilon}_\nu, m_\Sigma) = P_{\chi^2}\{y < \bar{\epsilon}_\nu\}, \qquad y \sim \chi^2_{m_\Sigma} \tag{4.54}$$

delivers a value in the range $Q_{\chi^2} \in [0,1]$, which can then be used as a normalized quality measure of the filter consistency over time. This quality measure can be combined with one of the first two described track score methods, which in contrast assess the quality of the object based on the existence of sensor observations.

The five different object confirmation methods are summarized below:

1. Increases the tracks score, if there are measurements from at least one sensor. Decreases the score, if there are no detections.

2. Increases the tracks score, if there are measurements from at least one sensor. Decreases the score, if there are no detections or if one sensor should have observed something, but did not.

3. Observes the estimated state covariance $\mathbf{P}$. Only if the diagonal elements P_{ii} are below a threshold, the object is confirmed. The track score is in this case a vector composed of the diagonal elements P_{ii}.

4. Assumes a uniform distribution of false alarms over the measurement space and a Gaussian distribution for true observations. The track score is the log likelihood ratio of the true object hypothesis and the false alarm hypothesis.

5. As track score serves the cumulative probability Q_{χ^2} of the time-averaged normalized innovation squared (ANIS).

The methods described above provide object quality measures for sensors which only produce detections with Gaussian distributed uncertainty and no other signal related features. In general however, the central classification is responsible for the derivation of measures upon which an application can decide if the object is relevant. These measures are then additionally based on signal related features provided by the sensors.

4.3.7 Conclusion

The main characteristics of the feature level approach are summarized in the following

- Data from other sensors support the sensor internal processing. For instance, the association gates are small due to the global state estimate. Additionally, multi modal measurement functions can be resolved using the information accumulated by all sensors.

- One sensor or several sensors can be used indirectly for the attention control of another sensor. No additional, individual interfaces are necessary, as the interface between sensors and the fusion module provides the necessary data.

- The object state must not be observable by each sensor individually in the overlapping FOV of the sensors. A video sensor providing angle only measurements can be easily fused with a sensor delivering distances and angles.

- Classification processes only slightly preprocessed data and can account for statistical dependence between sensor signals.

- Widely sensor model independent fusion process.

- Limited communication load.

In terms of the limited bandwidth of present communication systems in vehicles it might be necessary to implement the fusion process on the sensor producing the largest amount of feature data. This however does not change the fusion architecture as such.

4.4 Alternative Fusion Architectures

The feature-level fusion described above is a somewhat academic and puristic architecture, which is contrasted against a low- and high-level approach. It should, however, not be neglected that it might be advisable to use mixed-level architectures.

Multi-level fusion addresses the possibility to fuse different sensorial data in different levels [116]. Care has to be taken, that the same data is only used once in the fusion chain, in order to prohibit autocorrelated measurements. For instance, a low-level fusion of two video sensors is applied in stereo vision. In contrast, object data submitted by vehicle-to-vehicle communication would be incorporated on a high-level. Alternatively, it might be better to perform the state estimation on a feature-level and the classification on a high-level.

In order to further limit the communication load, it might be necessary to realize a hybrid fusion in which each sensor performs a classification and object tracking in order to reduce the false alarms. The features of likely objects are then forwarded to the central feature-level fusion processor. The threshold for the confirmation of objects is, however, much smaller than in purely high-level approaches and serves for the filtering of very unlikely object hypotheses.

4.5 System Latency

For very time-critical advanced driver assistance systems, such as Pre-Crash or automatic emergency brake, the overall system latency is of high interest. The advantage of such advanced driver assistance systems is drastically reduced if their reaction time is too long with respect to the high speed of passenger cars and therefore short time to contact in extreme situations.

Two different latency measures can be identified. The object detection latency defines the time span between the first detection by a sensor and the object confirmation. The state estimation latency describes, for already confirmed objects, the time period until a change in the sensor signal is reflected by a state estimate change at the output of the tracking system.

4.5.1 Object Detection Latency

The object detection latency, defined as the time span between the first detection by a sensor and the object confirmation, is mainly influenced by the object confirmation approach and the chosen parameters described in Chapter 4.3.6 or a more sophisticated object classification. As this latency usually depends on the object position, initial trajectory and appearance, standardized test objects and scenarios must be defined in order to enable a quantitative evaluation of different sensor setups and fusion approaches.

In general, the object detection latency is strongly correlated with the permitted false alarm rate. The lower the desired false alarm rate, the higher is the detection latency.

4.5.2 State Estimation Latency

The state estimation latency is defined in this context as the time span until an information measured by the sensor is reflected in the output of the fusion system — assuming an already confirmed object. This latency depends on the measurement process time, the sensor preprocessing time, the data communication speed to the fusion system and the processing time of the fusion.

The sensors in practical applications have often different sensor processing times and different measurement cycle times, which might even vary over time. Therefore, the measurements arrive at the fusion processor in general not in chronological order. Different methods for filtering these out-of-sequence observations are detailed in [13] and [6]. In practice the exact methods, which also account for the process noise effects, are difficult to realize.

Alternatively, a buffering of the measurements solves the problem of out-of-sequence observations [88, 143]. The filtering is then performed on the chronologically sorted data. This buffering however introduces an additional latency to the system, as the environment description is not updated as soon as the measurements are available. This latency depends mainly on the slowest sensor [88]. Sensors with high update rates must be buffered and therefore delayed until the measurements of the slowest sensor is available.

In this work the sensors are synchronized. The laser scanner and the video camera are both triggered by a central unit. Synchronized sensors imply several advantages:

- There are no out-of-sequence measurements.
- The prediction is performed only once per cycle. For asynchronous sensors the prediction must be performed for each sensor individually[6].
- The system latency is minimized and predictable.
- Low-level fusion on a raw data level is possible [86, 82].

Even if it is often argued, that synchronous sensors are not feasible for serial production, there are meanwhile automotive radar, lidar and video sensors in research stages available, which enable a synchronization [69, 132, 1].

Figure 4.7 shows the system schedule for the proposed feature-level fusion system with two synchronized sensors and one application (Chapter 4.3.1). The measurement process and additional preprocessing stages are denoted by "SP 1". Once the predicted object state P is transferred from the fusion unit to the sensors, the feature prediction, association and generation of new objects can be accomplished ("SP 2"). The associated features F are then transferred and fused. The resulting updated environment description U is propagated to the application where the application depending processing ("AP") takes place.

[6] In the proposed feature-level fusion architecture asynchronous sensors would lead additionally to an increased communication load as the predictions are propagated to the sensors.

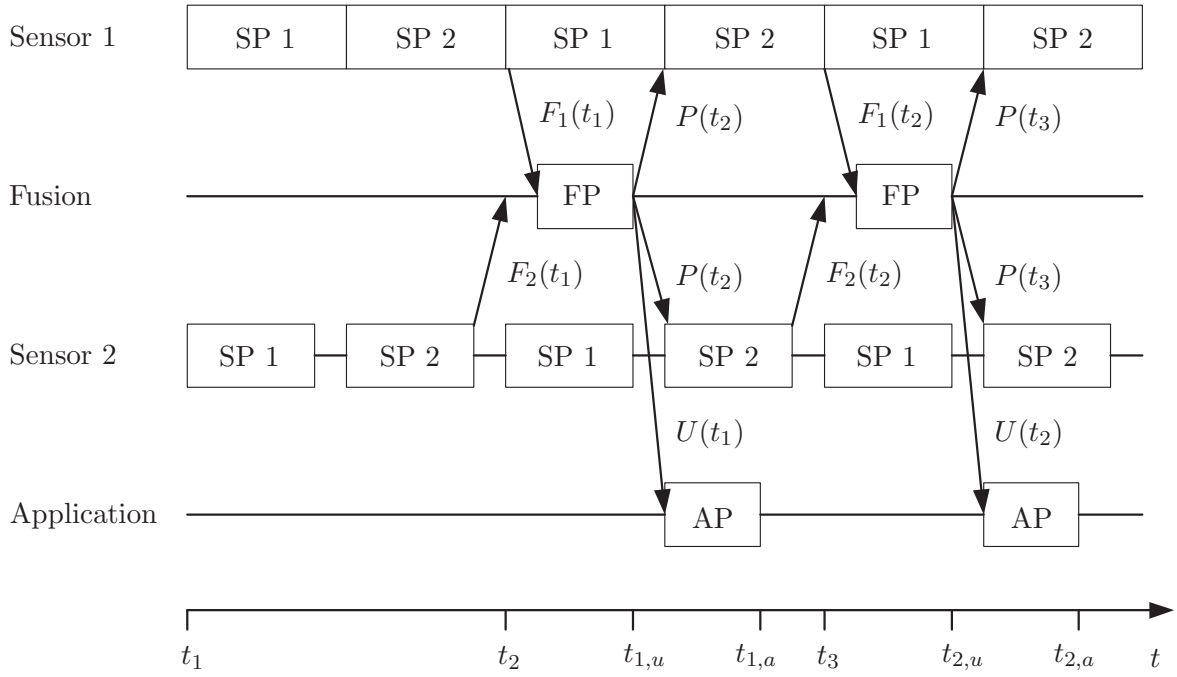


Figure 4.7: System schedule and resulting latency.

The minimal possible response time T_l, it takes a measurement to change the output of the application is therefore given by

$$T_l = t_{k,a} - t_k = \max(T(\text{SP 1})) + \max(T(\text{SP 2})) + T(\text{F}) + T(\text{AP}) + T_c \tag{4.55}$$

where t_k denotes the measurement time at the sampling interval k, $t_{k,a}$ is the time when the application processing of the cycle k is finished, $\max(T(\text{SP 1}))$ and $\max(T(\text{SP 2}))$ are the sensor processing times of the slowest sensor, $T(\text{F})$ and $T(\text{AP})$ are the fusion and the application processing times and T_c denotes the time, the communication needs.

4.6 Discussion

A feature-level fusion has been presented and contrasted to low- and high-level fusion architectures. The proposed feature-level fusion aims at an optimal tracking and classification performance, as the data is fused on a low abstraction level. The central fusion module is widely sensor unspecific and the required bandwidth for the communication between the sensors and the fusion module is limited. The advantages of the low- and the high-level approach are thus combined, while the disadvantages are kept within limits.

In the following chapters a laser scanner and a video module will be detailed. These sensor modules are integrated into the presented feature-level fusion system.

5 Laser Scanner Data based Feature Extraction and Association

The feature level fusion architecture shown in Figure 4.3 depicts a general model of a fusion approach, which comprises the sensor level data processing in several sensor modules. The central fusion process delivers the vehicle environment description to the applications. The sensor utilized in this work are a laser scanner and a video camera. The sensor modules are integrated into the feature-level fusion.

The laser scanner produces a distance profile of the environment. Apart from the distance measurements, the laser scanner provides a measure for the reflectivity of the observed object. This enables the detection of measurements of reflecting surfaces. This reflector classification is the first component of the laser scanner's preprocessing as shown in Figure 5.1. An often cited challenge for lidar sensors is the identification of measurements of the ground. These measurements must be suppressed, as the subsequent feature extraction would otherwise produce false object detections. Fortunately, the laser scanner performs measurements at different elevation levels, which increases the performance of the ground detection. The aim of the subsequent segmentation is the generation of measurement clusters which each represent an object in reality. Optimally there is exactly one segment per object and only one object per segment. The occlusion detection marks measurements, which are part of a potentially occluded object. At this early processing stage, without the a priori knowledge of the dimensions of objects, the occlusion detection can only produce hypotheses.

The feature extraction analyses each segment for characteristic features, describing a vehicle. The position of a corner of a car is searched for or the orientation, position and length of one side. These features are predicted using the objects delivered by the fusion processor (Figure 4.3). The sensor model is used in order to transform the object description into the feature space of the laser scanner. The predicted and measured features of the different objects are then associated. Extracted features which can not be associated to an object form a new object in the object generation module.

The main focus in this chapter lies on the feature extraction. Three new approaches have been developed. In the first approach, complete contour models are matched to the laser scanner measurement. The contour models describe one or two sides of a vehicle. A robust, iterative multiple-hypothesis matching approach finds possible positions and orientations of the contour models with respect to the measurements. The second approach exploits the measurements on reflecting surfaces. The number plate and the reflectors of the rear lights of cars can in general be identified as reflecting surfaces. This information enables a classification of the rear of vehicles and additionally an improvement of the orientation estimation. The third approach is applied to trucks which are passed by the host vehicle. As trucks are very long, it occurs that the rear of the truck is already

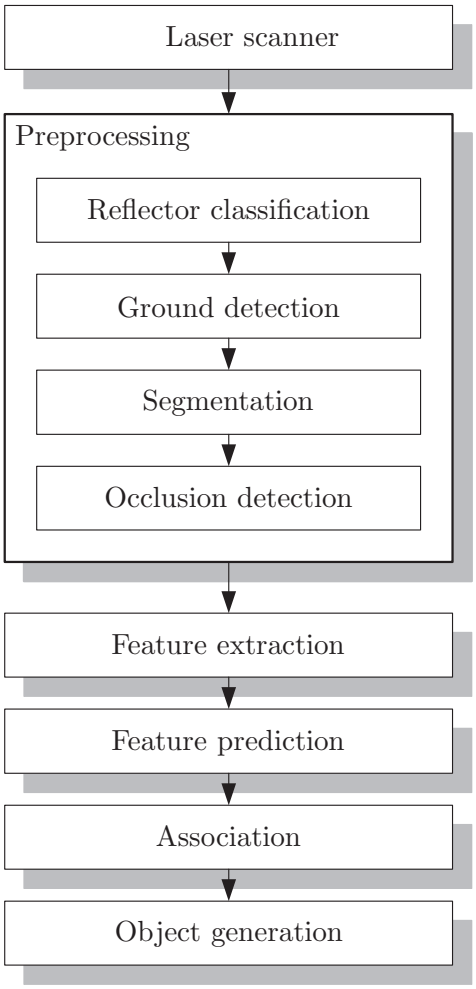


Figure 5.1: Laser scanner data processing components.

out of the field of view of the laser scanner, while the front of the truck can not be observed either. Therefore, there is no corner which provides a position measurement in the driving direction. Only the lateral position is observed. However, as the laser scanner measures partly on the wheels and the underbody of the truck, the resulting structure of all measurements exhibits the motion of the truck in the measurement sequence. This motion relative to the host vehicle supports the velocity and position estimates in the driving direction.

Following a short description of the laser scanner, each of the data processing components shown in Figure 5.1 is described in detail.

5.1 The Multi-Layer Laser Scanner

The laser scanner LD ML of the company IBEO Automobile Sensor GmbH utilized in this work has a field of view of up to 270°. The azimuth angular resolution is 0.5° at

20 Hz scan frequency [69]. The elevation aperture angle is 3.2° and divided into four measurement layers. The detection range is up to 128 m and the repetition accuracy +/− 3 cm. The laser scanner is eye-safe (laser class 1).

The laser scanner is based on the time-of-flight measurement principle. A short laser pulse is sent and the travel time of the pulse to the object and back to the sensor is measured. This time interval is proportional to the distance of the detected object. A rotating prism deflects the laser beam, thus performing a distance profile of the laser scanner's environment. Four photo diodes receive the incoming laser pulse. The spherical polar coordinates are transformed into cartesian coordinates for further data processing [52]. An exemplary scenario exhibiting the performance of the laser scanner is shown in Figure 3.9.

5.2 Preprocessing

The preprocessing performs the classification of measurements on reflectors, the ground detection, segmentation of measurements and the occlusion detection as shown in Figure 5.1. Each of these components will be briefly summarized in the following.

5.2.1 Reflector Classification

Additional to the distance measurement and the angular information, the laser scanner delivers the echo pulse width (EPW), which can be used as a measure for the irradiance of the incoming echo pulse at the receiver diode [121]. Assuming that the object surface is in all directions larger than the cross sectional area of the laser pulse, a distance depending threshold function $\theta_{epw}(d)$ can be found which classifies measurements on retroreflecting surfaces [121]. Measurements of the rear reflectors of vehicles, number plates, traffic signs or reflections post can thus be identified. The mathematical model of the threshold function is motivated from the physical properties of light emission, reflection and reception. The parameters of the model were determined from experiments with retro-reflecting objects and objects with a white surface. Figure 5.2 shows exemplarily the EPW distribution over the rear of a Volkswagen Passat at a distance of 20 m. The profile is taken at the height of the rear reflectors. The line in Figure 5.2(a) indicates the threshold $\theta_{epw}(d)$, with d =20 m. In Figure 5.2(b) the measurements are projected into the image domain. The color indicates the magnitude of the EPW. The echo pulse width of the measurements of the reflectors and the number plate is significantly higher than that of the measurements on the remaining part of the car.

In [24] the threshold function is extended for the application in the near field (< 5 m). Additionally, it could be shown in experiments that the EPW is widely independent of the overlap between the retroreflecting surface and the cross sectional area of the laser pulse. Even if the retroreflector is much smaller than the cross sectional area of the laser pulse, the same threshold function $\theta_{epw}(d)$ applies, which redundantizes the above stated assumption. This is especially useful for the classification of distant reflection posts or rear reflectors. Figure 5.3 shows the EPW of the rear reflectors of a car depending on the distance. For this experiment a rear light was separated from the car in order to prohibit an influence of the remaining car body on the EPW magnitude. The EPW is in this case

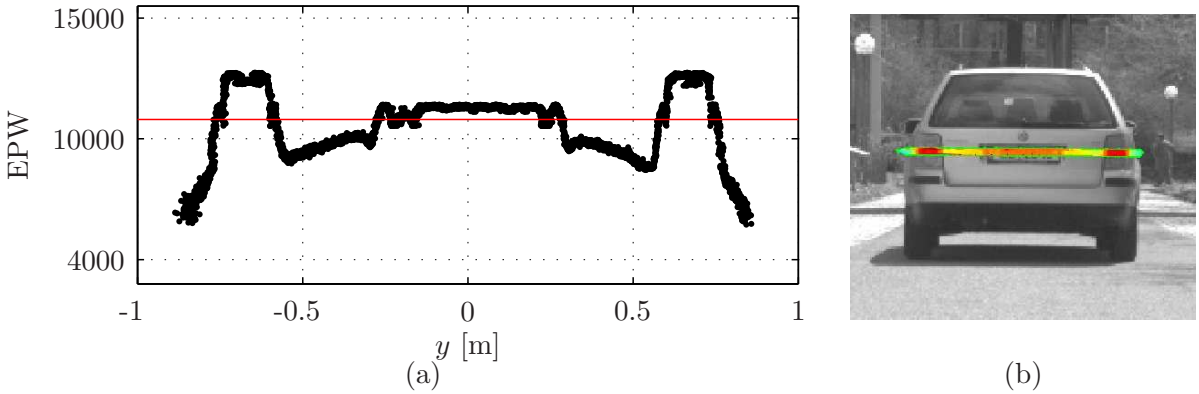


Figure 5.2: Echo pulse width (EPW) profile of the rear of a car. The red line in (a) indicates the threshold $\theta_{epw}(d)$. The measurements in (b) are color coded with respect to the EPW.

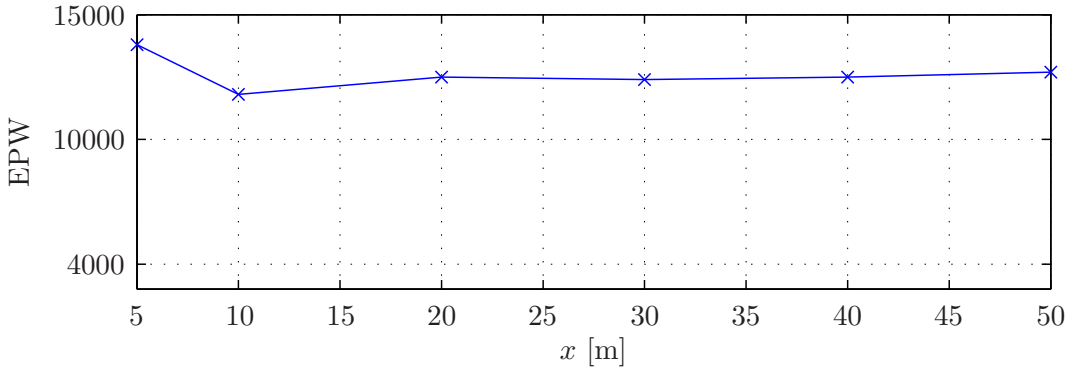
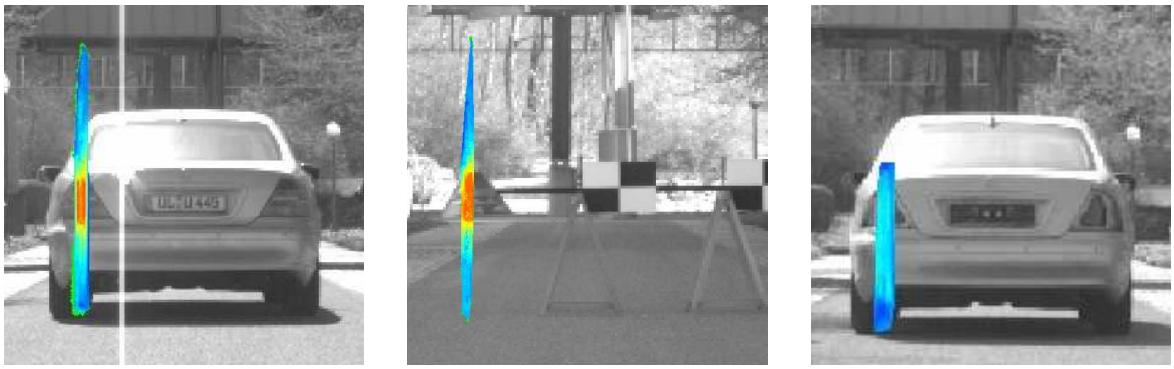


Figure 5.3: EPW of a car's rear reflector depending on the distance.

corrected in order to account for the dependence of the distance. It can be seen that the EPW is almost completely independent of the overlap between the retroreflecting surface and the cross sectional area of the laser pulse. At 50 m the cross sectional area of the laser pulse is approximately 5 times as large as the size of the reflector.

Further experiments exhibited, that objects in the field of view (FOV) of one receiving diode can provoke a signal from the other diodes. If there is a reflector in the FOV of one photo diode, the other photo diodes may produce a strong enough signal in order to enable a detection, even if there is no object in their FOV. However, this applies only for retroreflecting surfaces. The echo pulse of a Lambertian surface, which is in the FOV of only one receiver diode does not cause a false detection in the other channels. Figure 5.4 shows three pictures demonstrating this blooming effect. In Figure 5.4(a) a EPW profile in the elevation is shown. Blue regions indicate a low EPW and red regions a high EPW. As expected the EPW is high at the reflecting rear lights and low on the rest of the vehicle. However, there are measurements from above the vehicle, but with the same distance measurement. In Figure 5.4(b) the rear lights are separated from the car. The experiment clearly shows, that the rear lights produce the false detections. The car without its rear light does not exhibit any false detections (Figure 5.4(c)). In order to



(a) (b) (c)

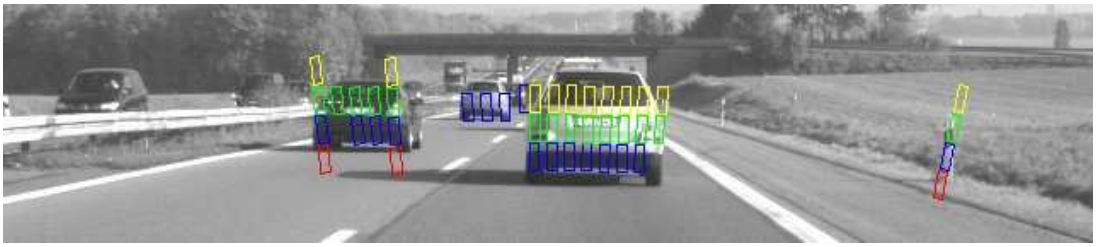
Figure 5.4: Echo pulse width profiles. The measurements are color coded with respect to the EPW. (a) shows the rear of a Mercedes S-Class. There are false detections above the roof of the car. In (b) one rear light is mounted on a horizontal pole. The false detections above and below the rear light are clearly visible. (c) shows the same car without its rear lights. There is are no false detections observable.

prohibit false detections, all detections which are in the same pulse as a measurement on a reflector are invalidated. Figure 5.5 visualizes this procedure. Figure 5.5(a) exhibits the blooming effect at the left car and the reflection post on the right. Figure 5.5(b) shows the measurements, which are classified as measurements of reflectors and in (c) the result of the blooming suppression is shown. The suppression of false detections is especially important for the determination of the height of an object.

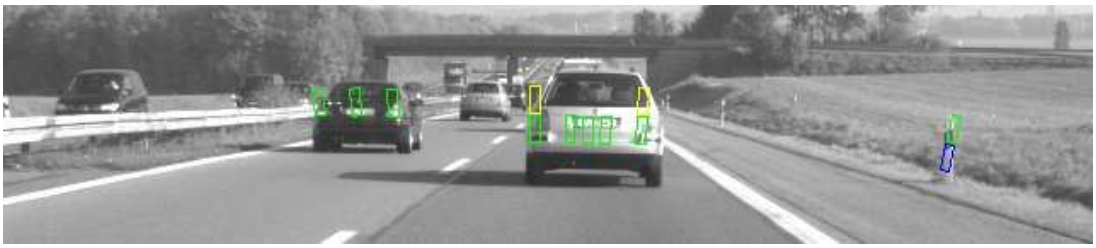
5.2.2 Ground Detection

The laser scanner performs a detection if the amplitude of the incoming laser pulse exceeds a predefined threshold. In general the amplitude of a pulse reflected from the ground is too low, as the angle of incidence between the laser pulse and the normal of the road surface is large. However, due to the pitching of the host vehicle, an uneven road, the angle of incidence becomes smaller and the amplitude of the reflected laser pulse strong enough in order to enable a detection. Other reasons for a detection might be the higher reflectivity of lane markings on the ground or the higher reflectivity of a rough surface, such as gravel.

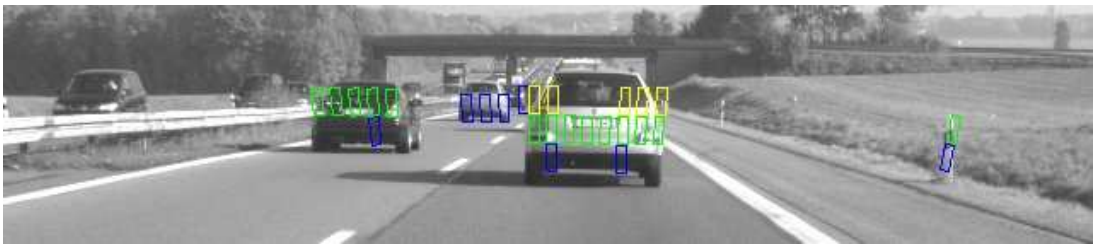
A software module provided by IBEO AS classifies each detection and marks those classified as a measurement on the ground (Figure 5.6). The parameters of the classification algorithm are chosen so as to realize a high sensitivity (detection rate). However, this leads to some detections on relevant objects which are wrongly classified as measurements on the ground. Choosing the parameters so as to realize a high precision (few false ground detections) does not solve the problem, as the false negative rate increases and some measurements on the ground are not detected.



(a)



(b)



(c)

Figure 5.5: (a) All laser scanner measurements. False detections can be observed on the car on the left lane and the reflection post on the right. (b) Only measurements of reflectors. (c) Measurements after the blooming suppression.

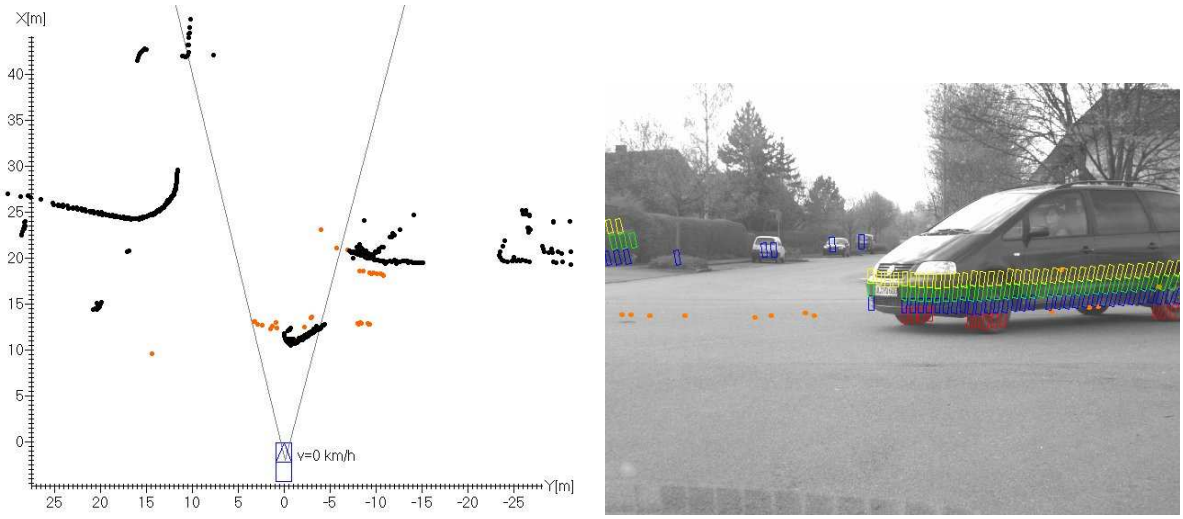
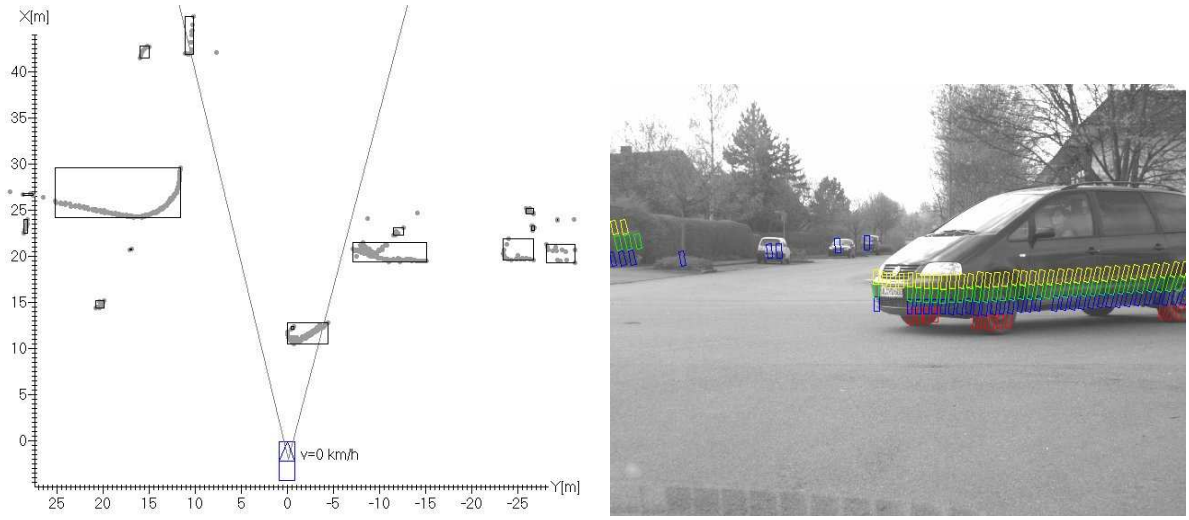


Figure 5.6: The laser scanner measurements classified as ground measurements are marked as orange dots.



(a) (b)

Figure 5.7: The rectangles in (a) indicate the bounding boxes of the segments.

5.2.3 Segmentation

In order to reduce the amount of data which has to be processed in the subsequent feature extraction modules, the laser scanner measurements are combined to segments. The aim of the segmentation is to generate clusters which each represent an object in reality. Optimally there is exactly one segment per object and only one object per segment.

The segments are created based on a distance criterion. Measurements with a small distance to neighboring measurements are included in the same segment. Both the x and y components of the distance d_x and d_y have to be below thresholds $\theta_{s,x}(x)$ and $\theta_{s,y}(y)$, which depend on the parameters $o_{s,x}$, $o_{s,y}$, $m_{s,x}$ and $m_{s,x}$ and the mean coordinates $\bar{x}$ and $\bar{y}$ of the examined measurements

$$\theta_{s,x}(\bar{x}) = \max\left(o_{s,x}, m_{s,x}\bar{x}\right) \tag{5.1}$$
$$\theta_{s,y}(\bar{y}) = \max\left(o_{s,y}, m_{s,y}\bar{y}\right) \tag{5.2}$$

The choice of the parameters is critical for the object generation. Too small thresholds $\theta_{s,x}$ and $\theta_{s,y}$ lead to a high degree of object disintegration. Single objects might be represented by several segments. However, choosing high thresholds leads to a reduced object separation. Therefore, close objects tend to be combined to only one segment. Segmentation parameters should in general be optimized with respect to the application.

A bounding box is calculated from all measurements which are assigned to a segment. The corners of the bounding box are determined from the minimal and maximal coordinates over all measurements (Figure 5.7). The height d_z^{ls} and the z-coordinate z^{ls} of the object are calculated from this bounding box. Additionally the mean position of all measurements is determined and referred as the center of gravity in the following.

5.2.4 Occlusion Detection

Another module provided by IBEO AS analyses the scan for possible partial occlusions. Individual measurements are marked if they represent the first measurement on a possibly occluded part of an object. Consequent modules which extract edges and corners of objects, can thus determine, if individual corners of the object are possibly occluded. An alternative method analyses the geometric arrangement of the segments instead of the raw measurements.

5.3 Feature Model and Feature Measurement Equation

For data association and state estimation purposes the predicted object state $\hat{\mathbf{x}}(k|k-1)$ is projected into the measurement space of the laser scanner using the measurement function $\mathbf{h}_{ls}$. The discrete-time measurement equation follows from Equation (4.5)

$$\mathbf{z}_{ls}(k) \quad = \quad \mathbf{h}_{ls}[k, \mathbf{x}(k)] + \mathbf{w}_{ls}(k) \tag{5.3}$$

The complete measurement function $\mathbf{h}_{ls}$ is constructed from the individual measurement functions described below, which depend on the observed features. The Jacobi matrix, necessary for the association and the state update, follows from Equation (4.21).

5.3.1 Contour Model Features

The position of a cuboidal object is given by the origin of the object coordinate system, which lies centrally a quarter of the object length away from the rear. (Chapter 4.3.2). As this point lies within the object contour, it is not directly observable. If the object is partially occluded, the reference point may not even be indirectly observable from a single scan. Therefore a feature model is suggested. Features like corners or sides of an object can be observed and provide the necessary measurements for a subsequent state estimation using a Kalman filter approach [23, 87].

A laser scanner observes either one or two sides of a cuboidal object. The position of one side $\mathbf{p} = \begin{bmatrix} x & y & z & 1 \end{bmatrix}'$ is represented by a point $\mathbf{p}_{r,m,m}$, $\mathbf{p}_{m,l,m}$, $\mathbf{p}_{f,m,m}$ or $\mathbf{p}_{m,r,m}$ (Figure 5.8). Furthermore the side is described by its length l^{ls}, height d_z^{ls} and its orientation ψ^{ls} given in the laser scanner coordinate system. If two sides of the object are observed, then the corner is taken as a feature. A corner feature is defined by its position $\mathbf{p}_{r,l,m}$, $\mathbf{p}_{f,l,m}$, $\mathbf{p}_{f,r,m}$ or $\mathbf{p}_{r,r,m}$, orientation ψ^{ls}, the length of the two sides l_1^{ls} and l_2^{ls} and the height d_z^{ls}.

The measurement function is composed of several transformations. First, the feature position $\mathbf{p}^{obj}$ from Table 5.1 is transformed from the object into the vehicle coordinate system using the homogeneous matrix $\mathbf{H}_{obj \to veh}$ from Equation (4.2). This transformation is followed by the inverse homogeneous matrix $\mathbf{H}_{veh \to ls} = \mathbf{H}_{ls \to veh}^{-1}$ from Equation (3.2) which results in the measurement function $\mathbf{h}_p^{ls}$ for the feature position $\mathbf{p}^{obj}$

$$\mathbf{h}_p^{ls}[k, \mathbf{x}(k), \mathbf{p}^{obj}] = \mathbf{H}_{veh \to ls} \mathbf{H}_{obj \to veh} \mathbf{p}^{obj} \tag{5.4}$$

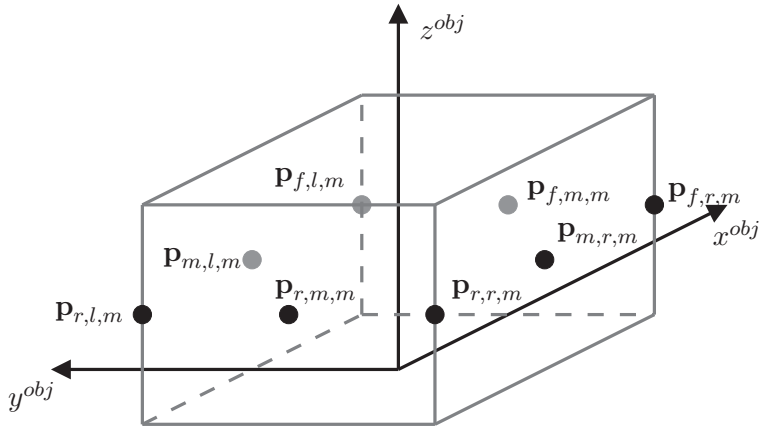


Figure 5.8: Laserscanner feature positions based on the cuboidal object model.

Table 5.1: Feature coordinates in the object coordinate system.

Feature	x^{obj}	y^{obj}	z^{obj}
$\mathbf{p}^{obj}_{r,l,m}$	$-0.25d_x$	$0.5d_y$	$0.5d_z$
$\mathbf{p}^{obj}_{r,m,m}$	$-0.25d_x$	0	$0.5d_z$
$\mathbf{p}^{obj}_{r,r,m}$	$-0.25d_x$	$-0.5d_y$	$0.5d_z$
$\mathbf{p}^{obj}_{m,l,m}$	$0.25d_x$	$0.5d_y$	$0.5d_z$
$\mathbf{p}^{obj}_{m,r,m}$	$0.25d_x$	$-0.5d_y$	$0.5d_z$
$\mathbf{p}^{obj}_{f,l,m}$	$0.75d_x$	$0.5d_y$	$0.5d_z$
$\mathbf{p}^{obj}_{f,m,m}$	$0.75d_x$	0	$0.5d_z$
$\mathbf{p}^{obj}_{f,r,m}$	$0.75d_x$	$-0.5d_y$	$0.5d_z$

The measurement function h^{ψ}_{ls} for the yaw angle ψ is given by the object orientation ψ from the state vector and the yaw angle ψ_{ls} of the calibrated laser scanner

$$h^{\psi}_{ls}[k, \mathbf{x}(k)] = \psi(k) - \psi_{ls} \tag{5.5}$$

The length, width and height measurement functions neglect the pitch and roll angles of the laser scanner, the host vehicle and the object

$$h^{d_x}_{ls}[k, \mathbf{x}(k)] = d_x \tag{5.6}$$
$$h^{d_y}_{ls}[k, \mathbf{x}(k)] = d_y \tag{5.7}$$
$$h^{d_z}_{ls}[k, \mathbf{x}(k)] = d_z \tag{5.8}$$

As it is not a priori known which feature the laser scanner observes in the current time step, the association determines a posteriori the appropriate measurement functions and assigns the measured lengths of the observed sides l_1 and l_2 to the appropriate measurement functions h^{d_x} and h^{d_y}.

5.3.2 Displacement Feature

Additionally to the position, orientation, length, width and height it is possible to measure the displacement $(\Delta x^{ls}, \Delta y^{ls})$ of the object by a matching of the point clouds of two consecutive scans. The advantage of this measurement is its independence of a cuboidal object model. It is only assumed, that the observed object is rigid.

The measurement function is composed of the difference between the projected feature position at the current measurement cycle $\mathbf{h}_p^{ls}[k, \mathbf{x}(k), \mathbf{p}^{obj}(k)]$ and the feature position at the last cycle $\mathbf{h}_p^{ls}[k-1, \mathbf{x}(k-1), \mathbf{p}^{obj}(k-1)]$.

$$\begin{aligned} h_{\Delta x}^{ls}[\mathbf{x}(k), \mathbf{x}(k-1), \mathbf{p}^{obj}(k), \mathbf{p}^{obj}(k-1)] &= h_{p,1}^{ls}[\mathbf{x}(k), \mathbf{p}^{obj}(k)] \\ &\quad - h_{p,1}^{ls}[\mathbf{x}(k-1), \mathbf{p}^{obj}(k-1)] \qquad (5.9) \\ h_{\Delta y}^{ls}[\mathbf{x}(k), \mathbf{x}(k-1), \mathbf{p}^{obj}(k), \mathbf{p}^{obj}(k-1)] &= h_{p,2}^{ls}[\mathbf{x}(k), \mathbf{p}^{obj}(k)] \\ &\quad - h_{p,2}^{ls}[\mathbf{x}(k-1), \mathbf{p}^{obj}(k-1)] \qquad (5.10) \end{aligned}$$

As feature position $\mathbf{p}^{obj}$ the observed contour feature should be chosen. If there is no contour feature observed, the origin of the object coordinate system can be chosen, with $\mathbf{p}^{obj} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}'$.

5.4 Feature Extraction

The feature extraction analyses the segments for predefined features and identifies their parameters. The features and their parameters are then used as measurements for the object state estimation and the classification. Considering cuboidal objects, features are the position of a side or a corner of an object, the orientation, the length and width. These features are extracted by matching predefined contour models to the laser scanner data.

If the echo pulse width information is available, the accuracy of the orientation estimation of vehicles ahead can be increased, by using only measurements on reflectors. As the feature extraction is then based on measurements of the same part of the object in consecutive scans, inaccuracies caused by a violation of the cuboidal object model assumption can thus be reduced.

The third part of the feature extraction estimates the displacement of objects in consecutive scans. This model free approach provides estimates of the relative motion of the object with respect to the host vehicle.

5.4.1 Matching of Contour Models

A laser scanner observes either one or two sides of a cuboidal object. Recent approaches estimate the position and orientation of the two sides individually [140, 143, 46] and then identify object corners or construct object boxes. A new approach is suggested, which approximates the whole observed contour by predefined contour models [23, 87]. This imposes additional constraints on the matching process, which increases the accuracy of the feature extraction process.

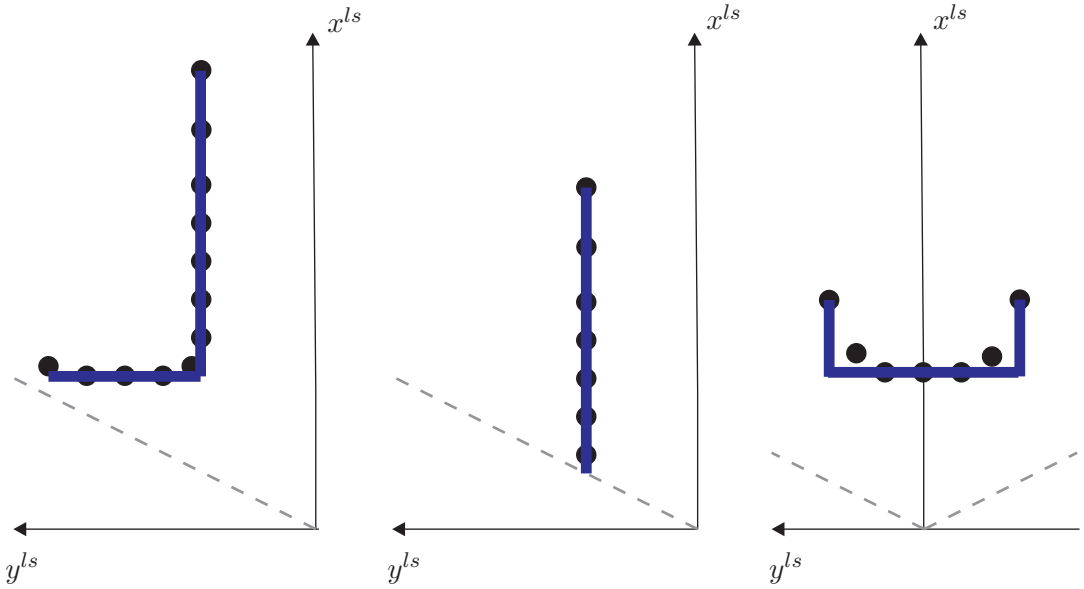


Figure 5.9: Contour models for different laser scanner measurement distributions. From left to right the L-shape, the I-shape and the U-shape are shown. The dashed lines indicate the FOV of the laser scanner.

The proposed contour models are shown in Figure 5.9. One ore two observed sides are represented by an I-shape or L-shape. The U-shape has its advantages for the representation of rears or fronts of vehicles with round corners. The usage of the U-shape can inhibit an incorrect application of an L-shape. Each model $\mathcal{M}_i \in \mathbb{M}$ is composed of L line elements l_j, which are defined by the two end points $\mathbf{a}_j$ and $\mathbf{b}_j$.

Model Preselection

In order to reduce the computational burden of the model matching, a preselection of possible contour models is performed. Applying a principal component analysis, the statistical properties of the measurement distribution included in a segment can be derived. The statistical properties indicate which contour models can be excluded from the further processing steps. A covariance matrix $\mathbf{M}$ is calculated from the x- and y-coordinates of the laser scanner measurements of one segment and their sample means $(\bar{x}, \bar{y})$

$$\mathbf{M} = \frac{1}{N-1} \sum_{i=1}^{N} \begin{bmatrix} x_i - \bar{x} \\ y_i - \bar{y} \end{bmatrix} \begin{bmatrix} x_i - \bar{x} & y_i - \bar{y} \end{bmatrix} \tag{5.11}$$

where N is the number of measurements. The eigenvalues λ_1 and λ_2 (with $\lambda_1 \geq \lambda_2$) of the covariance matrix $\mathbf{M}$ represent the variance of the analysed measurement distribution in the direction of the corresponding eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$. The quotient of the eigenvalues $\frac{\lambda_1}{\lambda_2}$ indicates if the data varies only in one or in two directions. If the quotient is above a predefined threshold $\frac{\lambda_1}{\lambda_2} \geq \theta_\lambda$, only the I-shape is tested. For $\frac{\lambda_1}{\lambda_2} < \theta_\lambda$ all three models are considered in the model matching. The threshold is chosen empirically and serves mainly as a reduction of the computational burden.

Model Matching

The model matching tries to fit a model optimally to a given set of data. A transformation, given by a rotation $\mathbf{R}_\psi$ around the z-axis and a translation $\mathbf{t} = \begin{bmatrix} t_x & t_y \end{bmatrix}'$, is optimized, which transforms the point set $\mathbb{P}$, consisting of the measurements $\mathbf{p}_i$ of one segment

$$\mathbf{p}_i^* = \mathbf{R}_\psi \mathbf{p}_i + \mathbf{t} \tag{5.12}$$

The optimization criterion is the mean $\bar{d}$ of the distances $d\left(\mathbf{p}_i^*, \mathcal{M}\right)$ between a transformed point $\mathbf{p}_i^* \in \mathbb{P}'$ and the model $\mathcal{M}$

$$\bar{d} = \frac{1}{N} \sum_{i=1}^{N} d\left(\mathbf{p}_i^*, \mathcal{M}\right) \rightarrow \min \tag{5.13}$$

As an optimization algorithm the Iterative-Closest-Point (ICP) is a well known method [10]. However, this approach does not account for outliers[1] in the data, which are common in the measurements of vehicles. The Weighted-ICP (WICP) assigns a weight w_i to the measurements according to their distance to the data. Through the choice of the weighting function, which will be explained later, outliers can be detected.

The WICP is composed of the following steps:

1. Determination of the initial transformation $\mathbf{R}_\psi(k=0)$, $\mathbf{t}(k=0)$ and application of (5.12).
2. Calculation of the corresponding point set $\mathbb{Q}$ for $\mathbb{P}'$ which minimizes $d\left(\mathbf{p}_i^*, \mathcal{M}\right)$.
3. Identification of the weights w_i for each $\mathbf{p}_i^* \in \mathbb{P}'$.
4. Optimization of the transformation $\mathbf{R}_\psi(k)$, $\mathbf{t}(k)$ and its application using (5.12).
5. Calculation of mean distance $\bar{d}(k)$ according to Equation (5.13).
6. Repeat from **step 2** for a maximal number of iterations or until the change of the mean distance is smaller than a predefined threshold $\|\bar{d}(k) - \bar{d}(k-1)\| < \epsilon$.

A known problem of ICP algorithms is, that they are prone to end up in local minima. Therefore, the choice of a good initial transformation $\mathbf{R}_\psi(k=0)$, $\mathbf{t}(k=0)$ in **step 1** is crucial. For the I-shape, the translation $\mathbf{t}(k=0)$ is determined by the mean of the measurements in the segment and the orientation by the first eigenvector $\mathbf{v}_1$ of the principle component analysis. The same applies for the U-shape. An L-shape applies only, if the angle between the laser beam and the sides of the observed object are approximately 45°. The rotation is chosen accordingly. The translation is determined so that the corner of the L-shape fits with the mean of the measurements. This choice of the initial transformation exhibits good performance applied to real laser scanner data. In order to further increase the probability of finding the global optimum, a Monte Carlo approach using several different initial transformations was considered. However, due to the computational effort implied by Monte Carlo methods, this approach was abandoned.

[1] Outliers are those measurements, which violate the cuboidal object model assumptions. Measurements of the exhaust of cars or inside the wheel guard are therefore treated as outliers.

The size of the I-shape and L-shape are initially infinite. Only after the matching process, the length of the I-shape and the length of the two sides of the L-shape are determined. The U-shape is matched to the data after the I-shape. The length of the I-shape is taken as the length of the base line of the U-shape. The two remaining sides are initially infinitely long and shortened when the final transformation is found.

In **step 2** a corresponding point on the contour model is calculated for each measurement. The optimal transformation, in the least square sense, between the two point sets is then given as a closed form solution. The corresponding point $\mathbf{q}_i$ is the point on the model $\mathcal{M}$ with the minimal distance $d\left(\mathbf{p}_i^*, \mathbf{q}_i\right)$ to the measurement $\mathbf{p}_i^*$. As the model is composed of several line segments, the distance between the measurement and the model $d\left(\mathbf{p}_i^*, \mathcal{M}\right)$ is given by the minimum of the distances $d\left(\mathbf{p}_i^*, l_j\right)$ between the point $\mathbf{p}_i^*$ and the line segments l_j

$$d\left(\mathbf{p}_i^*, \mathcal{M}\right) = d\left(\mathbf{p}_i^*, \mathbf{q}_i\right) = \min_{j \in \{1,\ldots,L\}} d\left(\mathbf{p}_i^*, l_j\right) = \min_{j \in \{1,\ldots,L\}} d\left(\mathbf{p}_i^*, \mathbf{q}_{i,j}\right) \tag{5.14}$$

The correspondence points $\mathbf{q}_{i,j}$ are calculated for each measurement $\mathbf{p}_i^*$ and line segment l_j, with

$$\mathbf{q}_{i,j} = \mathbf{a}_j + \lambda_{i,j}\left(\mathbf{b}_j - \mathbf{a}_j\right) \tag{5.15}$$

$$\lambda_{i,j} = \begin{cases} \lambda'_{i,j} & 0 \leq \lambda'_{i,j} \leq 1 \\ 0 & \lambda'_{i,j} < 0 \\ 1 & \lambda'_{i,j} > 1 \end{cases} \tag{5.16}$$

$$\lambda'_{i,j} = \frac{\left(\mathbf{p}_i^* - \mathbf{a}_j\right)\left(\mathbf{b}_j - \mathbf{a}_j\right)}{\left\|\mathbf{b}_j - \mathbf{a}_j\right\|^2} \tag{5.17}$$

where $\mathbf{a}_j$ and $\mathbf{b}_j$ denote the two end points on the line segment [173, 174].

In **step 3** a weight w_i is assigned to each point $\mathbf{p}_i^*$ according to its distance $d\left(\mathbf{p}_i^*, \mathcal{M}\right)$ to the model $\mathcal{M}$. In [97] a linear weighting function is suggested. If the distance exceeds a threshold d_{max} the point is classified as an outlier

$$w_i = \begin{cases} 1 - \frac{d(\mathbf{p}_i^*, \mathcal{M})}{d_{max}} & \text{if } d(\mathbf{p}_i^*, \mathcal{M}) \leq d_{max} \\ 0 & \text{if } d(\mathbf{p}_i^*, \mathcal{M}) > d_{max} \end{cases} \tag{5.18}$$

In [173] an adaptive choice of the distance threshold d_{max} is suggested, which depends on the mean distance $\bar{d}(k)$ and the expected mean distance $E\{\bar{d}\}$ in the global optimum given a Gaussian distributed measurement noise. Furthermore statistics of the distances $d(\mathbf{p}_i^*, \mathcal{M})$ are used, which are the sample deviation σ_d and median ζ_d.

$$d_{max} = \begin{cases} \overline{d} + 3\sigma_d & \text{if } \bar{d} < E\{\bar{d}\} \\ \overline{d} + 2\sigma_d & \text{if } E\{\bar{d}\} \leq \bar{d} < 3E\{\bar{d}\} \\ \overline{d} + \sigma_d & \text{if } 3E\{\bar{d}\} \leq \bar{d} < 6E\{\bar{d}\} \\ \zeta_d & \text{if } 6E\{\bar{d}\} \leq \bar{d} \end{cases} \tag{5.19}$$

Once the correspondence point set is found, the optimal transformation between the measurement point set $\mathbb{P}$ and correspondence point set $\mathbb{Q}$ is calculated in **step 4**. With the mean of the measurements $\bar{\mathbf{p}}$, the correspondence points $\bar{\mathbf{q}}$ and the sum of the weights w_Σ, the covariance matrix $\mathbf{N}$ is given by

$$\mathbf{N} = \begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix} = \frac{1}{w_\Sigma} \sum_{i=1}^{N} w_i \left(\mathbf{p}_i - \bar{\mathbf{p}}\right) \left(\mathbf{q}_i - \bar{\mathbf{q}}\right)' \tag{5.20}$$

The rotation ψ and the corresponding rotation matrix $\mathbf{R}_\psi(k)$ are derived from the covariance matrix

$$\psi = \arctan \frac{n_{12} - n_{21}}{n_{11} + n_{22}} \tag{5.21}$$

Applying the rotation matrix, the translation $\mathbf{t}(k)$ is determined with the weighted means $\bar{\mathbf{p}}_w$ and $\bar{\mathbf{q}}_w$

$$\mathbf{t}(k) = \bar{\mathbf{q}}_w - \mathbf{R}_\psi(k) \bar{\mathbf{p}}_w \tag{5.22}$$

with

$$\bar{\mathbf{p}}_w = \frac{1}{w_\Sigma} \sum_{i=1}^{N} w_i \mathbf{p}_i \qquad \bar{\mathbf{q}}_w = \frac{1}{w_\Sigma} \sum_{i=1}^{N} w_i \mathbf{q}_i \tag{5.23}$$

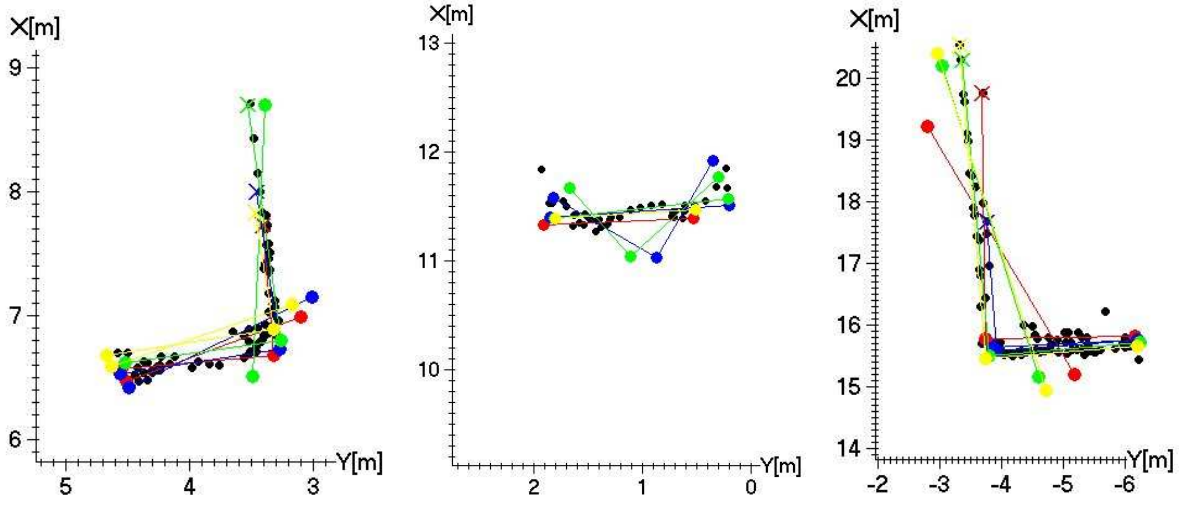
After the model matching process with the WICP, the length of the line segments is shorten in order to fit the measurements. The first and the last correspondence points determine the beginning and the end of the contour.

Model Selection

The three contour models are matched to the measurements of each of the four scan layers separately. Therefore a maximum of 12 hypotheses are generated (Figure 5.10). Different criteria were developed in order to select one hypothesis:

- The mean distance error $\bar{d}$.
- The number of measurements assigned to the hypothesis. Too few measurements of one hypothesis compared to other hypotheses indicate a weak data support.
- The length of the model contour. Long model contours are more likely to represent the whole object contour.

These parameters are based on the measurement process itself. However, these parameters do not necessarily provide one significantly best solution. Therefore several shape hypotheses are preserved and temporal evidence is incorporated in the association (Chapter 5.5.3). The Mahalanobis-distance in Equation (5.44) between the predicted feature $\mathbf{z}(k|k-1)$ and the measured feature $\mathbf{z}(k)$ provides a criterion, which is used in the association in order to solve the ambiguities between several shape hypotheses.



(a) (b) (c)

Figure 5.10: Different hypotheses of contour model matches. The laser scanner measurements are indicated in black. In (a) the green L-shape is finally selected in (b) the blue I-shape and in (c) the yellow L-shape.

Feature Parameter Observation

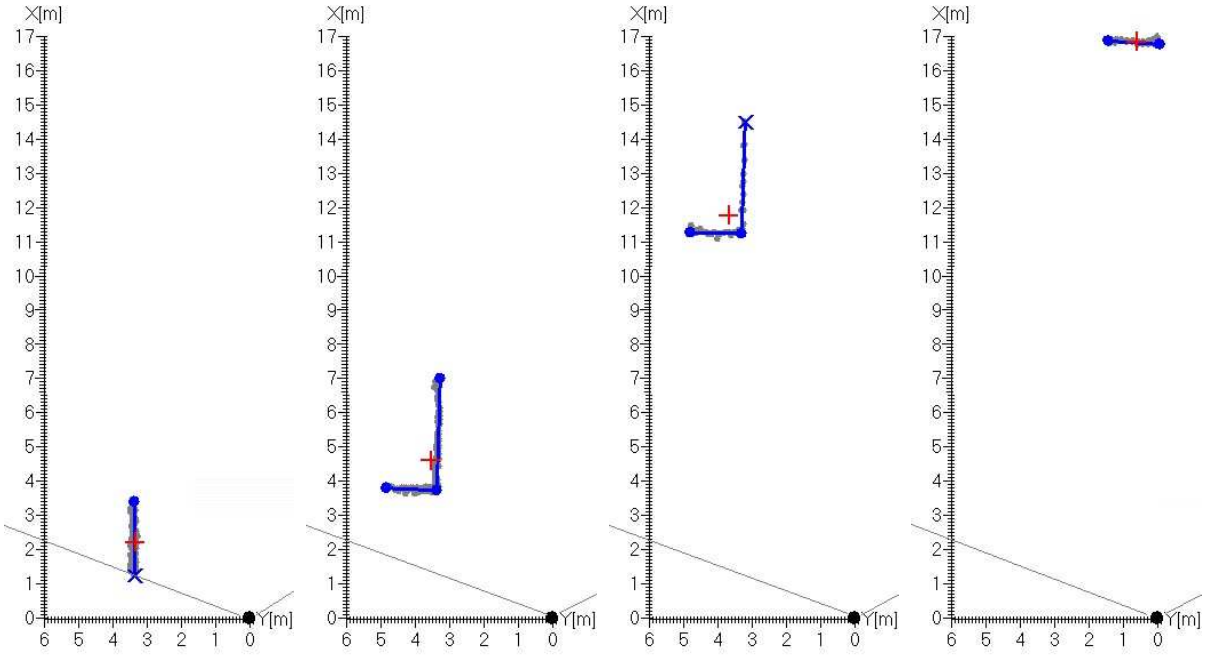
The measurement feature parameters are not always all observed. There are several reasons:

- An object is not completely in the sensor's field of view (Figure 5.11(a)).
- An object is partially occluded by other objects (similar to the case above).
- The sensor characteristics do not allow the observation (Figure 5.11(c)).
- Self-occlusion of an object (Figure 5.11(d)).

Figure 5.11 shows laser scanner data at different time instances of a car passing the host vehicle on the left side. In Figure 5.11(a) only the right corner of the front and part of the right side are seen. The position of the right corner of the front $\mathbf{p}_{f,r,m}$ and the orientation are therefore provided as measurements. The width w and length l can not be measured in this time instance. In Figure 5.11(b) the measurement vector is fully observed. The right corner of the rear $\mathbf{p}_{r,r,m}$ is used as the feature position. In Figure 5.11(c) the right corner of the front can not be observed as the angle of incidence is too large and the laser scanner pulses may be totally reflected. In Figure 5.11(d) the object occludes itself and the length cannot be observed. In the last two cases, the position of the feature $\mathbf{p}_{r,m,m}$, the orientation and the width are measured.

5.4.2 Exploiting the Echo Pulse Width Information

The echo pulse width information enables the classification of measurements on retroreflecting surfaces (Chapter 5.2.1). Measurements on the rear reflectors of vehicles or num-



(a) (b) (c) (d)

Figure 5.11: Laser scanner data at different time instances of a car passing the host vehicle on the left side. I-shapes and L-shapes are fitted to the data. An "X" at the end of line segments indicates that the corner is not observed. The center of gravity is shown as a red cross. The field of view of the laser scanner is indicated.

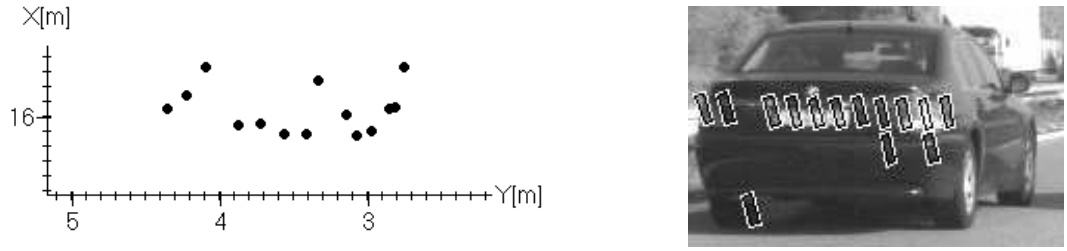
ber plates can thus be identified. A new approach uses exclusively the measurements on reflectors for an orientation estimation of vehicles seen from the rear [24]. A line is fitted into the measurements assuming a cuboidal object model. The advantage of this approach over a feature extraction using all measurements is, that the reflectors of a vehicle usually define a line, which is perpendicular to the vehicle's orientation. Measurements on the exhaust, inside the car or on the round parts of a rear of modern cars, which violate the cuboidal object model, are inhibited (Figures 5.12 and 5.13).

The parameters of the linear model are estimated by the method of the least squares incorporating the measurement error covariance matrix [51]. The line and the measurement model are given by

$$x_i = \beta_0 + \beta_1 y_i \tag{5.24}$$

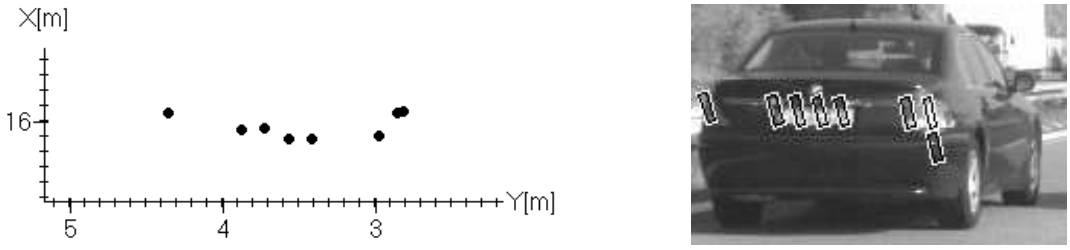
$$\begin{bmatrix} x_{m,i} \\ y_{m,i} \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} e_i \\ u_i \end{bmatrix} \tag{5.25}$$

where $\begin{bmatrix} x_{m,i} & y_{m,i} \end{bmatrix}'$ are the coordinates of the measurement i, $\begin{bmatrix} x_i & y_i \end{bmatrix}'$ denote the true values and $\begin{bmatrix} e_i & u_i \end{bmatrix}'$ is the vector of the normally distributed measurement errors. The parameters β_0 and β_1 describe the line.



(a) (b)

Figure 5.12: All measurements of the laser scanner are shown in (a). The corresponding image region is shown in (b) with the laser scanner measurements projected into the image domain. The measurement diverging from a straight line originate mainly from the exhaust or the vehicles side.



(a) (b)

Figure 5.13: Only laser scanner measurements on reflectors are shown in (a). In (b) the corresponding image region is shown with the remaining laser scanner measurements on reflectors. Using only measurement on reflectors, the measurement on the exhaust or the vehicles side can be omitted.

Measurement Error Covariance Matrix

The measurement error covariance matrix in polar coordinates $\mathbf{R}_p^{ls}$ is composed of the error variance in radial direction σ_r^2 and in angular direction σ_α^2

$$\mathbf{R}_p^{ls} = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix} \tag{5.26}$$

By linearization of the transformation at the operating point (r_0, α_0) the cartesian covariance matrix $\mathbf{R}_c^{ls}$ can be deduced

$$\mathbf{R}_c^{ls} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{bmatrix} \tag{5.27}$$

with

$$\sigma_x^2 = \sigma_r^2 \cos^2(\alpha_0) + \sigma_\alpha^2 r_0^2 \sin^2(\alpha_0) \tag{5.28}$$

$$\sigma_y^2 = \sigma_r^2 \sin^2(\alpha_0) + \sigma_\alpha^2 r_0^2 \cos^2(\alpha_0) \tag{5.29}$$

$$\sigma_{xy}^2 = \sigma_r^2 \sin(\alpha_0)\cos(\alpha_0) - \sigma_\alpha^2 r_0^2 \sin(\alpha_0)\cos(\alpha_0) \tag{5.30}$$

Least Squares Parameter Estimation Incorporating Measurement Errors

The following calculation of the regression line is only briefly outlined. A detailed derivation can be found in [51]. For the estimation of the parameters describing the model in Equation (5.24) it is assumed that all measurements have the same measurement error covariance matrix. The mean of all measurements $\begin{bmatrix}\bar{x}_m & \bar{y}_m\end{bmatrix}'$ is calculated. The measurement error covariance matrix $\mathbf{R}_c^{ls}$ is linearized at this point.

The least squares problem can be formulated by the sum of the Mahalanobis distances over all N measurements

$$\sum_{i=1}^{N} \begin{bmatrix} x_{m,i} - \hat{\beta}_0 - \hat{\beta}_1 \hat{y}_i & y_{m,i} - \hat{y}_i \end{bmatrix} (\mathbf{R}_c^{ls})^{-1} \begin{bmatrix} x_{m,i} - \hat{\beta}_0 - \hat{\beta}_1 \hat{y}_i \\ y_{m,i} - \hat{y}_i \end{bmatrix} \rightarrow \min \tag{5.31}$$

where $[\hat{\beta}_0, \hat{\beta}_1]$ are the estimated parameters and $\hat{y}_i$ denotes the estimate of the true y-coordinate of the measurement i.

For the solution the sample covariance matrix is calculated

$$\mathbf{M}_z = \begin{bmatrix} m_x^2 & m_{xy}^2 \\ m_{xy}^2 & m_y^2 \end{bmatrix} = \frac{1}{N-1} \sum_{i=1}^{N} \begin{bmatrix} x_{m,i} - \bar{x}_m \\ y_{m,i} - \bar{y}_m \end{bmatrix} \begin{bmatrix} x_{m,i} - \bar{x}_m & y_{m,i} - \bar{y}_m \end{bmatrix} \tag{5.32}$$

With the smaller value $\hat{\lambda}$ of the solution to the quadratic equation

$$\left| \mathbf{M}_z - \lambda \mathbf{R}_c^{ls} \right| = 0 \tag{5.33}$$

the slope of the line $\hat{\beta}_1$ is given by

$$\hat{\beta}_1 = \frac{m_{xy}^2 - \hat{\lambda}\sigma_{xy}^2}{m_y^2 - \hat{\lambda}\sigma_x^2} \tag{5.34}$$

Additionally the variance $\hat{\sigma}_{\beta_1}^2$ of $\hat{\beta}_1$ can be estimated

$$\hat{\sigma}_{\beta_1}^2 = \frac{\hat{\mathbf{h}}_2' \left(\mathbf{M}_z - \mathbf{R}_c^{ls}\right) \hat{\mathbf{h}}_2 \hat{\sigma}_v^2 + \sigma_x^2 \hat{\sigma}_v^2 - \hat{\sigma}_{xv}^2}{(N-1) \left[\hat{\mathbf{h}}_2' \left(\mathbf{M}_z - \mathbf{R}_c^{ls}\right) \hat{\mathbf{h}}_2\right]^2} \tag{5.35}$$

with

$$\hat{\sigma}_{xv}^2 = \sigma_{xy}^2 - \hat{\beta}_1 \sigma_x^2, \tag{5.36}$$

$$\hat{\sigma}_v^2 = \begin{bmatrix} 1 & \hat{\beta}_1 \end{bmatrix} \mathbf{R}_c^{ls} \begin{bmatrix} 1 \\ -\hat{\beta}_1 \end{bmatrix} \tag{5.37}$$

$$\hat{\mathbf{h}}_2' = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ -\hat{\beta}_1 \end{bmatrix} \hat{\sigma}_v^{-2} \hat{\sigma}_{uv}^2 \tag{5.38}$$

The estimated orientation $\hat{\psi}$ is calculated from the estimated slope $\hat{\beta}_1$

$$\hat{\psi} = -\arctan(\hat{\beta}_1) \tag{5.39}$$

The estimated variance $\hat{\sigma}_{\psi}^2$ can be deduced from $\sigma_{\beta_1}^2$ by linearization of the transformation at the operation point

$$\hat{\sigma}_{\psi}^2 = \left(\frac{1}{1 + \hat{\beta}_1^2} \right)^2 \hat{\sigma}_{\beta_1}^2 \tag{5.40}$$

The advantage of this least square approach over standard linear regression is twofold. First, it accounts for the measurement errors which affect the observation in both coordinates. Secondly, the variance of the estimated orientation of the line can be calculated. Therefore, the number of measurements and the variance of the measurements along the line influence the estimated variance in the orientation.

However, the reflectors of the rear of vehicles are not always all observed. In large distances the EPW might be too low due to dirty reflectors. At higher measurement frequencies the reflectors might be missed, due to the discrete sampling of the laser scanner, or in close range the vertical aperture can be too small. Therefore, the orientation, estimated by the contour model based feature extraction, is replaced by the orientation of the reflecting surface only if reflectors are observed and if the estimated uncertainty is lower.

Classification of the Rear of Vehicles

A Classification usually aims at a differentiation between relevant and irrelevant objects. Multi-class problems additionally distinguish between several classes of relevant objects, like cars, trucks, pedestrians and bikes [169]. In contrast the classification of the rear of vehicles is part of the dynamic state estimation. It delivers an unambiguous orientation for stationary vehicles observed from the rear.

The orientation of moving objects can be determined by their motion vector. However, the orientation of stationary objects can not be estimated using the velocity. The orientation of stationary objects is solely given by the contour based feature extraction methods, which only delivers an ambiguous orientation. The feature extraction can not differentiate between a vehicle seen from the rear, front or from the side. The classification of the rear of vehicles therefore fills this gap and delivers an unambiguous orientation. This is especially useful for a situation assessment in rear end collisions with stationary vehicles, as discussed in Chapter 10.

The classification exploits the presence of reflectors on the rear of vehicles, which can be identified due to their higher EPW (Chapter 5.2.1). The classification is composed of a rule-based part and a probabilistic multi-feature classification based on the belief theory proposed by Dempster-Shafer [95, 24].

The rule-based part rejects those segments, which do not contain any measurements on reflectors or of which the dimensions do not match the size of a rear of a car or truck.

For the probabilistic part of the classification, the orientation and its sample deviation are used. Additional features are extracted from the distribution of the measurements [24]:

- Width of the reflecting surface.
- Width over all measurements.



(a) (b)

Figure 5.14: Classification result for a situation on a freeway. (a) Bird's eye view of only the measurements of reflectors. The blue boxes indicate the classified rears of the vehicles. (b) Image with projected laser scanner measurements.

- Quotient of the width of the reflecting surface and the width of the complete surface.
- Normalized measurement density, which is the quotient of the number of measurements on the surface and the maximally expected number of measurements. The measurement density is calculated for each scan layer separately and the highest value is selected.

For each feature i the probability masses of the class "rear" $m_i(a_1)$, the class "other objects" $m_i(a_2)$ and the uncertainty $m_i(\Theta)$ are calculated. For each segment the probability masses of the features are then combined to the total probability masses $m(a_1)$, $m(a_2)$ and $m(\Theta)$ using the Dempster-Shafer combination rule [95]. A detailed description of the classification scheme, an empirical derivation of the underlying parameters and an evaluation can be found in [24].

Figure 5.14 shows exemplarily a classification result for a situation on a freeway. The red dots represent measurements on reflectors. The green dots in Figure 5.14(b) depict the remaining laser scanner measurements which are omitted in the bird's eye view. The blue boxes indicate the classified rears of the vehicles with their measured orientation. The classification and orientation estimation is performed on a single frame basis. It should be noted, that in this image all vehicles, which are not occluded, are correctly

classified. The breakdown triangle on the side-strip in a distance of 16 m is not classified as a vehicle's rear even if there are measurements of reflectors.

In this situation the classification exhibits its potential, as especially the stationary vehicles on the side-strip are correctly classified. The unambiguous orientation of these stationary vehicles can only be determined by the classification of the rear and not by the velocity vector.

5.4.3 Estimation of Object Displacement in Consecutive Scans

The contour model based feature extraction described in Chapter 5.4.1 produces an estimate for each scan. The object motion is then calculated by a derivation over time using a Kalman filter approach. However, object motion can be extracted without assumptions of a geometric object model. The suggested method compares the measurements of two consecutive scans [23, 87]. For each object a displacement vector is calculated by a matching process of the measurements assigned to the object from two consecutive time instances. The resulting estimated motion is however relative to the host vehicle.

In some situations this displacement information is not only an additional feature but a necessary measurement in order to keep the velocity of an object observable. Consider the scenario on a highway, where the host vehicle overtakes a truck (Figure 5.15). When the host vehicle is beside the truck, the rear of the truck is out of the FOV of the laser scanner and only the side of the truck is observed. The contour model based feature extraction now matches an I-shape with the scan data, which represent the side of the truck (Figure 5.15(b)). However, none of the two ends of the side can be observed. The left rear corner of the truck is not in the field of view and the left front corner can not be observed because of possible total reflections. Therefore, only the orientation and a one dimensional position, perpendicular to the orientation, are observed. The position, velocity and acceleration in the direction of the orientation are therefore locally unobservable.

However, it is possible to estimate the longitudinal motion of the truck in the scan data sequence, by analyzing the motion of the structure of the scan data. Figure 5.15(c) shows the matched scans of two consecutive measurement intervals. The structure of the measurements on the truck is exploited in order to find an unambiguous object displacement.

Point Matching using the WICP

The laser scanner measurement sets assigned to an object from the two consecutive measurement intervals are denoted $\mathbb{P}(k)$ and $\mathbb{P}(k-1)$. The aim is the precise matching of these point sets using the WICP described in Chapter 5.4.1. Only the translation is determined and the rotation change neglected[2]. Therefore, the transformation in Equation (5.12) is now given by

$$\mathbf{p}_i^* = \mathbf{p}_i + \mathbf{t} \tag{5.41}$$

[2]The feature extraction was tested for trucks on highways, where the change of orientation is so small, that the performance did not decrease noticeably when neglecting the rotation. However, the orientation estimation of the contour model based approach can be used as a given orientation for the WICP, if necessary.

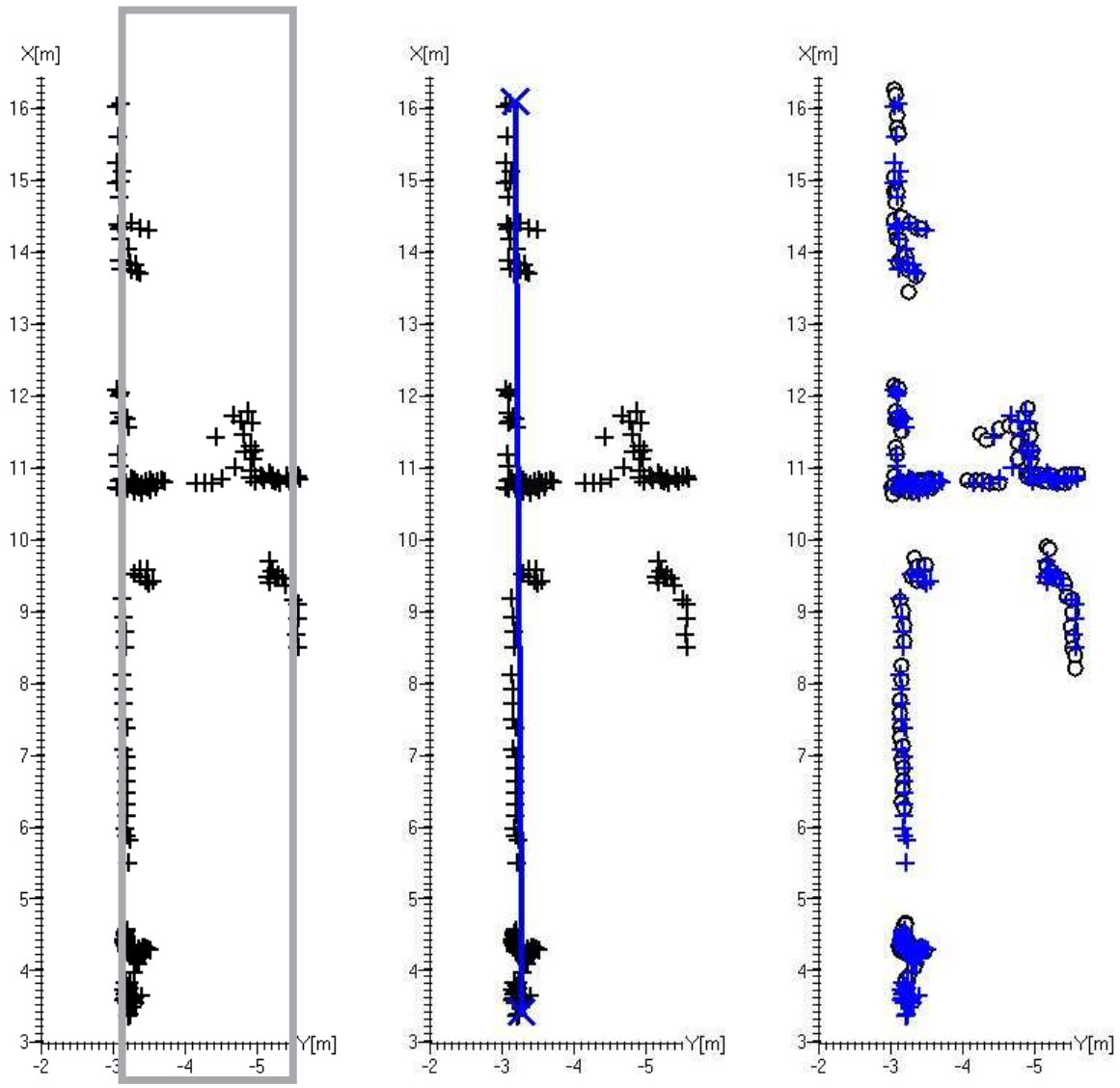


Figure 5.15: Truck which is being overtaken by the host vehicle. The truck is indicated in (a) together with the laser scanner measurements. (b) shows an I-shape fitted into the measurements. The "X" at both ends of the line indicate that no corner is observed. (c) shows the matched scans of two consecutive measurement intervals (circles and crosses).

The corresponding point $\mathbf{q}_i$ to $\mathbf{p}'_i(k) \in \mathbb{P}(k)$ is that point from $\mathbb{P}(k-1)$ which is the closest to $\mathbf{p}'_i(k)$. The displacement estimation is performed for each scan layer separately in order to save computational resources.

Performing the WICP on an object of which the laser scanner data exposes a dominant structure in one direction showed, that the resulting displacement estimation strongly depends on the initial transformation. Therefore, the estimation of the global minimum comprises three steps:

- The ICP is performed for few iterations from several initial positions, which are equally distributed over the border of the 3σ-gate calculated from the residual covariance $\mathbf{S}$ (Equation (4.29)). The resulting positions then form a new region of interest (ROI) where the global minimum is assumed to be located. Usually this new ROI is much smaller than the original 3σ-gate.
- The new ROI is then sampled and the mean distance error of the ICP $\bar{d}$ calculated for each equidistant displacement sample. Local minima are determined by comparison with the direct neighbours.
- The ICP is performed for each of these local minima. The resulting local minima are compared and a global minimum identified.

The first step serves mainly as a reduction of the computational burden, by reducing the search space for the detailed analysis. In the second and third step, the global minimum is searched for by sampling the search space and then analyzing local minima. Alternative search strategies sample the search space depending on the a posteriori probability density function utilizing Monte Carlo methods [158].

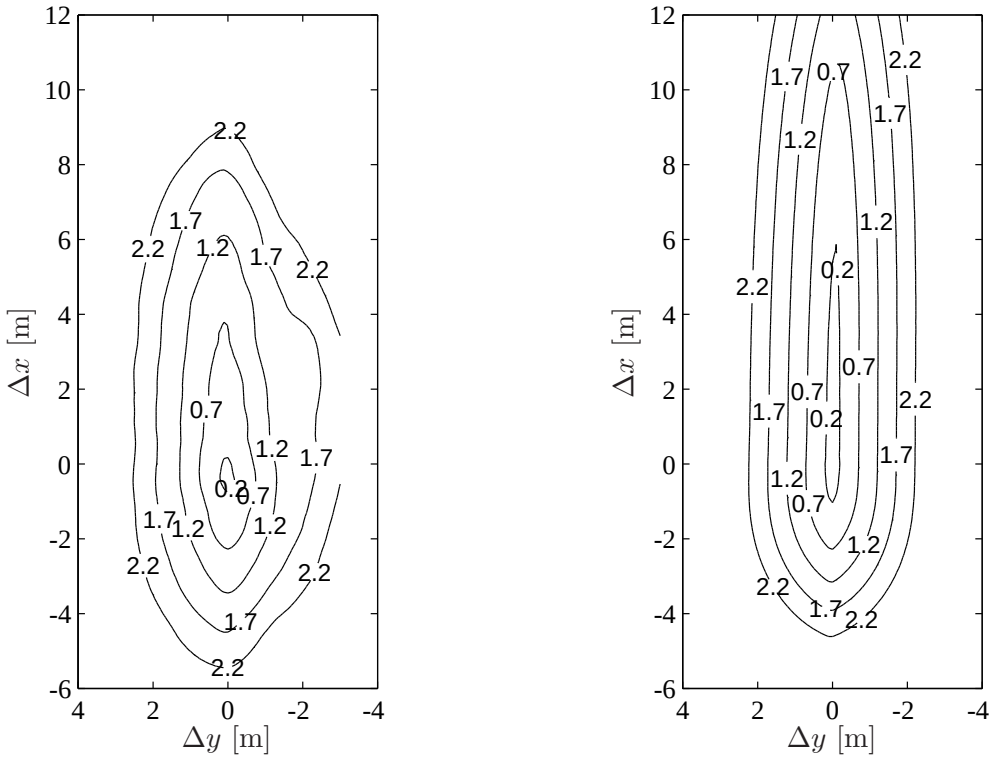
Estimation and Analysis of the Global Minimum

The mean residual distance $\bar{d}_g$ of the estimated global minimum is first analyzed for its absolute quality ($\bar{d}_g \overset{!}{<} \theta_d$) and its relative quality compared to the residuals $\bar{d}_{l,i}$ of the local minima ($2\bar{d}_g \overset{!}{<} \min(\bar{d}_{l,i})$).

Additionally the shape of the estimated global minimum is analyzed. Consider the two minima of the mean residual distance $\bar{d}(\Delta x, \Delta y)$ shown in Figure 5.16. Figure 5.16(a) shows the residual distribution of the ICP applied to the data of a truck. A sharp global optimum is exhibited, where Figure 5.16(b) shows a long minimum in x-direction which is due to the object with a dominant structure in x-direction. In this case the displacement can be estimated in y-direction but not in x-direction, in which the minimum is ambiguous.

Therefore the gradient of the residual distance distribution is calculated at the estimated global minimum in several directions[3]. The minimal gradient g_{min} provides information, if the estimated global minimum is sharply localized. If the minimal gradient is below a predefined threshold θ_g, the displacement measurement is discarded. Additionally the gradient is utilized for the online calculation of the measurement noise (Chapter 7.2.6).

[3] In general, many gradient directions must be considered. Here only the gradients in x- and the y-directions are calculated, as the objects of interest are assumed to be oriented parallel to the host vehicle.



(a) (b)

Figure 5.16: Mean residual distance $\bar{d}(\Delta x, \Delta y)$ for different displacements. (a) scan data of a truck was analyzed (Figure 5.15). (b) shows the distribution of the residuals for an object with a dominant structure in x-direction, such as a crash barrier.

The complete point matching algorithm is performed for each scan layer separately in order to reduce the CPU consumption. The result of one layer is considered to be an outlier and the mean of the remaining results is taken as the overall estimate for the displacement $(\Delta x, \Delta y)$.

5.5 Association

In a multi-object tracking system the association between existing object tracks and the observations produced by the sensors is in general ambiguous. Therefore the most probable assignment is chosen. This leads to an optimization problem, in which the overall cost of the assignments is minimized. A good overview of frequently cited algorithms can be found in [13, 61, 143].

The association is composed of two stages. First, for each object only the feasible sensor observations are selected by a gating process. The second stage then solves the ambiguities, if one observation is in the gate of several objects or several observations are in the gate of one object.

5.5.1 Gating

For each object a gate is constructed from the residual covariance $\mathbf{S}(k)$ stated in Equation (4.29). For a selection the elements $\nu_i(k)$ of the observation residual $\boldsymbol{\nu}(k)$, given in Equation (4.22), must be in the rectangular 3σ-gate with

$$|\nu_i(k)| \leq 3\sqrt{S_{ii}(k)} \tag{5.42}$$

5.5.2 Segment to Track Association

It is necessary to define a cost function for the subsequent optimization of the assignment. One approach is to define a statistical distance for the assignment of an observation to a track

$$d_G^2(k) = \boldsymbol{\nu}'(k)\mathbf{S}^{-1}(k)\boldsymbol{\nu}(k) + \ln\left(|\mathbf{S}(k)|\right) \tag{5.43}$$

which can be derived from Gaussian likelihood calculations [12]. The last term represents a penalty for newly created object tracks with high state uncertainty. In conflict situations the preference is given to already established objects. An iterative-nearest-neighbor algorithm then successively assigns observations to tracks, choosing always the combination with the lowest distance criterion.

The association of a segment to an object is based only on the center of gravity assigned to the segment and the predicted center of gravity given the object state. This increases the robustness especially in the initialization phase of an object track, compared to an assignment criterion based on the complete predicted and measured feature vectors. The prediction of the center of gravity can be derived by a calculation of the expected measurements given the object state and the sensor model of the laser scanner, using ray tracing techniques.

5.5.3 Shape Hypothesis Selection

The feature extraction described in Chapter 5.4.1, which matches contour models with point sets, constitutes a multi-modal measurement function. If several shape hypotheses are provided by the feature extraction, the association selects the shape hypothesis which fits best the prediction, based on the Mahalanobis-distance

$$d^2(k) = \boldsymbol{\nu}'(k)\mathbf{S}^{-1}(k)\boldsymbol{\nu}(k) \tag{5.44}$$

where $\boldsymbol{\nu}(k)$ is composed of the residual position of the observed edge or side, width, length and orientation, or a sub-set, depending on the observed feature.

5.6 Object Generation

Features which can not be associated to an already existing object constitute a new object. As the sensor model is necessary in order to create a new object from the given feature vector, the object generation is part of the sensorial tasks and not part of the central fusion system (Figure 4.3).

The aim of the object generation is the determination of the state estimate $\mathbf{x} = \begin{bmatrix} x & y & z & \psi & v_x & v_y & d_x & d_y & d_z \end{bmatrix}'$ and its covariance $\mathbf{P}$ from the extracted features. Assuming the measurement vector

$$\mathbf{z}_{ls}(k) = \begin{bmatrix} x^{ls} & y^{ls} & z^{ls} & \psi^{ls} & d_x^{ls} & d_y^{ls} & d_z^{ls} \end{bmatrix}' \tag{5.45}$$

the yaw angle ψ, the length d_x, width d_y and height d_z of the estimated state result directly from the inverse measurement functions (5.5)–(5.8), with

$$\begin{aligned} \psi &= h_{\psi}^{ls,inv}[\mathbf{z}_{ls}(k)] = \psi^{ls} + \psi_{ls} && (5.46)\\ d_x &= h_{d_x}^{ls,inv}[\mathbf{z}_{ls}(k)] = d_x^{ls} && (5.47)\\ d_y &= h_{d_y}^{ls,inv}[\mathbf{z}_{ls}(k)] = d_y^{ls} && (5.48)\\ d_z &= h_{d_z}^{ls,inv}[\mathbf{z}_{ls}(k)] = d_z^{ls} && (5.49) \end{aligned}$$

where ψ_{ls} denotes the calibrated yaw angle of the laser scanner. The estimated position follows from the measurement function $\mathbf{h}_p^{ls}$ for the feature position (Equation (5.4)), with

$$\mathbf{p}^{ls} = \mathbf{H}_{veh\to ls}\mathbf{H}_{obj\to veh}\mathbf{p}^{obj}$$

With Equation (4.2) the term can be reformulated

$$\begin{aligned} \mathbf{p}^{ls} &= \mathbf{H}_{veh\to ls}\begin{bmatrix} \cos\psi & -\sin\psi & 0 & x \\ \sin\psi & \cos\psi & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{bmatrix}\mathbf{p}^{obj} && (5.50)\\ &= \mathbf{H}_{veh\to ls}\left(\mathbf{R}_{obj\to veh}\mathbf{p}^{obj} + \mathbf{t}_{obj}^{veh}\right) && (5.51) \end{aligned}$$

where

$$\mathbf{R}_{obj\to veh} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 & 0 \\ \sin\psi & \cos\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{t}_{obj}^{veh} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{5.52}$$

Equation (5.51) is then solved with respect to $\mathbf{t}_{obj}^{veh}$, which results in the inverse measurement function $\mathbf{h}_p^{ls,inv}$ for the object position $\mathbf{t}_{obj}^{veh}$

$$\mathbf{t}_{obj}^{veh} = \mathbf{h}_p^{ls,inv}[k, \mathbf{z}_{ls}, \mathbf{p}^{obj}] = \mathbf{H}_{ls\to veh}\mathbf{p}^{ls} - \mathbf{R}_{obj\to veh}\mathbf{p}^{obj} \tag{5.53}$$

The complete inverse measurement function $\mathbf{h}_{ls,inv}$ is then built from Equations (5.46)–(5.49) and (5.53).

For the estimation of the initial covariance matrix $\mathbf{P}$, the Jacobi matrix $\mathbf{H}_{ls,inv}$ of $\mathbf{h}_{ls,inv}$ is calculated, with

$$\mathbf{H}_{ls,inv}(k) = \left[\nabla_{\mathbf{z}}\mathbf{h}'_{ls,inv}\right]'\Big|_{\mathbf{z}=\hat{\mathbf{z}}(k)} \tag{5.54}$$

The covariance matrix is then derived from the measurement covariance matrix $\mathbf{R}$ with

$$\mathbf{P}(0) = \mathbf{H}'_{ls,inv}(k)\mathbf{R}(k)\mathbf{H}_{ls,inv}(k) \tag{5.55}$$

5.7 Discussion

Novel approaches for the laser scanner data based vehicle detection and localization have been presented. The algorithms are combined to form a laser scanner module, which is connected to the previously described feature-level fusion module.

One approach models vehicles as cuboids. Different contour shapes were designed which are matched as a whole to the laser scanner data. Thus, the position, orientation, length and width of the observed vehicle can be estimated. Special care has been taken to account for object occlusions and sensor specific limitations. The approach can handle situations, in which the object state is not completely observable.

A novel orientation estimation for vehicles seen from the rear has been presented. It exploits the echo pulse width information of the laser scanner. The echo pulse width enables the detection of reflecting surfaces, such as the number plate or the rear reflectors of a vehicle. A regression line through the measurements on these reflectors inhibits the influence of measurement outliers inside the car or on the exhaust of the vehicle. Additionally, the classification of the rear of vehicles is indicated. This classification is useful for an identification of vehicles and for an unambiguous determination of their orientation.

The displacement tracking estimates the object motion without the assumption of a geometric object model. The object displacement, calculated by a matching process of the measurements from two consecutive time instances, provides a motion estimate relative to the host vehicle. The displacement information is especially useful in situation, in which no corner of the tracked vehicle can be observed. In these situations the position of the object is unobservable. The displacement however provides the necessary information about the object's motion.

6 Video Data based Feature Extraction

As shown in the preceding chapter it is possible to detect objects with a laser scanner, localize them and even estimate their width, length and orientation. However, even if the angular resolution of the laser scanner is high compared to radar sensors, it is still limited to 0.5° at 20 Hz scan frequency. In comparison, the angular resolution of a camera with 28° horizontal aperture angle and 640x480 pixel resolution, which is applied in this work, is more than 10 times higher. Even if the strength of a camera usually lies in lane recognition, traffic sign recognition or object classification, it is utilized in this work to optimize the precision of the dynamic state estimation of vehicles. Especially in the middle and far range, the higher resolution is exploited to improve the estimation of the objects' location and dimensions. Together with a lane recognition, this would enable precise lane assignments and early lane change recognition.

The camera sensor is integrated into the feature-level fusion as shown in Figure 6.1. The purpose of the camera is the improvement of the object state estimates. As the laser scanner is very reliable in object detection and has a very large field of view of 150° in azimuth[1] and 3.2° in elevation and up to 128 m in distance, the fusion relies solely on the object detection of the laser scanner. Therefore, the object generation and the association are omitted in the camera module. The objects which are already established by the laser scanner are forwarded from the fusion unit to the camera unit. From the object positions the camera module predicts the features in the image. Features are for instance the position of the left or right edge of the rear of a preceding car. The feature extraction then searches for the predicted features in the image and delivers new estimates. The advantage of this attention control of the image processing lies in the a priori information provided by the fusion unit about the dimensions and position of the object in the image. This a priori information reduces drastically the three dimensional search space, given by the two position coordinates and the dimensions of the object. Therefore, more complex image processing techniques can be exploited, than in a purely video data based tracking approach, in which the whole search space must be considered.

There are two different alternatives for the attention control shown in Figure 6.1.

- Predicted state estimates $\hat{\mathbf{x}}(k|k-1)$ and their covariances $\mathbf{P}(k|k-1)$ provided by the fusion module.
- Measurements provided by the laser scanner (dashed blue line)

The first alternative has the advantage that the data is provided by a standardized interface from the fusion unit. This supports the scalability of the system and exchangeability of the sensors. The second approach provides the most actual data and might be

[1] In general the field of view of the laser scanner is about 270°. Integrated into the vehicle's front, the field of view is reduced to 150°.

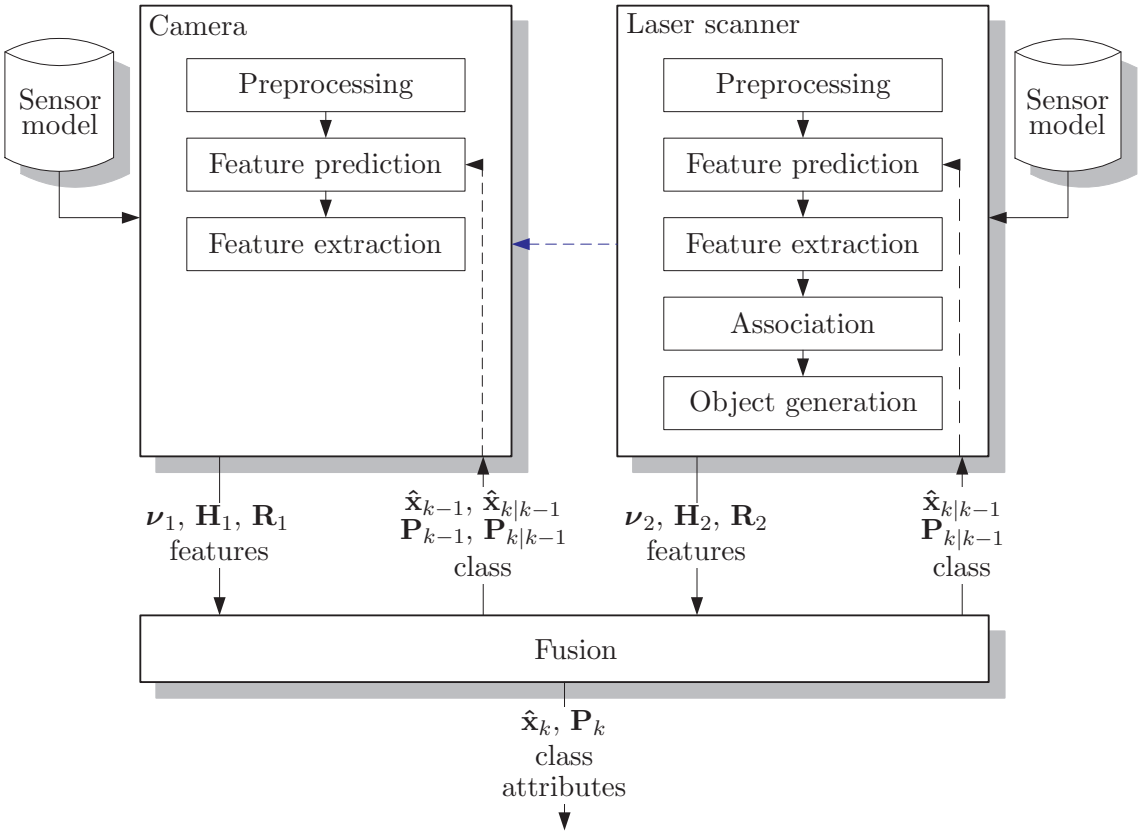
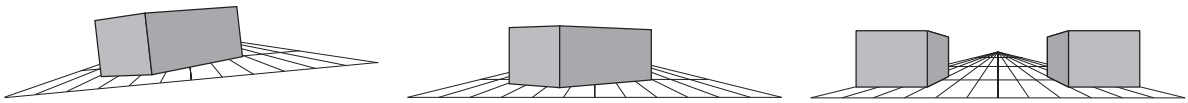


Figure 6.1: Camera sensor integration into the feature-level fusion architecture.

more accurate than the predicted object state. However, it implies a dependence between the laser scanner and the camera module. Therefore, the first approach is suggested in this work in order to keep the fusion architecture as simple and as general as possible.

Some image processing algorithms detailed in this chapter need the previous state estimates $\hat{\mathbf{x}}(k-1)$ and their covariances $\mathbf{P}(k-1)$ additional to the predictions. They are provided by the fusion module (Figure 6.1).

Following the description of the feature model and the measurement equation, classical model based feature extraction methods will be detailed in this chapter. They exploit image edges or dark regions originating from the shadow underneath the vehicles. First implementations of these algorithms showed, that the accuracy in the position and width estimation is significantly improved with respect to a purely laser scanner based state estimation [131, 82, 83]. However, the algorithms produce biased estimates in certain situations, the availability is rather low, they fail for several vehicle types and they imply strong constraints on the position and orientation of the observed vehicles. Therefore, a model-free template tracking approach was finally chosen which is applicable in a wide range of situations and does not imply any constraint concerning the vehicle type. Even if the evaluation in Chapter 8 only investigates the template tracking approach, the classical feature-model based methods are shortly outlined in the following to demonstrate the possibility of a seamless integration into the feature-level fusion framework.



(a) (b) (c)

Figure 6.2: Appearance of cuboidal objects in the image under different assumptions. (a) No assumption. (b) Camera roll and pitch close to zero. (c) Camera roll and pitch close to zero and object orientation same as host vehicle's orientation.

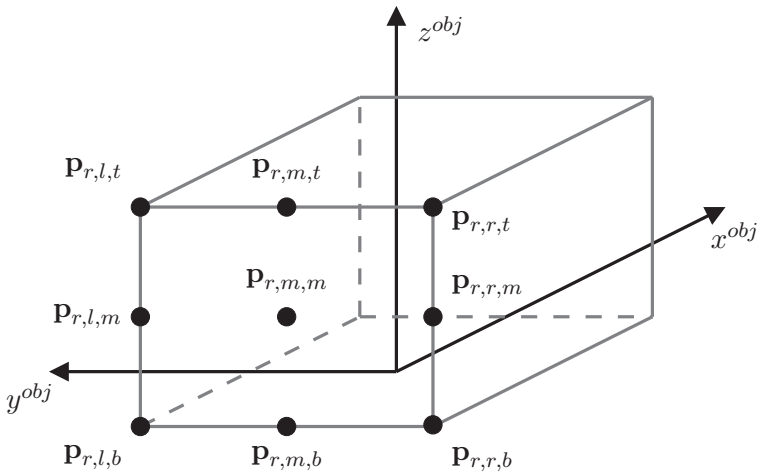


Figure 6.3: Camera feature positions based on the cuboidal object model.

6.1 Feature Model and Measurement Equation

The shape of objects is approximated in this work by a cuboid. The object state, which is defined in the host vehicle coordinate system, contains the position $\begin{bmatrix} x & y & z \end{bmatrix}'$, the yaw angle ψ, the velocity $\begin{bmatrix} v_x & v_y \end{bmatrix}'$ and the length, width and height $\begin{bmatrix} d_x & d_y & d_z \end{bmatrix}'$ of the cuboid (Chapter 4.3.2). The pitch and roll angles of the object are assumed to be identical with those of the host vehicle. An example of such a cuboidal object projected into the image domain is shown in Figure 6.2(a). It is worth noticing that in general the contour edges of objects are not horizontally or vertically aligned, as the camera can be mounted with a significant pitch, roll or yaw angle, estimated by the calibration procedure.

However, the roll and pitch angles are usually so small, that an object exhibits vertical edges in the image as shown in Figure 6.2(b). As the horizontal lines of the cuboid can still appear in an angle, the projection of one side of the cuboidal object has the form of a trapezoid.

If additionally the yaw angle of the camera is negligible and the objects' orientation is either aligned with the host vehicle's orientation or perpendicular to it ($\psi \approx n\frac{\pi}{2}$, with $n \in \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$) then the object's side facing the host vehicle appears as a rectangle in the image as shown in Figure 6.2(c).

Figure 6.3 depicts the features used by the video data feature extraction based on the cuboidal object model defined in Chapter 4.3.2. Features are for instance the left and right

Table 6.1: Feature coordinates in the object coordinate system.

Feature	x^{obj}	y^{obj}	z^{obj}
$\mathbf{p}_{r,l,b}^{obj}$	$-0.25d_x$	$0.5d_y$	0
$\mathbf{p}_{r,m,b}^{obj}$	$-0.25d_x$	0	0
$\mathbf{p}_{r,r,b}^{obj}$	$-0.25d_x$	$-0.5d_y$	0
$\mathbf{p}_{r,l,m}^{obj}$	$-0.25d_x$	$0.5d_y$	$0.5d_z$
$\mathbf{p}_{r,m,m}^{obj}$	$-0.25d_x$	0	$0.5d_z$
$\mathbf{p}_{r,r,m}^{obj}$	$-0.25d_x$	$-0.5d_y$	$0.5d_z$
$\mathbf{p}_{r,l,t}^{obj}$	$-0.25d_x$	$0.5d_y$	d_z
$\mathbf{p}_{r,m,t}^{obj}$	$-0.25d_x$	0	d_z
$\mathbf{p}_{r,r,t}^{obj}$	$-0.25d_x$	$-0.5d_y$	d_z

edge of the rear $\mathbf{p}_{r,l,m}$ and $\mathbf{p}_{r,r,m}$ or the position of the rear given by its center $\mathbf{p}_{r,m,m}$. The coordinates of a feature position in the object coordinate system are given in Table 6.1, where d_x denotes the length of the object, d_y the width and d_z the height.

The discrete-time measurement equation is

$$\mathbf{z}_{img}(k) = \mathbf{h}_{img}[k, \mathbf{x}(k)] + \mathbf{w}_{img}(k) \tag{6.1}$$

where $\mathbf{z}_{img}$ is the measurement at the measurement interval k, $\mathbf{h}_{img}$ the nonlinear measurement function and $\mathbf{w}_{img}$ the measurement noise.

The measurement function for a feature position $\mathbf{h}_p^{img}$ includes the homogeneous matrix $\mathbf{H}_{obj\to veh}$ from Equation (4.1), which transforms a feature position $\mathbf{p}^{obj}$ from the object coordinate system into the vehicle coordinate system. This is followed by the projection $\mathbf{H}_{veh\to img}$ from Equation (3.30) into the image coordinate system

$$\begin{bmatrix} h_1^{img}(k) \\ h_2^{img}(k) \\ h_3^{img}(k) \end{bmatrix} = \mathbf{H}_{veh\to img}\mathbf{H}_{obj\to veh}(k)\mathbf{p}^{obj}(k) \tag{6.2}$$

with

$$\mathbf{h}_p^{img}[k, \mathbf{x}(k), \mathbf{p}^{obj}] = \begin{pmatrix} h_u^{img}[\mathbf{x}(k), \mathbf{p}^{obj}(k)] \\ h_v^{img}[\mathbf{x}(k), \mathbf{p}^{obj}(k)] \end{pmatrix} = \begin{bmatrix} \frac{h_1^{img}(k)}{h_3^{img}(k)} \\ \frac{h_2^{img}(k)}{h_3^{img}(k)} \end{bmatrix} \tag{6.3}$$

This measurement function represents the most general projection of the cuboid features into the image as shown in Figure 6.2(a).

6.2 Classical Model based Feature Extraction Algorithms

A standard procedure for the detection and localization of vehicles in the image domain relies on features such as the symmetry, color cues, shadows, corners, image edges, local

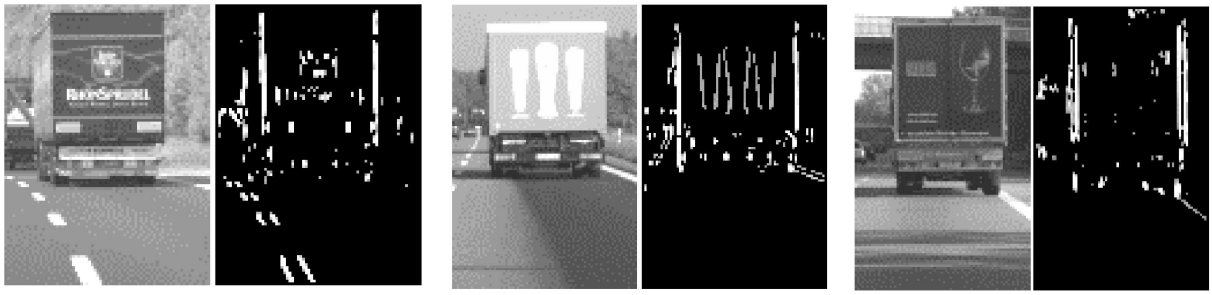


Figure 6.4: Images of trucks and the corresponding vertical edges. Most trucks exhibit strong vertical edges.



(a) (b) (c)

Figure 6.5: Edge feature extraction. (b) The bottom line is determined from the dark areas underneath the truck. (c) Hypotheses of the left and right side of the rear generated from vertical edges.

entropy and bright spots. In the following an edge feature based approach for the localization of trucks will be detailed. For the localization of passenger cars, the dark regions underneath the vehicle are exploited.

Even if these classical feature-model based methods are not utilized in the final evaluation, the following description demonstrates the possibility of an integration into the feature-level fusion and provides the necessary measurement equations.

6.2.1 Edge Features

A common assumption is, that the contour of a vehicle in the image exhibits strong edges. A vehicle seen from the rear can then be decomposed into two vertical edges on either side and a horizontal edge at the bottom. This method therefore delivers two uncorrelated edge feature positions for both sides and one for the bottom of the vehicle's rear.

In contrast to passenger cars the rear of trucks usually exhibits strong vertical edges in the image, even if the truck is seen from the side and rear (Figure 6.4). With image processing techniques the vertical edges of the trucks rear can be extracted as shown in Figure 6.5(c). The image processing steps described in [131] can be summarized as:

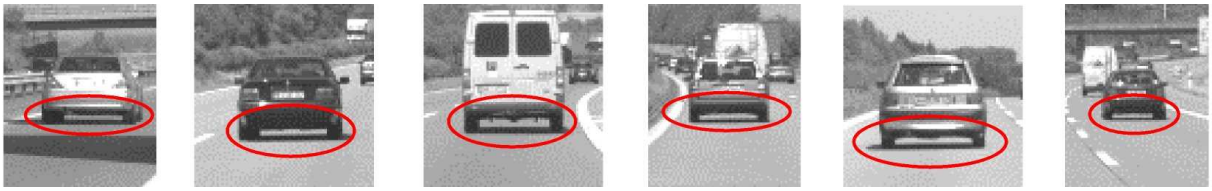


Figure 6.6: The dark regions underneath the vehicles are widely independent of the type of car and the illumination.

- Edge detection for vertical edges.
- Histogram of vertical edges over the columns of the image region.
- The peaks in the histogram represent the edge hypotheses.
- Selection of the maximum likelihood edges, incorporating the a priori information from the feature prediction.

The u-coordinate of the position of these edges is taken as a measurement for the left and right rear edge of the cuboidal object $\mathbf{p}_{r,r,m}^{obj}$ and $\mathbf{p}_{r,l,m}^{obj}$. Additionally the image row can be identified which represents the bottom of the trucks rear $\mathbf{p}_{r,m,b}^{obj}$. Figure 6.5(b) shows a binary image of the 10 % darkest pixels in the image patch. In this binary image the lowest horizontal edge is taken as the measurement for the bottom of the trucks rear. The measurement function is therefore

$$\mathbf{h}_{img}[k, \mathbf{x}(k)] = \begin{pmatrix} h_u^{img}[\mathbf{x}(k), \mathbf{p}_{r,r,m}^{obj}(k)] \\ h_u^{img}[\mathbf{x}(k), \mathbf{p}_{r,l,m}^{obj}(k)] \\ h_v^{img}[\mathbf{x}(k), \mathbf{p}_{r,m,b}^{obj}(k)] \end{pmatrix} \tag{6.4}$$

It is worth noticing that in case of a partial occlusion of the truck on one of the sides, one edge is not visible for the camera. Alternatively, the bottom of the rear might not be seen due to spray behind the truck. Then the corresponding entry in the measurement function is neglected. Therefore, this approach can handle partial occlusions and missing image features. Additionally, this edge feature extraction approach fully exploits the advantages of the general derivation of the measurement function, as it can be applied to vehicles with orientations different from the host vehicle's. However, even if most trucks exhibit strong vertical edges, there are some trucks, such as gasoline or car transporters which cause the described algorithms to fail.

6.2.2 Matching of a Deformable Synthetic Template

In this approach a deformable synthetic template of a vehicle is matched to the image [131, 82, 83]. The template matching method extracts the position of the bottom of the rear in the image and delivers additionally the width of the vehicle in the image. This method exploits the appearance of the wheels and the shadow underneath the vehicles, which is widely independent of the type of car (Figure 6.6). However, this method is limited to passenger cars, which are located in the far range ($x > 50$ m) and seen from the



(a) (b) (c) (d)

Figure 6.7: Deformable synthetic templates. (a) shows a typical template which contains the bumper, the wheels and the shadow from the vehicle. (b) Same template with larger wheels, (c) different shadow shapes are considered, (d) the size of the dark region beneath the bumper can vary.



(a) (b) (c)

Figure 6.8: (a) Image region. (b) Binary image of the 10 % darkest pixels. (c) The template match is indicated.

rear. Therefore the template matching is only applied to passenger cars which are already detected by the laser scanner and of which the estimated orientation is approximately the same as the host vehicle's ($\psi \approx 0$).

The deformable template shown in Figure 6.7 represents the dark region underneath passenger cars and is composed of the bumper, the wheels and the shadow underneath the vehicle. The parameters of the template include the size of the template, the size of the wheels, the shape of the shadow and the size of the bumper. The image processing steps described in [131] can be summarized as:

- Generation of binary image from the 10 % darkest pixel in the image region (Figure 6.8(b)).
- Matching of the deformable template (Figure 6.8(c)).

The measurement function $h_{d_u}^{img}$ for the width d_u^{img} of the object in the image is given by the difference in the u-coordinates of the left and right features at the bottom of the vehicle's rear $\mathbf{p}_{r,l,b}^{obj}$ and $\mathbf{p}_{r,r,b}^{obj}$.

$$h_{d_u}^{img}[k, \mathbf{x}(k)] = h_u^{img}[\mathbf{x}(k), \mathbf{p}_{r,r,b}^{obj}(k)] - h_u^{img}[\mathbf{x}(k), \mathbf{p}_{r,l,b}^{obj}(k)] \tag{6.5}$$

The complete measurement function for the template matching follows from the combi-

nation of Equations (6.3) and (6.5)

$$\mathbf{h}_{img}[k, \mathbf{x}(k)] = \begin{pmatrix} h_u^{img}[\mathbf{x}(k), \mathbf{p}_{r,m,b}^{obj}(k)] \\ h_v^{img}[\mathbf{x}(k), \mathbf{p}_{r,m,b}^{obj}(k)] \\ h_u^{img}[\mathbf{x}(k), \mathbf{p}_{r,r,b}^{obj}(k)] - h_u^{img}[\mathbf{x}(k), \mathbf{p}_{r,l,b}^{obj}(k)] \end{pmatrix} \tag{6.6}$$

Preliminary results showed, that this approach works in general reasonably well. However, it is limited to cars in the far range, observed straight from the rear and with little lateral offset. For vehicles on adjacent lanes or in distances below 50 m the algorithm often fails or does not produces significant better estimates for the lateral position and width than the already accurate laser scanner.

6.3 Kalman Filter based Template Tracking

Since the feature-model based algorithms described above have limited availability, are not applicable to all vehicles and impose constraints on the position of the observed objects, an appearance-model free approach is finally chosen, which works in a wide range of situations.

6.3.1 Feature Extraction

The motion of an object in the image is estimated over several successive images with a template tracking approach. The flow chart of the template tracking is shown in Figure 6.9.

A template is initialized when a tracked object is in two consecutive images completely in the field of view of the camera. The template, which represents one side of an object, is generated from the previous image.

Thereafter and in all subsequent cycles the template is searched for in the current image. For the correlation based search, the template is resized, depending on the object distance and the distance at the template initialization. Depending on the correlation coefficient, the search is considered to be successful. Two thresholds $\theta_{r,l}$ and $\theta_{r,h}$, with $\theta_{r,l} < \theta_{r,h}$, define the quality of the match. If the correlation coefficient r_k is above $\theta_{r,h}$, the search is considered to be successful and the position of the template in the current image taken as a measurement for the subsequent Kalman filter. A correlation coefficient r_k between the two thresholds $\theta_{r,l} \leq r_k < \theta_{r,h}$ indicates that the match is of decreased quality. The measurement is still used for the state estimation, however the template is deleted and a new template generated in the next cycle. If the correlation coefficient drops below the threshold $\theta_{r,l}$, the search failed and the measurement is discarded. Correlation coefficients of $\theta_{r,l} = 0.7$ and $\theta_{r,h} = 0.8$ were found to provide good performance.

Template Initialization

A new template is generated, if an object tracked by the laser scanner is twice completely in the field of view of the camera. A template is generated from the previous image and subsequently searched for in the current image. For a preceding car the best suited feature

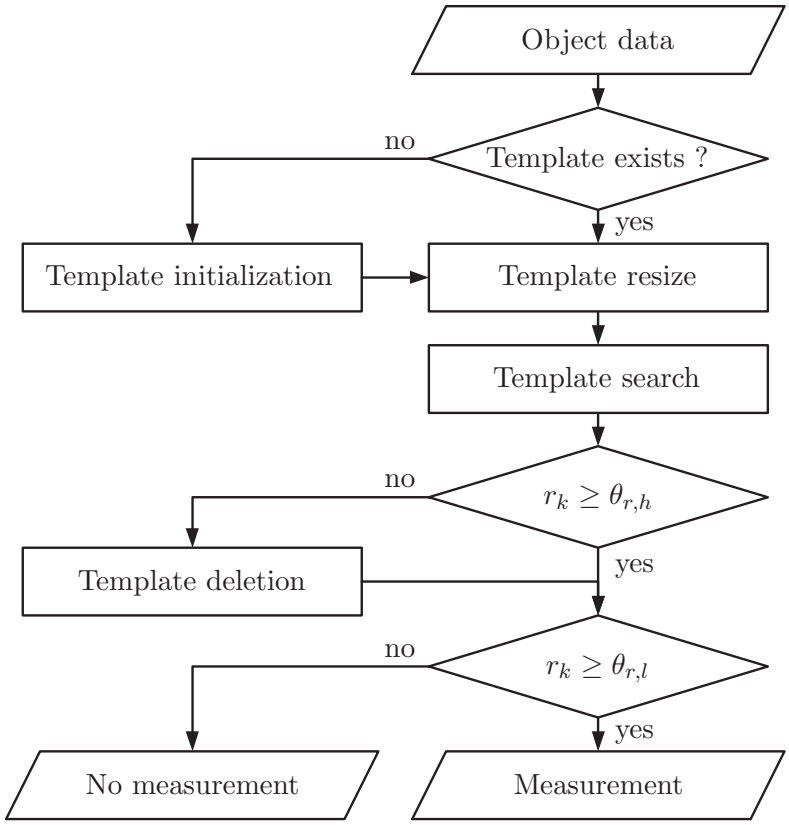


Figure 6.9: Flow chart of the template tracking algorithm. The correlation coefficient of the measurement interval k is denoted by r_k. The parameters $\theta_{r,l}$ and $\theta_{r,h}$ define a lower and higher threshold for the correlation coefficient. A good choice is $\theta_{r,l} = 0.7$ and $\theta_{r,h} = 0.8$.

is its rear, as it is well observed by the camera independent of the lane the preceding vehicle is driving on[2].

The a posteriori state estimate of the object at the last measurement interval $\hat{\mathbf{x}}(k-1)$ delivers the best estimate of the position of the object at $k-1$ in terms of the minimum mean square error. Therefore the position of the center of the rear is given by

$$\hat{\mathbf{p}}^{img}_{r,m,m}(k-1) \quad = \quad \begin{pmatrix} h^{img}_u[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,m,m}(k-1)] \\ h^{img}_v[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,m,m}(k-1)] \end{pmatrix} \tag{6.7}$$

The template is given by the image region, which results from the projection of the rear

[2] A preceding car is here taken as an example. For a crossing car the best suited feature would be the visible side or two sides could be taken for a vehicle seen from an angle. The described algorithm applies analogously.

of the cuboid into the last image. The image region is defined by its four corners:

$$\hat{\mathbf{p}}^{img}_{r,l,b}(k-1) = \begin{pmatrix} h^{img}_u[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,l,b}(k-1)] \\ h^{img}_v[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,l,b}(k-1)] \end{pmatrix} \tag{6.8}$$

$$\hat{\mathbf{p}}^{img}_{r,r,b}(k-1) = \begin{pmatrix} h^{img}_u[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,r,b}(k-1)] \\ h^{img}_v[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,r,b}(k-1)] \end{pmatrix} \tag{6.9}$$

$$\hat{\mathbf{p}}^{img}_{r,r,t}(k-1) = \begin{pmatrix} h^{img}_u[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,r,t}(k-1)] \\ h^{img}_v[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,r,t}(k-1)] \end{pmatrix} \tag{6.10}$$

$$\hat{\mathbf{p}}^{img}_{r,l,t}(k-1) = \begin{pmatrix} h^{img}_u[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,l,t}(k-1)] \\ h^{img}_v[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,l,t}(k-1)] \end{pmatrix} \tag{6.11}$$

The quadrangle generated by these corners is in general not a rectangle (Figure 6.2(a)). However, subsequent image processing algorithms are more efficient for rectangular image regions. One possibility is to warp the quadrangle into a rectangle, which compensates for the perspective and therefore rectifies the image patch to a normalized view [166].

However, another less time consuming approach, which applies for small perspective distortions, approximates the image region by a rectangle. This approach uses the points in the middle of the cuboid edges $\hat{\mathbf{p}}^{obj}_{r,m,b}(k-1)$, $\hat{\mathbf{p}}^{obj}_{r,r,m}(k-1)$, $\hat{\mathbf{p}}^{obj}_{r,m,t}(k-1)$ and $\hat{\mathbf{p}}^{obj}_{r,l,m}(k-1)$ and generates a rectangle from their projections in the image. The position of the rectangle is then given by $\hat{\mathbf{p}}^{img}_{r,m,m}(k-1)$ from Equation (6.7) and the width $\hat{d}_u(k-1)$ and height $\hat{d}_v(k-1)$ of the image region defined by

$$\hat{d}_u(k-1) = h^{img}_u[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,r,m}(k-1)] - h^{img}_u[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,l,m}(k-1)] \tag{6.12}$$

$$\hat{d}_v(k-1) = h^{img}_v[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,m,b}(k-1)] - h^{img}_v[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,m,t}(k-1)] \tag{6.13}$$

For the following image processing steps, only those object sides which are (almost) perpendicular to the optical axis of the camera are considered, such as the rear of a preceding car. The quadrangular image region is then well approximated by a rectangle. In most cases on highways, country roads and even in urban areas, the vehicles of interest are either driving in the same or the opposite direction. Due to the limited field of view of the camera, crossing cars are only seen for a very short time.

The resulting image region, which contains one side of an object is taken as a template I_t. For a subsequent scaling of the template the distance to the object in the camera coordinate system is assigned to the template. The x-coordinate of the rear of the object in the camera coordinate system $x^{cam}_t = \hat{x}^{cam}(k-1)$ is given by center of the object side $\hat{\mathbf{p}}^{obj}_{r,m,m}(k-1)$ transformed into the camera coordinate system, with

$$\hat{\mathbf{p}}^{cam}_{r,m,m}(k-1) = \begin{bmatrix} \hat{x}^{cam}(k-1) \\ \hat{y}^{cam}(k-1) \\ \hat{z}^{cam}(k-1) \\ 1 \end{bmatrix} = \mathbf{H}_{veh\to cam}\mathbf{H}_{obj\to veh}[\hat{\mathbf{x}}(k-1), \hat{\mathbf{p}}^{obj}_{r,m,m}(k-1)] \tag{6.14}$$

Template Resize

Once a template is initialized, it is searched for in all subsequent images. However, with a changing distance to the tracked object, the dimensions of the object projected into the image domain change. In order to compensate for this effect, the template is resized before the matching process. The scaling factor s is calculated from the x-coordinate of the rear of the object in the camera coordinate system assigned to the template x_t^{cam} and the predicted x-coordinate $\hat{x}^{cam}(k|k-1)$

$$s(k) = \frac{x_t^{cam}}{\hat{x}^{cam}(k|k-1)} \tag{6.15}$$

The predicted x-coordinate $\hat{x}^{cam}(k|k-1)$ is calculated analogous to $\hat{x}^{cam}(k-1)$ in Equation (6.14). The precision of the scaling factor s depends on the uncertainty in the distance assigned to the template x_t^{cam} and the predicted x-coordinate $\hat{x}^{cam}(k|k-1)$. However, as the precision of the position measurement is mainly influenced by the accurate laser scanner measurements in the proposed fusion system, the error in the scaling factor is neglected and assumed to have no influence on the accuracy of the subsequent template matching. For inaccurate position estimates the range of possible scaling factors could be calculated from the uncertainties of x_t^{cam} and $\hat{x}^{cam}(k|k-1)$. The template search could then include the scaling factor along with the displacement into the set of parameters to be optimized.

The scaled template I_t^* is calculated by a bilinear interpolation from the original template I_t. If the scaling factor leaves the interval $0.5 < s < 2$ a new template is initialized.

Template Search

For the search of the scaled template in the current image, there are many different algorithms, including optical flow methods [8]. A correlation based search of the image region is chosen as adequate. The normalized correlation coefficient $r(\Delta\hat{u}, \Delta\hat{v})$ represents the similarity of two image regions I_1 and I_2, where $\bar{g_1}$ and $\bar{g_2}$ denote the mean gray value of the image regions and $g_i(u, v)$ the gray value at the position $\begin{bmatrix} u & v \end{bmatrix}'$ in the image region i

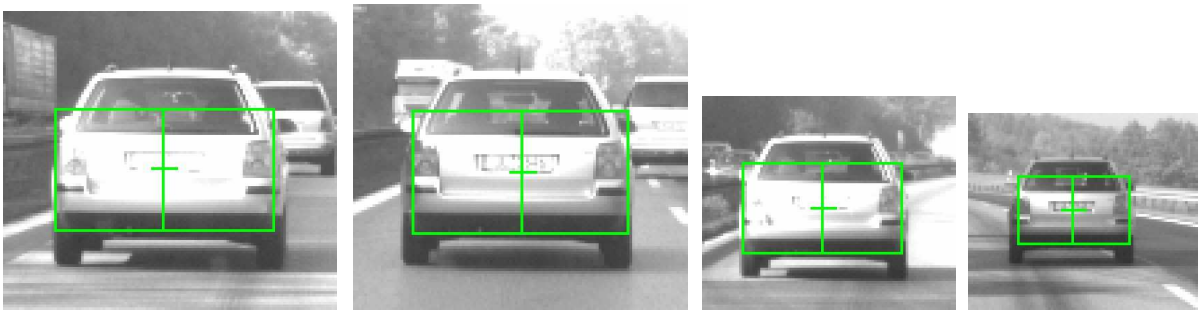
$$r(\Delta\hat{u}, \Delta\hat{v}) = \frac{\sum\limits_{(u,v)' \in I_1, I_2(\Delta\hat{u}, \Delta\hat{v})} \left(g_1(u,v) - \bar{g_1}\right)\left(g_2(u,v) - \bar{g_2}\right)}{\sqrt{\sum\limits_{(u,v)' \in I_1} \left(g_1(u,v) - \bar{g_1}\right)^2} \sqrt{\sum\limits_{(u,v)' \in I_2(\Delta\hat{u}, \Delta\hat{v})} \left(g_2(u,v) - \bar{g_2}\right)^2}} \tag{6.16}$$

It delivers a measure in the interval $r \in [-1, 1]$ and is not sensitive to global illumination changes. The image region $I_1 = I_t^*$ represents the scaled template. It is compared to different image regions $I_2(\Delta\hat{u}, \Delta\hat{v})$ of the current image $I(k-1)$. The vector $\begin{bmatrix} \Delta\hat{u} & \Delta\hat{v} \end{bmatrix}'$ denotes the position displacement between the two image regions. The best displacement $\begin{bmatrix} \Delta\hat{u}^* & \Delta\hat{v}^* \end{bmatrix}'$ is chosen by maximizing the correlation coefficient $r(\Delta\hat{u}, \Delta\hat{v})$. A common approach in order to retrieve sub-pixel accuracy is a quadratic interpolation. A parabola is fitted into the correlation coefficients in both coordinate directions. The position of the maximum of the parabola represents the final displacement estimate $\begin{bmatrix} \Delta u & \Delta v \end{bmatrix}'$ with sub-pixel accuracy

$$\begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} \Delta\hat{u}^* + \Delta\hat{u}_{sb} \\ \Delta\hat{v}^* + \Delta\hat{v}_{sb} \end{bmatrix} \tag{6.17}$$



Template at 28 m



28 m 28 m 36 m 47 m

Figure 6.10: Original template and the positions of the resized template in subsequent images.

where the sub-pixel correction vector $\begin{bmatrix}\Delta\hat{u}_{sb} & \Delta\hat{v}_{sb}\end{bmatrix}'$ is given by (Appendix A.3)

$$\begin{bmatrix}\Delta\hat{u}_{sb}\\ \Delta\hat{v}_{sb}\end{bmatrix} = \begin{bmatrix}\frac{r(\Delta\hat{u}^*-1,\Delta\hat{v}^*)-r(\Delta\hat{u}^*+1,\Delta\hat{v}^*)}{2[r(\Delta\hat{u}^*+1,\Delta\hat{v}^*)+r(\Delta\hat{u}^*-1,\Delta\hat{v}^*)]-4r(\Delta\hat{u}^*,\Delta\hat{v}^*)}\\ \frac{r(\Delta\hat{u}^*,\Delta\hat{v}^*-1)-r(\Delta\hat{u}^*,\Delta\hat{v}^*+1)}{2[r(\Delta\hat{u}^*,\Delta\hat{v}^*+1)+r(\Delta\hat{u}^*,\Delta\hat{v}^*-1)]-4r(\Delta\hat{u}^*,\Delta\hat{v}^*)}\end{bmatrix}$$

The mean of the parabola maxima is used as a final correlation coefficient $r(\Delta u, \Delta v)$ (Appendix A.3)

$$\begin{aligned} r(\Delta u,\Delta v) &= r(\Delta\hat{u}^*,\Delta\hat{v}^*)[1+\tfrac{1}{2}(\Delta\hat{v}_{sb}^2+\Delta\hat{u}_{sb}^2)] \\ &\quad -\tfrac{1}{4}\left[r(\Delta\hat{u}^*+1,\Delta\hat{v}^*)+r(\Delta\hat{u}^*-1,\Delta\hat{v}^*)\right]\Delta\hat{u}_{sb}^2 \\ &\quad -\tfrac{1}{4}\left[r(\Delta\hat{u}^*,\Delta\hat{v}^*+1)+r(\Delta\hat{u}^*,\Delta\hat{v}^*-1)\right]\Delta\hat{v}_{sb}^2 \end{aligned} \tag{6.18}$$

The resulting position of the center of the scaled template $\begin{bmatrix}u=\Delta u & v=\Delta v\end{bmatrix}'$ is then taken as a measurement. The size of the search region is derived from the predicted uncertainty of the template position $\hat{\sigma}_u(k|k-1)$, with

$$\begin{bmatrix}\hat{\sigma}_u^2(k|k-1) & \hat{\sigma}_{uv}(k|k-1)\\ \hat{\sigma}_{uv}(k|k-1) & \hat{\sigma}_v^2(k|k-1)\end{bmatrix} = \mathbf{H}_{img}(k)\mathbf{P}(k|k-1)\mathbf{H}'_{img}(k) \tag{6.19}$$

where $\mathbf{H}_{img}$ is the Jacobi matrix of measurement function $h_{img}[k, \hat{\mathbf{x}}_a(k|k-1)]$ from Equation (6.24). A good choice is a 3σ-gate for the search region.

Figure 6.10 shows an exemplary template and the positions of the resized template in subsequent images. It can be observed, that the correlation based search is widely insensitive to illumination changes and distance changes. After the last cycle the template is updated, as the correlation coefficient becomes too low.

6.3.2 Augmented State Representation

In this approach, an image region defines initially a template which is searched for in all subsequent images. Unfortunately, the object positions in the subsequent images cannot

be used directly as measurements for a Kalman filter. The error of the object positions is strongly correlated with the error in the initial template position, which is derived from the state estimate. The measurements are all biased, which violates the assumptions of the Kalman filter of zero mean measurement noise. Therefore, this offset in the initial template position must be incorporated into the state estimation.

The error in the initial template position is given by

$$\begin{bmatrix} o_u(k_t) \\ o_v(k_t) \end{bmatrix} = \begin{bmatrix} h_u^{img}[\hat{\mathbf{x}}(k_t), \hat{\mathbf{p}}_{r,m,m}^{obj}(k_t)] - h_u^{img}[\mathbf{x}(k_t), \mathbf{p}_{r,m,m}^{obj}(k_t)] \\ h_v^{img}[\hat{\mathbf{x}}(k_t), \hat{\mathbf{p}}_{r,m,m}^{obj}(k_t)] - h_v^{img}[\mathbf{x}(k_t), \mathbf{p}_{r,m,m}^{obj}(k_t)] \end{bmatrix} \tag{6.20}$$

where k_t is the measurement interval in which the template is initialized. Once a template is initialized, the state is augmented by $\begin{bmatrix} o_u & o_v \end{bmatrix}'$ resulting in the augmented state representation $\mathbf{x_a}$, with

$$\mathbf{x}_a(k) = \begin{bmatrix} x & y & z & \psi & v_x & v_y & d_x & d_y & d_z & o_u & o_v \end{bmatrix}' \tag{6.21}$$

The offset $\begin{bmatrix} o_u(k_t) & o_v(k_t) \end{bmatrix}'$ is handled as a parameter. The process covariance is therefore zero. The offset $\begin{bmatrix} o_u(k_t) & o_v(k_t) \end{bmatrix}'$ is set to zero, when the template is generated and subsequently estimated in the Kalman filter by the fused laser scanner and video data. The associated initial variances $\sigma_{o_u}^2(k_t)$ and $\sigma_{o_v}^2(k_t)$ are calculated from the uncertainty of the coordinates of the rear feature, which are given by

$$\begin{bmatrix} \sigma_{o_u}^2(k_t) & \sigma_{o_u o_v}(k_t) \\ \sigma_{o_u o_v}(k_t) & \sigma_{o_v}^2(k_t) \end{bmatrix} = \mathbf{H}_{img}(k_t)\mathbf{P}(k_t)\mathbf{H}'_{img}(k_t) \tag{6.22}$$

where $\mathbf{H}_{img}(k_t)$ is the Jacobi matrix of measurement function $\mathbf{h}^{img} = \begin{bmatrix} h_u^{img} & h_v^{img} \end{bmatrix}'$ from Equation (6.3)

$$\mathbf{H}_{img}(k_t) = \left\{ \nabla_{\mathbf{x}} \left[\mathbf{h}^{img}(\mathbf{p}_{r,m,m}^{obj}) \right]' \right\}' \Bigg|_{\mathbf{x}=\hat{\mathbf{x}}(k_t)} \tag{6.23}$$

6.3.3 Feature Measurement Equation

The measurement equation for the center of the template image region follows from Equation (6.20). It is given by the measurement function of the center of the rear of the vehicle corrected by the offset of the template $\begin{bmatrix} o_u & o_v \end{bmatrix}'$

$$\mathbf{h}_{img}[k, \mathbf{x}_a(k)] = \begin{bmatrix} h_u^{img}[\mathbf{x}_a(k), \mathbf{p}_{r,m,m}^{obj}(k)] + o_u(k) \\ h_v^{img}[\mathbf{x}_a(k), \mathbf{p}_{r,m,m}^{obj}(k)] + o_v(k) \end{bmatrix} \tag{6.24}$$

6.4 Discussion

In this chapter classical model based feature extraction algorithms are outlined. The possibility of their integration into the feature-level fusion has been demonstrated and the measurement equations have been derived.

In order to improve the performance of the image processing a novel feature-level fusion of an image template tracking and a laser scanner data based object detection has been developed. While the objects are detected and localized by the laser scanner, the template

tracking aims at an improvement of the estimates, exploiting the high angular resolution of the image. The advantage of the template tracking lies in its robustness against different object types and the applicability in a wide range of traffic scenarios.

In the following chapter the measurement accuracy of the template tracking will be determined empirically. The evaluation in Chapter 8 finally demonstrates the performance gain of the fusion approach with respect to a purely laser scanner based object tracking.

7 Parameter Identification

The measurement uncertainty of the laser scanner and the camera are characterized in different experiments in this chapter. This analysis is important for an efficient parameterization of the Kalman filter. The determined measurement uncertainties can be directly used for the generation of the measurement covariance matrix.

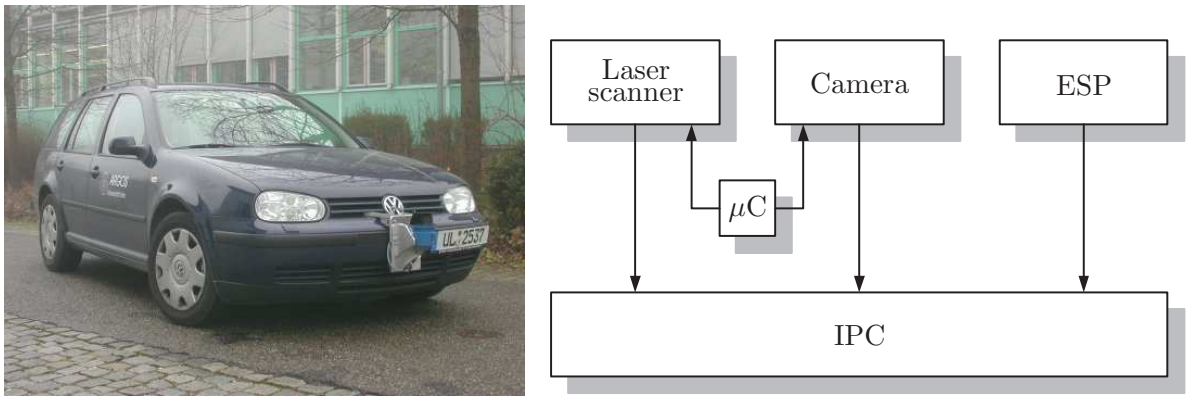
7.1 Experimental Platform

The Department of Measurement, Control and Microtechnology at the University of Ulm possesses several cars equipped with sensors for an environment perception. Throughout this thesis a Volkswagen Golf equipped with a laser scanner at the front bumper and a video camera behind the windshield has been used as test vehicle (Figure 7.1(a)). The multi-layer laser scanner, produced by the company IBEO Automobile Sensor GmbH, is described in detail in Chapter 5.1. The camera is a digital 12 bit CCD pixelfly camera manufactured by PCO [115]. A 12 mm lens enables a horizontal aperture of 28°. A microcontroller triggers the laser scanner and the camera synchronously at a frequency of 20 Hz (Figure 7.1(b)). A controller in the laser scanner varies the speed of the motor which turns the laser scanner head, so that the head points forwards when the trigger signal is received. This procedure minimizes the time difference between the image and the laser scanner measurements. A central industrial personal computer records the measurements from the laser scanner, camera and the electronic stability program controller (ESP). The ESP delivers wheel speeds, the yaw rate, the steering wheel angle and the cross acceleration, which are used for the ego-motion estimation. The mounting position of the laser scanner and the camera are measured manually and the sensors orientations are determined by the calibration described in Chapter 3. The tracking algorithms described throughout this work are applied offline to the previously recorded data.

7.2 Identification of the Measurement Error Probability Distributions

7.2.1 Laser Scanner — Azimuth

The azimuth measurement accuracy of the laser scanner depends on the measurement feature and therefore on the feature extraction process. The different methods described in Chapter 5.4 can be roughly classified in line fitting processes for the rear, front or side of a vehicle and the L-shape fitting, if more than one side of a vehicle is observed simultaneously. The U-shape falls for this examination into the category of the line segments. The theoretical and experimental analysis of the measurement uncertainty



(a) (b)

Figure 7.1: (a) Experimental vehicle. The laser scanner is mounted at the front bumper and the camera behind the windshield beside the inner rear mirror. (b) System architecture. The microcontroller triggers the laser scanner and the camera in order to enable a synchronization.

is described exemplarily in the following for the most prominent case of the line fitting process.

A line segment describes the position of one side of a vehicle in the laser scanner coordinate system. The two ends of the line segment correspond to the observed corners of the vehicle. The accuracy of the azimuthal position of the line segment, which is defined by the point in the center between the two ends, depends mainly on the azimuthal accuracy of the measurements delimiting the line segment. For line segments which are perpendicular to the observing laser scanner, the variance of the line segments azimuthal position is given by

$$\sigma_{\alpha^*}^2 = \frac{1}{2}\sigma_\alpha^2 \tag{7.1}$$

where σ_α^2 is the variance of the azimuth measurement of a corner.

Theoretical Measurement Uncertainty

The theoretical measurement uncertainty of the azimuth measurement σ_α^2 of an object's corner can be derived from the measurement principle of the laser scanner. Figure 7.2(a) shows four detections on a vehicle's rear. The detections are depicted as rectangular regions, because of the beam width of the laser pulse. As can be seen at the left corner of the vehicle's rear, a detection can in general be established even if the cross section of the laser pulse only partly overlaps with the object. As the degree of overlap can not be measured, the exact position of the corner is not known. The azimuthal measurement of the corner of the object is therefore uncertain and might be biased.

The uncertainty and bias of the corner measurement depend on the azimuthal resolution $\Delta\alpha$ of the laser scanner, the azimuthal beam width β (Figure 7.2(b)) and the amount of overlap $\lambda \in [0, 1]$, which suffice for a detection. For the calculation of the bias and the uncertainty, it is assumed that the probability density of the real position of the vehicle's



(a) (b)

Figure 7.2: Four detections on a vehicle's rear. The azimuthal laser beam width is denoted β, the angular resolution $\Delta\alpha$ and the angle of the detection at the right corner α_r.

corner is uniformly distributed over a certain angle. The minimal and maximal angles α_{min} and α_{max}, where a right corner might be located, are given by

$$\alpha_{min} = \alpha_r - \frac{\beta}{2} + \lambda\beta \tag{7.2}$$

$$\alpha_{max} = \alpha_r + \Delta\alpha - \frac{\beta}{2} + \lambda\beta \tag{7.3}$$

where α_r denotes the azimuth angle of the detection at the corner (Figure 7.2(b)). The angle α_{min} denotes the minimal angle, at which the corner overlaps sufficiently with the laser pulse, in order to trigger a detection at α_r. The angle α_{max} corresponds to the maximal angle, which just does not trigger a detection in the next pulse $\alpha_r + \Delta\alpha$. Without loss of generality it is assumed in the following that

$$\alpha_r = 0 \tag{7.4}$$

The bias μ_α in the azimuth of a right corner measurement is given by the integral over the angular probability distribution

$$\begin{aligned}
\mu_\alpha(\Delta\alpha, \beta, \lambda) &= \int_{\alpha_{min}}^{\alpha_{max}} \frac{1}{\Delta\alpha}\alpha d\alpha = \int_{-\frac{\beta}{2}+\lambda\beta}^{\Delta\alpha-\frac{\beta}{2}+\lambda\beta} \frac{1}{\Delta\alpha}\alpha d\alpha \\
&= \frac{1}{2\Delta\alpha}\alpha^2 \Big|_{-\frac{\beta}{2}+\lambda\beta}^{\Delta\alpha-\frac{\beta}{2}+\lambda\beta} \\
&= \frac{1}{2\Delta\alpha}\left[\left(\Delta\alpha - \frac{\beta}{2} + \lambda\beta\right)^2 - \left(-\frac{\beta}{2} + \lambda\beta\right)^2\right] \\
&= \frac{1}{2\Delta\alpha}\left[\Delta\alpha^2 + 2\Delta\alpha\left(\lambda\beta - \frac{\beta}{2}\right)\right] \\
&= \frac{1}{2}(\Delta\alpha + 2\lambda\beta - \beta)
\end{aligned} \tag{7.5}$$

It is interesting to notice, that the bias μ_α is independent of the azimuthal beam width β for $\lambda = 0.5$

$$\mu_\alpha(\Delta\alpha, \beta, \lambda = 0.5) = \frac{1}{2}\Delta\alpha \tag{7.6}$$

If a 50 % overlap is sufficient for a detection, which could be considered the standard case, the bias is therefore strongly significant and amounts to half of the angular resolution $\Delta\alpha$. The true azimuth angle of a right corner given a corner measurement at $\alpha_r = 0$ is on average $0.5\Delta\alpha$ further to the right and of a left corner $0.5\Delta\alpha$ further to the left, independent of the beam width. For the calculation of the position of a line segment, which is given by the mean position of two corners, this bias is not of importance, as the mean is in theory offset free. However, the angle difference, representing the width of the observed object side, is on average significantly underestimated.

In practice, the parameters β and λ are difficult to parameterize. Even if the sensor manufacturer provides values for the beam width β, it is known that the intensity of the infrared light is not uniformly distributed within the beam and might be significant beyond the specified limits. Additionally, it can be shown in experiments, that the degree of overlap λ, sufficient for a detection, depends on the reflectivity of the object.

The variance of the azimuth measurement of a corner σ_α^2 can be calculated analogously to Equation (7.5)

$$\begin{aligned}
\sigma_\alpha^2(\Delta\alpha,\beta,\lambda) &= \int_{-\frac{\beta}{2}+\lambda\beta}^{\Delta\alpha-\frac{\beta}{2}+\lambda\beta} \frac{1}{\Delta\alpha}\left[\alpha-\mu_\alpha(\Delta\alpha,\beta,\lambda)\right]^2 d\alpha \\
&= \frac{1}{3\Delta\alpha}\left[\alpha-\mu_\alpha(\Delta\alpha,\beta,\lambda)\right]^3\Big|_{-\frac{\beta}{2}+\lambda\beta}^{\Delta\alpha-\frac{\beta}{2}+\lambda\beta} \\
&= \frac{1}{3\Delta\alpha}\left\{\left[\Delta\alpha-\frac{\beta}{2}+\lambda\beta-\frac{1}{2}(\Delta\alpha+2\lambda\beta-\beta)\right]^3\right. \\
&\quad \left.-\left[-\frac{\beta}{2}+\lambda\beta-\frac{1}{2}(\Delta\alpha+2\lambda\beta-\beta)\right]^3\right\} \\
&= \frac{1}{3\Delta\alpha}\left[\left(\frac{1}{2}\Delta\alpha\right)^3-\left(-\frac{1}{2}\Delta\alpha\right)^3\right] \\
&= \frac{1}{12}\Delta\alpha^2 \qquad (7.7)
\end{aligned}$$

Therefore the measurement uncertainty of the azimuth is independent of the beam width β and λ and depends only on the angular resolution $\Delta\alpha$.

It should be mentioned that the probability distribution of the azimuth of the line segments position, which is the mean position of the two corners of a side, has the shape of a triangular function, as the corners azimuth angle is assumed to be uniformly distributed. Its variance is given in Equation (7.1). Figure 7.3(a) shows a histogram of a simulated triangular probability density function and a Gaussian approximation with equal variance. In Figure 7.3(b) a certain percentage of outliers were additionally considered. Outliers denote in this context missed detections at the corners of the object. It can be seen, that the probability density distribution obtains additional tails, and the Gaussian approximation becomes wider.

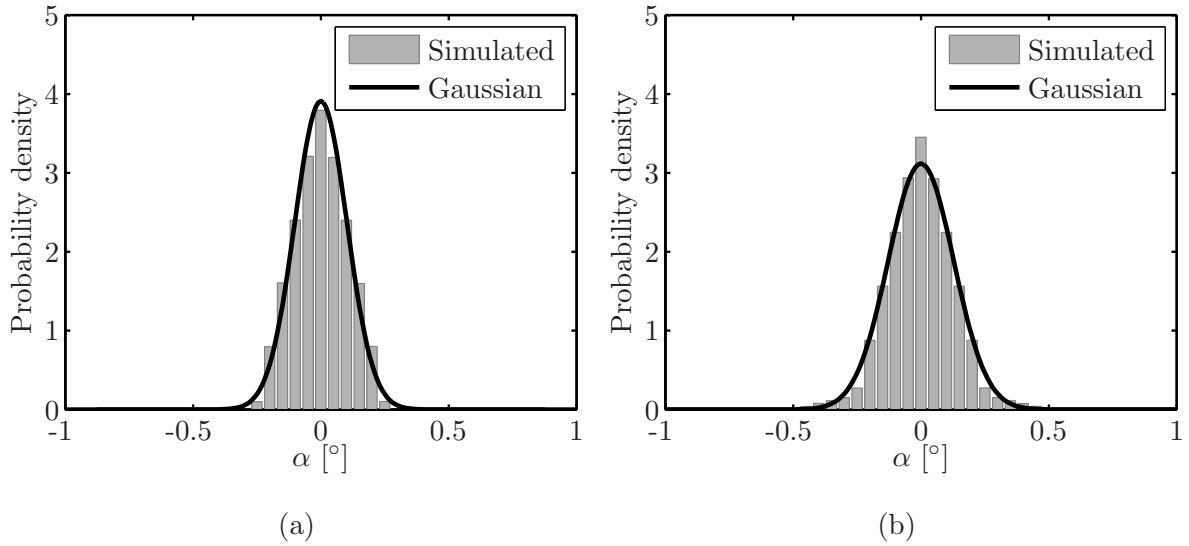


Figure 7.3: (a) Histogram of a simulated triangular probability density function (pdf) for $\Delta\alpha = 0.5°$ and the corresponding Gaussian pdf with $\sigma^2_{\alpha^*} = \frac{1}{24}\Delta\alpha^2$. (b) Same plot as in (a), but accounting for some missed detections at the corners which leads to a the heavy tails of the simulated function.

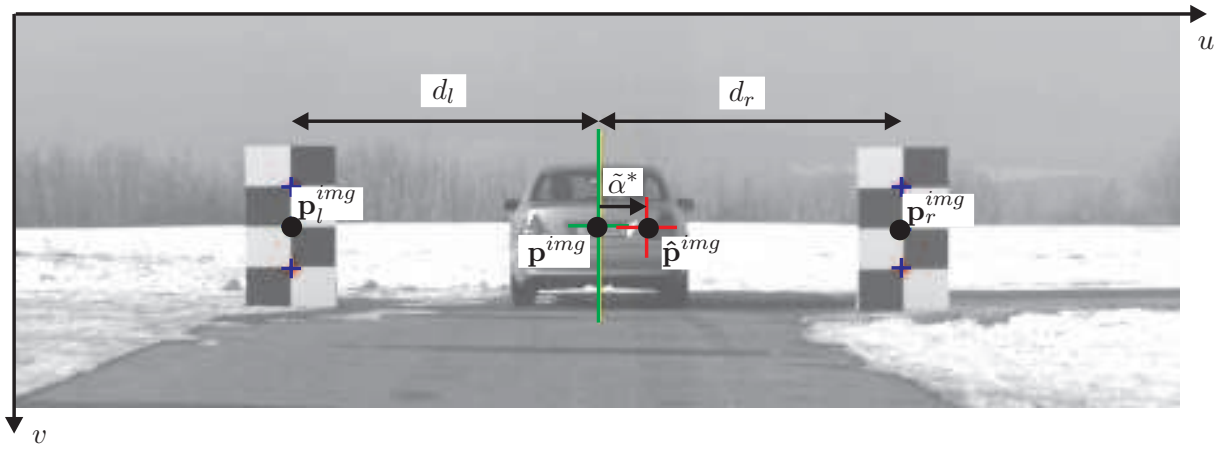


Figure 7.4: Experimental setup for the calculation of azimuthal measurement and estimation errors. The distances d_l and d_r are measured manually. The points $\mathbf{p}_l^{img}$ and $\mathbf{p}_r^{img}$ are determined using image processing techniques. The reference position in the image $\mathbf{p}^{img}$ results from $\mathbf{p}_l^{img}$ and $\mathbf{p}_r^{img}$ and the knowledge of d_l and d_r. $\hat{\mathbf{p}}^{img}$ denotes the estimated position and $\tilde{\alpha}^*$ the error in the azimuth.

Experimental Setup

The experimental setup for a verification of the theoretically derived measurement uncertainty is illustrated in Figure 7.4. Different vehicles are positioned between two reference objects. The rear of the vehicles is aligned with a plane through the two checker pattern surfaces. The host vehicle, equipped with a calibrated laser scanner and video camera, is driven from large distances towards the object vehicle. Measurements are taken synchro-



(a) (b)

Figure 7.5: Templates, which are correlated with the images, in order to retrieve the position of the reference objects.

nously with the laser scanner and the video camera at 20 Hz measurement frequency. For each measurement cycle the image position of the two reference objects is calculated using image processing techniques. From these positions and the known real world position of the vehicle between the reference objects, the position of the center of the rear can be derived in image coordinates. This reference position $\mathbf{p}^{img}$ is then transformed into a ray in the laser scanner coordinate system. The angle between this ray and the measured position $\hat{\mathbf{p}}$ results in the azimuthal measurement error $\tilde{\alpha}^*$ in the polar laser scanner coordinate system. For illustration purposes the projection of the measured position $\hat{\mathbf{p}}^{img}$ is depicted in Figure 7.4 with the resulting measurement error $\tilde{\alpha}^*$.

Cars were chosen as objects for the derivation of the measurement errors as they represent the most important class of objects for the environment perception described in this work. As the measurement errors are most likely to depend on the object class, the following error statistics apply strictly speaking only for cars observed from the rear.

The image position of the left and the right reference objects $\mathbf{p}_l^{img}$ and $\mathbf{p}_r^{img}$ are obtained with template matching techniques, utilizing the templates shown in Figure 7.5. The image is searched for the vertical arrangement of the templates (b)-(a)-(b) from Figure 7.5. The points $\mathbf{p}_l^{img}$ and $\mathbf{p}_r^{img}$ are calculated with a linear regression from the three template positions ((b)-(a)-(b)).

The reference position of the vehicle in the image $\mathbf{p}^{img}$ is then derived from the real world distance of the center of the vehicle to the left and right reference object d_l and d_r.

$$\mathbf{p}^{img} = \frac{d_l}{d_l + d_r} \left(\mathbf{p}_r^{img} - \mathbf{p}_l^{img} \right) + \mathbf{p}_l^{img} \tag{7.8}$$

This assumes, that the image plane is parallel to a plane through the two reference objects. Additionally the rear of the vehicle must lie in this plane. The distances d_l and d_r must be measured once for each vehicle.

Once the reference image position is calculated, the corresponding feature position of the vehicle's rear, measured by the laser scanner, must be associated. One way could be, to project all measured feature positions into the image and assign the closest in image coordinates. This procedure however proved to be susceptible to measurement errors, which are close to the reference position in the image, but deviated strongly in their distance from the distance to the rear of the vehicle. Therefore, the distance to the rear of the vehicle is estimated from the image and incorporated into the association.

The distance to the rear of the vehicle z^{cam^*} in the pinhole camera coordinate system $\mathcal{C}_{cam^*}$ can be derived from the distance of the reference objects in the image and the known distance in the world coordinate system. As it is assumed, that the image plane is

parallel to a plane through the two reference objects, the z-coordinates z^{cam^*} of the two reference objects in the pinhole camera coordinate system $\mathcal{C}_{cam^*}$ are equal. The projection in Equation (3.27) delivers for two points in the pinhole camera coordinate system with identical z-coordinate

$$\begin{aligned} s(u_1 - u_2) &= \alpha_u \left(x_1^{cam^*} - x_2^{cam^*} \right) && (7.9) \\ s &= z^{cam^*} && (7.10) \end{aligned}$$

Combining Equation (7.9) and (7.10) results after rearrangement in

$$z^{cam^*} = \frac{\alpha_u \left(x_1^{cam^*} - x_2^{cam^*} \right)}{(u_1 - u_2)} \tag{7.11}$$

Substituting $\left(x_1^{cam^*} - x_2^{cam^*} \right)$ by the distance of the two reference objects $(d_l + d_r)$ and u_1 and u_2 by the u-coordinates of the reference objects in the image u_r and u_l, with $\mathbf{p}_l^{img} = \begin{bmatrix} u_l & v_l \end{bmatrix}'$ and $\mathbf{p}_r^{img} = \begin{bmatrix} u_r & v_r \end{bmatrix}'$, yields

$$z^{cam^*} = \frac{\alpha_u \left(d_l + d_r \right)}{u_r - u_l} \tag{7.12}$$

The thus estimated distance z^{cam^*} to the rear of the vehicle is not very accurate in large distances compared to a laser scanner distance measurement. The accuracy, however, suffices for an association. The distance z^{cam^*} and the position in the image $\mathbf{p}^{img}$ is compared to the different laser scanner measurements. Of those laser scanner measurement, which fall into a set gate, the closest measurement $\hat{\mathbf{p}}^{ls}$ is chosen.

After the association is achieved, the error of the measurement $\hat{\mathbf{p}}^{ls}$, associated to the reference position, can be determined. The angular error $\tilde{\alpha}^*$ of the laser scanner azimuth measurement is calculated in the xy-plane of the laser scanner coordinate system as depicted in Figure 7.6. For this purpose the reference position $\mathbf{p}^{ls}$ must be constructed from the reference position in the image $\mathbf{p}^{img}$. However, the reference position in the image $\mathbf{p}^{img}$ delivers only a ray $\mathbf{g}$ in the laser scanner coordinate system, on which the true position of the vehicle's rear must be located. As the distance is not known with sufficient accuracy from the image processing especially for large distances, the radial distance r of the associated laser scanner rear measurement $\hat{\mathbf{p}}^{ls}$ is taken as the true distance. The reference position of the rear is then given by the point along the ray $\mathbf{g}$, which has a distance of r to the origin of the laser scanner coordinate system. The reference position $\mathbf{p}^{ls}$ results from the intersection of the ray $\mathbf{g}$ and a circle of radius r with its center in the origin of the laser scanner coordinate system (Figure 7.6). The azimuth measurement error $\tilde{\alpha}^*$ is then calculated from $\mathbf{p}^{ls}$ and $\hat{\mathbf{p}}^{ls}$, with

$$\tilde{\alpha}^* = \arctan\left(\frac{\hat{y}^{ls}}{\hat{x}^{ls}} \right) - \arctan\left(\frac{y^{ls}}{x^{ls}} \right) \tag{7.13}$$

A detailed derivation of the reference position $\mathbf{p}^{ls}$ in the laser scanner coordinate system is provided in Appendix A.4.

For a reference system it is of major importance, that its accuracy is one order of magnitude higher than the evaluated measurement system. Therefore, it is estimated in

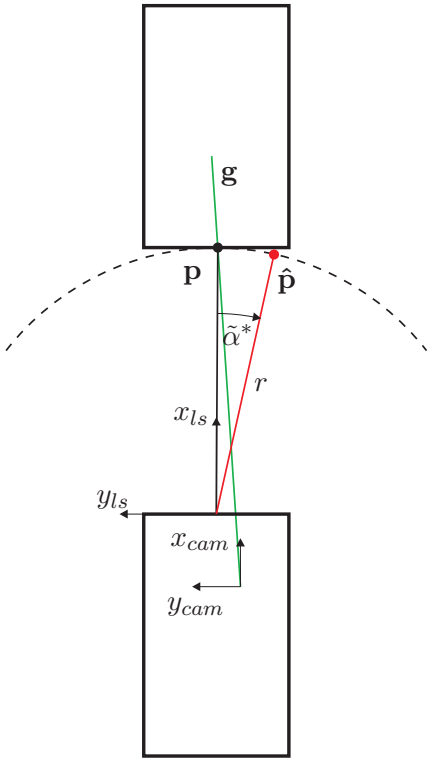


Figure 7.6: Azimuth measurement error $\tilde{\alpha}^*$ in the xy-plane of the laser scanner coordinate system. The reference position is denoted by $\mathbf{p}$ and $\hat{\mathbf{p}}$ is the estimate.

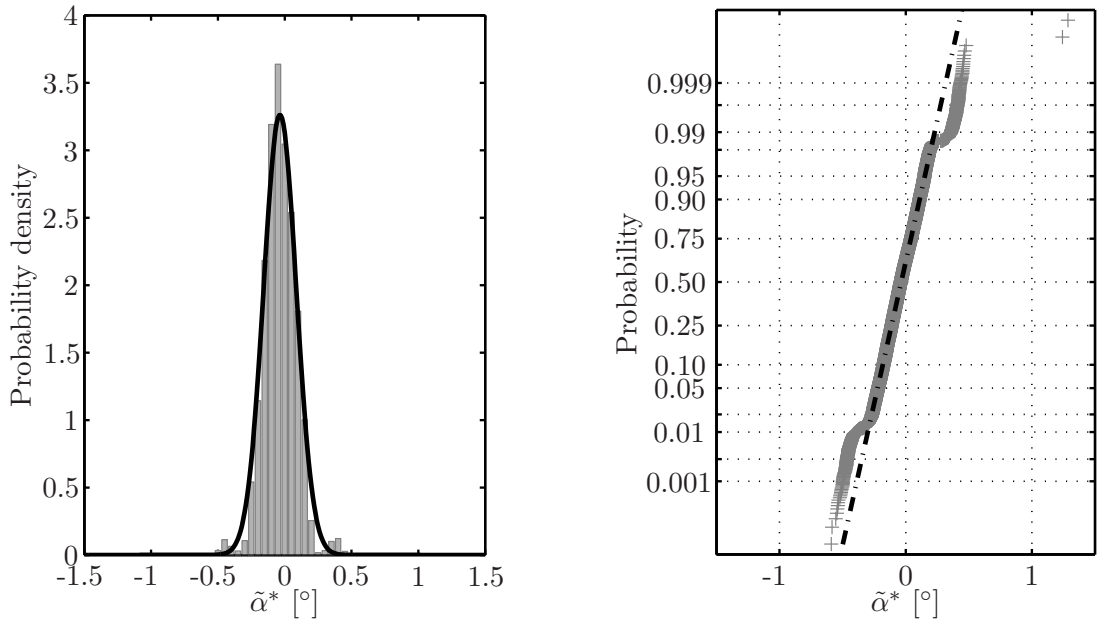
the following. The accuracy of the reference angle α_r, with

$$\alpha_r = \arctan\left(\frac{y^{ls}}{x^{ls}}\right) \tag{7.14}$$

is affected by three different error sources:

- The image processing delivers the reference point in the image only up to a certain accuracy.
- The transformation of the reference point in the image into a ray in the laser scanner coordinate system is influenced by the precision of the calibration process.
- The distance measurement of the laser scanner taken as the ground truth for the calculation of the angular error is impaired by a distance dependent bias and a measurement noise.

As the variance of the measurement uncertainty is of interest in this experiment, the calibration errors can be neglected, as they impose mainly a bias. The error in the distance measurement of the laser scanner is not negligible in close distances. However, the influence on the reference angle decreases with the distance of the vehicle's rear. Therefore a minimal distance was calculated, for which the error in the reference angle α_r, caused by the error in the distance measurement and the image processing, is on average approximately $\frac{1}{10}$ of the expected measurement error $\tilde{\alpha}^*$. This minimal distance is 30 m.



(a) (b)

Figure 7.7: (a) Empirical probability density function and its Gaussian approximation of the azimuth measurement error $\tilde{\alpha}^*$. (b) Normal probability plot (npp) of the azimuth error $\tilde{\alpha}^*$. Superimposed an extrapolated line joining the first and third quartiles.

Realization of the Experiment

Overall 95 drives towards or backwards away from a vehicle positioned between the two reference objects were conducted. As object cars, a Toyota Carina, a Volkswagen Passat and a Mercedes S-Class were used. The laser scanner and the camera were triggered by a central microcontroller at 20 Hz measurement frequency. A total of 20,874 measurement cycles were recorded. The host vehicle was driven in a straight line at speeds varying between 10–40 km/h.

Experimental Results

For the statistical evaluation only measurements in a range further than 30 m are considered. The laser scanner feature extraction matches contour models to the segmented laser scanner raw measurements as described in Chapter 5.4.1. The middle of the line segment corresponding to the rear of the observed car is taken as measured position of the vehicle's rear $\hat{\mathbf{p}}^{ls}$. In large distances the vehicle's rear triggers sometimes only one raw measurement. In those cases this single measurement is refused for this evaluation, as the expected variance of a single measurement is twice the variance of the average position of two points as shown above.

For each measurement cycle the azimuth measurement error $\tilde{\alpha}^*$ is calculated from Equation (7.13). The resulting empirical probability density function (pdf) and its Gaussian approximation is shown in Figure 7.7(a). In the center, the pdf seems to be similar to the

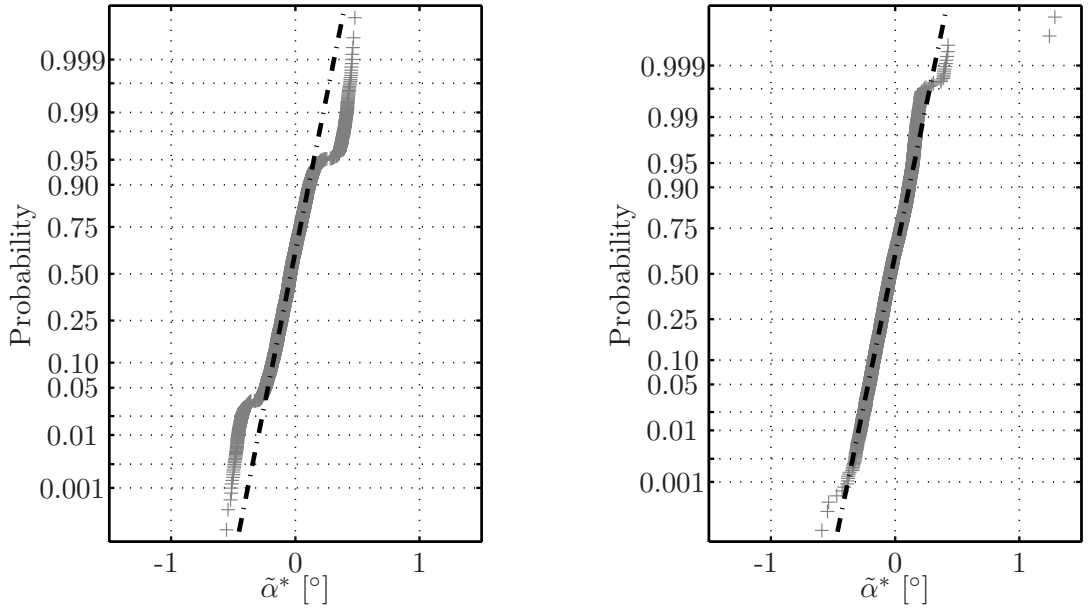


Figure 7.8: (a) Normal probability plot (npp) of the azimuth error $\tilde{\alpha}^*$ of the Carina sample set. (b) Normal probability plot (npp) of the S-Class sample set.

Gaussian, it however exhibits heavy tails. These phenomena can be better observed in the normal probability plot (npp) in Figure 7.7(b). In a normal probability plot the data is drawn against a theoretical normal distribution in such a way that the points should form an approximate straight line. Deviations from this straight line indicate deviations from normality [29]. Figure 7.7(b) shows the sample data and superimposed a line joining the first and third quartiles. This line is extrapolated out to the ends of the sample [108]. The distribution is very linear in the probability range of $[0.05, 0.95]$, which indicates a normal distribution of the data in the center. However, at the tails the distribution diverges from the ideal line. This is due to the heavy tails exhibited in the pdf. Comparing the probability density with the expected function in Figure 7.3, a good similarity can be observed. Therefore, the heavy tails are most likely caused by missed detections at the corners of the vehicle's rear as indicated in the simulated pdf in Figure 7.3(b).

Figure 7.8 contrasts the normal probability plots (npp) for the Toyota Carina and the S-Class sample data. While the npp of the S-Class is linear over almost the complete data range, the npp of the Carina exhibits strong deviations at the tails. It can therefore be deduced that the amount of missed detections at the corners of the vehicle's rear depends on the observed vehicle. The size and the quality of the reflectors and the amount of dirt covering the observed car certainly influence the detection characteristics of the laser scanner.

Finally, Table 7.1 concludes the sample mean $\mu_{\tilde{\alpha}^*}$, the sample standard deviation $s_{\tilde{\alpha}^*}$ and the root mean square ($RMS = \sqrt{\mu_{\tilde{\alpha}^*}^2 + s_{\tilde{\alpha}^*}^2}$) of the different samples of the azimuth measurement error $\tilde{\alpha}^*$. The standard deviation $s_{\tilde{\alpha}^*}$ of the Carina sample is significantly larger than the theoretical value, which can be explained by the heavy tails of the pdf.

Table 7.1: Sample mean $\mu_{\tilde{\alpha}^*}$, sample standard deviation $s_{\tilde{\alpha}^*}$ and the RMS of the different samples of the azimuth measurement error $\tilde{\alpha}^*$.

	# Meas.	$\mu_{\tilde{\alpha}^*}$ [°]	$s_{\tilde{\alpha}^*}$ [°]	$\text{RMS}_{\tilde{\alpha}^*}$ [°]
Theoretical	–	0	0.10	0.10
Carina	4,317	-0.03	0.15	0.15
Passat	9,157	-0.04	0.12	0.12
S-Class	7,189	-0.03	0.11	0.12
All	20,874	-0.04	**0.12**	0.13

The values for the Passat deviates from the theoretical value only in the magnitude of twice the measurement accuracy of the reference system of 0.01°. The sample standard deviation of the S-Class sample is within that accuracy. The sample means $\mu_{\tilde{\alpha}^*}$ are of the same small magnitude. They are most likely caused by the residual error from the calibration procedure.

As an overall measurement uncertainty for the parameterization of the Kalman filter a standard deviation for the azimuth measurement of the laser scanner of $\sigma_{\alpha^*} = 0.12°$ is chosen, which is the average sample standard deviation given in Table 7.1. This standard deviation assumes, that both corners of a vehicle are observed. The azimuth measurement corresponds to the mean position of both corner measurements. If only one corner of the vehicle is observed, a standard deviation of $\sigma_\alpha = \sqrt{2}\sigma_{\alpha^*} = 0.17°$ is chosen according to Equation (7.1).

7.2.2 Laser Scanner — Distance

A detailed analysis of the sample standard deviation and bias of the laser scanner's distance measurements can be found in [53]. The standard deviation and bias depend on the reflectivity of the observed object. For cars and passengers this influence can be widely neglected in a distances above 2 m. Under ideal circumstances the standard deviation of measurements on the rear of a stationary vehicle is smaller than 0.06 m. The bias varies between 0.05 and 0.2 m.

These experiments however neglect the influence of measurements underneath or inside the vehicle in dynamic scenarios, which increase the standard deviation. Additionally, as the bias is distance dependent, it can change abruptly with the distance. This behavior is best modeled by an increased standard deviation. As it is hard to quantify these measurement errors, a measurement uncertainty with standard deviation of $\sigma_r = 0.15$ m is assumed.

7.2.3 Laser Scanner — Position

The position of the laser scanner features are provided by the feature extraction in cartesian coordinates. Therefore, the azimuthal and radial uncertainties, derived in the previous sections, are transformed into a covariance matrix in cartesian coordinates as ex-

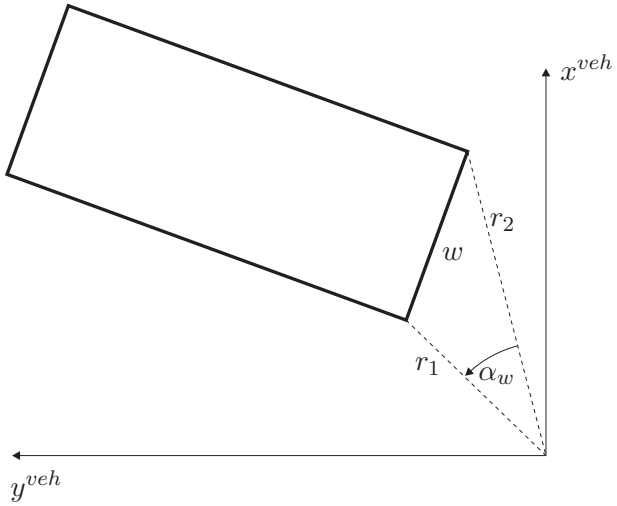


Figure 7.9: The width of an object w results from the distances to the corners r_1 and r_2 and the angle α_w between the two corner points.

plained in Chapter 5.4.2. This transformation is performed for all features, of which the distance r_0 is larger than 30 m. At shorter distances the violation of the cuboidal object model is not negligible compared to the angular resolution. Round corners of cars, different width of the car at different heights become noticeable. These not observable object model induced uncertainties are here modeled by an increased measurement uncertainty in cartesian coordinates. Standard deviations of $\sigma_x = \sigma_y = 0.15$ m are chosen heuristically for distances below 30 m.

7.2.4 Laser Scanner — Width and Length

Theoretical Measurement Uncertainty

The width and length of the observed car is provided by the laser scanner feature extraction as the distance between the two end points of the corresponding line segment. As shown in Figure 7.9 the law of cosine can be used to calculate the width from the angle α_w between the two points, delimiting the line segment, and the distances r_1 and r_2

$$w = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\alpha_w)} \tag{7.15}$$

The uncertainty of the width, represented by its variance σ_w^2, therefore depends on the uncertainty of the distance measurements and of the angle difference. The variance of the angle difference $\sigma_{\alpha_w}^2$ amounts theoretically to twice the variance of the individual angle measurement σ_α^2 given in Equation (7.7)

$$\sigma_{\alpha_w}^2 = 2\sigma_\alpha^2 = \frac{1}{6}\Delta\alpha^2 \tag{7.16}$$

where $\Delta\alpha$ is the azimuthal measurement resolution. The measurement variance of the width σ_w results then from the variance of the distance measurements σ_r^2 and the variance

of the angle difference $\sigma^2_{\alpha_w}$

$$\sigma^2_w = W \begin{bmatrix} \sigma^2_r & 0 & 0 \\ 0 & \sigma^2_r & 0 \\ 0 & 0 & \sigma^2_{\alpha_w} \end{bmatrix} W' = W \begin{bmatrix} \sigma^2_r & 0 & 0 \\ 0 & \sigma^2_r & 0 \\ 0 & 0 & 2\sigma^2_\alpha \end{bmatrix} W' \tag{7.17}$$

with the Jacobian of the width given by

$$W = \begin{bmatrix} \frac{\partial w}{\partial r_1} & \frac{\partial w}{\partial r_2} & \frac{\partial w}{\partial \alpha_w} \end{bmatrix} \tag{7.18}$$

$$= \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\alpha_w)}} \begin{bmatrix} r_1 - r_2 \cos(\alpha_w) \\ r_2 - r_1 \cos(\alpha_w) \\ r_1 r_2 \sin(\alpha_w) \end{bmatrix}' \tag{7.19}$$

The expected probability density function of the angle difference α_w has theoretically a triangular shape, assuming an uniform distribution for the individual angles. In practice however, the two angles are strongly correlated for many real world situations. Consider the situation depicted in Figure 7.2(a). There are four detections on the vehicle. Assuming, that the necessary overlap of the laser pulse with the object must be $\lambda = 0.5$, and the distance to the car does not change, there are either three or four detections on the vehicle's rear. Consequentially, there are only two discrete possibilities for the angle difference: $\alpha_w \in \{2\Delta\alpha, 3\Delta\alpha\}$. However, in this experiment the distance to the observed vehicle changes continuously over time, so that the measurement error probability density distribution is assumed to be of triangular shape[1].

As already mentioned, the width of a vehicle is underestimated by the angle difference α_w. The theoretical bias μ_{α_w} is twice the bias for the measurement of a corner μ_α from Equation (7.5)

$$\mu_{\alpha_w} = -\Delta\alpha - 2\lambda\beta + \beta \tag{7.20}$$

where $\Delta\alpha$ denotes the azimuthal resolution of the laser scanner, β the azimuthal beam width and λ the amount of overlap, which suffice for a detection. For $\lambda = 0.5$ the bias μ_{α_w} depends only on the resolution

$$\mu_{\alpha_w} = -\Delta\alpha \tag{7.21}$$

This bias is not negligible especially for large distances.

Experimental Setup

The experimental setup is the same as shown in Figure 7.4. The width of the rear of the cars w is measured beforehand. As the longitudinal axis of the host vehicle is aligned with the object vehicle, the width is given by

$$w = 2r \tan\left(\frac{\alpha_w}{2}\right) \tag{7.22}$$

where r is the distance to the middle of the rear of the vehicle. The reference angle α_w is therefore given by

$$\alpha_w = 2 \arctan\left(\frac{w}{2r}\right) \tag{7.23}$$

[1] In collision mitigation or avoidance applications the distance to the car changes as well, which further supports this assumption.

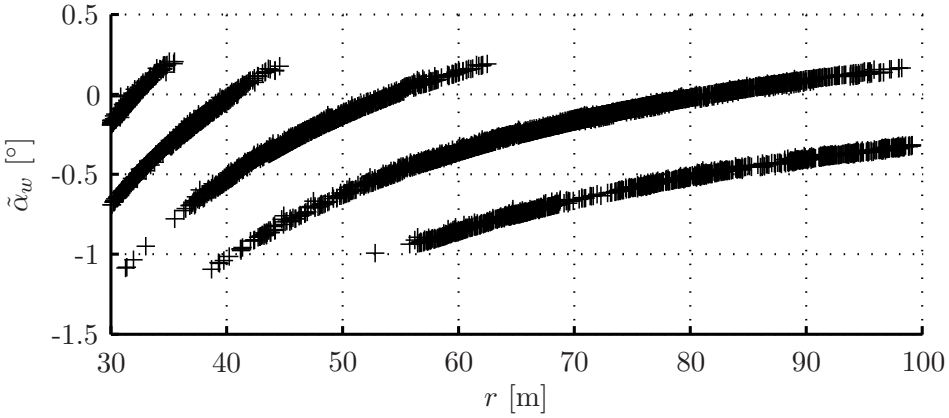


Figure 7.10: Errors in the angular width $\tilde{\alpha}_w$ of the Carina sample, plotted against the distance to the rear of the observed vehicle r. At any distance there are only two to three discrete width measurements possible due to the discrete angular measurement principle of the laser scanner.

The distance r is taken from the laser scanner rear measurement $\hat{\mathbf{p}}^{ls}$. The accuracy of the reference angle α_w is therefore restricted by the accuracy of the distance measurement of the laser scanner. In distances above 30 m the resulting maximal error of the reference angle α_w is 0.02°. The accuracy is increasing for larger distances.

The measured angle $\hat{\alpha}_w$ is given by the angle difference between the two points delimiting the line segment, which corresponds to the rear of the observed vehicle (Figure 7.9). The error of the angle measurement is then calculated from the reference angle α_w and the measured angle $\hat{\alpha}_w$

$$\tilde{\alpha}_w = \hat{\alpha}_w - \alpha_w \tag{7.24}$$

Realization of the Experiment and Results

The same experiments were conducted and the same data samples used as for the evaluation of the azimuth measurement described in Chapter 7.2.1.

For each measurement cycle the angular error $\tilde{\alpha}_w$ is calculated from Equation (7.24). In Figure 7.10 the error of the complete Carina sample is plotted against the distance to the rear of the observed vehicle. It exhibits very clearly the already mentioned correlation of the angular error and the distance. For a set distance and a given vehicle there are always two and for some distances three discrete values for the angular error $\tilde{\alpha}_w$. Interestingly, the angular width is not only underestimated. The overestimation amounts up to $\tilde{\alpha}_w = 0.25°$ for some measurements in the Carina sample set. An overestimation can only occur, if the necessary overlap between the laser pulse cross section and the object λ is less than 50 %. For instance, if the laser cross section in azimuth were $\beta = 0.5°$ and $\lambda = 0.25$, the maximal overestimation would be 0.25°. Even, if it can be stated, that the necessary overlap between the laser pulse cross section and the object λ is less then 50 % at least for reflecting surfaces, the exact combination of λ and β can not be determined with the given experiments. Additionally, it can be assumed, that the parameters vary depending on the distance and the reflectivity of the observed surface.

Therefore, the aim of this experiment is the determination of an approximate average

Table 7.2: Sample mean $\mu_{\tilde{\alpha}_w}$ and the sample standard deviation $s_{\tilde{\alpha}_w}$ of the measurement error of the angular width $\tilde{\alpha}_w$ for the different data sets.

	# Meas.	$\mu_{\tilde{\alpha}_w}$ [°]	$s_{\tilde{\alpha}_w}$ [°]
Theoretical ($\lambda = 0.5$)	–	-0.50	0.20
Theoretical ($\lambda = 0.25, \beta = 0.5°$)	–	-0.25	0.20
Carina	4,528	-0.29	0.25
Passat	9,157	-0.28	0.22
S-Class	7,186	-0.24	0.21
All	20,871	**-0.27**	**0.23**

bias and the variance of the angular width measurement for passenger cars. Table 7.2 concludes the results for the different data samples. The sample standard deviation $s_{\tilde{\alpha}_w}$ of the Carina sample deviates from the theoretical value, which is most likely due to an increased number of missed detections at the corners of the vehicle's rear. The standard deviation of the Passat and S-Class sample are very close to the theoretically derived value. All offsets indicate, that the necessary overlap between the laser pulse cross section and the object λ must be less than 50 %.

As an overall measurement uncertainty for the parameterization of the Kalman filter a standard deviation for the angular width measurement $\sigma_{\alpha_w} = 0.23°$ is chosen, which is the average sample standard deviation given in Table 7.2. The bias is therefore $\mu_{\alpha_w} = -0.27°$.

All width measurements are offset compensated in the laser scanner feature extraction module, before they are forwarded to the fusion unit. For that purpose the observed line segment is extended on both ends by $-\mu_{\alpha_w}/2$. The end points and the resulting width are recalculated (Equation (7.15)). The final variance of the width σ_w^2 is given by Equation (7.17).

7.2.5 Laser Scanner — Orientation

Experimental Setup

Figure 7.11 shows the experimental setup. The host vehicle is located close to a straight section of a highway, where cars drive approximately 80 km/h (22 m/s). Measurements were taken from passing and approaching cars. The true orientation of the observed cars is assumed to be 0° or 180°.

Realization of the Experiment

Overall 3,756 orientation measurements were taken from 57 cars passing the host vehicle and 2,580 measurements from 86 approaching cars[2]. The minimal accepted length of a contour is 0.8 m. As the orientation of a car determined from the contour is ambiguous

[2]The number measurements on approaching cars is smaller, as they were sometimes occluded by the passing vehicles. Additionally the detection range of vehicles seen from the front is lower.

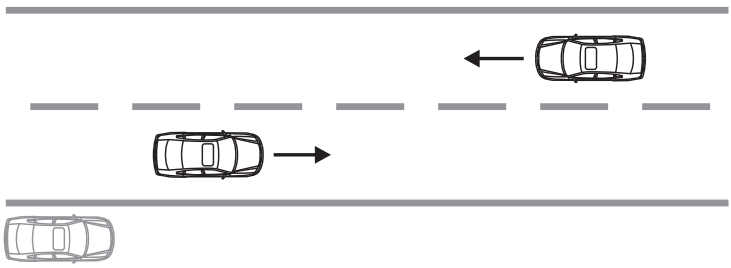


Figure 7.11: Experimental setup for the evaluation of the orientation measurement. The host vehicle is parked beside a straight section of a road and observes passing vehicles.

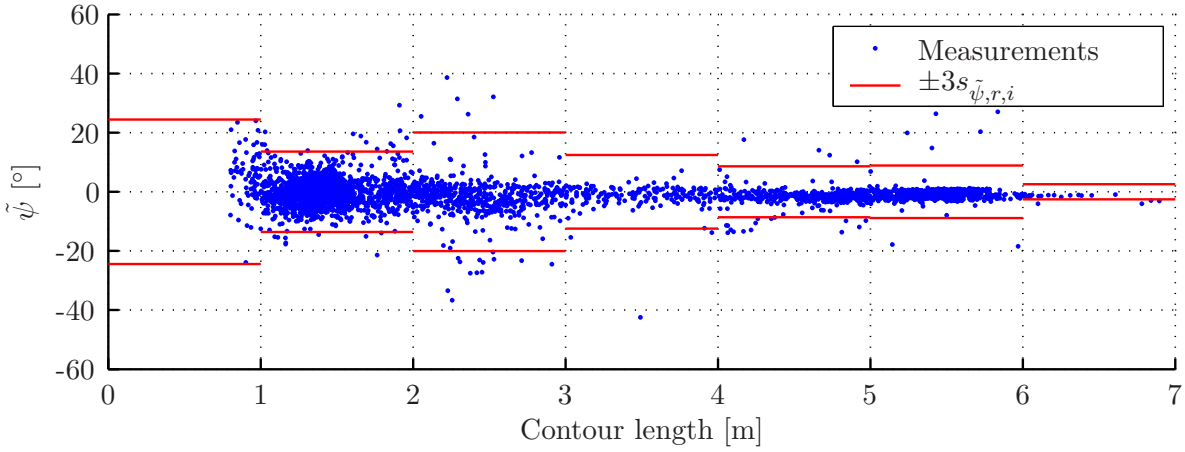


Figure 7.12: Error in the orientation $\tilde{\psi}$ for cars observed from the rear and the side plotted against the length of the measured contour. The $3s_{\tilde{\psi},r,i}$ limits indicate the uncertainty of the different data subsets.

without the velocity vector, the measured orientation is always corrected by $n\frac{\pi}{2}$ to be in the range of $[-\frac{\pi}{4}, \frac{\pi}{4}]$. As the true orientation is zero, the orientation error $\tilde{\psi}$ is identical with the measured orientation ψ

$$\tilde{\psi} = \psi \tag{7.25}$$

Experimental Results

The precision of the orientation given by the contour models, which are fitted to the laser scanner raw data as described in Chapter 5.4.1, depends on the observed shape of the object. The front of a car is usually very rounded at the corners and does not fit the contour models as well as the rear of cars. Therefore, the evaluation is performed separately for passing and approaching cars.

Figure 7.12 and 7.13 show the error in the orientation measurements for passing and approaching cars plotted against the contour length, which is the sum of the length of the individual line segments. The observed length of the cars is about 4–5 m and the width around 1.5 m. The complete contour length of a side and the front or rear is therefore 5.5–

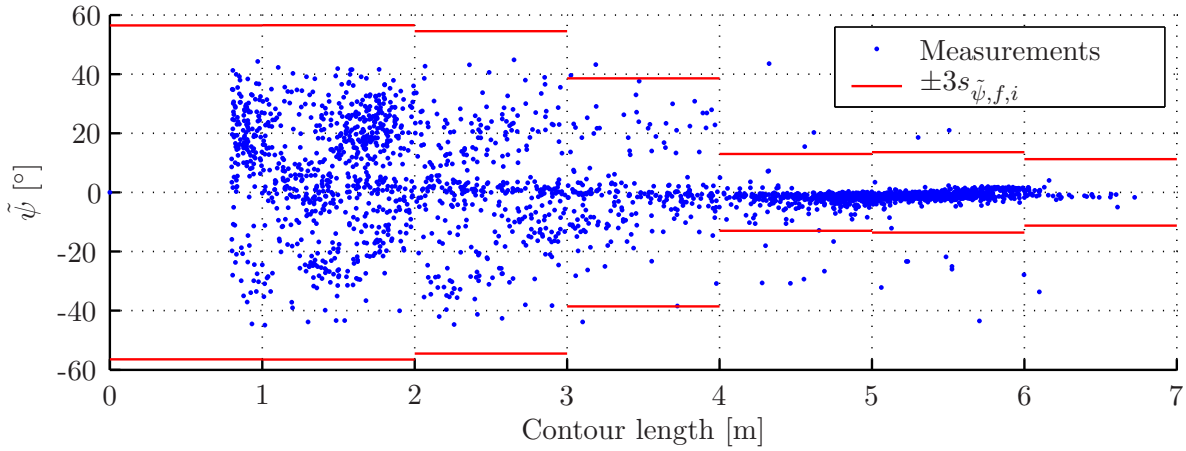


Figure 7.13: Error in the orientation $\tilde{\psi}$ for cars observed from the front and the side plotted against the length of the measured contour. The $3s_{\tilde{\psi},r,i}$ limits indicate the uncertainty of the different data subsets.

6.5 m. For cars observed from the front the orientation measurements are very inaccurate, as the round form of the front is often wrongly approximated by an L-shape and not an I-shape. If the side is observed additionally, the orientation becomes significantly more accurate. The measured orientation of passing cars is generally of higher accuracy.

Table 7.3 concludes the evaluation. The sample standard deviation for cars observed from the front $s_{\tilde{\psi},f,i}$ and cars observed from the rear $s_{\tilde{\psi},r,i}$ is given depending on the contour length. These sample standard deviations are used as measurement uncertainties in the Kalman filter. By default the higher uncertainties of an approaching car are used. If the unambiguous orientation, known from the orientation of the velocity vector or the classification of the rear exploiting its reflectors, indicates the observation of a rear, the

Table 7.3: Sample standard deviation of the orientation measurement error $s_{\tilde{\psi}}$ subject to the contour length and separated into cars observed from the front and from the rear.

		Rear and side		Front and side	
i	Contour length [m]	# Meas.	$s_{\tilde{\psi},r,i}$ [°]	# Meas.	$s_{\tilde{\psi},f,i}$ [°]
1	0–1	78	8.1	157	18.8
2	1–2	1,808	4.5	647	18.9
3	2–3	508	6.7	354	18.2
4	3–4	180	4.2	223	12.9
5	4–5	419	2.9	439	4.3
6	> 5	763	2.8	760	4.4
all		3,756		2,580	

Table 7.4: Sample standard deviation of the orientation measurement error $s_{\tilde{\psi}}$ depending on the utilized method. One method uses all measurements and the other method only those of reflectors.

# Meas.	$s_{\tilde{\psi},all}$ [°]	$s_{\tilde{\psi},reflector}$ [°]
539	5.8	3.8

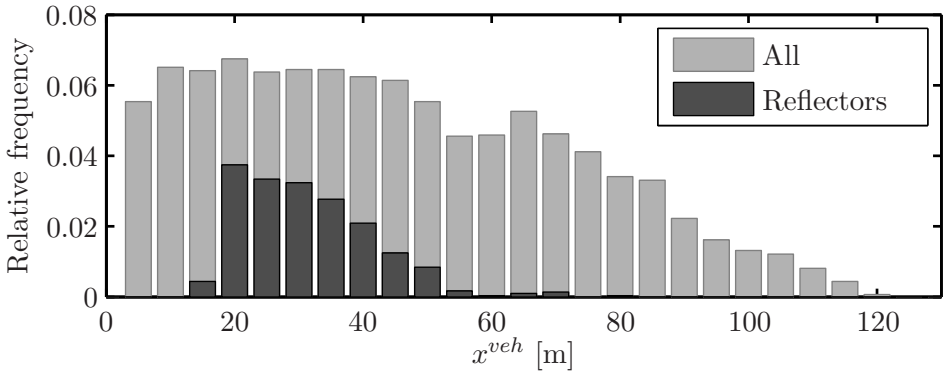


Figure 7.14: Relative frequency of orientation measurements depending on the utilized method plotted against the distance to the observed vehicle x^{veh}. One method uses all measurements and the other method only those of reflectors.

significantly smaller values $s_{\tilde{\psi},r,i}$ are used.

The sample standard deviation $s_{\tilde{\psi},r,3}$ for vehicles observed from the rear and a contour length of 2–3 m is remarkably high with respect to contour lengths of 1–2 m and 3–4 m. For a contour length of 1–2 m mainly the rear is observed only. A line is fitted into the measurements with good accuracy. In cased in which the contour length is 3–4 m the rear and a large part of the side are observed. In these cases a L-shape is fitted into the measurements with good accuracy. If, however, only a fraction of the side is observed in addition to the rear, the algorithms has difficulties to decide between an I- and a L-shape. An I-shape fit leads, however, to large errors.

The second method for the estimation of the orientation addresses this deficiency. It uses only laser scanner detections from reflectors (Chapter 5.4.2). This approach applies only for the rear of vehicles, which exhibit a reflecting number plate and the rear reflectors on either side. The raw measurements from a frontal number plate are usually not sufficient for a linear regression. Table 7.4 compares the sample standard deviation $s_{\tilde{\psi}}$ of the two methods. The first uses all raw measurements and the other employs only the measurements on reflectors. For comparison purposes only those orientation measurements are considered, which are provided by both methods. The error of the orientation can be significantly improved, if the orientation estimate is only based on the reflector measurements.

However, the reflectors are often missed by the laser pulse and can only be detected up to a certain distance. Figure 7.14 shows the relative frequency of orientation measurements depending on the distance to the observed vehicle for both methods. The frequency

is normalized to the sum of the orientation measurements of the first method, which uses all raw measurements. It can be clearly observed, that the method employing the reflectors can not be used alone. Instead, the orientation from all raw measurements is applied by default and replaced by the orientation from reflectors if it is available. The sample standard deviations $s_{\tilde{\psi},reflector}$ from Table 7.4 then represents the measurement uncertainty for the Kalman filter.

7.2.6 Laser Scanner — Displacement

The displacement calculation from point clouds in subsequent scans, explained in Chapter 5.4.3, facilitates the direct estimation of the motion. It is particularly useful in situations, in which the position of a vehicle is not observable. This applies mainly, when the host vehicle overtakes a truck and the length of the truck inhibits temporarily the observation of any of the trucks corners.

Experimental Setup

For the estimation of the displacement measurement uncertainty, stationary trucks are passed by the host vehicle. The true displacement results from the known ego-motion. In practice the predicted displacement from the Kalman filter is used as true displacement, using a stationary motion model.

Realization of the Experiment

Overall 22 measurement sequences were recorded from different trucks with 1,903 cycles. The gradient of the residual distance distribution is calculated in x and y direction (Chapter 5.4.3). For high gradients, there is a distinct minimum and the displacement is unambiguous. Low gradients indicate a high degree of uncertainty about the estimated displacement.

Experimental Results

For the estimation of the measurement uncertainty it is assumed that the variance of the displacement error is identical in x and y direction and depends on the gradients g_x and g_y. The measurement errors of both displacement components are therefore combined. Table 7.5 summarizes the resulting sample standard deviations of the displacement error with respect to the gradient.

7.2.7 Camera — Template Tracking

The image template tracking described in Chapter 6.3 acquires a template from the image content at an image position given by the laser scanner based object tracking. This template is then searched for in subsequent images and its position determined. The resulting displacements are subject to stochastic errors.

Table 7.5: Sample standard deviation of the displacement measurement error $s_{\Delta x}$ and $s_{\Delta y}$ depending on the gradients of the residual distance distribution g_x and g_y.

g_x, g_y	# Meas.	$s_{\Delta x}, s_{\Delta y}$ [m]
0.14–0.4	1,411	0.039
0.4–0.7	1,451	0.030
0.7–1.0	771	0.026

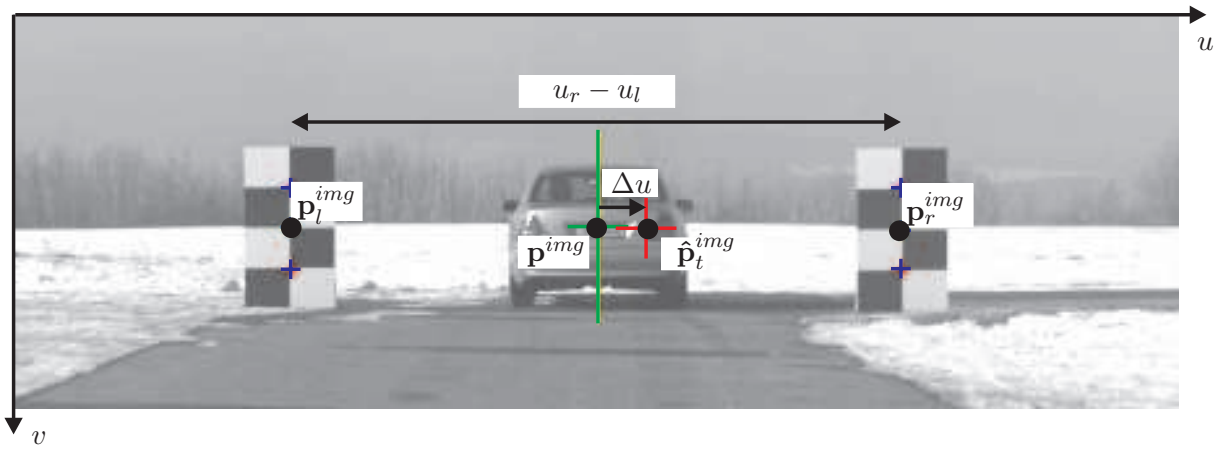


Figure 7.15: Experimental setup for the calculation of the displacement errors of the template tracking. The points $\mathbf{p}_l^{img} = [u_l, v_l]'$ and $\mathbf{p}_r^{img} = [u_r, v_r]'$ are determined using image processing techniques. The reference position in the image is denoted by $\mathbf{p}^{img}$, $\hat{\mathbf{p}}_t^{img}$ is the template position and Δu the offset between the reference and template position.

Experimental Setup

For the characterization of the displacement error, the same experiment was conducted as for the evaluation of the azimuth measurement error described in Chapter 7.2.1. Figure 7.15 shows the setup.

At the template initialization, the offset in the u-coordinate Δu between the reference image position $\mathbf{p}^{img} = \begin{bmatrix} u & v \end{bmatrix}'$ and template position $\mathbf{p}_t^{img} = \hat{\mathbf{p}}_t^{img}(k_0) = \begin{bmatrix} u_t & v_t \end{bmatrix}'$ is calculated

$$\Delta u(k_0) = u_t(k_0) - u(k_0) \tag{7.26}$$

where k_0 denotes the cycle of the template initialization. This initial offset $\Delta u(k_0)$ of the template tracking is estimated by the Kalman filter via the fusion of the laser scanner data and the template tracking.

The offset $\Delta u(k_0)$ is however not normalized to the distance to the object. This can be implicitly achieved by a normalization to the distance between the two reference objects in the image

$$\Delta u_n(k_0) = \frac{\Delta u(k_0)}{u_r(k_0) - u_l(k_0)} = \frac{u_t(k_0) - u(k_0)}{u_r(k_0) - u_l(k_0)} \tag{7.27}$$

where u_l and u_r are the u-coordinates of the reference object positions $\mathbf{p}_l^{img}$ and $\mathbf{p}_r^{img}$ and k_0 the cycle of the template initialization.

For a perfect template tracking the normalized offset $\Delta u_n(k)$ should be constant for all subsequent cycles k. The expected offset compensated template position $u_t(k)$ at cycle k can therefore be calculated from the normalized offset with

$$\begin{aligned} \Delta u_n(k) &= \Delta u_n(k_0) \\ \frac{u_t(k) - u(k)}{u_r(k) - u_l(k)} &= \Delta u_n(k_0) \\ u_t(k) &= u(k) + \Delta u_n(k_0)\,[u_r(k) - u_l(k)] \end{aligned} \tag{7.28}$$

where $u(k)$ denotes the reference position of the vehicle at cycle k, u_l and u_r are the u-coordinates of the reference object positions $\mathbf{p}_l^{img}$ and $\mathbf{p}_r^{img}$ and $\Delta u_n(k_0)$ is the normalized offset of the initial template.

As the template tracking is not optimal, the u-coordinate $\hat{u}_t(k)$ of the estimated template position $\hat{\mathbf{p}}_t^{img} = \begin{bmatrix}\hat{u}_t & \hat{v}_t\end{bmatrix}'$ is affected by an error $\tilde{u}_t(k)$

$$\tilde{u}_t(k) = \hat{u}_t(k) - u_t(k) \tag{7.29}$$

where the true template position $u_t(k)$ is given by Equation (7.28). Inserting Equations (7.27) and (7.28) into (7.2.7) yields

$$\tilde{u}_t(k) = \hat{u}_t(k) - u(k) - \Delta u_n(k_0)\,[u_r(k) - u_l(k)] \tag{7.30}$$

$$= \hat{u}_t(k) - u(k) - [u_t(k_0) - u(k_0)]\,\frac{u_r(k) - u_l(k)}{u_r(k_0) - u_l(k_0)} \tag{7.31}$$

The accuracy of the true template position depends on the error in the position estimate of the reference objects. Even if the accuracy is very high ($\sigma_u \approx 0.05$ pixel) it is not negligible. However, it can be assumed that the errors in the reference positions are uncorrelated with the template tracking estimates. The sample standard deviation of the measurement error $\tilde{u}_t$ provides thus an upper limit for the true uncertainty.

Realization of the Experiment

The same experiments were conducted as for the evaluation of the azimuth measurement described in Chapter 7.2.1. However, the template tracking was applied, even if there was only one raw measurement from the laser scanner. This explains the higher number of utilized measurements in Table 7.6 compared to Table 7.1.

Experimental Results

The sample means $\mu_{\tilde{u}_t}$ and sample standard deviations of the measurement error $\tilde{u}_t$ for the different data sets are concluded in Table 7.6. As an overall standard deviation of the measurement uncertainty for the Kalman filter $\sigma_{u_t} = 0.22$ pixel is used. The standard deviation σ_{v_t} for the v coordinate is assumed to be identical.

Table 7.6: Sample means $\mu_{\tilde{u}_t}$ and sample standard deviations of the measurement error $\tilde{u}_t$ of the template position.

	# Meas.	$\mu_{\tilde{u}_t}$ [pixel]	$s_{\tilde{u}_t}$ [pixel]
Carina	4,659	0.04	0.24
Passat	9,325	-0.00	0.22
S-Class	7,234	-0.02	0.22
All	21,218	-0.01	**0.22**

7.3 Discussion

The presented experiments suggest a methodology for a quantitative evaluation of the laser scanner's measurement characteristics, the laser scanner feature extraction performance and the image processing accuracy. The results are supported by theoretical derivations of the measurement uncertainties. The empirically determined measurement uncertainties provide a good basis for a parameterization of the Kalman filter.

The evaluation of the novel orientation estimation, which exploits the reflecting surfaces on vehicle rears, proved to be significantly superior compared to the contour matching method.

However, for a statistically sound conclusion, the suggested experiments ought to be conducted with several more vehicles and in a larger variety of situations. Additionally, the influence of different weather conditions on the measurement quality, such as rain, snow and fog, would be of interest. The effort of such extensive experiments was however beyond the scope of this work. Nevertheless the derived methodology provides the necessary basis for the suggested experiments.

8 Evaluation

In the first part of this chapter the parameters and the quality measures used throughout the evaluation are described. The second part comprises the evaluation of the environment description for passenger cars. Three different tracking algorithms for passenger cars are contrasted with respect to different traffic scenarios. In the third part the evaluation of the laser scanner based tracking of trucks is presented.

8.1 Motion Model Parameters

For the Kalman filter based fusion of measurements and the tracking of multiple objects, the process model and process noise covariance must be chosen. The process model is represented in the linear case by the state transition matrix $\mathbf{F}$. The measurement functions are detailed in the chapters describing the sensor modules. In the previous chapter a detailed theoretical and experimental derivation of the measurement noise covariance is presented.

For the process a relatively simple white noise acceleration model is utilized as a second order kinematic motion model, with the state transition matrix and the process noise covariance matrix described in Chapter 4.3.4. However, for performance reasons, the z-coordinate and the height d_z are decoupled from the remaining state description

$$\mathbf{x} = \begin{bmatrix} x & y & \psi & v_x & v_y & d_x & d_y \end{bmatrix}' \tag{8.1}$$

According to Equation (4.9) the noise gain $\mathbf{\Gamma}$ and the process noise covariances $\mathbf{Q}_d$ are given by

$$\mathbf{\Gamma} = \begin{bmatrix} \frac{1}{2}T^2 & 0 & 0 \\ 0 & \frac{1}{2}T^2 & 0 \\ 0 & 0 & T \\ T & 0 & 0 \\ 0 & T & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{Q}_d = \begin{bmatrix} \sigma_{a_x}^2 & 0 & 0 \\ 0 & \sigma_{a_y}^2 & 0 \\ 0 & 0 & \sigma_{\psi}^2 \end{bmatrix} \tag{8.2}$$

The standard deviation of the process noise modeling the acceleration σ_{a_x} and σ_{a_y} are chosen as

$$\sigma_{a_x} = 4 \ \frac{\mathrm{m}}{\mathrm{s}^2} \qquad \sigma_{a_y} = 4 \ \frac{\mathrm{m}}{\mathrm{s}^2} \tag{8.3}$$

Maximal decelerations of passenger cars of 11 $\mathrm{m/s^2}$ are therefore between the 2- and 3-sigma values of the process noise assuming that the maximal deceleration can be obtained

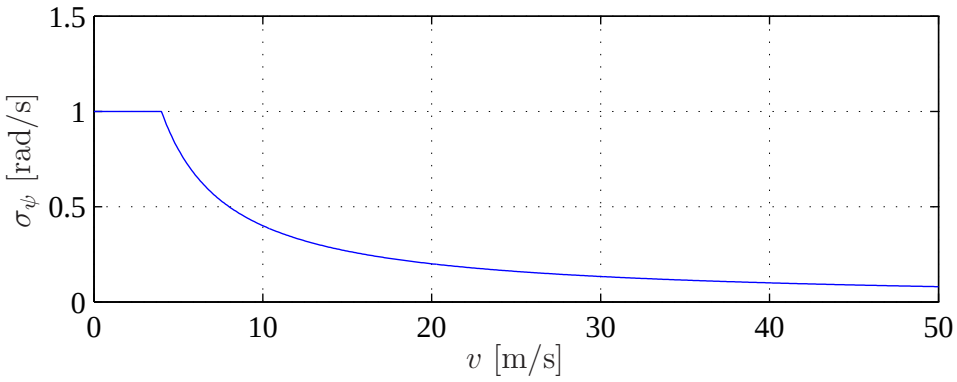


Figure 8.1: The standard deviation of the process noise for the yaw angle σ_ψ depends on the current velocity estimate v.

within one measurement interval. The process noise of the yaw angle is chosen depending on the velocity

$$\sigma_\psi = \begin{cases} 1.0 \ \frac{\text{rad}}{\text{s}} & \text{if} \quad v < 4 \ \frac{\text{m}}{\text{s}} \\ 4 \ \frac{\text{rad}\cdot\text{m}}{\text{s}^2} v^{-1} & \text{if} \quad v \geq 4 \ \frac{\text{m}}{\text{s}} \end{cases} \tag{8.4}$$

The yaw rate ω is in general limited by the cross acceleration $a_c = |\omega v|$. As the process noise of the yaw angle models the unknown yaw rate and the maximal expected cross acceleration is $a_c = 4 \ \frac{\text{m}}{\text{s}^2}$, the standard deviation of the process noise of the yaw angle is chosen as $\sigma_\psi = 4 \ \frac{\text{m}}{\text{s}^2} v^{-1}$. For low speed situations below $4 \ \frac{\text{m}}{\text{s}}$ a different process noise is chosen, as the maximal yaw rate is limited by the velocity and the turn radius of the car. This would lead to a standard deviation which converges to zero for a speed close to zero. In practice a constant low speed process noise, independent of the velocity performs better, as it compensates for a temporarily biased measurement noise. The process noise of the yaw angle is shown in Figure 8.1 as a function of the velocity.

The width and length are constant parameters. The standard deviation of the process noise is therefore zero

$$\sigma_{d_x} = 0 \ \frac{\text{m}}{\text{s}} \qquad \sigma_{d_y} = 0 \ \frac{\text{m}}{\text{s}} \tag{8.5}$$

In practise it is advisable to use very small values as pseudo-noise in order to account for the errors introduced by the linearizations of the non-optimal extended Kalman filter approach.

8.2 Center of Gravity Tracking

The standard tracking approaches model the objects as a point, neglecting the dimensions of the observed objects. The first approaches described in the literature for laser scanner based vehicle tracking for driver assistance systems therefore calculate the mean position of all laser scanner measurements associated with the object [93]. This center of gravity (COG) represents the measured position of the (point-)object. The newly developed tracking methods described in this thesis are evaluated with respect to this standard

approach. The state vector of the COG tracking consists of the position and the velocity

$$\mathbf{x} = \begin{bmatrix} x & y & v_x & v_y \end{bmatrix}' \tag{8.6}$$

The laser scanner measurement vector contains the position of the center of gravity

$$\mathbf{z}_{ls} = \begin{bmatrix} x & y \end{bmatrix}' \tag{8.7}$$

The uncertainty in the distance of the center of gravity is modeled by a normally distributed error with a standard deviation of $\sigma_r = 0.15$ m (Chapter 7.2.2). The measurement uncertainty of the azimuth depends mainly on the accuracy of the measurements at the corners of the observed vehicle. As shown in Chapter 7.2.1 the standard deviation of the mean of two corner measurement, representing the position of a vehicle's rear is $\sigma_\alpha = 0.12°$. The resulting covariance matrix in polar coordinates is transformed into cartesian coordinates as shown in Chapter 5.4.2.

In the close range below 30 meters the measurement error of the center of gravity is however larger, as the position of the center of gravity changes with a varying perspective on the observed vehicle. For a vehicle crossing at an intersection, the center of gravity is first situated at the front and moves over the side to the rear. As these errors depend on many factors, such as the vehicle form, the speed and the distance, it is hard to derive the uncertainty empirically. Therefore a heuristically determined standard deviation of

$$\sigma_x = 0.3 \text{ m}, \qquad \sigma_y = 0.3 \text{ m} \tag{8.8}$$

is chosen.

8.3 Performance Measures for Environment Descriptions

For an evaluation and comparison of different vehicle environment perception systems, the definition of performance measures for the resulting vehicle environment descriptions is indispensable. Such measures enable a quantitative and not only qualitative description of the performance on a statistical basis. A standardization enables the comparison of systems across different research teams. Furthermore performance measures facilitate an optimization of the system under development. It is then possible to conduct an automated, global and objective optimization. The prerequisite is however a statistically unbiased set of test data from representative scenarios, which can be hard to accomplish.

The definition of the performance measures is driven by the requirements of the driver assistance applications. However, the requirements vary between different applications. As there is usually a trade-off between the performance measures, it is desirable to characterize the performance of the system by a curve or a set of curves in an application independent fashion. Each application then chooses the operation point in this diagram which best satisfies its needs.

Throughout this work, it is assumed that the sensor data fusion can be decomposed into a part, which fuses the data concerning the state estimate and another part which fuses the classification information. Both parts however interact in general and are not independent.

Table 8.1: Performance measures for vehicle environment perception systems.

Performance measure	Module	Detailed description
Detection rate	classification	[167]
False alarm rate	classification	[167]
Estimation error	state estimation	Chapter 8.3.1
Consistency	state estimation	Chapter 8.3.2
Detection latency	classification, state estimation	Chapter 4.5.1

A common application independent representation of the performance of the classification is the receiver operating characteristic curve. In this diagram the detection rate is plotted over the false alarm rate for different parameter values of the classifier. It depends on the application which particular combination of false alarm and detection rate produces the best results.

The measurement noise reduction performance of the state estimation is well described by the mean state error in a non-maneuver case [4]. For road participants a non-maneuver means a motion on a straight line with zero acceleration. This mean error is then compared to the peak error during a maneuver. In general there is a trade-off between these performance measures. Comfort systems, such as an Adaptive Cruise Control (ACC) prefer small mean errors and accept high peak errors in highly dynamic scenarios. Safety relevant systems may not tolerate high peak errors and need the errors to be limited. For pre-crash or emergency braking systems (EB) the consistency of the state estimate should be of high interest. If the error of the estimated state is often higher than the estimated uncertainty, the probability raises that these applications falsely disengage.

In order to reduce the false alarm rate, a fusion system often delays the object output to the application after the first detection by the sensor [13]. In this initial object verification period, information about the object is accumulated so as to differentiate true object detections from clutter. This object detection latency is not as relevant for slow reacting comfort applications. It however limits the set of driving situations, a future driver assistance system can handle. Some safety critical applications might only be educible with a short object detection latency.

Table 8.1 concludes the performance measures and provides references to the chapters, in which a detailed description of the definitions are given. Further definitions can be found in [160].

An evaluation of the environment description is frequently neglected and the performance of a subsequent application quantified instead. This indirect approach however often reflects the performance of the underlying vehicle environment perception only very superficially. An evaluation of the applications is indispensable, but it can not replace the evaluation of the vehicle environment description.

8.3.1 Error Measures

The **state estimation error** $\tilde{\mathbf{x}}$ is defined by the true state $\mathbf{x}$ and the current state estimate $\hat{\mathbf{x}}$

$$\tilde{\mathbf{x}}(k) \equiv \hat{\mathbf{x}}(k) - \mathbf{x}(k) \tag{8.9}$$

If a whole sequence of state estimates is evaluated, the **sample mean error** $\mu_{\tilde{\mathbf{x}}}$, the **sample standard deviation** of the error $s_{\tilde{\mathbf{x}}}$ and the **root mean square** RMS error can be calculated

$$\mu_{\tilde{\mathbf{x}}} = \frac{1}{K}\sum_{k=1}^{K}\tilde{\mathbf{x}}(k) \tag{8.10}$$

$$s_{\tilde{\mathbf{x}}} = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}\left(\tilde{\mathbf{x}}(k) - \mu_{\tilde{\mathbf{x}}}\right)^2} \tag{8.11}$$

$$RMS = \sqrt{\mu_{\tilde{\mathbf{x}}}^2 + s_{\tilde{\mathbf{x}}}^2} \tag{8.12}$$

8.3.2 Filter Consistency

The consistency criteria of a filter are [4]:

- The state errors should be acceptable as zero mean and have magnitude commensurate with the state covariance as yielded by the filter.
- The innovations should have the same property.
- The innovations should be acceptable as white.

The first criterion is tested with the **normalized (state) estimation error squared (NEES)**

$$\epsilon(k) = \tilde{\mathbf{x}}'(k)\mathbf{P}(k)^{-1}\tilde{\mathbf{x}}(k) \tag{8.13}$$

Under the hypothesis H_0, that the filter is consistent, ϵ is chi-square distributed with n_x degrees of freedom, where n_x is the dimension of the state vector. In **Monte Carlo simulations** the **average NEES** is calculated

$$\bar{\epsilon}(k) = \frac{1}{N}\sum_{i=1}^{N}\epsilon_i(k) \tag{8.14}$$

Then $N\bar{\epsilon}$ should be chi-square distributed with Nn_x degrees of freedom. For consistency $\bar{\epsilon}$ should lie in an acceptance interval

$$\bar{\epsilon}(k) \in [r_1, r_2] \tag{8.15}$$

where the acceptance interval is determined such that

$$P\left\{\bar{\epsilon}(k) \in [r_1, r_2] | H_0\right\} = 1 - \alpha \tag{8.16}$$

where the significance level is often chosen as $\alpha = 0.05$. If the NEES is larger than the upper limit r_2, then the estimation errors are on average too high, compared to the estimated uncertainty given by $\mathbf{P}$. For automotive safety application, which take the estimated state uncertainty $\mathbf{P}$ into account for their collision risk estimation, an underestimated state uncertainty could lead to false reactions. There are several reasons for a state inconsistency. The process noise or the measurement noise covariance could be too small with respect to the true noise or there is a bias in the state estimates. In the opposite case the NEES is smaller than the lower limit r_1. This overestimation of the state uncertainty is more vital for comfort applications, as it exhibits an inefficient filter. Smaller process noise or measurement noise covariances should be chosen.

In real-world applications there is usually only a single run ($N = 1$) and the acceptance interval becomes relatively large. Then the average NEES could be calculated over the sequence. This is not completely correct, as the NEES is locally correlated and the individual samples for the calculation of the average NEES should be uncorrelated. However, for large sequences the small local correlation becomes insignificant. Alternatively, the cumulative probability Q_{χ^2} of the chi-square density function for ϵ with n_x degrees of freedom can be calculated

$$Q_{\chi^2}(\epsilon, n_x) = P_{\chi^2}\{y < \epsilon\}, \qquad y \sim \chi^2_{n_x} \tag{8.17}$$

Cumulative probability values above $1 - \alpha$ should only occur in α percent of the time. Exploiting this approach, a chi-square probability plot can be formed, which shows for each NEES value the corresponding normalized frequency over the whole sequence, which should be optimally the same as the cumulative probability.

The second criterion is tested using the **normalized innovation squared (NIS)**

$$\epsilon_\nu(k) = \boldsymbol{\nu}'(k)\mathbf{S}(k)^{-1}\boldsymbol{\nu}(k) \tag{8.18}$$

where $\boldsymbol{\nu}$ is the measurement residual from Equation (4.26) and $\mathbf{S}$ is the residual covariance from Equation (4.29). If the innovations are white and with zero mean, then ϵ_ν has a chi-square distribution with m degrees of freedom, where m denotes the dimension of the measurement residual vector $\boldsymbol{\nu}$. The NIS is a process intrinsic quality measure. As this measure is independent of the true state estimates, it can be calculated online. As already mentioned in Chapter 4.3.6 the **time-averaged NIS (ANIS)** provides useful insights of the current quality of the filters estimation process, with

$$\bar{\epsilon}_\nu = \sum_{k=1}^{K} \boldsymbol{\nu}'(k)\mathbf{S}^{-1}(k)\boldsymbol{\nu}(k), \qquad m_\sum = \sum_{k=1}^{K} m(k) \tag{8.19}$$

where m is the time varying dimension of the measurement vector and $m_\sum$ denotes the degrees of freedom of the chi-square distribution of $\bar{\epsilon}_\nu$. Calculating the cumulative probability Q_{χ^2} of the chi-square distribution for $\bar{\epsilon}_\nu$ with $m_\sum$ degrees of freedom

$$Q_{\chi^2}(\bar{\epsilon}_\nu, m_\sum) = P_{\chi^2}\{y < \bar{\epsilon}_\nu\}, \qquad y \sim \chi^2_{m_\sum} \tag{8.20}$$

delivers a value in the range $Q_{\chi^2} \in [0, 1]$, which can then be used as a normalized quality measure of the process intrinsic filter consistency over time. An application can refuse for instance the state estimate, if the ANIS exceeds a certain threshold and thus reduce the number of inconsistent estimates.

8.4 Evaluation of the Environment Description for Passenger Cars

The performance of the object tracking of cars depends strongly on the quality of the feature extraction, the appropriateness of the object model and the accuracy of the measurements. Three different feature extraction and tracking approaches will be contrasted in the following:

- The center of gravity tracking (COG) described in Chapter 8.2.
- The cuboid tracking (CT) approximates a car by a cuboidal object model. Contour shapes are fitted to the laser scanner measurements for a determination of the position, width, length and orientation of the vehicle as described in Chapter 5.4.1 and 5.4.2.
- A feature-level fusion approach combines the cuboid tracking with the video based template tracking (CT+TT) described in Chapter 6.3. The template tracking is supposed to enhance the accuracy of the lateral position and velocity estimates of preceding vehicles.

8.4.1 Experimental Setup

Cuboid Tracking

The cuboid tracking (CT) is solely based on the laser scanner data. The features of the cuboid in the laser scanner measurement space are its corners and its sides. Different contour models, representing a corner, a side or two sides, are fitted into the laser scanner data. The position $\begin{bmatrix} x & y & z \end{bmatrix}'$, the orientation ψ, the length of the sides d_x and d_y and the height d_z are estimated by the signal processing of the laser scanner module, resulting in the measurement vector

$$\mathbf{z}_{ls} = \begin{bmatrix} x & y & z & \psi & d_x & d_y & d_z \end{bmatrix}' \tag{8.21}$$

The measurement vector varies depending on the situation, as the parameters are not always all observable. For instance, the length d_x of a preceding vehicle can not be measured. The coordinates $\begin{bmatrix} x & y & z \end{bmatrix}'$ describe different locations on the cuboid, in function of the observed feature. Therefore, the nonlinear measurement function must be adapted in each cycle and for every object to the observed feature and its parameters. The different measurement functions are described in detail in Chapter 5.3.1.

Fusion of Cuboid Tracking and Image Template Tracking

The cuboid tracking is combined with the video data based template tracking utilizing the feature-level fusion architecture. The laser scanner module and the fusion unit are the same as for the cuboid tracking. The system is extended by the camera module. The measurement vector of the camera module is given by

$$\mathbf{z}_{cam} = \begin{bmatrix} u_t & v_t \end{bmatrix}' \tag{8.22}$$

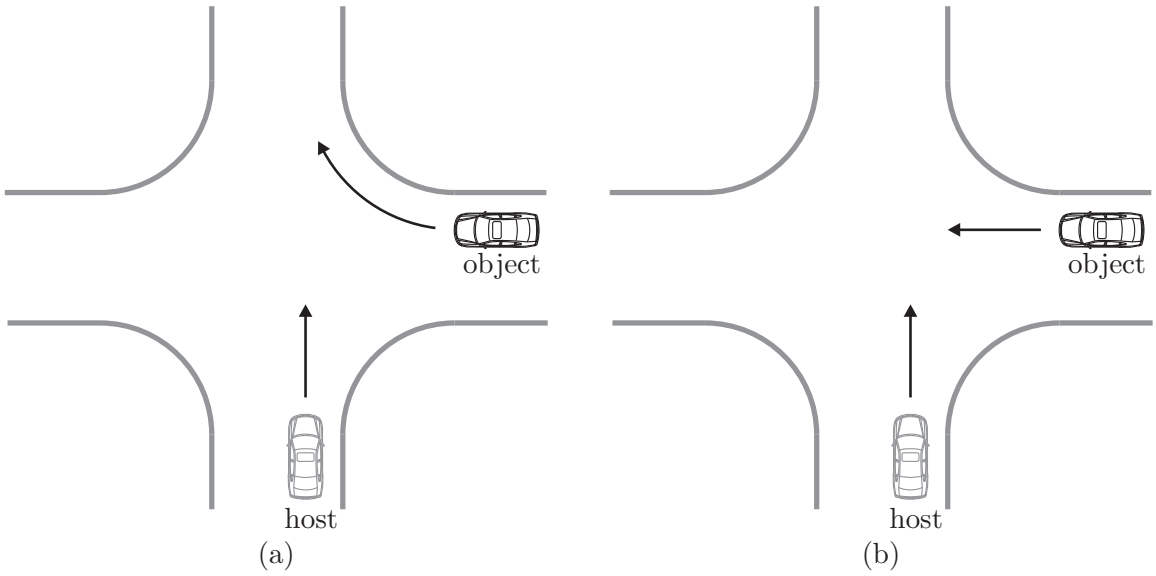


Figure 8.2: Prominent accident scenarios at intersections.

where $\begin{bmatrix} u_t & v_t \end{bmatrix}'$ denotes the template position. The measurement function is described in Chapter 6.3.3.

8.4.2 Position Estimate

The advantage of the cuboid tracking over the center of gravity tracking becomes apparent mainly in the close range (< 30 m), whereas the accuracy of the fusion approach is more significant in the medium and far range (> 30 m). Therefore an intersection scenario demonstrates the characteristics of the center of gravity and the cuboid tracking approaches in the close range. The reference setup already exploited for the determination of the azimuth measurement accuracy in Chapter 7.2.1 is used for a comparison of the fusion approach with the two laser scanner based tracking approaches for vehicles in the far range.

Intersection Scenario

Figure 8.2 shows two of the most prominent accident scenarios at German and French intersections. In 38.4–46.8 % of the intersection accidents in Germany and 55 % in France, depending on the statistics, the opponent car comes either from the left or right and crosses the path with the host vehicle [119]. For the feature extraction and tracking there is no difference if the opponent is coming from the left or right or if the object vehicle drives straight over the intersection or turns into the street where the host vehicle comes from. Therefore, the diverse combinations of vehicle paths are represented in the following by the two scenarios depicted in Figure 8.2.

In order to evaluate the center of gravity (COG) and the cuboid tracking (CT), both intersection scenarios were re-enacted with two test vehicles. The first row of Figure 8.3 shows the camera image for the scenario from Figure 8.2(a) at three time intervals. The

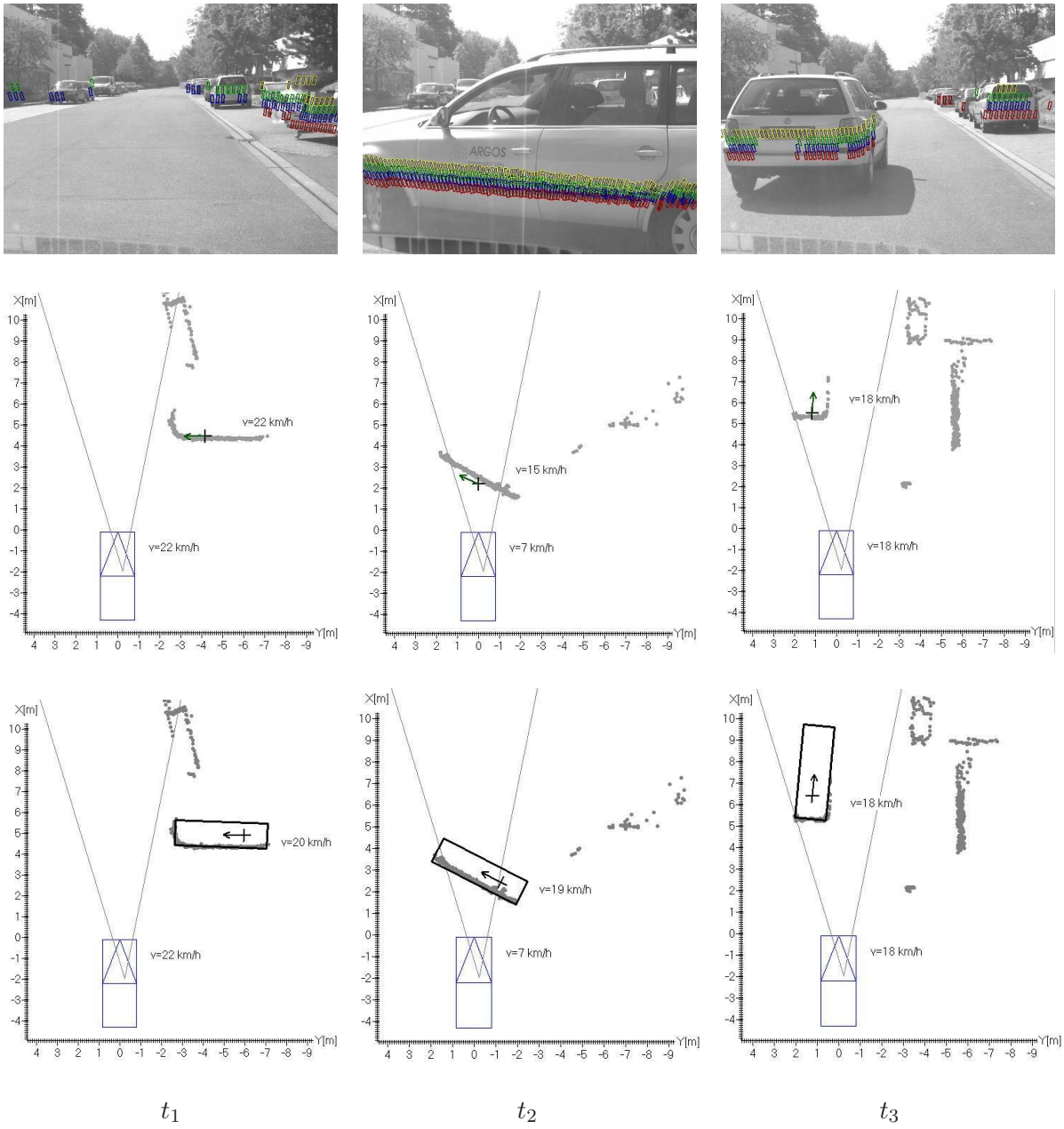


Figure 8.3: Three time intervals of the scenario shown in Figure 8.2(a). At the top an image with the laser scanner data. The second row shows the center of gravity (COG) tracking results and the last row the cuboid tracking (CT) data. The laser scanner data and the field of view of the camera are indicated in grey. It is evident that the CT approach (bottom row) produces significantly better position estimates than the COG tracking (second row).

laser scanner measurements are projected into the image. The bird view in the second row exhibits the results of the COG tracking. The cross depicts the estimated object position. Characteristically for the COG tracking the reference point moves along the object contour. At the beginning it lies close to the frontal left corner moves to the middle of the left side and finally end up at the rear of the observed car. Therefore the position of the vehicle can not be determined precisely in a fixed object coordinate system. The assumption, that the estimated position represents the true center of gravity of the vehicle would lead for instance to high position errors.

The third row in Figure 8.3 presents the estimates from the cuboid tracking (CT). The CT estimates the position, width, length and orientation of the object assuming a cuboidal object shape. In this case the reference point lies in a fixed position to the object contour measured by the laser scanner and delivers therefore a precise localization of the observed object.

Additionally, the CT provides a more detailed description than the COG, as the object dimensions are part of the estimate. Subsequent safety applications can incorporate the object dimensions into their accident risk calculations and reduce the reaction times (Chapter 10).

However, it should be noted that the motion of the reference point along the object contour, observed in the COG tracking, is significantly less prominent in situations, where the object is mainly observed only from one side. It depends on the application if the precision of the COG tracking is sufficient.

Preceding Vehicle

The feature-level fusion approach of the cuboid tracking and the video based template tracking (CT+TT) is applied in scenarios, in which the host vehicle follows another vehicle. It aims at an improved precision of the lateral position and velocity estimates for directly preceding vehicles, cutting-in or cutting-out vehicles.

The accuracy is evaluated with the reference setup detailed in Chapter 7.2.1, which provides ground truth data for the position of the object vehicle with respect to the host vehicle.

Table 8.2 shows the performance of the azimuth estimation error $\tilde{\alpha}$. The sample standard deviation $s_{\tilde{\alpha}}$ of the estimation error for the COG tracking and the CT are very similar, whereas the feature-level fusion with the video template tracking (CT+TT) im-

Table 8.2: Position error statistics for preceding vehicles. Sample mean $\mu_{\tilde{\alpha}}$, sample standard deviation $s_{\tilde{\alpha}}$ and the RMS of the azimuth estimation error $\tilde{\alpha}$. The number of measurements differ as the image based template tracking enables a continuous track, even if there are no laser scanner measurements.

	# Meas.	$\mu_{\tilde{\alpha}}$ [°]	$s_{\tilde{\alpha}}$ [°]	$\mathrm{RMS}_{\tilde{\alpha}}$ [°]
COG	20,656	-0.03	0.12	0.12
CT	20,656	-0.03	0.11	0.12
CT+TT	21,101	-0.04	**0.05**	**0.07**

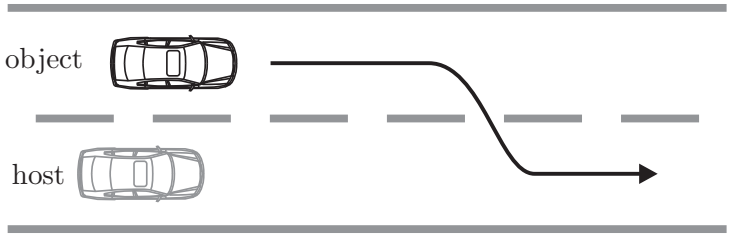


Figure 8.4: Cut-in situation.

proves the accuracy of the azimuth estimate significantly. The accuracy is more than twice as high.

The higher number of measurements for the CT+TT approach results from an improved object detection rate due to the fusion of the two sensors. In some time intervals the video tracking produces a measurement even if the laser scanner does not detect the object.

Cut-in Situation

Present Automatic Cruise Control (ACC) systems recognize cut-in and cut-out situations late and therefore react inadequately, which sometimes forces the driver to interact. The systems have difficulties, as they miss information about the drivers intentions and the lane course. Additionally, the usually deployed radar or lidar sensors have a poor angular resolution compared to a camera. For an immediate reaction these systems would therefore need additional and more accurate sensor information.

The previous section presented the improved accuracy of the CT+TT approach. The aim of this section is a visualization of the improvement in the state estimate for a cut-in situation depicted in Figure 8.4.

Figure 8.5 shows results of the cuboid tracking approach (CT) for three time intervals. The error ellipses in the bird view representing the estimated position uncertainty are enlarged by a factor of 50. The lateral uncertainty is much higher for the object in the distance than for the cutting-in vehicle. Analogously the images contain the estimated object position and uncertainty. At t_2 and t_3 the lateral position is slightly off-center for the cutting-in vehicle and the distant car.

In contrast, Figure 8.6 presents the results of the laser scanner and video fusion approach (CT+TT). The estimated uncertainty in the azimuth is much smaller for the vehicles in the FOV of the camera, even in large distances. The estimated positions of the cars shown in the image domain are accurate throughout the sequence.

The fusion approach of the cuboid tracking and the image template tracking delivers very accurate position estimates. However, it remains to be shown, that the higher accuracy improves an ACC system significantly. This might only be accomplished in combination with an accurate lane detection system.

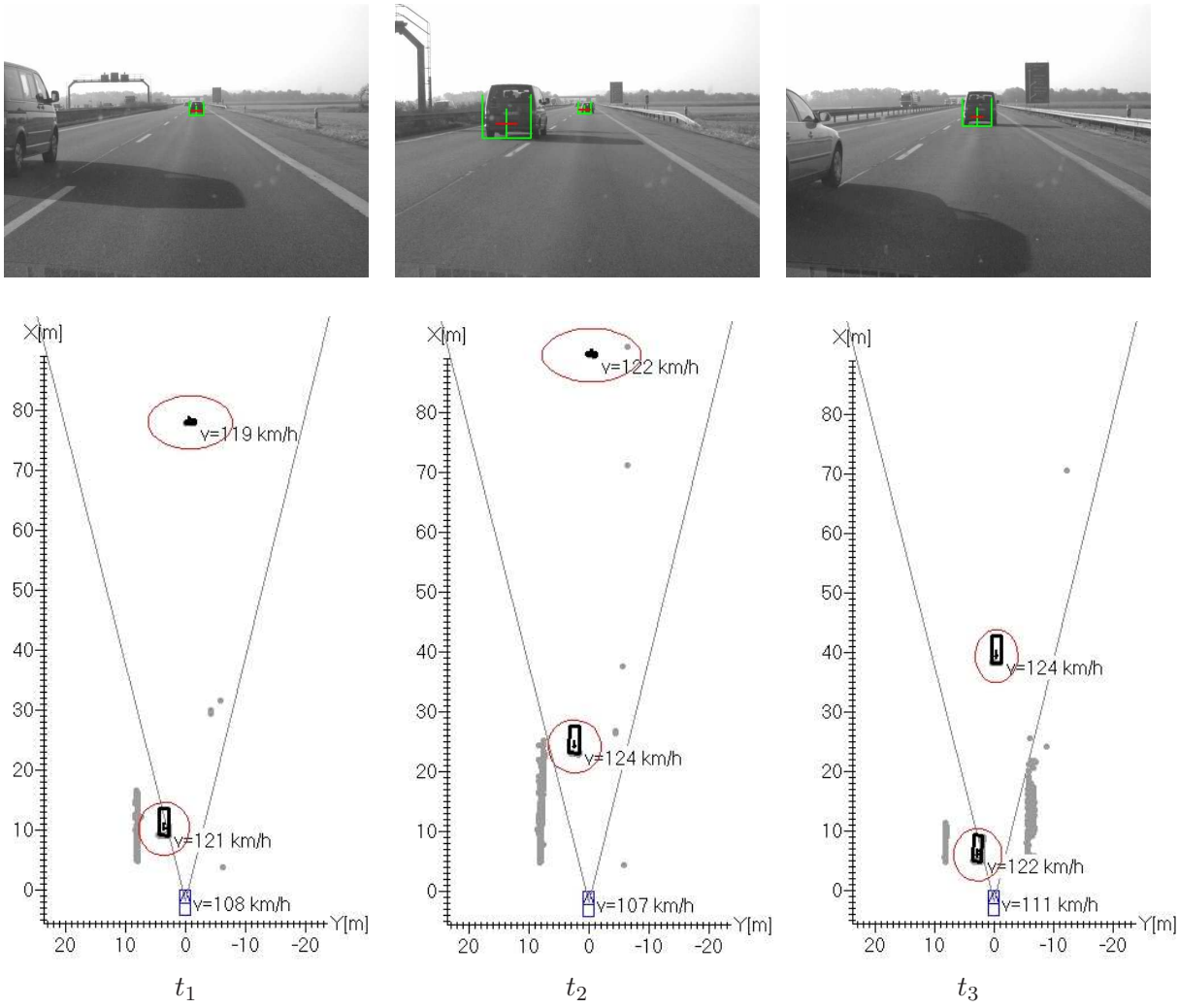


Figure 8.5: Laser scanner approach. Three time instances of a cut-in situation for the cuboid tracking (CT). The top row shows the estimated vehicle positions in green projected into the image. The red bar indicates the 3-sigma uncertainty in the lateral position. At the bottom, the state estimates are shown in the bird's eye view. The laser scanner data and the FOV of the camera are indicated in grey. The 50-sigma ellipses of the estimated position covariances are shown in red.

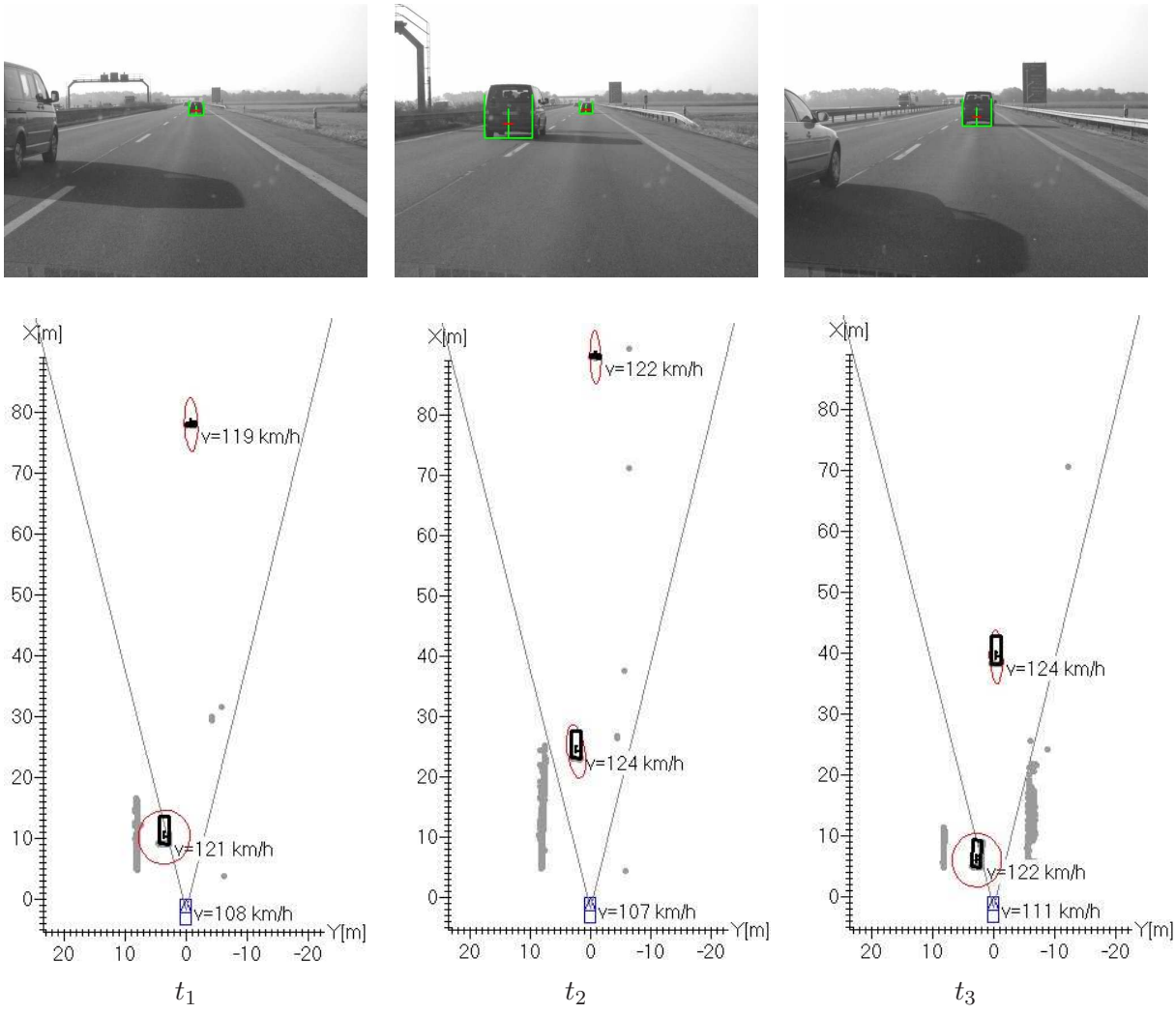


Figure 8.6: Fusion approach. Three time instances of a cut-in situation for the cuboid tracking with template tracking (CT+TT). The top row shows the estimated vehicle positions in green projected into the image. The red bar indicates the 3-sigma uncertainty in the lateral position. At the bottom, the state estimates are shown in the bird's eye view. The laser scanner data and the FOV of the camera are indicated in grey. The 50-sigma ellipses of the estimated position covariances are shown in red.

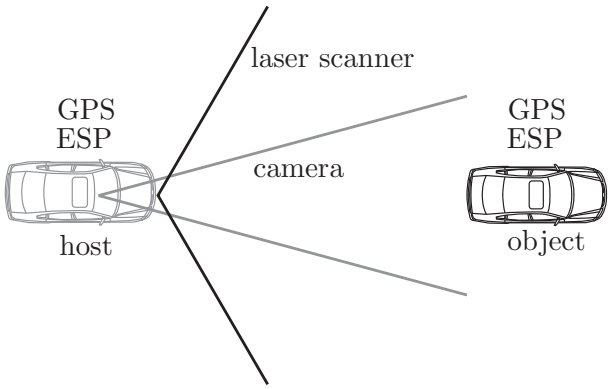


Figure 8.7: Experimental setup for an evaluation of the velocity estimation of the object tracking approaches. The host vehicle is equipped with a laser scanner, camera and an ESP system. The object vehicle records is own ego-motion delivered by the ESP which is used as ground truth data. The GPS of both vehicles serves for a synchronization of the data.

8.4.3 Velocity Estimate

A precise velocity estimate is important for most driver assistance systems. ACC applications determine the necessary host vehicle velocity from the relative velocity in order to keep the distance to the preceding vehicle, and intelligent break assist systems or emergency break applications calculate the time to collision from the estimated velocity in order to determine the activation time.

Experimental Setup

Figure 8.7 depicts the experimental setup. The host vehicle is equipped with the vehicle environment sensors, which are the laser scanner and the camera and additionally with the standard ESP system and a GPS receiver. The object tracks are ego-motion compensated as described in Chapter 4.3.4. The estimated object velocity results therefore in the absolute velocity over ground. At each cycle a global time stamp is calculated from the GPS time and assigned to the tracks. The reference vehicle records its own velocity provided by the ESP system. The velocity from the ESP system of the object car serves then as the reference velocity for the estimated object tracks. The reference velocity and the estimated velocity are synchronized utilizing the GPS time stamps.

The accuracy of the reference velocity is estimated in a separate experiment. In this experiment the object vehicle and the host vehicle were both connected by a rigid tow rope. Different driving maneuvers were conducted, with acceleration and deceleration phases. The comparison of the ESP velocities of both vehicles, which are synchronized with the GPS time stamps, delivers an upper bound for an error statistics of each velocity signal, assuming uncorrelated measurement errors[1]. The statistics reveal a standard deviation of the velocity difference of $s = 0.04$ m/s and a bias of $\mu = 0.01$ m/s, which is sufficient for the following experiments, especially as the evaluation aims mainly at a comparison between different algorithms.

[1]The speed of the two vehicles was measured from the rear wheels only. This reduces the velocity bias in acceleration phases, as the vehicles have both front-wheel drive.

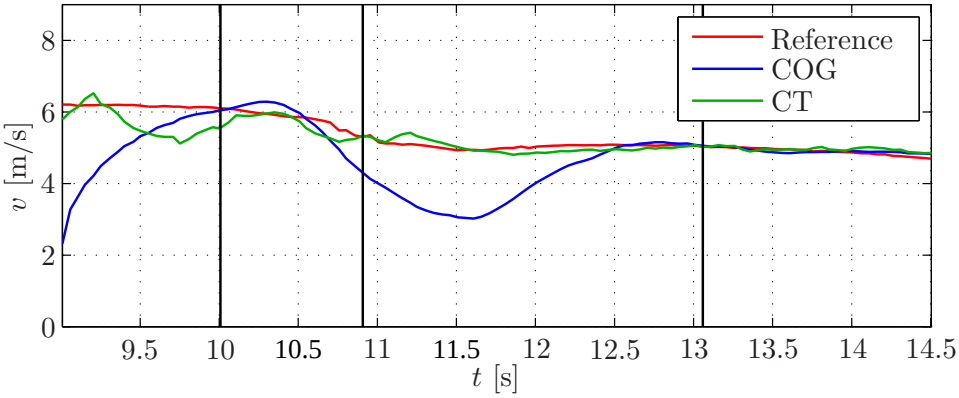


Figure 8.8: Velocity of a vehicle turning into the road in front of the host vehicle. The reference data is shown and the results from the center of gravity (COG) and the cuboid tracking (CT).

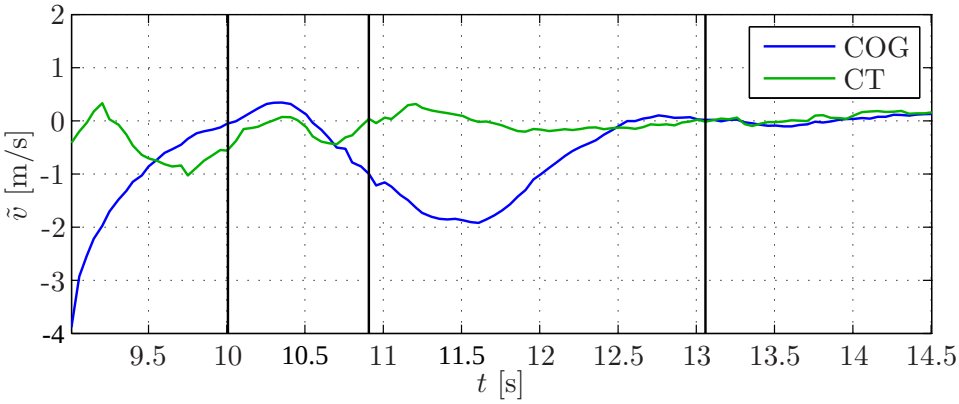


Figure 8.9: Velocity estimation errors for a vehicle turning into the road in front of the host vehicle. The results from the center of gravity (COG) and the cuboid tracking (CT) are shown.

Intersection

As already shown, the position estimate of the standard center of gravity tracking (COG) moves along the object contour if the vehicle is observed from different perspectives (Figure 8.3). Especially at intersection scenarios as depicted in Figure 8.2 this characteristic of the COG tracking becomes apparent. The motion of the object position along the contour induces an additional virtual velocity, which falsifies the estimated object velocity. Figure 8.8 shows the true reference velocity of a vehicle turning into the road in front of the host vehicle (Figure 8.2(a)). The estimated velocities of the COG and the cuboid tracking (CT) approach exhibit significant deviations from the reference velocity. Figure 8.9 presents the velocity estimation error. The time intervals t_1, t_2 and t_3 from Figure 8.3 are indicated in both plots.

The scenario can be roughly separated into three phases: The approach ($t < 10.5$ s), the turning (10.5 s $\leq t <$ 12.5 s) and the following phase (12.5 s $\leq t$). Both tracking approaches exhibit large errors in the initial approach phase. In the case of the COG tracking, the errors are induced by a motion of the position measurements from the

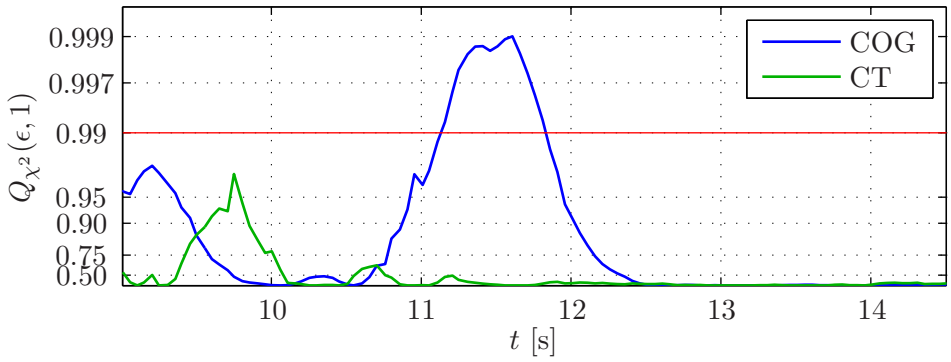


Figure 8.10: NEES of the velocity estimate for a vehicle turning into the road in front of the host vehicle exhibiting the consistency of the estimate. The results from the center of gravity (COG) and the cuboid tracking (CT) are shown.

Table 8.3: Velocity error statistics for the intersection scenario. Sample standard deviation of the velocity estimation error $s_{\tilde{v}}$, the bias $\mu_{\tilde{v}}$ and the RMS with respect to the the center of gravity (COG) and the cuboid tracking (CT) approaches.

		$\mu_{\tilde{v}}$ [m/s]		$s_{\tilde{v}}$ [m/s]		$\text{RMS}_{\tilde{v}}$ [m/s]	
Section	# Meas.	COG	CT	COG	CT	COG	CT
Approach	299	-1.0	**-0.6**	1.2	**0.6**	1.6	**0.9**
Crossing/Turning	398	-1.0	**-0.1**	0.9	**0.4**	1.3	**0.4**
Following	200	0.1	0.1	0.1	0.2	0.2	0.2

object's front over the corner to the left side. This motion results in an decrease of the estimated velocity. The same effect can be observed in the turning phase, when the center of gravity moves from the side to the rear. In the turning phase the CT approach exhibits only small errors.

Figure 8.10 presents the normalized estimation error squared (NEES). The NEES is a measure for the consistency of the velocity estimates with the estimated uncertainty. Large errors which are not reflected by a high estimated uncertainty lead to a high NEES value. As expected the plot exhibits a high degree of inconsistency of the COG tracking in the turning phase. The CT stays below the set 99 % probability region throughout the scenario and therefore indicates a consistent performance with respect to the one-sided probability region.

For each of the two scenarios depicted in Figure 8.2 five test drives were performed and an overall number of 897 measurement cycles recorded. Table 8.3 concludes the evaluation results for the two tracking approaches. In the approach and the crossing or turning phase, the CT performs significantly better than the COG tracking. Especially in the crossing/turning phase the sample standard deviation of the CT method is less than half and the RMS only a third compared to the COG tracking. The high bias of -1.0 m/s of the COG tracking especially in the crossing/turning phases underlines the necessity



Urban area Rural road Freeway

Figure 8.11: Overall 26,229 measurement cycles were recorded partly in urban areas, on rural roads and on a freeway. The laser scanner measurement are projected into the image.

of a tracking system, which accounts for the object's extent. The low bias of the cuboid tracking in the crossing/turning phase proves the good performance of this approach in this challenging scenario. In contrast, the relatively high bias of the CT in the initial phase is rather unexpected and can only be partly explained by the initial uncertainty in the velocity estimate. In the following phase both tracking approaches perform equally well. In this phase the center of gravity stays stable on the center of the rear and enables a good tracking accuracy.

Preceding Vehicle

For an evaluation of the tracking approaches under different driving conditions, a test drive was performed, which led partly through several urban areas, contained several sections on rural roads and a section on a freeway. Overall 26,229 measurement cycles were recorded, which corresponds to 22 minutes at a measurement frequency of 20 Hz. Figure 8.11 shows typical examples of the three different scenarios.

Over the complete sequence a car equipped with GPS and ESP preceded the host vehicle and recorded its velocity. In Figure 8.12 the reference velocity of the preceding vehicle is shown. Sections in which the velocity is below 15 m/s indicate an urban area, velocities between 15 m/s $\leq v <$ 28 m/s a rural road and velocities above 28 m/s the section on the freeway.

The standard deviation of the velocity error shown in Table 8.4 is 0.9 m/s for the COG, 0.8 m/s for the CT and with 0.5 m/s much lower for the CT+TT approach. The bias is in all three cases identical and most probable caused by biased ESP speed sensors[2]. All three tracking approaches reach a high detection rate. The CT+TT approach manages to increase the detection rate, as the video based tracking provides measurements in time instances, in which the laser scanner does not observe the vehicle. In the data sequence the preceding car leaves four times the FOV of the laser scanner in the elevation for 15, 3, 4 and 5 subsequent measurement cycles (overall 27), caused by a pitching of the host vehicle. The video based template tracking bridges those measurement gaps and reduces the number of missing observations to 1 cycle.

[2]The speed is measured by the ESP indirectly using the wheel speed. Different wheel diameters, caused for instance by the air pressure in the tire, induce a scale factor in the velocity estimate.

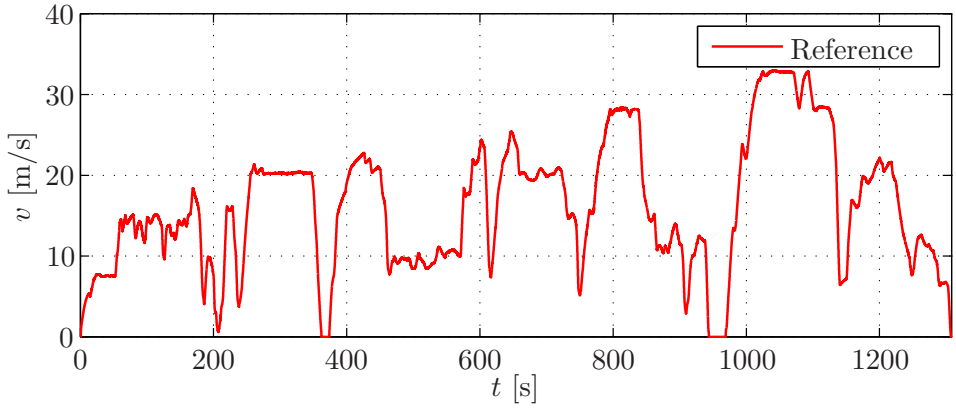


Figure 8.12: Velocity profile of the reference vehicle. Sections in which the velocity is below 15 m/s indicate an urban area, velocities between 15 m/s $\leq v <$ 28 m/s a rural road and velocities above 28 m/s the section on the freeway.

Table 8.4: Velocity error statistics for the tracking of a preceding vehicle. Sample standard deviation of the velocity estimation error $s_{\tilde{v}}$, the bias $\mu_{\tilde{v}}$ and the RMS with respect to the center of gravity (COG), the cuboid tracking (CT) and the fusion (CT+TT) approaches. Additionally the detection rate and the consistency rate are provided. The statistics are given for the complete data set (all) and for confirmed objects only (conf.).

	$\mu_{\tilde{v}}$ [m/s]		$s_{\tilde{v}}$ [m/s]		RMS$_{\tilde{v}}$ [m/s]		Det [%]		Cons [%]	
Method	all	conf.	all	conf.	all	conf.	all	conf.	all	conf.
COG	-0.1	-0.1	0.9	0.3	0.9	0.3	99.9	99.4	99.9	99.9
CT	-0.1	-0.1	0.8	0.2	0.8	0.2	99.9	99.2	99.9	100.0
CT+TT	-0.1	-0.1	0.5	0.2	0.5	0.2	100.0	99.4	99.9	100.0

The consistency of the filter is analyzed in terms of the NEES of the velocity. Figure 8.13 and Figure 8.14 show the cumulative probabilities of the corresponding NEES values assuming a chi-square distribution with one degree of freedom[3]. In these plots a horizontal line can be drawn and the corresponding probability value on the ordinate indicates the percentage of data points which should be below this threshold for the estimates to be consistent. For the calculation of the consistency rates in Table 8.4 a threshold of $Q = 0.99$ is chosen. All three tracking approaches are with a rate of 99.9 % therefore consistent with respect to the 99 % probability region. However, for pre-crash or emergency brake applications it is preferable, if the errors are bounded. The 99 % probability region is therefore, admittedly rather arbitrarily, defined as the error bound and every NEES value above should be prohibited. Two sections shortly before $t = 200$ s and at the end of the sequence exhibit a NEES value outside the 99 % probability region. For the CT+TT approach an additional third section shortly after $t = 600$ s can be identified as inconsistent.

[3]The plots for the CT approach are very similar to the CT+TT and are therefore omitted.

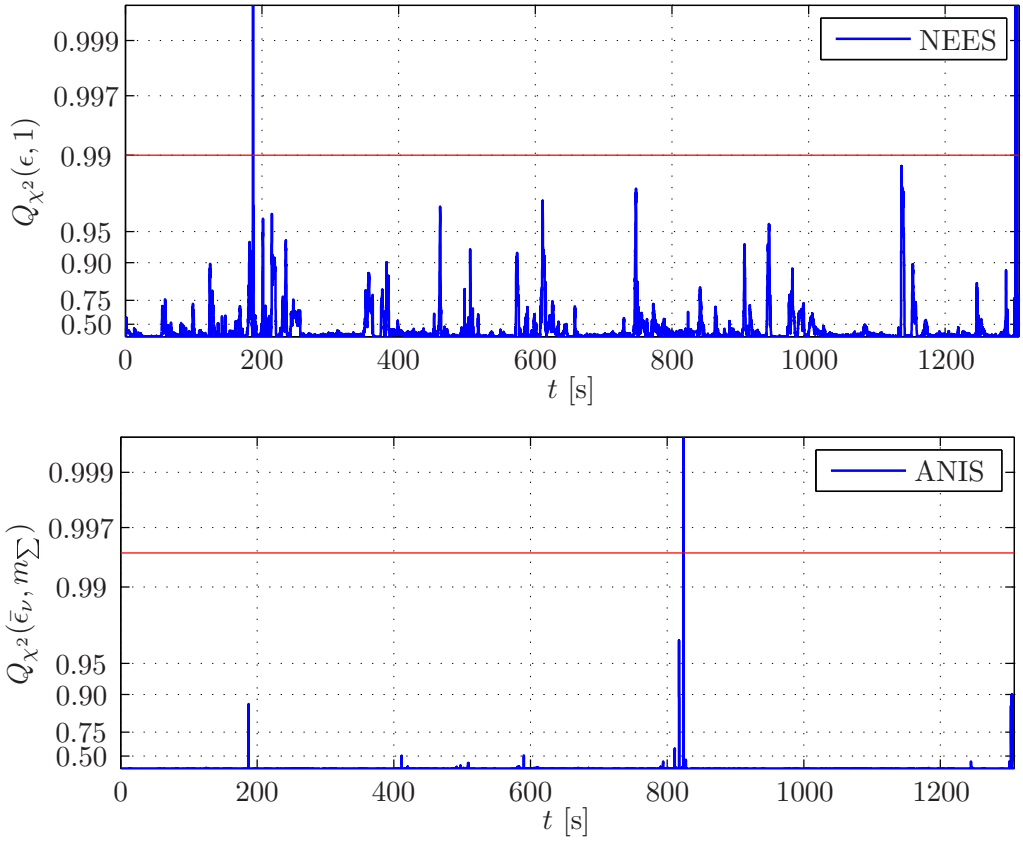


Figure 8.13: NEES and ANIS for the laser scanner based center of gravity (COG) tracking.

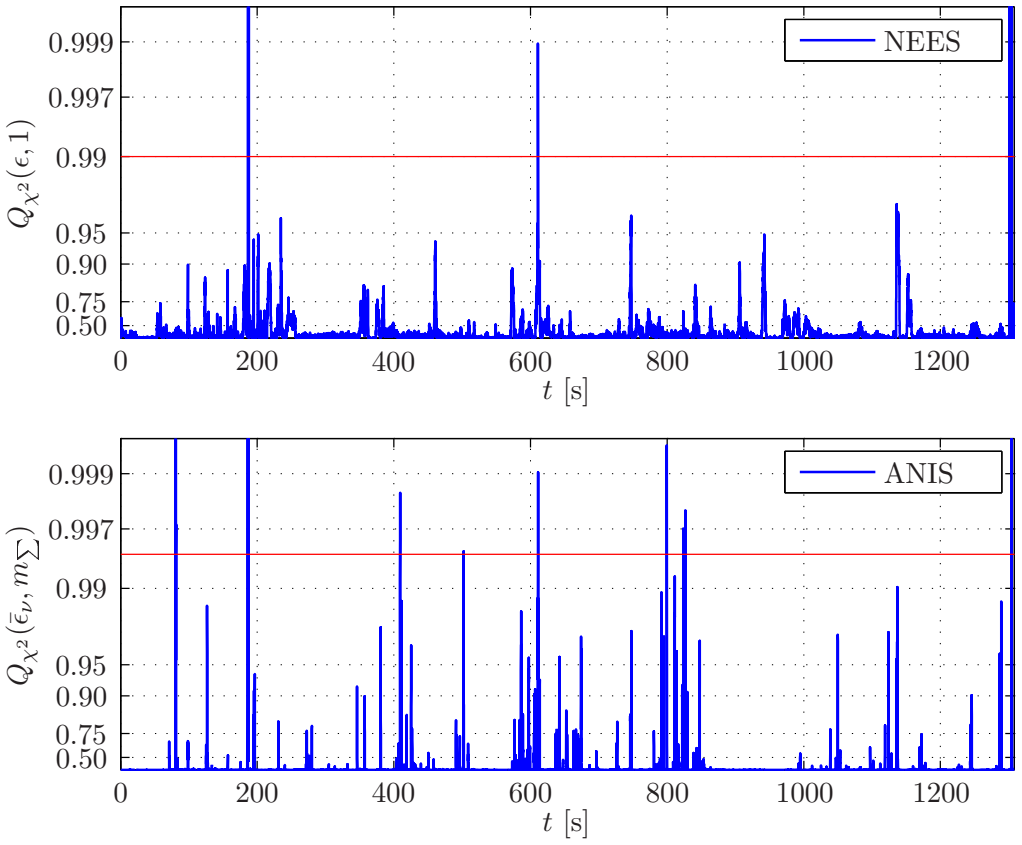


Figure 8.14: NEES and ANIS for the fusion approach (CT+TT).

Another consistency measure is the time-averaged normalized innovation error squared (ANIS). The ANIS averages the NIS over several measurement cycles (here over 8 cycles), which has the effect of a low-pass filter. The corresponding cumulative probability of the ANIS assuming a chi-square distribution is shown in Figures 8.13 and 8.14 for the COG and the CT+TT approaches. The ANIS of the COG tracking has in average a lower level and exhibits less peaks outside the 99.5 % probability region. Even if the ANIS of the CT+TT approach exhibits over the whole sequence statistically still consistency, the lower ANIS of the COG tracking seems to be preferable.

The difference in the ANIS of the two approaches results mainly from the different dimensions of the measurement vectors. The measurement vector of the COG tracking contains only the position, whereas a full measurement vector of the CT+TT approach comprises the position, width, length, orientation and image position. The ANIS of the CT+TT approach can therefore signal an inconsistency of the measurements if the orientation, width, length or image position deviate disproportionately from the predicted values, even if the measured position is consistent.

The velocity error statistics, detection rates and consistency rates of the velocity estimates discussed so far are calculated with respect to the complete data set (Table 8.4, columns *all*). In general, however, an application does not use the object state estimates at all time intervals. At the object creation, the state estimates are usually too uncertain to be usable. An application waits for a certain time period, until the uncertainty is within limits. For the evaluation the estimated standard deviation of the velocity components must be below $\sigma_{v,t} = 0.8$ m/s, for the orientation below $\sigma_{\psi,t} = 0.15$ rad and either the uncertainty of the width or the length must stay below $\sigma_{wl,t} = 0.15$ m. For the COG approach only the covariance of the velocity can be tested, as the COG does not estimate the orientation, width nor the length.

Additionally, an application might only want to utilize the state estimates if they are consistent in terms of the NEES. The NEES is in general, however, not available, as it is calculated with reference data. Therefore, the process intrinsic consistency ANIS can only be used, which is calculated online. If the ANIS exceeds a certain threshold, the object state is rejected. This object confirmation process is discussed in detail in Chapter 4.3.6. For the evaluation the ANIS must be within the 99.5 % probability region.

A confirmed object must meet both criteria. It must be consistent with respect to the ANIS and the estimated uncertainty must be within limits. Table 8.4 concludes the results for the confirmed (conf.) object state estimates. The sample standard deviation and the RMS of the velocity estimate of the confirmed objects is significantly reduced, as the initial periods of high uncertainty are filtered out. As expected the detection rate decreases, as the initial periods and the inconsistent sections are filtered out. The consistency check based on the online calculated ANIS effects the consistency rate (which is calculated from the NEES consistency of the velocity estimate) of the CT and CT+TT approaches. The consistency rate of the CT and CT+TT approaches can be increased to 100.0 %, while there is no significant change in the COG tracking. This can be best visualized considering Figures 8.13 and 8.14. The ANIS of the COG tracking identifies one section shortly after $t = 800$ s as inconsistent. However, the inconsistent time segments shown in the NEES plot are not identified. In contrast, the ANIS of the CT+TT approach exhibits many periods of inconsistent innovations. It identifies the three inconsistent sections shown in

the NEES. The higher sensitivity of the ANIS for the CT+TT approach can therefore be used to its advantage, as the CT+TT approach reaches a consistency rate of 100.0 % at the same detection rate of 99.4 % as the COG, which reaches a consistency rate of only 99.9 %.

It should be mentioned, that the parameters for the confirmation strategy are in general application specific. The fusion system delivers the state estimates, the uncertainties and the NIS values as part of the environment description. Each application can then perform its own object confirmation based on these values.

The process intrinsic consistency of the CT approach represented by the ANIS considers the consistency of the position, orientation, length and width measurements. Additionally, the quality of the contour matching could be incorporated. An object would only be confirmed, if the contour shape is well represented by the associated measurements. This would enable an additional useful consistency of the geometric object properties.

Cut-in Situation

At a first glance the results for the standard deviation of the velocity error in Table 8.4 are astonishing. There is no significant difference between the accuracy of the CT and the CT+TT approach considering only the confirmed objects. This is astonishing, as the CT+TT improves the lateral position accuracy significantly. However, the statistics reflect the uncertainty of the magnitude of the velocity vector. The longitudinal component of the velocity vector dominates the lateral component, as the object vehicle's orientation is mostly aligned with the host vehicle. In this case a noisy lateral velocity estimate has a negligible influence on the magnitude of the velocity vector. The accuracy of the velocity is mainly influenced by the measurement noise in the longitudinal direction, which is in both tracking approaches identical[4]. The statistics therefore provide useful insights into the tracking performance relevant for the longitudinal control of ACC applications and to some extend for safety applications such as EB.

For object selection algorithms within ACC systems or the object trajectory prediction in EB applications the accuracy of the lateral velocity is of interest. Consider a cut-in scenario as depicted in Figure 8.4. Figure 8.15 shows the lateral velocity component for the CT and the CT+TT approaches. The time intervals t_1, t_2 and t_3 from Figures 8.5 and 8.6 are depicted as well.

Initially the lateral velocity estimates of the two approaches are identical. Only when the vehicle's rear is completely in the FOV of the camera (at $t = 3.4$ s) the noise in the velocity signal is significantly reduced by the video based template tracking in the fusion approach.

Even if this plot implies a higher accuracy in the lateral velocity estimate of the CT+TT approach, the lateral velocity could still be biased. A bias in the host vehicle's yaw rate sensor, for instance, imposes a bias on the absolute lateral velocity estimate, which increases with the distance to the object.

[4]For the situations in which the object vehicle is standing and the noise in the lateral velocity component has a significant influence, the CT+TT is not applied, as it requires a certain minimum distance to the preceding vehicle.

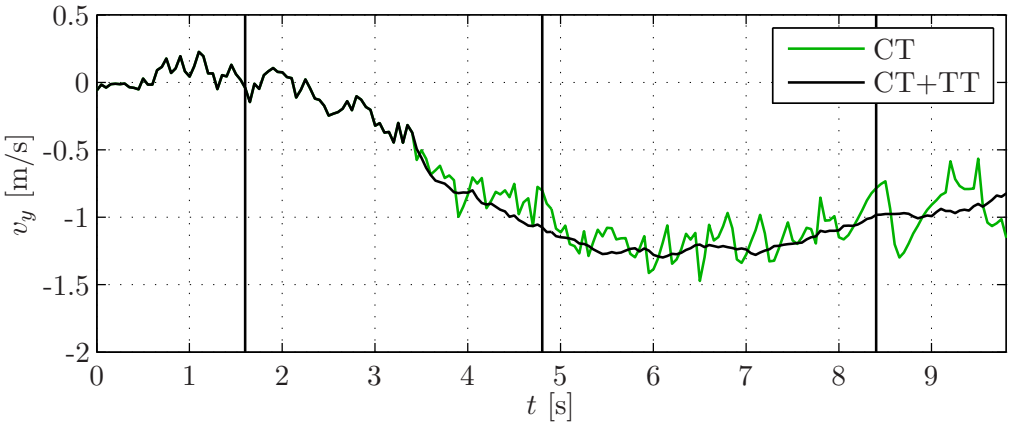


Figure 8.15: Lateral velocity estimates of the laser scanner based cuboid tracking (CT) compared to the cuboid tracking with image based template tracking (CT+TT).

8.4.4 Velocity Estimation Error for Stationary Roadside Objects

The evaluation described in the previous sections concentrate on the object tracking of cars and trucks. For most applications, the state estimate of the numerous stationary objects beside the road are of equal importance. Present ACC applications do not react on any stationary object. For the classification of a stationary object, the velocity estimate must be accurate. ACC systems based on radar sensors have the advantage of a direct radial velocity measurement. For systems based on a lidar sensor the velocity is estimated over time by associated position measurements.

For the evaluation of stationary objects, the 22 minute measurement sequence described in Chapter 8.4.3 is labeled as indicated in Figure 8.16 and detailed in [169]. The laser scanner measurements of the reference vehicle and all other cars, trucks, bikes and pedestrians receive a label. The remaining laser scanner measurements are assumed to originate from stationary objects.

The object lists of the different tracking approaches are then searched for the labeled objects in each cycle. The remaining objects in the object lists are the stationary objects. The velocity error of these stationary objects is directly given by the estimated velocity, as the true velocity is zero.

Table 8.5 presents the resulting statistics for the velocity error $\tilde{v}$. The RMS of the velocity estimates considering all objects is high, which is partly caused by the high initial uncertainty of numerous newly created objects. The consistency rate exhibits a high degree of inconsistency. The velocity estimates of only 77.2 % of the object states for the CT approaches and 81.3 % for the COG tracking are consistent with the estimated uncertainty. However, these rather low values can be significantly increased with the confirmation strategy. About 92 % of the confirmed objects are consistent for the CT approaches. Even though this increase is rather promising, the remaining 8 % inconsistent velocity estimates should be subject to further research. A detection rate can not be provided in this case, as the stationary objects were not labeled. Therefore, the true number of objects is not known. The consistency rate results from the number of objects which are created by the tracking and which are not labeled.

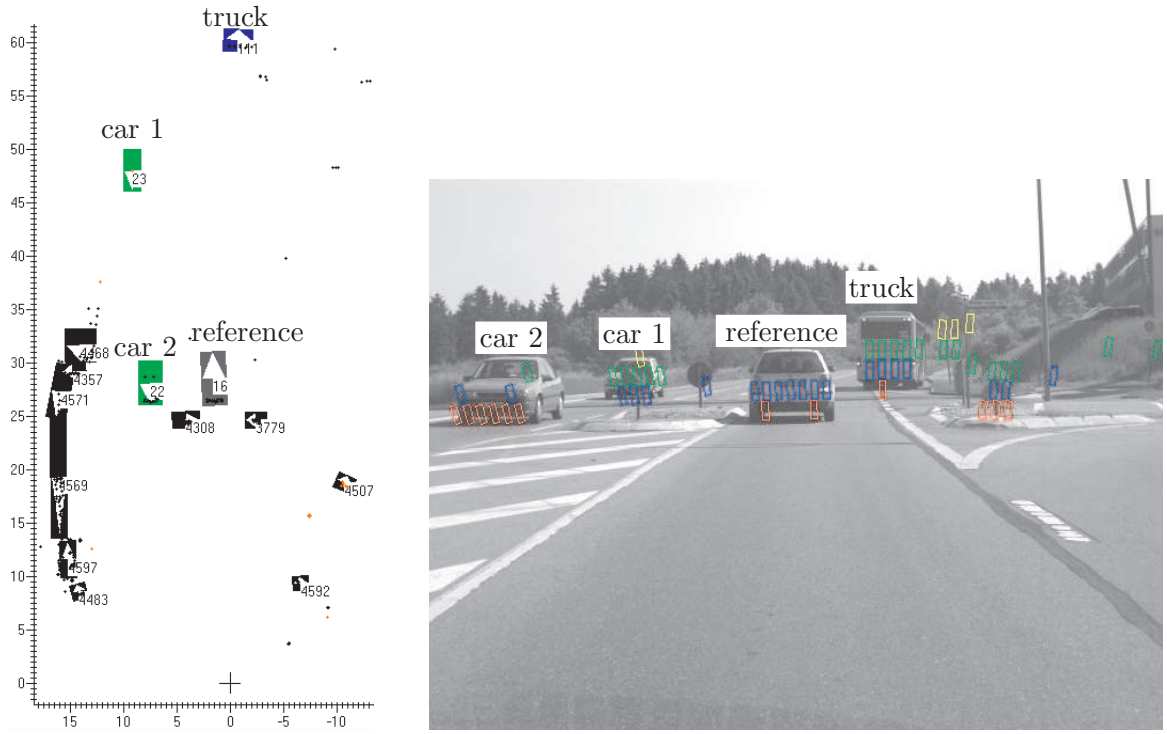


Figure 8.16: Labeled laser scanner measurements. The reference vehicle and all other cars, trucks, pedestrians and (motor-)bikes are labeled.

Table 8.5: Statistics for stationary roadside objects. The $\mathrm{RMS}_{\tilde{v}}$ of the velocity estimation error $\tilde{v}$ with respect to the laser scanner based center of gravity tracking (COG), the cuboid tracking (CT) and the fusion approach (CT+TT). Additionally the consistency rate is provided. The statistics are given for the complete data set (all) and for confirmed objects only (conf).

	$\mathrm{RMS}_{\tilde{v}}$ [m/s]		Consistency rate [%]	
Method	all	confirmed	all	confirmed
COG	**17.6**	7.6	**81.3**	78.6
CT	21.0	**3.0**	77.2	**92.1**
CT+TT	20.9	3.2	77.2	91.8

Table 8.6: Detection (Det) and consistency (Cons) rates for the preceding car and stationary objects with respect to the laser scanner based center of gravity tracking (COG), the cuboid tracking (CT) and the fusion approach (CT+TT). Additionally the number of objects is provided, which are detected in the driving corridor (DC) between the host vehicle and the preceding vehicle. The statistics are given for the complete data set (all) and for confirmed objects only (conf).

			Stationary obj.		Preceding car			
	in DC [1/h]		Cons [%]		Det [%]		Cons [%]	
Method	all	conf.	all	conf.	all	conf.	all	conf.
COG	**214**	44	**81.3**	78.6	99.9	**99.4**	99.9	99.9
CT	925	**30**	77.2	**92.1**	99.9	99.2	99.9	**100.0**
CT+TT	933	**30**	77.2	91.8	**100.0**	**99.4**	99.9	**100.0**

Comparing these results with other environment perception systems, it should be noted, that the FOV of the utilized laser scanner is many times larger than the FOV of most other sensors. As most stationary objects are situated on the roadside or beyond, the laser scanner observes proportionally more area, which is crowded by stationary, largely unstructured obstacles.

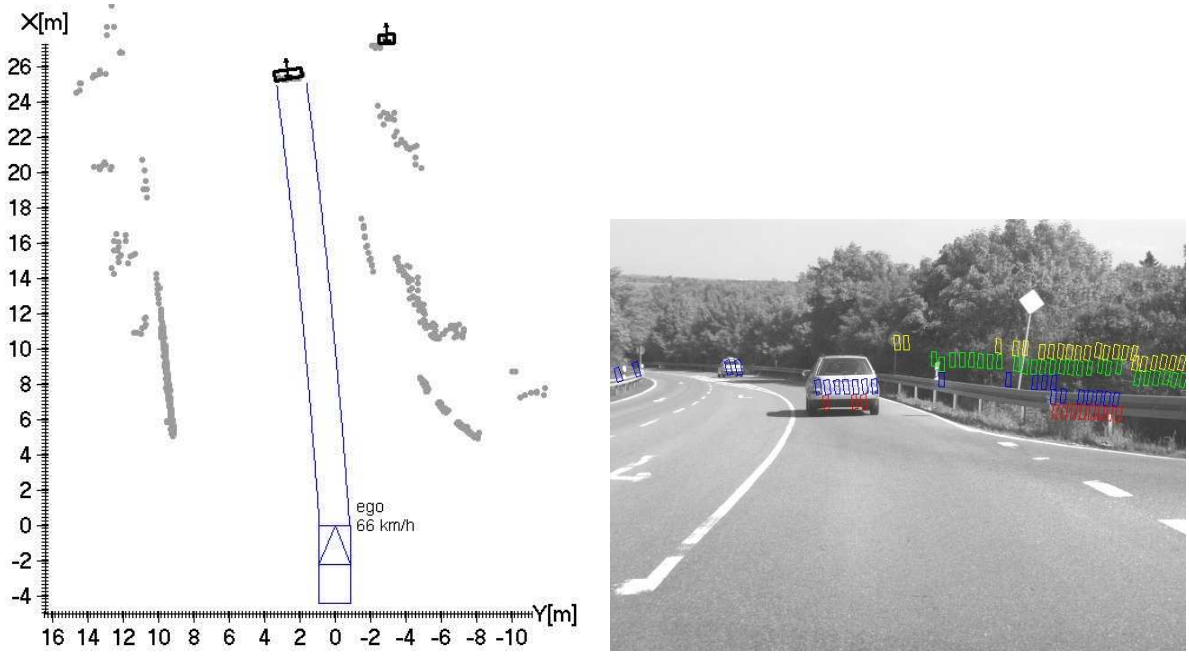
8.4.5 Detection Statistics

Table 8.6 concludes the detection and consistency rates for the 22 minutes sequences recorded on freeways, rural roads and urban areas described in Chapter 8.4.3 and 8.4.4. A third measure describes the number of objects, which are detected in the driving corridor (DC) between the host vehicle and the preceding vehicle (Figure 8.17).

Detections in the Driving Corridor

Considering all detections, the objects in the driving corridor are rather numerous for the CT approaches, which can be explained by the cuboid object model. All objects are approximated by a cuboid, even those which are rather poorly described by this model. If the orientation of a large object on the side of the road becomes inaccurate over time, part of this object can move into the driving corridor. It is usually that part of an object, which is not supported by measurements at that time instance.

However, the powerful consistency checks of the CT approaches identify almost all of the objects in the driving corridor as inconsistent and reject them. For the COG tracking the remaining 16 confirmed objects (44 per hour) of the 26,229 measurement cycles are caused by dust behind the preceding car shown in Figure 8.18. For the CT approaches 9 of the remaining 11 confirmed objects (30 per hour) in the driving corridor result from the dusty air. Two object detections are caused by an object representing a bush very close the preceding vehicle on a winding road in a village which moves slightly into the driving corridor.



(a) (b)

Figure 8.17: Driving corridor shown in the bird's view (a).

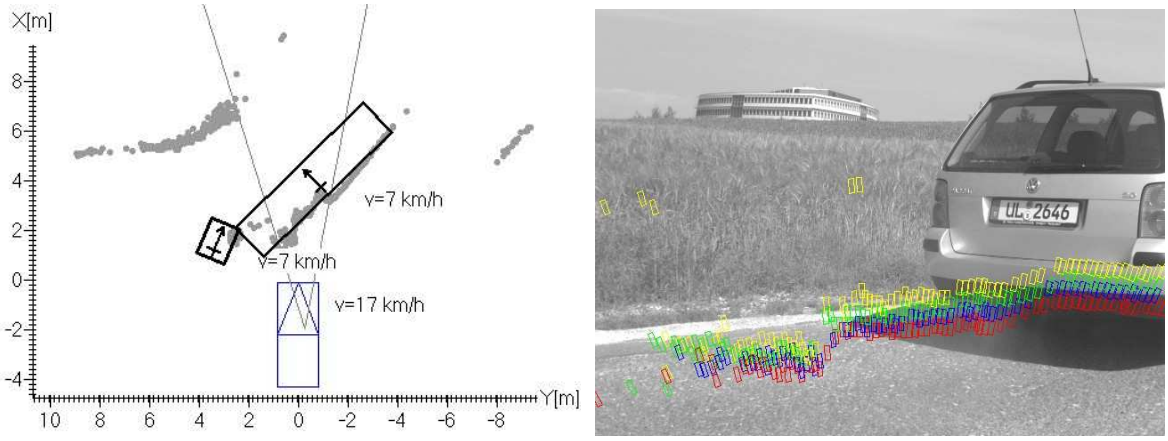


Figure 8.18: Dust behind the preceding car leads to incorrect object states, inconsistent measurements and to ghost objects.

Table 8.7: Situations in which there is no confirmed object for the preceding vehicle and the corresponding (relative) frequency for the fusion approach (CT+TT).

Situation	not confirmed (number)	not confirmed [%]	not confirmed [%] (whole sequence)
Initialization period	$5 \cdot 7 = 35$	22	0.13
Feature extraction laser scanner	5	3	0.02
Feature extraction camera	27	17	0.10
Dust	42	26	0.16
Pitching of the host vehicle	15	9	0.06
Object merged with roadside	37	23	0.14
Sum	161	100	0.61

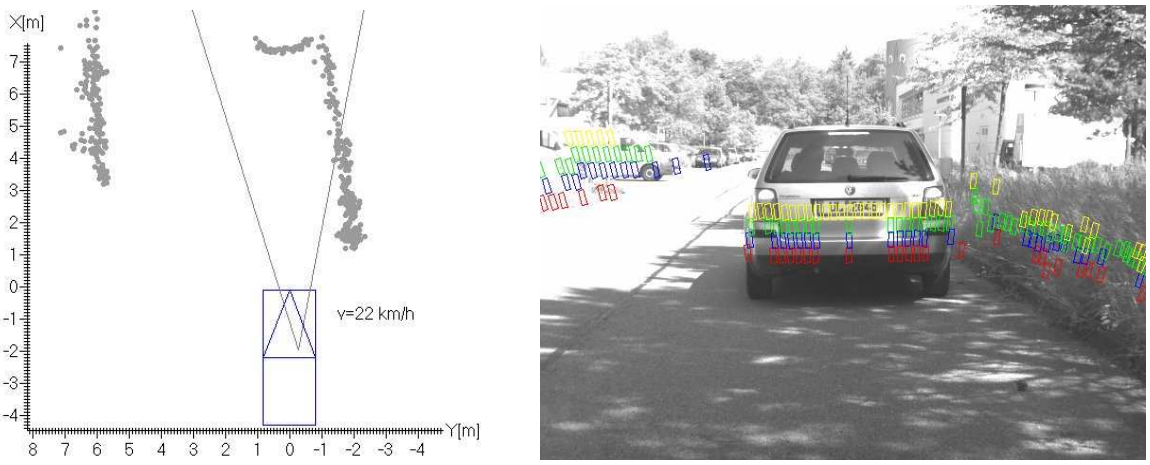


Figure 8.19: Preceding vehicle is standing close to the roadside.

Not Confirmed Objects

Considering the detection rates of the CT+TT in Table 8.6, the question arises, why in 0.6 % of the time the preceding object is not confirmed, even if the overall detection rate is 100 % and the corresponding consistency rate 99.9 %. Table 8.7 presents the results of a detailed analysis of the 161 time instances (0.6 %), in which the object is not confirmed.

The most prominent reason for a rejection of the tracked object by the object confirmation check is the process intrinsic inconsistency caused by the dust behind the preceding vehicle (Figure 8.18). This situation last more than two seconds. In 23 % of the 161 time intervals, the preceding vehicle comes to a standstill very close to the road border (Figure 8.19). The laser scanner measurements on the vehicle and the roadside grass are so close, that the segmentation merges the observations to a single segment and the subsequent feature extraction produces erroneous measurements. The object confirmation module recognizes these inconsistent measurements and rejects the object state estimate. 22 % or 35 of the 161 cycles are at the beginning of a data set. As the whole 22 minute

sequence contains 5 separate data sets and the initialization phase of an object until the first object confirmation takes 7 measurement cycles, which corresponds to 350 milliseconds, the overall number of cycles for the object initialization is therefore 35. 20 % of the 161 cycles of rejected object tracks are caused by an incorrect feature extraction. The template tracking of the camera is sometimes mislead by shadows over large parts of the vehicle in large distances. The laser scanner feature extraction produces in 99.98 % of the overall time correct measurements, which are consistent with the predicted uncertainties. Only in 5 cycles the laser scanner feature extraction produces incorrect measurements even if the raw measurements exhibit no significant irregularities. In 15 cycles the object is not confirmed as measurements are missing because of a strong pitching of the host vehicle. In most cases of strong pitching the camera template tracking bridges the cycles of missing laser scanner measurement. The number of cycles, which can be bridged, is however limited, as the uncertainty of the radial velocity raises with time without the distance measurements of the laser scanner. For long laser scanner measurement gaps the object is finally rejected by the object confirmation.

Inconsistent Confirmed Objects

Table 8.6 presents the consistency rates of the tracking approaches with respect to the 99 % probability region of the NEES calculated from the velocity. The consistency rates are 99.9 % for the COG and 100.0 % for the CT approaches. Table 8.8 shows the situations in which the track of the preceding vehicle is confirmed and forwarded to the applications even if the velocity estimates are inconsistent.

Dust behind the vehicle (Figure 8.18) impairs the position and velocity estimation and leads to inconsistent object state estimates. The process intrinsic consistency checks of the object confirmation recognize most of the inconsistent laser scanner features. However, for the COG tracking 13 measurement cycles remain and for the CT approaches only 3. For 4 time instances an incorrect laser scanner feature extraction of the cuboid tracking leads to inconsistent velocity estimates. In 10 measurement cycles the COG tracking is disturbed by the segments, which merge the road side grass with the measurements from the preceding vehicle (Figure 8.19). The object confirmation of the CT approaches recognizes this situation and rejects the inconsistent estimates.

Table 8.8: Situations of a confirmed track representing the preceding vehicle with inconsistent velocity estimates and the corresponding frequency. Results are shown for the laser scanner based center of gravity tracking (COG), the cuboid tracking (CT) and the fusion approach (CT+TT).

Situation	COG	CT	CT+TT
Dust	13	3	3
Feature extraction laser scanner	–	4	4
Object merged with roadside	10	–	–
Sum	23	7	7



(a) (b)

Figure 8.20: The host vehicle passes a stationary truck. (a) The bird's eye view with the laser scanner data and the current state estimate. (b) The corresponding image with the laser scanner projected into the image domain.

8.5 Evaluation of the Environment Description for Trucks

The position and motion estimation of a truck observed from the rear is fairly straight forward using the CT or CT+TT approach. The tracking becomes a challenge, if the truck is passed and only seen from the side. The segmentation often fails, as there are large gaps between the measurements and the object disintegrates. Even if the object is represented by a single segment, the position of the truck is often not observable, as the rear corners are out of the FOV and the angle of incident to the frontal left corner too close to 90° for a reliable measurement. Both challenges are addressed by the displacement estimation explained in Chapter 5.4.3, which facilitates the direct estimation of the motion. The following experiments investigate the accuracy and robustness of the displacement tracking (DT) compared to the center of gravity tracking (COG), which is the standard tracking used in the literature (Chapter 8.2).

8.5.1 Stationary Truck

Experimental Setup

For the determination of the estimation errors, the host vehicle passes a stationary truck as shown in Figure 8.20. The estimation error is equal to the estimated velocity, as the true velocity is zero.

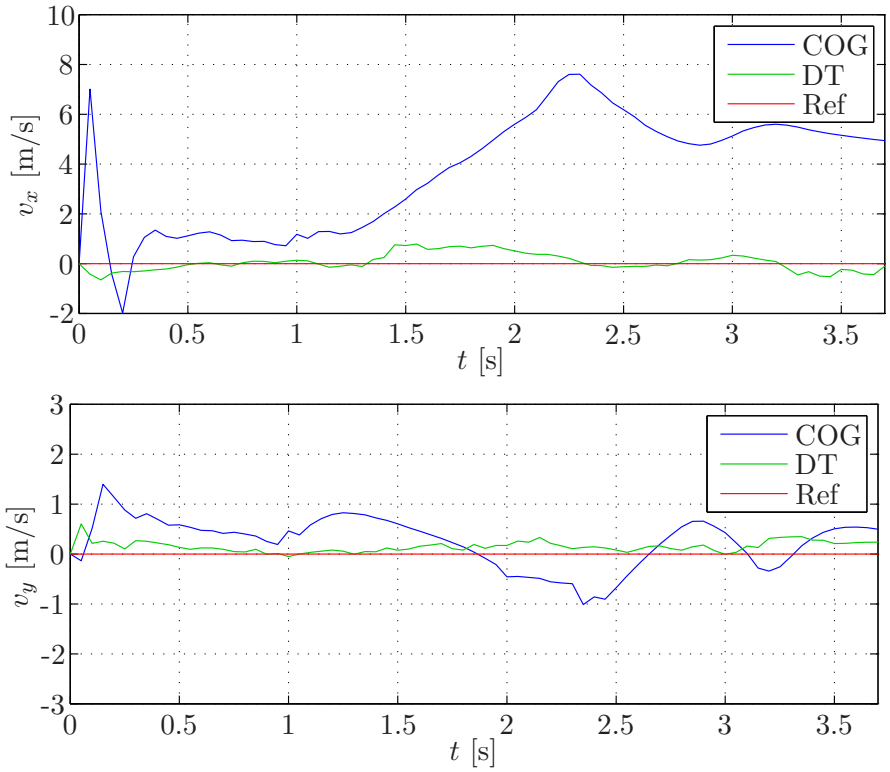


Figure 8.21: Estimated velocity of a stationary truck for the center of gravity (COG) and the displacement tracking (DT). The reference velocity is zero.

In general, the displacement tracking is combined with the cuboid tracking. If the host vehicle is beside the truck, the lateral position and orientation of the track can still be measured by an I-shape, even if the longitudinal position is not observable. However, after an initialization with the cuboid tracking, only the laser scanner displacement measurements are utilized in this experiment, in order to demonstrate the performance of the DT approach. The segmentation parameters were adapted for the COG tracking, in order to enable a continuous object track.

Experimental Results

Figure 8.21 contrasts the estimated velocity of the displacement tracking and the center of gravity tracking with respect to the reference (Ref). The COG tracking exhibits strong estimation errors, which are caused by a moving center of gravity. While the host vehicle passes the truck, the rear of the truck leaves the FOV and the center of gravity moves along the contour of the truck, which induces an additional virtual velocity. In contrast, the estimated velocity of the displacement tracking is very close to the reference for both velocity components. The estimated uncertainty of the velocity components is depicted in Figure 8.22. The displacement tracking exhibits a significantly smaller estimated standard deviations and a faster convergence than the COG tracking.

The filter consistency is shown in Figures 8.23 and 8.24 in terms of the normalized innovation error squared (NIS) ϵ_ν and the normalized (velocity) estimation error squared

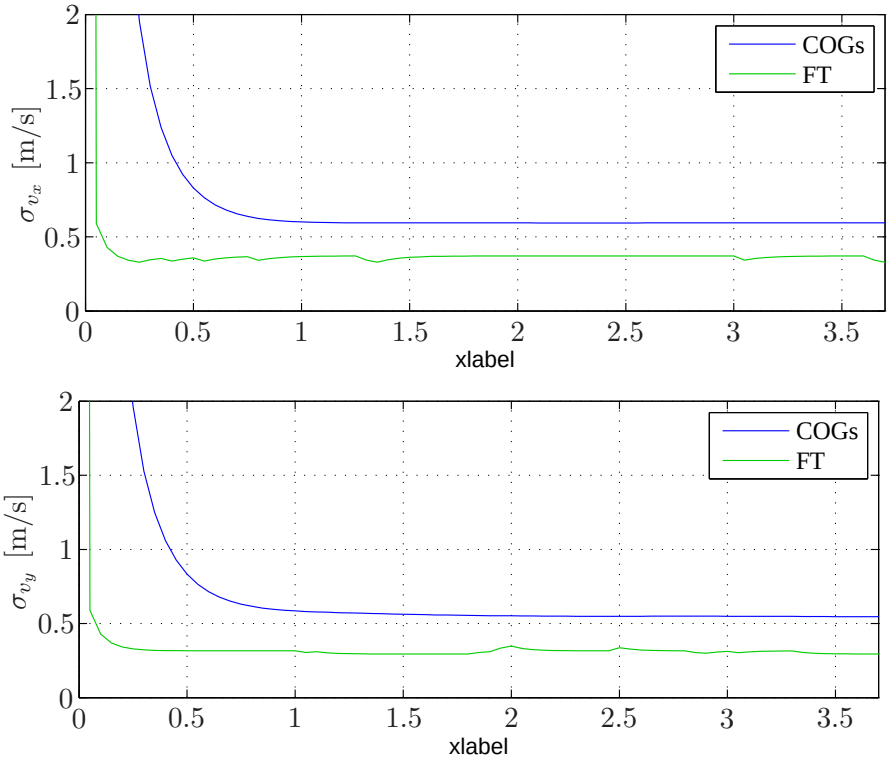


Figure 8.22: Velocity uncertainty estimates of a stationary truck for the center of gravity (COG) and the displacement tracking (DT).

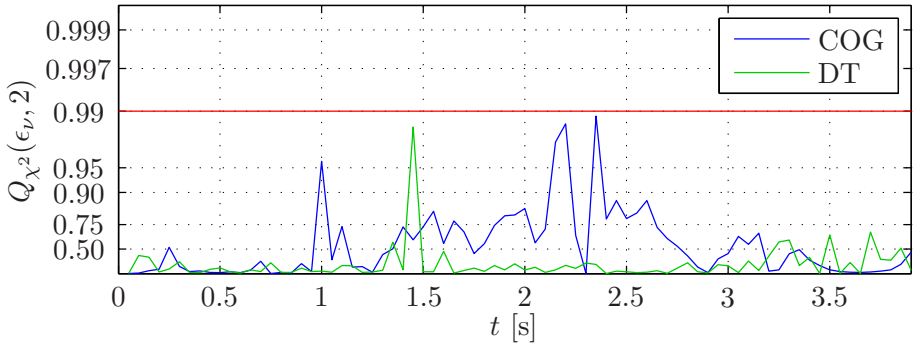


Figure 8.23: NIS of a stationary truck for the center of gravity (COG) and the displacement tracking (DT).

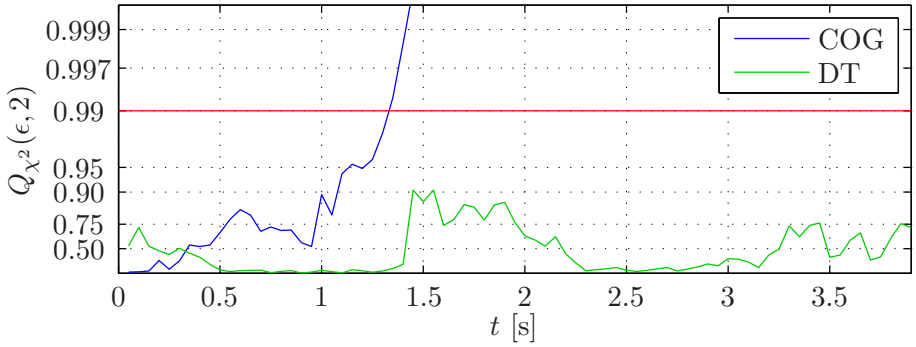


Figure 8.24: NEES of the velocity of a stationary truck for the center of gravity (COG) and the displacement tracking (DT).

(NEES) ϵ. The plots actually exhibit the cumulative probability $Q_{\chi^2}(\epsilon, 2)$ assuming a chi-square distribution with two degrees of freedom. A horizontal line can then be drawn at an arbitrary value of Q. If the filter is consistent, the relative frequency of data points below any such line must be higher than the corresponding cumulative probability Q. Exemplary, a line indicates a commonly used probability region of 99 %. Both tracking schemes perform well in terms of the NIS. However, the strong velocity estimation error of the COG tracking is clearly seen in Figure 8.24. The error in the velocity of the COG tracking is for a long period significantly higher than the estimated uncertainty allows. The COG velocity estimate is therefore not consistent. In contrast, the NEES of the displacement tracking is always within the probability region of 90 %.

For this exemplary situation, the displacement tracking delivers consistent velocity estimates with a low estimated uncertainty. In contrast the COG tracking produces inconsistent velocity estimates, even if the estimated uncertainty is significantly higher and the filter therefore less efficient.

8.5.2 Moving Truck

Even if situations with a moving truck as depicted in Figure 8.25 do not provide a reference velocity, they enable an evaluation over a longer period of time.

Experimental Setup

The host vehicle was driven beside a truck on a freeway. As can be seen in Figure 8.26 the velocity of the host vehicle varied with respect to the trucks velocity, so that different parts of the truck were observed. A sequence of almost 86 seconds was recorded with 1,713 measurements.

Experimental Results

The estimated uncertainties of the velocity shown in Figure 8.26 are again very small. The NIS in Figure 8.27 exhibits in this experiment several values above the 99 % line and many outside the 90 % probability region. However, it is hard to tell if the consistency criterion is met for every probability. Figure 8.28 visualizes the consistency depending on the chosen probability threshold. This chi-square probability plot is the equivalent to the normal probability plot described in Chapter 7.2.1 for a chi-square distribution. For each NIS value its corresponding empirical cumulative probability is calculated assuming a chi-square distribution with two degrees of freedom. The line represents the theoretical cumulative probability function for the same chi-square distribution. Data points above or on the line indicate, that the filter is consistent with respect to its innovation error. For a consistency independent of the chosen probability region, all data points should be above or on the line. This plot is therefore a powerful tool in order to visualize the consistency without the necessity to chose a certain probability region.

Additionally, the plot indicates, that the innovation error squared is not chi-square distributed with two degrees of freedom, as the data points do not lie directly on the line. As the data is significantly above the line, it can be deduced, that the innovation covariance is on average larger than the innovation error. An overestimated measurement

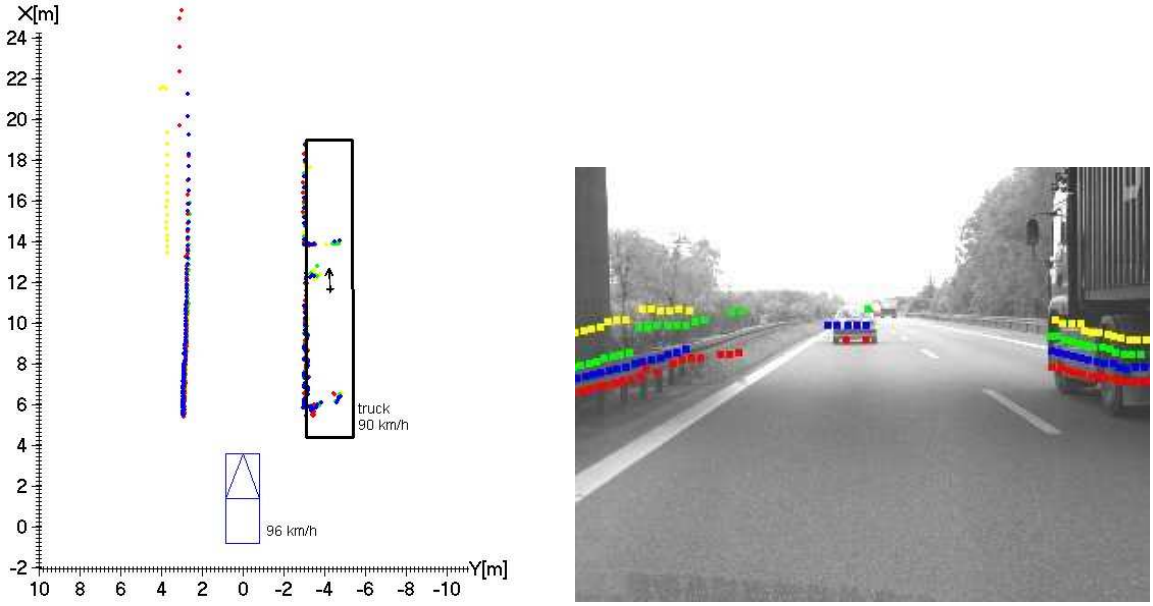


Figure 8.25: The host vehicle passes a truck on a freeway.

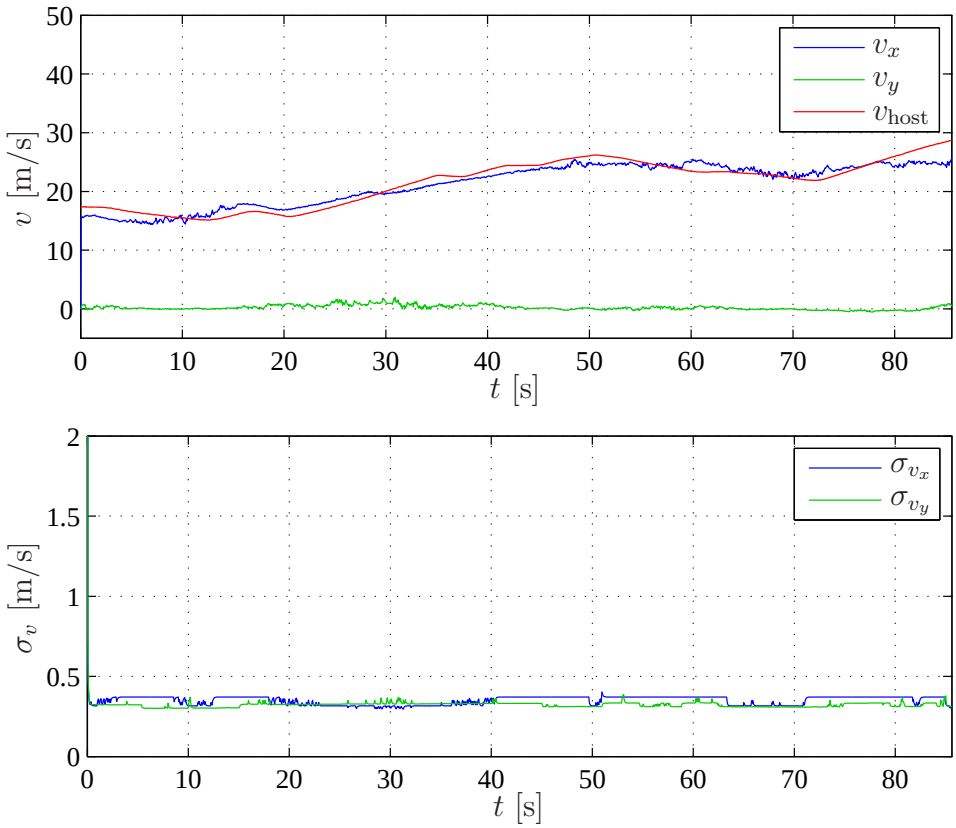


Figure 8.26: Estimated velocity and the corresponding estimated standard deviation of a truck which is passed by the host vehicle for the displacement tracking (DT). The velocity of the host vehicle is provided for comparison.

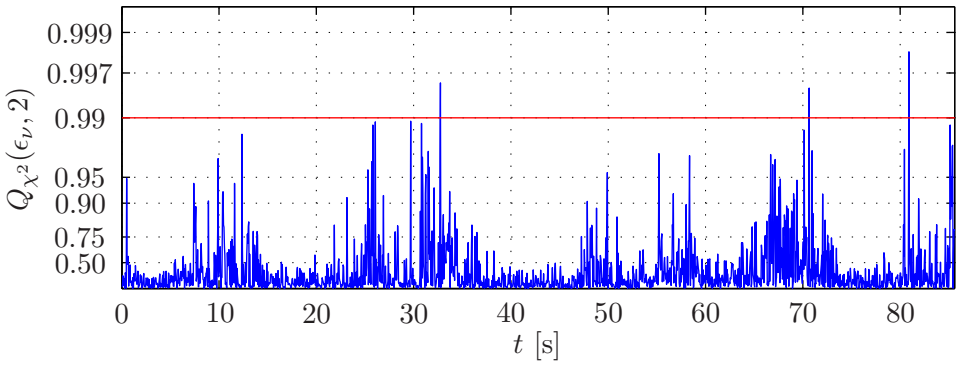


Figure 8.27: NIS of a truck which is passed by the host vehicle for the displacement tracking (DT).

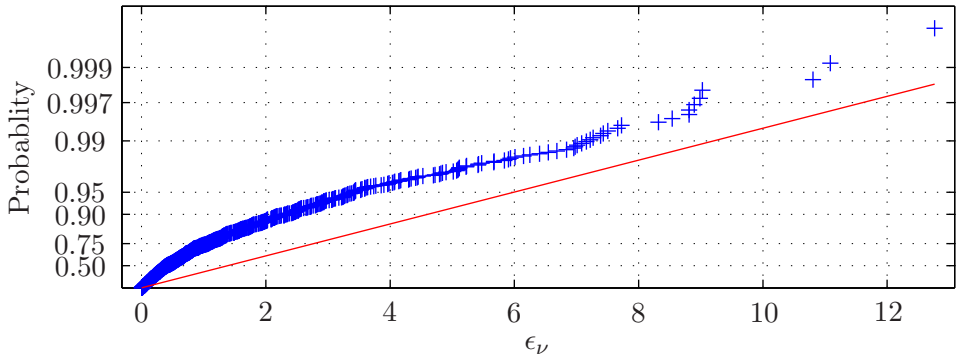


Figure 8.28: Chi-square probability plot for the NIS of a truck which is passed by the host vehicle. The red line indicates the theoretical cumulative probability function.

covariance or predicted state covariance are the two possible reasons. In this case, the predicted state covariance, designed for moderate object velocity changes, is most likely too large, as in the scenario the velocity of the object varies only little. In general, the data should be as close as possible to the line, as the efficiency of an estimator increases with the similarity of its empirical cumulative probability distribution with the theoretical function. It depends however on the application, whether the consistency or the efficiency is of higher interest.

8.5.3 Detection Rate and Precision

In practice, laser scanner measurements from a crash barrier beside the host vehicle appear in a similar shape as a truck driving beside the host vehicle. It can be difficult for a classifier to differentiate between these two object types without a priori knowledge, such as the track history of an already classified truck. Therefore, the displacement estimation must be robust against observations from other object types.

The less significant the structure of the observations is, the more difficult it becomes to estimate the displacement. Therefore, the displacement measurement is discarded, if the gradient of the residual distance distribution is below a threshold of 0.14 (Chapter 5.4.3). As the gradient is calculated in x and in y direction, a one dimensional displacement

Table 8.9: Detection rate D and precision P of the displacement tracking (DT) applied to trucks and roadside objects. The correct displacement measurements in one and two dimensions is denoted by $n_{1D,r}$ and $n_{2D,r}$, the incorrect displacement measurements by $n_{1D,f}$ and $n_{2D,f}$. The overall number of measurement cycles is given by n_{ges} and n_{0D} is the number of cycles, in which no measurement could be produced.

Object	n_{ges}	$n_{2D,r}$	$n_{2D,f}$	$n_{1D,r}$	$n_{1D,f}$	n_{0D}	D	P
Stationary truck	1903	1708	42	133	0	20	**97 %**	**98 %**
	100 %	90 %	2 %	7 %	0 %	1 %		
Roadside	5752	152	406	5011	145	38	**90 %**	**90 %**
	100 %	3 %	7 %	87 %	3 %	1 %		

can be measured, if only one gradient fails the test. If both gradients fail the test, the measurement is omitted.

The evaluation is based on measurements from 22 scenarios with stationary trucks and 8 sequences with different objects delimiting the road. Some roads are equipped with crash barriers. On others bushes and grass delimit the trafficable surface. A displacement is considered to be correct, if its error is within the three-sigma gate of the measurement uncertainty.

The detection rate follows from the correct displacement measurements in one and two dimensions $n_{1D,r}$ and $n_{2D,r}$ and the overall number of measurements n_{ges}, with

$$D = \frac{n_{2D,r} + n_{1D,r}}{n_{ges}} \tag{8.23}$$

$$\tag{8.24}$$

The precision is calculated from the correct displacement measurements and the incorrect displacement measurements $n_{1D,f}$ and $n_{2D,f}$, with

$$P = \frac{n_{2D,r} + n_{1D,r}}{n_{2D,r} + n_{1D,r} + n_{2D,f} + n_{1D,f}} \tag{8.25}$$

Table 8.9 concludes the results for trucks and roadside objects. The detection rate for trucks is with 97 % very high. The precision is with 98 % still acceptable. In 90 % a correct two-dimensional displacement is measured in an additional 7 % a correct one-dimensional displacement. The observations from roadside objects form very often a straight line with little structure. The low number of 3 % correct two-dimensional displacement estimates reflects this circumstance. However, in 87 % a correct one-dimensional displacement can be determined. The detection rate of 90 % for roadside objects is completely acceptable, however the precision of 90 % should be higher.

The results show, that the displacement estimation works very well for trucks. Applied to other objects, the method performs quite robust in terms of the identification of one-dimensional structures, even if the overall precision should be higher.

8.6 Discussion

The presented evaluation provides a thorough performance analysis of the developed algorithms for vehicle environment perception. It is the first work in the area of laser scanner based object tracking known to the author, which substantiates the empirical results of the state estimates with ground truth data. This enables a meaningful comparison among different algorithms. In particular, it provides a quantitative benchmark for future approaches.

Three tracking approaches for passenger cars are compared: The standard center of gravity tracking (COG) known from the literature, the novel laser scanner data based cuboid tracking (CT) and the fusion approach (CT+TT). The fusion approach combines the cuboid tracking with the video template tracking.

The cuboid tracking has been developed especially for intersection scenarios, in which the performance of the standard center of gravity tracking delivers large errors in the velocity estimates. The novel approach accounts for the object dimensions and orientation. Thus, it reduces the errors by a factor of more than three. This is of major importance for safety applications, such as an emergency brake for intersections which will be described in Chapter 10.

The fusion of the cuboid tracking with the template tracking applies mainly for the tracking of preceding vehicles. With its higher angular resolution, the camera can improve the accuracy of the azimuthal position estimate by a factor of more than two. In combination with a lane recognition the template tracking is especially useful for an early lane change detection of preceding vehicles.

The object detection statistics and the consistency of the estimates has been demonstrated based on more than 25,000 measurement frames in representative scenarios on freeways, rural roads and urban areas. With the novel fusion approach CT+TT the highest detection rate of 99.4 % at a consistency rate of 100.0 % is achieved. A high detection rate is important for ACC applications, where the availability of the system must be high. The consistency of the estimates is necessary for safety applications such as the emergency brake. The 0.6 % missed object detections are mainly caused by dust behind the preceding vehicle or a merging of the vehicle with the roadside grass. Only 0.12 % of missed objects are caused by an erroneous feature extraction of the camera or the laser scanner modules. The two other tracking approach perform similar well, which is at the first glance astonishing for the COG tracking. It can be explained by the set of investigated scenarios, in which the object vehicle is always followed by the host vehicle and therefore observed mainly from the rear. In these situations the COG delivers a stable position estimate of the center of the rear. For the tracking of preceding vehicles the computational efficient COG tracking is therefore completely adequate.

The analysis of the velocity estimation error of the stationary roadside objects delivers however not as promising results. The COG approach produces only 79 % consistent estimates. The velocity estimates of the stationary objects deviate often significantly from zero. ACC applications utilize the velocity estimate in order to suppress stationary objects for the target selection. Wrong velocity estimates can therefore lead to a wrong target selection and incorrect behavior. Significantly better are the velocity estimates of the novel cuboid tracking approaches. The estimates are more than twice as accurate and the number of consistent estimates is increased from 79 % to 92 %.

An additional useful performance measure is the number of ghost objects in the driving corridor of the vehicle. While the center of gravity approach produces ghost objects in 44 time instances per hour, the novel cuboid tracking approaches reduce the number to 30 per hour. The ghost objects are mainly caused by dust behind the preceding vehicle. This performance measure is of special interest for safety systems as ghost objects lead potentially to a false engagement of the system. This will be shown in Chapter 10 exemplarily for an emergency brake system.

The benefit of the displacement tracking has been demonstrated in traffic scenarios with trucks. In situations in which the center of gravity tracking produces completely useless velocity estimates, the novel displacement tracking delivers accurate and consistent estimates over the complete set of scenarios. The structure analysis of the displacement tracking identifies reliably ambiguous point cloud patterns, such as the line structure of crash barriers. In those cases only a one-dimensional motion orthogonal to the line is calculated.

Other performance measures of interest would be the detection range or the separability of close objects. Further investigations should evaluate the performance degradation in adverse weather conditions. Rain, fog and snow will most likely impair the performance of both sensors. The impact of challenging light conditions, such as at dust, dawn or at night, should be investigated with respect to the performance of the template tracking.

9 Interacting Multiple Model Filter

The traffic participants relevant for this work are cars, trucks and (motor-)bikes. These objects can be described by their position, velocity, dimensions and orientation. The aim of the state estimation is the optimal estimation of these state variables describing an object recursively over time given noisy or incomplete observations.

The general optimal solution of the recursive estimation problem is given by the recursive Bayesian estimation. The exact probability density function representing the state estimate is calculated recursively over time. The Bayesian recursive estimation is a model based approach. A Markovian state space model is used for the propagation of the state probability density over time. Most approaches use one dynamic model describing the state transition, such as a second order polynomial model, assuming the acceleration to be almost constant over time and only perturbed by random noise. However, in real world applications an object's motion can often be better described by several models. The different models apply at different times, depending on the characteristics of the present motion.

Multiple-model approaches account for different object behaviors. They use a bank of elementary filters based on a set of models that represent possible motion patterns [109, 102, 5]. The development of the cost-effective Interacting Multiple Model (IMM) [14] made the multiple model approaches practical for maneuvering target tracking.

An often stated problem of IMM approaches is the parameterization of the model set. This includes the model selection, the parameterization of the model and the determination of a model transition probability matrix. In all publications in the area of advanced driver assistance systems known to the author, the model set design is performed rather arbitrarily, especially the choice of the transition probability matrix.

In this chapter novel IMM applications to vehicle tracking will be presented. In contrast to previous work, the parameterization of the model set is performed statistically based on real vehicle trajectories.

The chapter begins with a description of different dynamic models. This is followed by a short overview of the IMM algorithm. The model set design is detailed and the performance of two optimized model sets is illustrated in traffic jam situations and intersection scenarios.

9.1 Dynamic Models

The dynamic modeling of the Kalman filter kernels used in the IMM method determines the constraints of the dynamic process. Two different general model groups are considered. The free motion model and the coordinated turn model. In the case of the coordinated turn model the motion is directed into the object's orientation. The wheel slip is neglected. The geometric object model is in both cases described by a point with

an orientation in a two-dimensional coordinate system. These simplifications, with respect to the 3D approach of a cuboid tracking described in the previous chapters, serve a clearer description of the IMM approaches. They help to focus on the interesting part of the dynamic modeling.

The dynamic models are briefly described in the following. Further details can be found in [4, 13].

Free Motion Model

The dynamic state of the system is given by

$$\mathbf{x} = \begin{bmatrix} x & y & \psi & v_x & v_y & a_x & a_y \end{bmatrix}' \tag{9.1}$$

where $[x, y]'$ is the position, ψ the yaw angle, $[v_x, v_y]'$ the velocity vector and $[a_x, a_y]'$ the acceleration vector. In the free motion models the position, velocity and acceleration are decoupled from the orientation.

A model which estimates position, velocity and acceleration is the piecewise constant Wiener process acceleration model. This third order kinematic motion model has a state transition matrix $\mathbf{F}$ as described in [4]

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & T & 0 & \frac{1}{2}T^2 & 0 \\ 0 & 1 & 0 & 0 & T & 0 & \frac{1}{2}T^2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & T & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & T \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{9.2}$$

where T is the sampling period. The noise gain $\mathbf{\Gamma}$ and the covariance of the process noise $\mathbf{Q}_d$ are given by

$$\mathbf{\Gamma} = \begin{bmatrix} \frac{1}{2}T^2 & 0 & 0 \\ 0 & \frac{1}{2}T^2 & 0 \\ 0 & 0 & 1 \\ T & 0 & 0 \\ 0 & T & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad \mathbf{Q}_d = \begin{bmatrix} \sigma_{a_x}^2 & 0 & 0 \\ 0 & \sigma_{a_y}^2 & 0 \\ 0 & 0 & \sigma_{\psi}^2 \end{bmatrix} \tag{9.3}$$

with the covariance matrix $\mathbf{Q}$

$$\mathbf{Q} = \mathbf{\Gamma} \mathbf{Q}_d \mathbf{\Gamma}' \tag{9.4}$$

where $\sigma_{a_x}^2$ and $\sigma_{a_y}^2$ are the variances of the process noise modeling an acceleration increment, σ_{ψ}^2 is the variance of the process noise of the yaw angle.

A second model is used to represent an almost constant velocity driving behavior. A piecewise constant white noise acceleration model is utilized as a second order kinematic motion model [4]. The state vector of the model is represented by

$$\mathbf{x} = \begin{bmatrix} x & y & \psi & v_x & v_y \end{bmatrix}' \tag{9.5}$$

The state transition matrix $\mathbf{F}$ is given by

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & T & 0 \\ 0 & 1 & 0 & 0 & T \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{9.6}$$

The noise gain $\mathbf{\Gamma}$ and the covariance of the process noise $\mathbf{Q}_d$ are given by

$$\mathbf{\Gamma} = \begin{bmatrix} \frac{1}{2}T^2 & 0 & 0 \\ 0 & \frac{1}{2}T^2 & 0 \\ 0 & 0 & 1 \\ T & 0 & 0 \\ 0 & T & 0 \end{bmatrix} \qquad \mathbf{Q}_d = \begin{bmatrix} \sigma_{a_x}^2 & 0 & 0 \\ 0 & \sigma_{a_y}^2 & 0 \\ 0 & 0 & \sigma_\psi^2 \end{bmatrix} \tag{9.7}$$

where $\sigma_{a_x}^2$ and $\sigma_{a_y}^2$ are the variances of the process noise modeling an acceleration increment and σ_ψ^2 is the variance of the process noise of the yaw angle.

A third model represents a stationary behavior. The state is given by

$$\mathbf{x} = \begin{bmatrix} x & y & \psi \end{bmatrix}' \tag{9.8}$$

The state transition matrix $\mathbf{F}$ is

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{9.9}$$

The noise gain $\mathbf{\Gamma}$ and the covariance of the process noise $\mathbf{Q}_d$ are given by

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{Q}_d = \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_\psi^2 \end{bmatrix} \tag{9.10}$$

where σ_x^2 and σ_y^2 are the variances of the process noise modeling a position change, σ_ψ^2 is the variance of the process noise of the yaw angle.

Coordinated Turn Model

Vehicles and (motor-)bikes usually drive in the direction of their orientation. This constraint can be utilized in the dynamic modeling. The dynamic state of the coordinated turn model is given by

$$\mathbf{x} = \begin{bmatrix} x & y & \psi & v & a \end{bmatrix}' \tag{9.11}$$

where $[x, y]'$ denotes the position, ψ the yaw angle, v the velocity, ω the yaw rate and a the acceleration. The orientation of the velocity and acceleration vectors are here defined by the yaw angle.

The differential equation describing a constant acceleration and constant yaw rate is given by

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} v\cos(\psi) \\ v\sin(\psi) \\ \omega \\ a \\ 0 \\ 0 \end{bmatrix} \tag{9.12}$$

The prediction is calculated by integration of the differential equation (9.12)

$$\begin{aligned} \hat{\mathbf{x}}(k|k-1) &= \mathbf{f}[\hat{\mathbf{x}}(k-1)] && (9.13)\\ &= \hat{\mathbf{x}}(k-1) + \int_0^T \dot{\mathbf{x}}(\tau)d\tau && (9.14)\\ &= \hat{\mathbf{x}}(k-1) + \begin{bmatrix} \frac{(v+aT)}{\omega}SW + \frac{a}{\omega^2}CW - \frac{v}{\omega}\sin(\psi) - \frac{a}{\omega^2}\cos(\psi) \\ -\frac{(v+aT)}{\omega}CW + \frac{a}{\omega^2}SW + \frac{v}{\omega}\cos(\psi) - \frac{a}{\omega^2}\sin(\psi) \\ \omega T \\ aT \\ 0 \\ 0 \end{bmatrix} && (9.15) \end{aligned}$$

where

$$SW = \sin(\psi + \omega T), \qquad CW = \cos(\psi + \omega T) \tag{9.16}$$

The process noise covariance matrix $\mathbf{Q}$ is calculated using the direct discrete approach as described in [4]:

$$\mathbf{Q} = \mathbf{\Gamma}\mathbf{D}\mathbf{Q}_d\mathbf{D}'\mathbf{\Gamma}' \tag{9.17}$$

where the noise gain matrix $\mathbf{\Gamma}$ is given by the linearized state transition matrix $\mathbf{F}$. The covariance of the process noise is denoted by $\mathbf{Q}_d$ and $\mathbf{D}$ is a mapping matrix

$$\mathbf{\Gamma} = \mathbf{F} = \left[\nabla_{\mathbf{x}}\mathbf{f}(\mathbf{x})'\right]'\Big|_{\mathbf{x}=\hat{\mathbf{x}}(k-1)} \qquad \mathbf{D} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \mathbf{Q}_d = \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_\omega^2 \end{bmatrix} \tag{9.18}$$

In phases of low estimated yaw rate $\omega \approx 0$ an approximation is chosen in order to account for the singularity in Equation (9.15) for $\omega = 0$

$$\lim_{\omega\to 0}\hat{\mathbf{x}}(k|k-1) = \hat{\mathbf{x}}(k-1) + \begin{bmatrix} (vT + \frac{1}{2}aT^2)\cos(\psi) \\ (vT + \frac{1}{2}aT^2)\sin(\psi) \\ \omega T \\ aT \\ 0 \\ 0 \end{bmatrix} \tag{9.19}$$

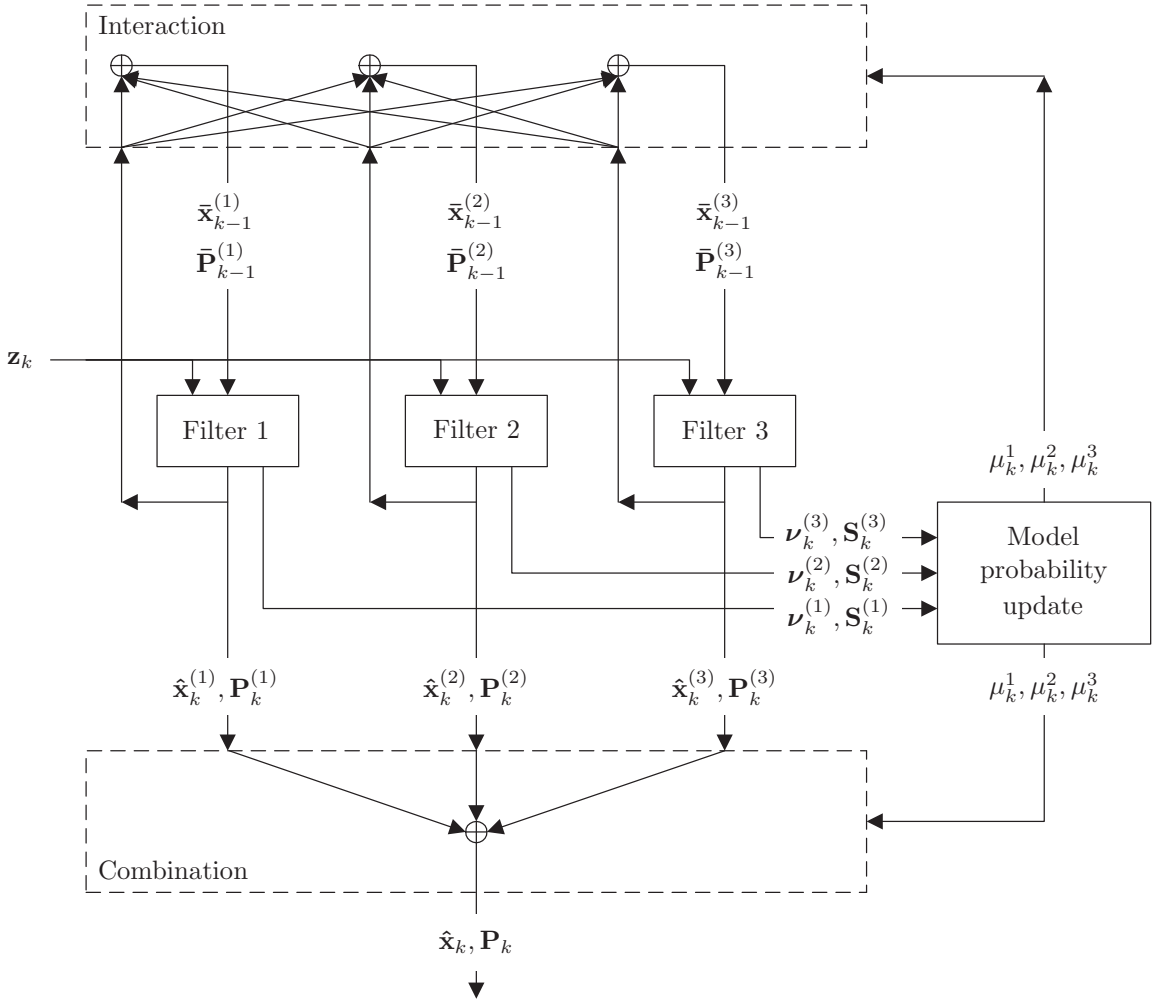


Figure 9.1: Architecture of the IMM estimator with three models.

9.2 The Interacting Multiple Model Algorithm

Multiple model approaches address hybrid systems, where the base state $\mathbf{x}$ varies continuously and s is the system mode, which has discrete values. A continuous-time hybrid system with additive noise is therefore described by the following dynamic and measurement equations

$$\dot{\mathbf{x}}(t) = \mathbf{f}\left[\mathbf{x}(t), s(t), t\right] + \mathbf{v}(t) \tag{9.20}$$

$$\mathbf{z}(t) = \mathbf{h}\left[\mathbf{x}(t), s(t), t\right] + \mathbf{w}(t) \tag{9.21}$$

The base state $\mathbf{x}$ contains as usual the position, velocity, acceleration, orientation and yaw rate. The mode s represent the dynamic behavior pattern, such as stationary, constant velocity or constant acceleration behavior.

There are different multiple model estimation methods described in the literature of which a very cost-efficient method is the interacting multiple model (IMM) [102].

The IMM method is composed of four main parts as shown in Figure 9.1:

- Interaction — the individual filter estimates are mixed with respect to the predicted model probabilities.
- Model specific filtering — each filter predicts and updates its state estimate using its dynamic model assumptions.
- Model probability update — the model probability of each model is updated with respect to the innovation error.
- Combination — for output purposes a combined state estimate is calculated from the state estimates weighted by the model probabilities.

Following [4, 13, 102, 5] the IMM algorithm is summarised below.

9.2.1 Interaction

The interaction between the different modes is modeled as an Markovian switching process. The entries of the Markov matrix π_{ji} are used with the model probability of the previous cycle $\mu_{k-1}^{(j)}$ to calculate the predicted model probability

$$\mu_{k|k-1}^{(i)} = \sum_j \pi_{ji} \mu_{k-1}^{(j)} \tag{9.22}$$

With the conditional model probability $\mu_{k-1}^{(j|i)}$, given the object is in state i that the transition occurred from state j

$$\mu_{k-1}^{(j|i)} = \frac{\pi_{ji} \mu_{k-1}^{(j)}}{\mu_{k|k-1}^{(i)}} \tag{9.23}$$

the mixing of the state estimates $\hat{\mathbf{x}}_{k-1}^{(j)}$ and their covariances $\mathbf{P}_{k-1}^{(j)}$ can be performed

$$\bar{\mathbf{x}}_{k-1}^{(i)} = \sum_j \mu_{k-1}^{(j|i)} \hat{\mathbf{x}}_{k-1}^{(j)} \tag{9.24}$$

$$\bar{\mathbf{P}}_{k-1}^{(i)} = \sum_j \mu_{k-1}^{(j|i)} [\mathbf{P}_{k-1}^{(j)} + (\bar{\mathbf{x}}_{k-1}^{(i)} - \hat{\mathbf{x}}_{k-1}^{(j)})(\bar{\mathbf{x}}_{k-1}^{(i)} - \hat{\mathbf{x}}_{k-1}^{(j)})'] \tag{9.25}$$

Each filter is therefore influenced by the other filters. The degree of influence depends on the estimated model probability $\mu^{(i)}$ and the state transition probabilities π_{ji} from the Markov matrix. The transition probabilities π_{ji} are design parameters of the IMM method. Their choice will be discussed in Chapter 9.3.2.

9.2.2 Model Specific Filtering

Several Kalman filters run in parallel as depicted in Figure 9.1. The predicted state estimates $\hat{\mathbf{x}}_{k|k-1}^{(i)}$ and their covariances $\mathbf{P}_{k|k-1}^{(i)}$ are calculated using the different dynamic models. In each filter i the innovation is performed by calculation of the measurement residual $\boldsymbol{\nu}_k^{(i)}$, the residual covariance $\mathbf{S}_k^{(i)}$, the updated state estimate $\hat{\mathbf{x}}_k^{(i)}$ and its covariance $\mathbf{P}_k^{(i)}$. Each filter therefore calculates its own estimates depending on the filter specific dynamic model.

9.2.3 Model Probability Update

The model probability is calculated in each cycle. It depends mainly on the likelihood for the observation. The likelihood is calculated from the residual $\boldsymbol{\nu}_k^{(i)}$ and the residual covariance $\mathbf{S}_k^{(i)}$assuming Gaussian statistics

$$L_k^{(i)} = \frac{\exp\left[-(1/2)(\boldsymbol{\nu}_k^{(i)})'(\mathbf{S}_k^{(i)})^{-1}\boldsymbol{\nu}_k^{(i)}\right]}{[(2\pi)^m|\mathbf{S}_k^{(i)}|]^{1/2}} \tag{9.26}$$

The dimension of the measurement vector is here denoted by m. The term in the exponent $\boldsymbol{\nu}'(\mathbf{S})^{-1}\boldsymbol{\nu}$ is the normalized innovation squared (NIS) from Chapter 8.3.2. It is a measure for the consistency of the filter.

Finally, the new model probabilities $\mu_k^{(i)}$ are calculated

$$\mu_k^{(i)} = \frac{\mu_{k|k-1}^{(i)} L_k^{(i)}}{\sum_j \mu_{k|k-1}^{(j)} L_k^{(j)}} \tag{9.27}$$

9.2.4 Combination

For output purposes only a combined state estimate is determined. The overall state estimate $\hat{\mathbf{x}}_k$ and its covariance $\mathbf{P}_k$ are a weighted combination of the model state estimates $\hat{\mathbf{x}}_k^{(i)}$ and their covariances $\mathbf{P}_k^{(i)}$

$$\hat{\mathbf{x}}_k = \textstyle\sum_i \mu_k^{(i)} \hat{\mathbf{x}}_k^{(i)} \tag{9.28}$$

$$\mathbf{P}_k = \textstyle\sum_i \mu_k^{(i)} [\mathbf{P}_k^{(i)} + (\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_k^{(i)})(\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_k^{(i)})'] \tag{9.29}$$

9.3 Model Set Design

The description of the IMM in the previous section revealed several parameters, which influence the behavior of the filter. First, the dynamic models of the different filter kernels and their process noise covariance must be chosen. The number of models is theoretically not restricted. In the literature usually two models are used, sometimes three. Too many models however deteriorate the performance [102]. Secondly, the Markov matrix, which models the model state transition probability, must be chosen.

The search for the best parameter set is basically an optimization process. Representative traffic scenarios must be defined and an optimization criterion determined. The optimization criterion is usually a function of the estimated state and the true state. Therefore, ground truth data must exist.

The best optimization is obtained with real sensor data acquired from representative scenarios. The vehicle equipped with the sensors is then driving on typical trajectories and is surrounded by other objects which follow representative motion patterns. An optimization of the state estimation of the tracked objects however requires the true object states of involved vehicles to be known. Precise ground truth measurement equipment, which is synchronized between the different traffic participants, is necessary.

As this kind of sophisticated measurement equipment is difficult to obtain, if not impossible for some applications, many researchers switch to purely synthetic data. Trajectories and the resulting sensor data are then produced based on mathematical models. For an optimization of the dynamic modeling of a state estimation algorithm, this approach is not advisable, as realistic trajectories of vehicles are very hard to produce synthetically.

Trajectories from real test drives should be utilized instead. Thus, a hybrid approach is suggested in this work. Realistic trajectories of vehicles can be obtained from the wheel speed, yaw rate and steering angle sensors build into modern cars as part of the ESP system. The trajectories are calculated by integration of the velocity and the yaw rate. Even if the position and the yaw angle might be impaired by a low frequency drift, the high and medium frequency pattern is realistic, which is most important for the comparison of state estimation algorithms based on different dynamic modeling. The sensor data is synthetically generated using a sensor model based on these realistic trajectories. The measurement uncertainty is represented by zero mean Gaussian noise which is not completely realistic. The approach is legitimate as the aim is the optimization of the model set for the IMM and not the optimization of the sensor model. It is hypothesized that the characteristics of the sensor model have minor influence on the optimization result. A realistic trajectory is however important.

For the optimization Monte Carlo simulations are performed. Different optimization criteria detailed in Chapter 8.3 are adapted for Monte Carlo simulations in the following.

The first and second order statistics are calculated from different simulation runs for the same time instance

$$\mu_{\tilde{\mathbf{x}}}(k) = \frac{1}{N}\sum_{i=1}^{N}\tilde{\mathbf{x}}^i(k) \tag{9.30}$$

$$s_{\tilde{\mathbf{x}}}(k) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(\tilde{\mathbf{x}}^i(k)-\mu_{\tilde{\mathbf{x}}}(k)\right)^2} \tag{9.31}$$

$$RMS(k) = \sqrt{\mu_{\tilde{\mathbf{x}}}(k)^2+s_{\tilde{\mathbf{x}}}(k)^2} \tag{9.32}$$

where N is the number of Monte Carlo runs. A sample mean deviating from zero exhibit a bias in the state estimate for the given interval k. The sample standard deviation $s_{\tilde{\mathbf{x}}}(k)$ related to the measurement noise covariance results in the noise reduction ability of the filter. Given a sequence of state estimates the **mean RMS** and the **peak RMS** provide useful measures

$$meanRMS = \frac{1}{K}\sum_{k=1}^{K}RMS(k) \tag{9.33}$$

$$peakRMS = \max_{k} RMS(k) \tag{9.34}$$

While the mean RMS represents the average performance of the filter, the peak RMS identifies the time instance with the worst performance, which is usually during a maneuver change when the process model assumptions are strongly violated.

An optimization over the complete parameter space would demand for an enormous computational power. Therefore the optimization problem is split into different sub-

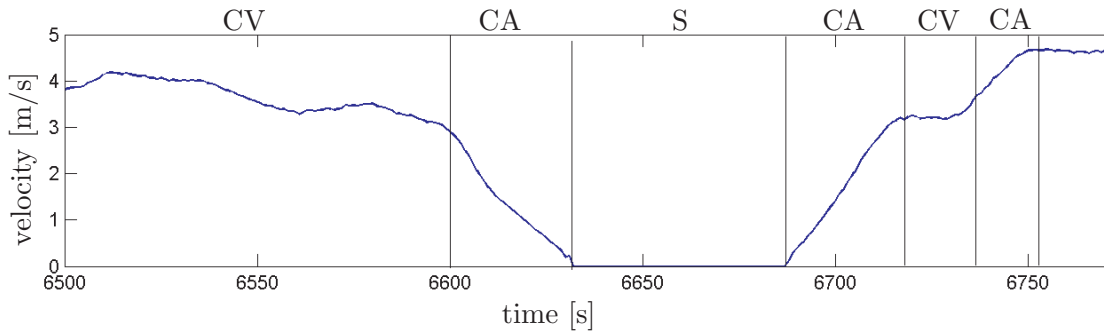


Figure 9.2: Analysis of a traffic jam situation. The sequence is divided into sections of stationary (S), constant velocity (CV) and constant acceleration (CA) dynamics.

problems. The choice of the dynamic models is separated from the choice of the Markov matrix.

9.3.1 Choice of Dynamic Models

The aim of the model set design is the determination of a model set which represents all different motion patterns performed by cars. The optimization was focused on traffic jam scenarios which cover a wide range of driving situations. A hybrid approach for the generation of ground truth data was chosen as described above. Test drives in traffic jams were conducted, accumulating more than an hour of sensor data.

Different model sets were created from the dynamic models described in Chapter 9.1. Two groups of model sets exhibited promising results. One group of model sets contains a stationary model (S), a constant velocity model (CV) and a constant acceleration model (CA). The other model set group is composed of two coordinated turn models (CT) with differing process noise levels.

Each model set group was optimized by variation of the process noise choice. As optimization criteria the mean RMS and the peak RMS were utilized. 500 Monte Carlo runs were performed for each model set choice. The resulting model sets are described in Chapter 9.4.1 and 9.5.1.

9.3.2 Choice of Markov Matrix

For the parameterization of the Markov matrix real data from traffic jam situations was analyzed. The traffic jam sequence is divided into sections of almost constant dynamic behavior. The classes of dynamic behavior are determined from the model set under investigation. For the first model set group the sequence is for instance divided into sections of stationary (S), constant velocity (CV) and constant acceleration (CA) dynamics as shown in Figure 9.2. In order to prevent a too frequent switching between the states, slight deviations from the dynamic constraints are allowed. The degree of deviation is determined by the choice of the model noise. From the number of transitions from one mode into another and the length of the sections a statistic is created which is used for the parameterization of the Markov matrix.

The diagonal elements of the transition matrix represent the probability that the object remains in mode i. Following [13] they are related to the mean sojourn time $E[\tau_i]$ and the sampling interval T

$$\pi_{ii} = 1 - \frac{T}{E[\tau_i]} \tag{9.35}$$

The mean sojourn time is the average length of a section of constant dynamic behavior in seconds

$$E[\tau_i] = \frac{1}{N_i} \sum_j \tau_{i,j} \tag{9.36}$$

where N_i is the number of sections assigned to mode i and $\tau_{i,j}$ the length of the section j of mode i.

The off-diagonal elements are derived from the statistics of the mode transitions of the traffic jam scenario. If n_{ji} is the counted number of transitions from mode j to mode i and n_j is the sum of all transitions from mode j, the transition probabilities are given by

$$\pi_{ji} = \frac{n_{ji}}{n_j}(1 - \pi_{ii}) \tag{9.37}$$

9.4 Evaluation of Traffic Jam Situations

9.4.1 Model Set

For highway scenarios the model set group with the free motion models performed in experiments similarity well as the group with the coordinated turn models. For computational reasons the linear free motion models were preferred to the nonlinear coordinated turn model. The optimization of the model set resulted in the parameters which will be described in the following.

The first model is a stationary model where the standard deviation of the noise is

$$\sigma_x^* = \sigma_y^* = 0.32 \ \frac{\mathrm{m}}{\mathrm{s}} \tag{9.38}$$

The process noise is presented with respect to an interval T of one second. The variances are adapted to the varying measurement interval as described in [6, Chapter 8.7.4]

$$\sigma^2(k) = [\sigma^* T(k)]^2 \tag{9.39}$$

The noise in the yaw angle is given by

$$\sigma_\psi^* = 0.04 \ \frac{\mathrm{rad}}{\mathrm{s}} \tag{9.40}$$

The constant velocity model is parametrized with

$$\sigma_{a_x}^* = \sigma_{a_y}^* = 0.89 \ \frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.41}$$

$$\sigma_\psi^* = 0.12 \ \frac{\mathrm{rad}}{\mathrm{s}} \tag{9.42}$$

The model of constant acceleration has a high model noise with standard deviations of

$$\sigma_{a_x}^* = 8 \frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.43}$$

$$\sigma_{a_y}^* = 2 \frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.44}$$

$$\sigma_{\psi}^* = 0.25 \frac{\mathrm{rad}}{\mathrm{s}} \tag{9.45}$$

This model is designed to fit phases of constant acceleration and acceleration changes, as well as lane changes.

The Markov matrix $\mathbf{\Pi}(T)$ is determined from the statistics of traffic jam situations:

$$\mathbf{\Pi}(0.1\ sec) = \begin{bmatrix} 0.980 & 0.000 & 0.020 \\ 0.000 & 0.970 & 0.030 \\ 0.003 & 0.017 & 0.980 \end{bmatrix} \tag{9.46}$$

The transition probabilities are adapted in each cycle with respect to the current sampling interval as described in [13]. The initial model probabilities are chosen as

$$\boldsymbol{\mu} = \begin{bmatrix} 0.333 & 0.333 & 0.334 \end{bmatrix}' \tag{9.47}$$

The IMM based on this model set is compared to a single constant velocity Kalman filter (SKF) with

$$\sigma_{a_x}^* = 2 \frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.48}$$

$$\sigma_{a_y}^* = 2 \frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.49}$$

$$\sigma_{\psi}^* = 0.2 \frac{\mathrm{rad}}{\mathrm{s}} \tag{9.50}$$

9.4.2 Experimental Setup

The IMM algorithm is integrated into the centralized event-triggered sensor fusion architecture described in [143]. The experimental setup for the evaluation of different IMM implementations is composed of two test vehicles [84, 85].

The host vehicle is equipped with an ESP system, a GPS receiver and several forward looking sensors. Two infrared laser scanners are mounted at the left and right side of the front bumper, thus covering the region left, right and in front of the host vehicle. The object recognition range is up to 60 m. The preprocessed measurements which are delivered to the fusion system are a three point contour description of the detected objects. A 77 GHz long range radar measures the relative position and radial velocity of objects. The field of view is limited to 6° azimuth angle and 120 m distance. Additionally a monocular vision system, mounted behind the frontal windshield, detects vehicles and estimates the lane position and curvature. The horizontal aperture is 22° and the detection range up to 70 m.

The target vehicle records its ESP data. The target vehicle and the host vehicle are both equipped with GPS receivers in order to synchronize the sensor measurements with respect to a global time frame.

The performance is evaluated using re-enacted driving maneuvers.

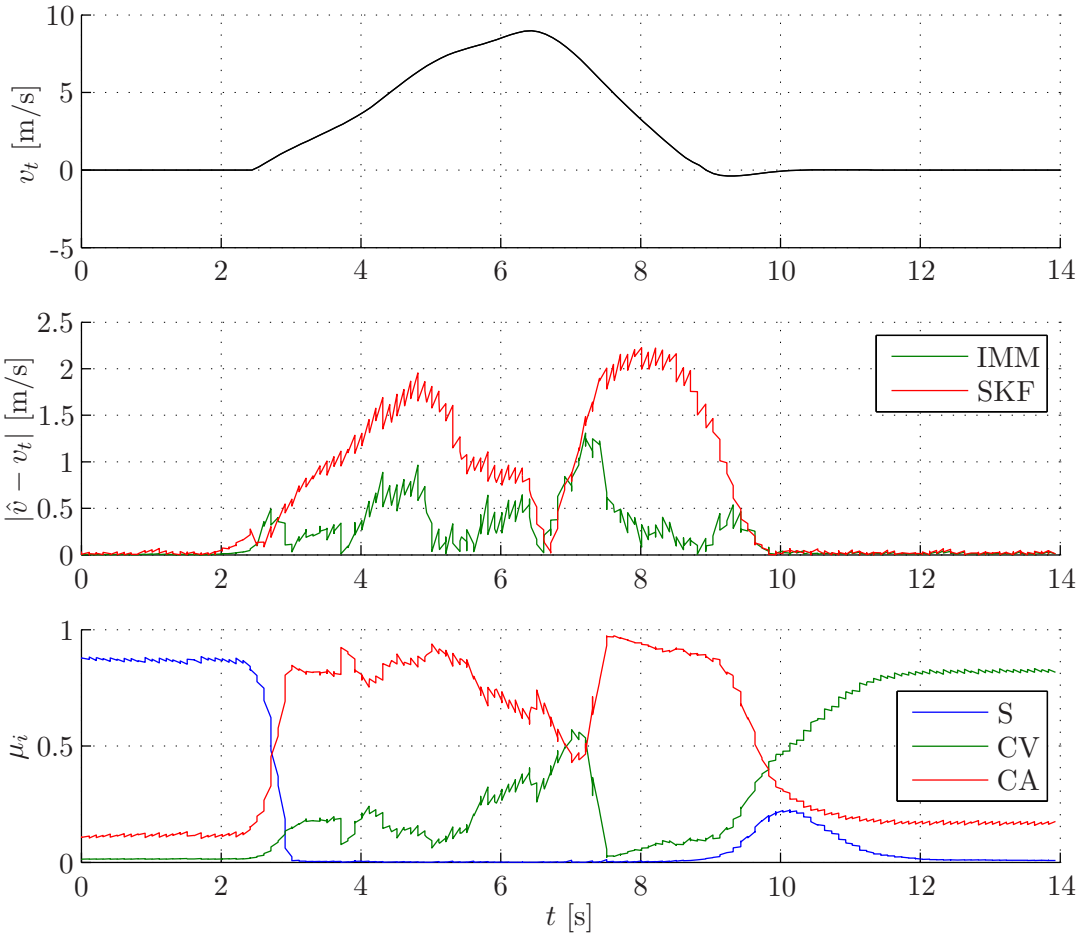


Figure 9.3: Situation of strong acceleration changes. Velocity v_t recorded in target vehicle. Errors in estimated velocity $\hat{v}$ of a single Kalman filter (SKF) and the IMM method. The model probabilities μ_i are shown at the bottom. S denotes the stationary model, CV the constant velocity and CA the almost constant acceleration model.

9.4.3 Results for Strong Acceleration Changes

Potentially very harmful situations occur at the end of traffic jams. Cars driving at high speed suddenly start breaking or even worse, have to brake from an acceleration phase.

In Figure 9.3 such a driving maneuver is shown. The true velocity v_t recorded from the target vehicle exhibits the strong acceleration and then strong deceleration behavior. As the velocity is slightly filtered, an overshooting can be observed subsequent to the deceleration phase. The error $|\hat{v} - v_t|$ in the estimated velocity $\hat{v}$ of the target vehicle observed from the host vehicle is shown below. As expected, the estimation error of a single constant velocity Kalman filter is high at acceleration phases as the model assumption does not meet the dynamics of the situation. The IMM method switches from the stationary to the constant acceleration model within few sampling intervals. Once the acceleration is finished the model probability for the model of constant velocity increases. However, as soon as the deceleration process starts the constant acceleration model again gains the highest probability. Thus, the IMM approach reduces the estimation error significantly

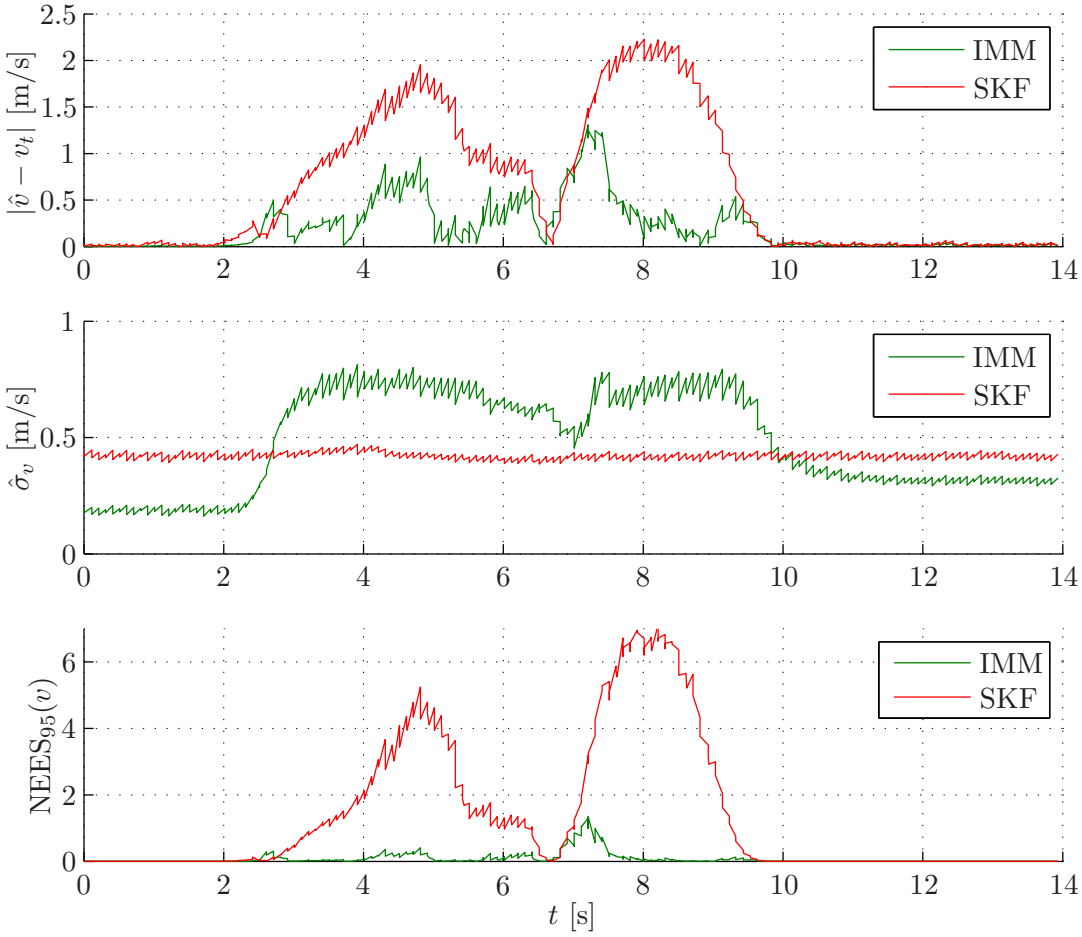


Figure 9.4: Situation of strong acceleration changes. Errors in estimated velocity $\hat{v}$ of a single Kalman filter (SKF) and of IMM. Estimated standard deviations of velocity $\hat{\sigma}_v$. NEES of velocity normalized to its 95 % probability region.

in acceleration phases compared to the single Kalman filter implementation.

Figure 9.4 shows the estimated standard deviation of the velocity $\hat{\sigma}_v$. The uncertainties in the estimation are almost constant for the Kalman filter of constant velocity. The estimated standard deviation of the IMM method however exhibits a strong variation. In phases of no motion, the uncertainties are quite low and as soon as the acceleration begins the standard deviation increases. This situation adaptive uncertainty in the estimation process is an important feature of the IMM approach as in tracking systems the measurement association gate depends on the estimated standard deviation. If the uncertainty is estimated too low, the object might be lost. However, using an IMM approach, the estimated standard deviation can be adapted to the driving situation. In this case, the standard deviation raises in phases of high dynamics. Therefore the association gate is enlarged and the object is more likely to be continuously tracked.

The normalized estimation error squared (NEES) expresses the filter consistency. Figure 9.4 shows the NEES of the velocity estimate normalized to a 95 % probability region, assuming a χ^2 distribution. The normalized NEES of the single Kalman filter is most of

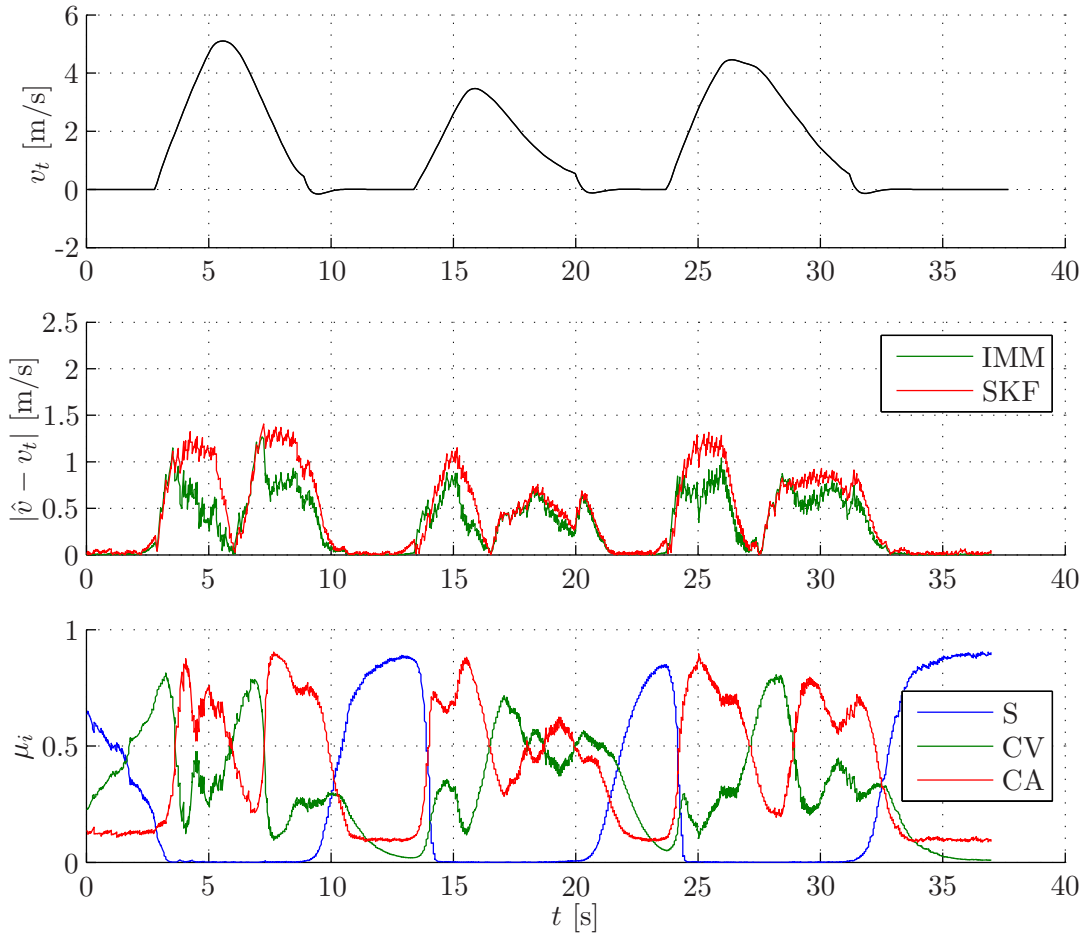


Figure 9.5: Stop and go situation. Velocity v_t recorded in target vehicle. Errors in estimated velocity $\hat{v}$ of a single Kalman filter (SKF) and the IMM method. The model probabilities μ_i are shown at the bottom. S denotes the stationary model, CV the constant velocity and CA the almost constant acceleration model.

the time above 1.0, i.e. the estimated filter uncertainty is not consistent with the true estimation error considering the 95 % probability region. Conversely, the estimation of the IMM method is consistent during almost the whole maneuver. The consistency is important for safety applications which consider the estimated uncertainty for their collision risk calculation. If the uncertainties are underestimated a safety system might falsely engage.

9.4.4 Results for Traffic Jam Situations

Traffic jam situations are characterized by cars driving at low speed, accelerating, decelerating and stopping. These different driving maneuvers can change in rapid succession. In Figure 9.5 such a driving maneuver is shown. The true velocity v_t recorded by the target vehicle exhibits first a standstill and an acceleration directly followed by a deceleration to a halt. This pattern is repeated twice.

With its four sensors the host vehicle measures position and velocity of the target

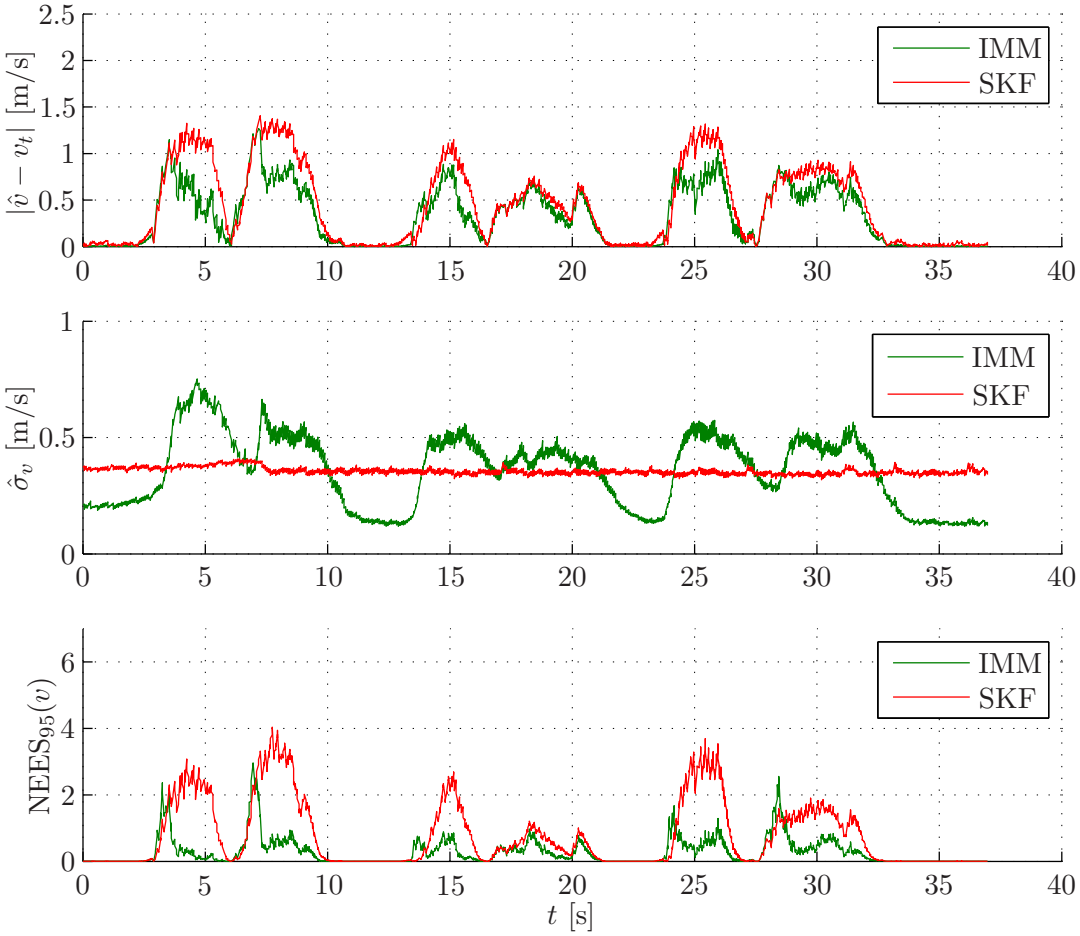


Figure 9.6: Stop and go situation. Errors in estimated velocity $\hat{v}$ of a single Kalman filter (SKF) and of IMM. Estimated standard deviations of velocity $\hat{\sigma}_v$. NEES of velocity normalized to its 95 % probability region.

vehicle. The error $|\hat{v} - v_t|$ in the estimated velocity $\hat{v}$ is shown below. The error of the single constant velocity Kalman filter (SKF) is high during acceleration and deceleration phases. The IMM implementation can however reduce the error in the acceleration phase after a short period of adaption. The model probabilities $\mu^{(i)}$ show that the IMM method switches at the beginning of acceleration phases from the stationary model to the constant acceleration model. As the acceleration is only about 2 $\frac{m}{s^2}$, the probability of the constant velocity model deviates from zero. This has the effect that the error of the IMM method compared to the single Kalman filter is lower, but not as significantly as expected. Using four models, with an additional model for constant acceleration with very low process noise did not improve the results, as the switching across the different models takes too long when compared to the duration of the maneuver.

Figure 9.6 shows the estimated standard deviation of the velocity $\hat{\sigma}_v$. The uncertainties in estimation are almost constant for the Kalman filter of constant velocity. The estimated standard deviation of the IMM method however exhibits a strong variation. In phases of standstill, the IMM estimated standard deviation is significantly lower compared to that

of the single Kalman filter. This enables a noise reduction of more than 50 % compared to the single Kalman filter. As soon as the acceleration or deceleration begins the standard deviation of the IMM method increases again.

The normalized NEES of the single Kalman filter is in phases of acceleration most of the time above 1.0, i.e. the estimated filter uncertainty is not consistent with the true estimation error considering the 95 % probability region. In contrast, the IMM estimation is consistent during almost the whole maneuver. Exceptions are at the beginning of the acceleration when the model probability of the acceleration model is still increasing. This is due to the maneuver detection delay of the IMM approach.

9.5 Evaluation of Intersection Scenarios

The performance of IMM using three models with stationary, constant velocity and constant acceleration Kalman filter kernels discussed in Chapter 9.4 proved to be superior to a single constant velocity Kalman filter. There are however some disadvantages in the choice of this model set.

Experiments exhibited that filter kernels with different polynomial orders (e.g. a model set containing constant acceleration and constant velocity models) result in an underestimation of higher order parameters (the acceleration) of the combined state vector in phases in which the dynamic situation is not exactly represented by one model. This is almost always true in traffic situations and in particular in traffic jams. It is therefore advisable that the model set should contain only models of the same order, if higher order parameters are of interest for instance for a controller. As an example, consider an IMM with a constant velocity and a constant acceleration filter kernel. If the observed object drives at constant speed with short weak acceleration intervals, the model probability of the constant acceleration model is usually low. Therefore, the mixed estimate of the acceleration parameter is always an underestimation of the true acceleration, even if the constant acceleration model has low estimation errors.

The above presented examples only show the performance in the longitudinal direction. The performance of the lateral estimation could not be determined based on real measurements, due to the lack of a full reference trajectory. In simulations it can be shown that the lateral estimate of the investigated model set performs well in traffic jams on highways.

However, in urban areas, cars may turn abruptly at crossings. This dynamic situation can be better represented by the above mentioned coordinated turn models. Based on these criteria a new model set was designed, based on the coordinated turn model.

9.5.1 Model Set

Three coordinated turn models are chosen, which differ only in the parameterization of the process noise.

The first model (F1) has a process noise with standard deviations of

$$\sigma^*_{a,1} = 2\ \frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.51}$$

$$\sigma^*_{\omega,1} = 0.6\ \frac{\mathrm{rad}}{\mathrm{s}^2} \tag{9.52}$$

This model is designed to fit phases of constant acceleration and constant turning.

The second model (F2) has a high process noise in the yaw rate with

$$\sigma^*_{a,2} = 2\ \frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.53}$$

$$\sigma^*_{\omega,2} = 3\ \frac{\mathrm{rad}}{\mathrm{s}^2} \tag{9.54}$$

which represents rapid transitions from straight driving to turning at crossings.

The third model (F3) of has a high process noise in the acceleration component

$$\sigma^*_{a,3} = 10\ \frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.55}$$

$$\sigma^*_{\omega,3} = 0.6\ \frac{\mathrm{rad}}{\mathrm{s}^2} \tag{9.56}$$

This model is developed for strong acceleration changes.

The process noise is here presented with respect to an interval T of one second. The variances are adapted to the varying measurement interval as described in [6, Chapter 8.7.4]

$$\sigma(k)^2 = [\sigma^* T(k)]^2 \tag{9.57}$$

The Markov matrix $\mathbf{\Pi}(T)$ is chosen as:

$$\mathbf{\Pi}(0.1\ sec) = \begin{bmatrix} 0.98 & 0.01 & 0.01 \\ 0.04 & 0.95 & 0.01 \\ 0.04 & 0.01 & 0.95 \end{bmatrix} \tag{9.58}$$

The initial model probabilities are chosen as $\boldsymbol{\mu} = \begin{bmatrix} 0.33 & 0.33 & 0.34 \end{bmatrix}'$.

This model set is compared to a single Kalman filter (SKF) which is parameterized equally to the first model of the IMM model set, with

$$\sigma^*_{a,\mathrm{SKF}} = 2\ \frac{\mathrm{m}}{\mathrm{s}^3} \tag{9.59}$$

$$\sigma^*_{\omega,\mathrm{SKF}} = 0.6\ \frac{\mathrm{rad}}{\mathrm{s}^2} \tag{9.60}$$

This model is designed to deliver smooth estimates and cope with normal driving phases, smooth acceleration changes, lane changes and turning at intersections.

9.5.2 Experimental Setup

A hybrid approach was chosen for the evaluation of the performance at intersections. The trajectories for the simulation were derived from the data acquired by the ESP of a test

vehicle. The resulting trajectories are the natural paths of vehicles driving in typical scenarios.

The measurement vector is composed of the position $[x, y]'$ and the yaw angle ψ. Zero mean, Gaussian noise was added to the data with $\sigma_x = \sigma_y = 0.12$ m and $\sigma_\psi = 0.07\,\text{rad} = 4°$, thus simulating the measurements of a single laser scanner at 10 Hz scan frequency. The measurement interval is 100 ms. The measurement noise covariance $\mathbf{R}$ for the Kalman filter estimation process has slightly increased variances in order to compensate for the measurement noise of the noisy ground truth data. A diagonal matrix is chosen with $\sigma_x = \sigma_y = 0.15$ m and $\sigma_\psi = 0.087\,\text{rad} = 5°$.

A single Kalman filter is compared to an IMM implementation. For each configuration 20 Monte Carlo runs are performed, each with a different measurement noise sequence. The results below show state estimates and performance measures averaged over these 20 Monte Carlo runs.

9.5.3 Results for Strong Acceleration Changes

Figure 9.7 shows the estimated velocity v and acceleration a of a vehicle which accelerates and then directly brakes. The estimates are averaged over 20 Monte Carlo runs. The gear change from first to second gear can be noted around $t = 8$ seconds. The single Kalman filter (SKF) exhibits a low RMS error in the velocity of 0.2 m/s in almost constant acceleration phases. The RMS error increases rather drastically to 2 m/s at the transition from the acceleration to the strong deceleration and when the vehicle comes finally to a standstill. The RMS error of the IMM is equally good in normal driving phases. However, in phases of strong acceleration changes, the RMS error can be significantly reduced compared to the SKF approach. This is caused by an increased probability of the third filter (F3) with a higher process noise in the acceleration, representing phases of strong acceleration changes.

The effects of the increased probability of the third filter (F3) is reflected in the estimated uncertainty of the velocity $\hat{\sigma}_v$. The uncertainty raises in phases of acceleration change. In contrast, the single Kalman filter exhibits a constant uncertainty after an initial period.

This situation adaptive covariance estimation by the IMM leads to an increased consistency of the filter as shown by the cumulative probability of the averaged NEES $Q_{\chi^2}(\bar{\epsilon}, 80)$. The NEES is calculated from the position, orientation and velocity, averaged over 20 Monte Carlo runs. Both approaches exhibit inconsistent estimates in phases of acceleration change. However, the IMM achieves consistent estimates after a short period of adaption. In contrast, the SKF delivers inconsistent estimates over long time periods especially towards the end of the scenario. The better consistency of the IMM results from the model switching to an appropriate model, which in consequence increases the estimated uncertainty and improves the accuracy of the state estimates.

The situation adaptive estimation of the IMM enabled by the model switching can be observed in detail in Figure 9.8, which shows the transition from the acceleration (2 m/s^2) to the strong deceleration (-7 m/s^2) phase.

The first time the IMM delivers an inconsistent estimate is at 13.3 seconds. The first subsequent consistent estimate is at 13.8 seconds. The IMM therefore needs 0.5 seconds to adapt the model probabilities. However, the benefit of the continuous change in the

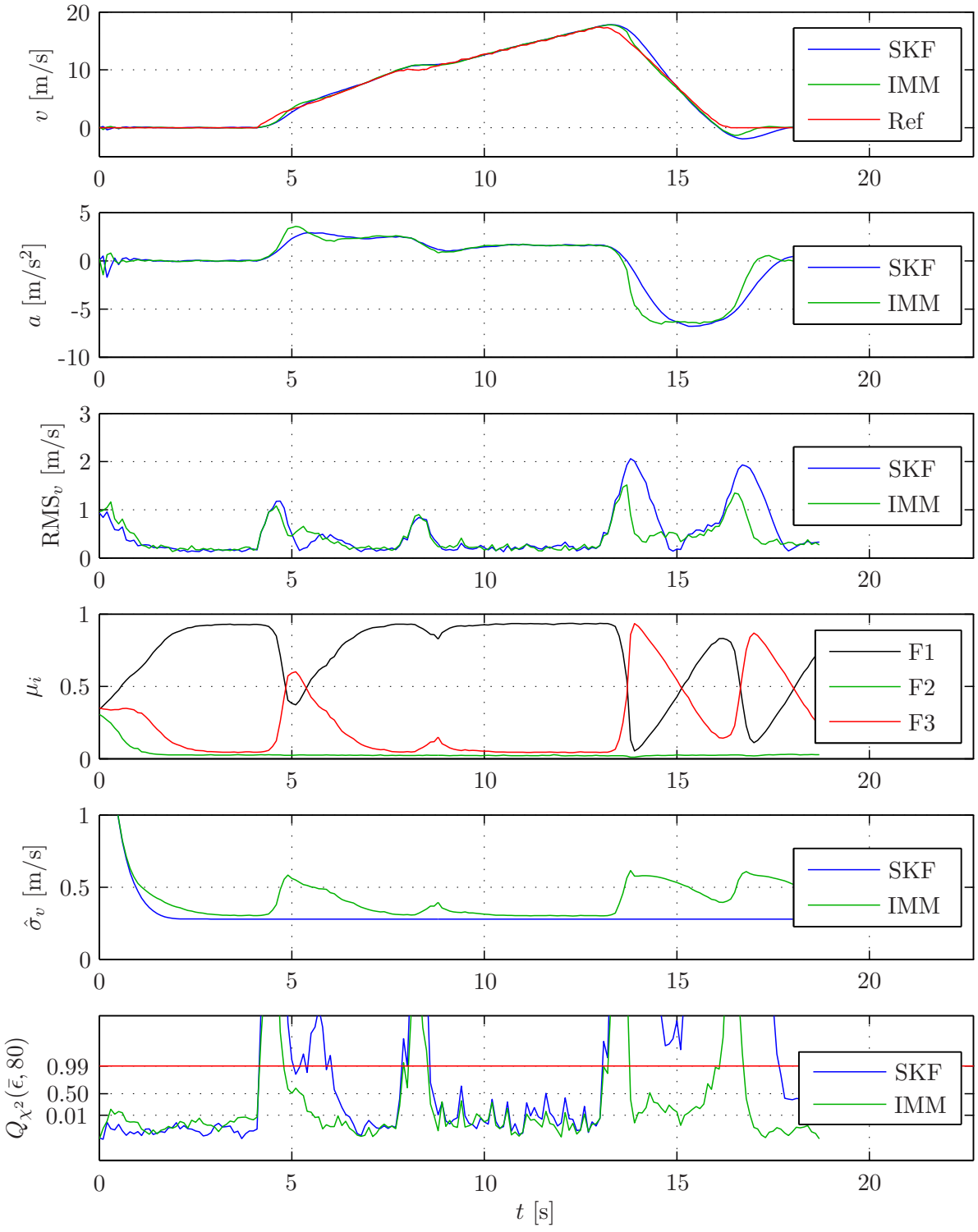


Figure 9.7: Acceleration followed by a strong deceleration. Comparison between single Kalman filter (SKF) and IMM. Averaged velocity v and acceleration estimates a, reference velocity (Ref), RMS of the velocity, model probabilities μ_i, estimated standard deviation of the velocity $\hat{\sigma}_v$ and cumulative probability of the averaged NEES $Q_{\chi^2}(\bar{\epsilon}, 80)$ which should stay below 99 %.

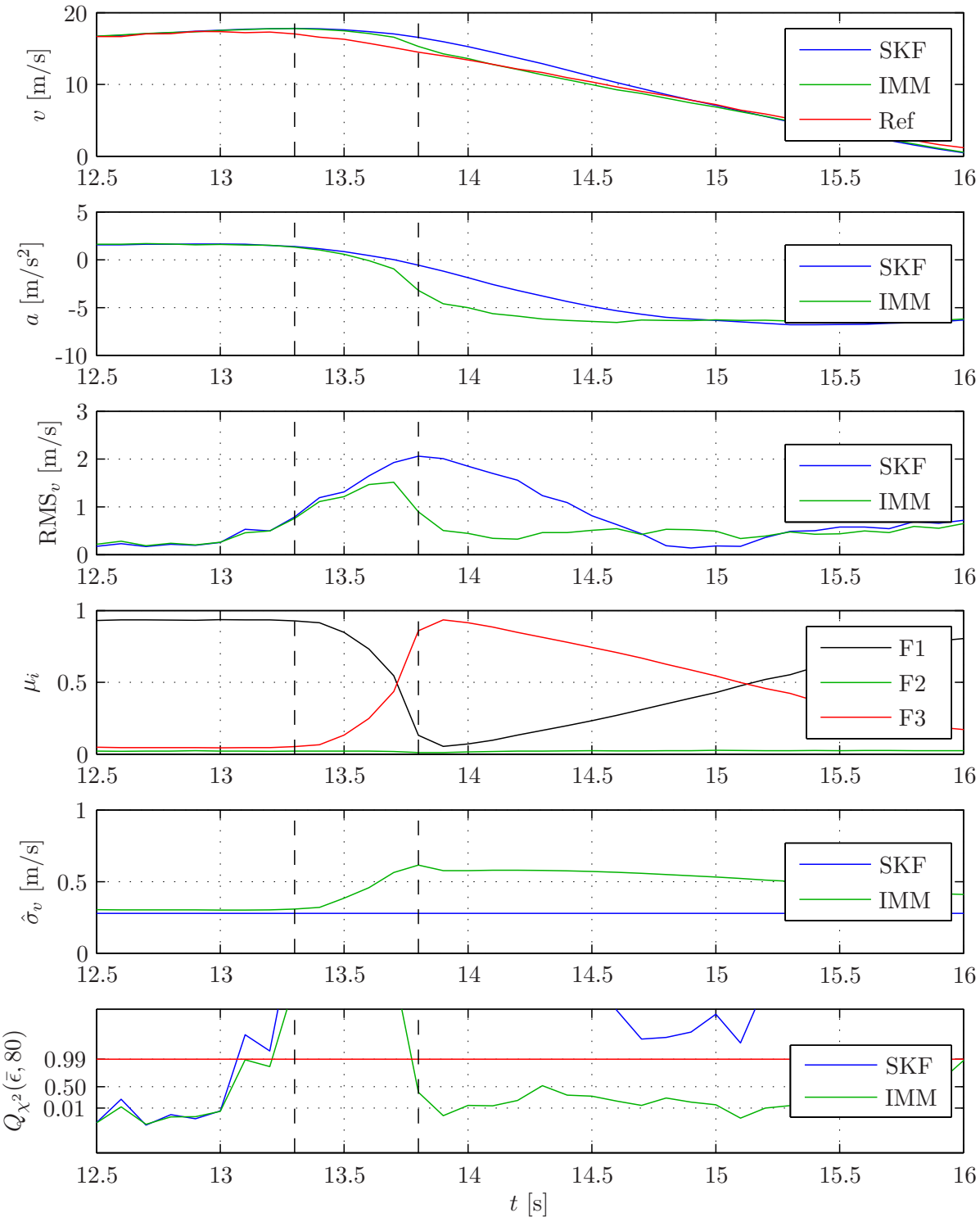


Figure 9.8: Transition from acceleration (2 m/s^2) to strong deceleration (-7 m/s^2). Comparison between single Kalman filter (SKF) and IMM. Averaged velocity v and acceleration estimates a, reference velocity (Ref), RMS of the velocity, model probabilities μ_i, estimated standard deviation of the velocity $\hat{\sigma}_v$ and cumulative probability of the averaged NEES $Q_{\chi^2}(\bar{\epsilon}, 80)$ which should stay below 99 %. Time intervals of first inconsistent estimate of the IMM at 13.3 s and first subsequent consistent estimate at 13.8 s are indicated.

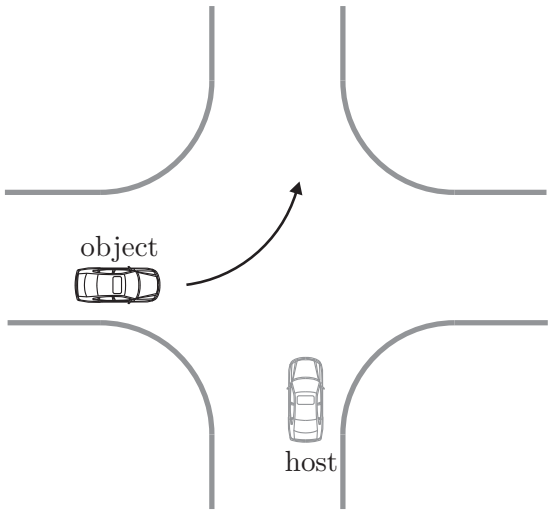


Figure 9.9: Left turn situation at an intersection.

model probabilities becomes apparent already earlier. The adaption of the acceleration is faster than in the SKF and the RMS of the velocity is reduced before the time interval of 13.8 seconds.

The improved acceleration estimation is especially noticeable. The single Kalman filter delivers an estimate of -5 m/s^2 at 14.5 seconds. The IMM produces the estimate of -5 m/s^2 at 14 seconds, thus reducing the latency by about 0.5 seconds.

It should however be mentioned that the increased model probability of the third filter and therefore increased estimated uncertainty of the longitudinal dynamics, leads to an increased noise level in the longitudinal estimates. The IMM needs approximately 3–4 seconds until the probability of the first filter (F1), representing normal driving phases, is restored. In the intermediate time the probability and therefore the estimated uncertainty are however continuously decreasing.

9.5.4 Results for Turning Situations

The results for situations of strong acceleration changes exhibit a promising performance of IMM compared to the single Kalman filter. In normal driving phases, the first filter kernel of the IMM has the highest probability. When the acceleration changes, the third model of increased process noise in the longitudinal object dynamics gains in probability. The second IMM filter kernel represents situations of rapid steering maneuvers. These situations will be presented in the next section.

The present section demonstrates the performance in normal turning situations at intersections as depicted in Figure 9.9 and is dedicated to the analysis of the coordinated turn model. Figure 9.10 shows the estimated yaw angle ψ and yaw rate ω averaged over 20 Monte Carlo simulations in a situation in which the observed car turns to the left at an intersection. IMM and single Kalman filter perform similarly well. The RMS error is of about 0.03 rad over the whole sequence, which is less than half the measurement noise, thus exhibiting a good yaw angle noise reduction. The error in the yaw angle does not

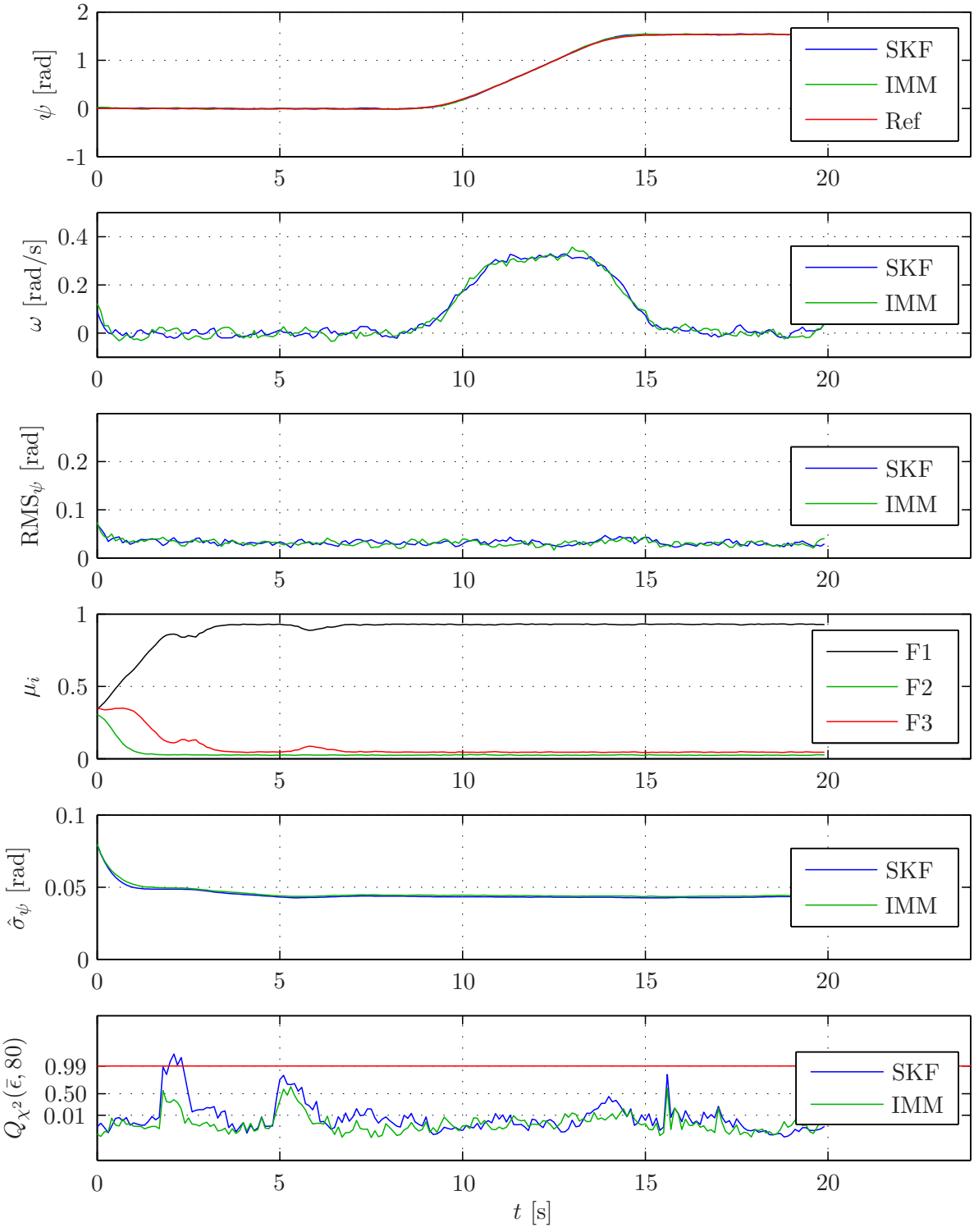


Figure 9.10: Left turn at an intersection. Comparison between single Kalman filter (SKF) and IMM. Averaged yaw angle ψ and yaw rate estimates ω, reference yaw angle (Ref), RMS of the yaw angle, model probabilities μ_i, estimated standard deviation of the yaw angle $\hat{\sigma}_\psi$ and cumulative probability of the averaged NEES $Q_{\chi^2}(\bar{\epsilon}, 80)$ which should stay below 99 %.

significantly increase in the turning phase or at the maneuver transition. Both approaches produce consistent estimates. Only at the beginning, the NEES of the SKF is outside the one-sided 99 % probability region. This can be explained by the acceleration of the observed vehicle (not shown here). The IMM reacts with an increased probability of the third filter and keeps the estimates consistent.

9.5.5 Results for Rapid Steering Maneuvers

The last experiment analyzes rapid steering maneuvers. Figure 9.11 shows the estimated yaw angle ψ and yaw rate ω averaged over 20 Monte Carlo runs for a car performing a rapid right followed by a rapid left turn. The estimated yaw rates exhibit the superiority of the IMM compared to the single Kalman filter (SKF) approach. The IMM yaw rate estimates are delivered up to 0.3 seconds earlier than the SKF estimates. This reduced latency has its effect on the error in the yaw angle. While the RMS error of the SKF is in a time period of about 1.5 seconds above 0.1 rad (5.7°) in phases of rapid steering, the IMM reduces the length of this period to about 0.3 seconds. This significant reduction of the estimation errors by the IMM is obtained by the model switching process. In transition phases from left to right turn (or vice versa) the probability of the second filter kernel (F2) increases. This filter models rapid yaw rate changes with a high process noise. The estimated uncertainty of the yaw angle $\hat{\sigma}_\psi$ reflects the probability changes of the IMM aproach. This situation adaptive uncertainty estimation and the improved state estimates lead to a significantly better consistency of the IMM approach compared to the SKF as shown by the cumulative probability of the averaged NEES $Q_{\chi^2}(\bar{\epsilon}, 80)$.

However, even if the IMM approach produces most of the time consistent and the SKF most of the time inconsistent estimates, the consistency of the IMM is still not satisfactory. Inconsistent estimates in phases of almost constant yaw angle can only be explained by the nonlinear coordinated turn model and the non-optimal solution of the extended Kalman filter kernels. Future research should investigate advanced Kalman filter approaches for nonlinear models, including sigma-point Kalman filters mentioned in Chapter 2.4.

9.6 Discussion

This chapter describes interacting multiple model approaches, which are effectively applied to traffic scenarios relevant to advanced driver assistance systems. The disadvantage of a single filter design is that there is an optimization trade-off between the filter response during normal driving and maneuver phases of the tracked object. Multiple model approaches address this deficiency.

Different model sets were parameterized for traffic scenarios based on real vehicle trajectories. Statistics from traffic jams were used for a determination of the Markov matrix. The evaluation showed that the chosen models represent the different driving modes accurately. The classification of the driving mode given by the model probabilities might be useful for a situation assessment.

Compared to a single filter the IMM approach provides more accurate state estimates. Using models with different process noise levels enables a situation adaptive estimation of the uncertainty. In situations in which the dynamic model assumption is violated, the

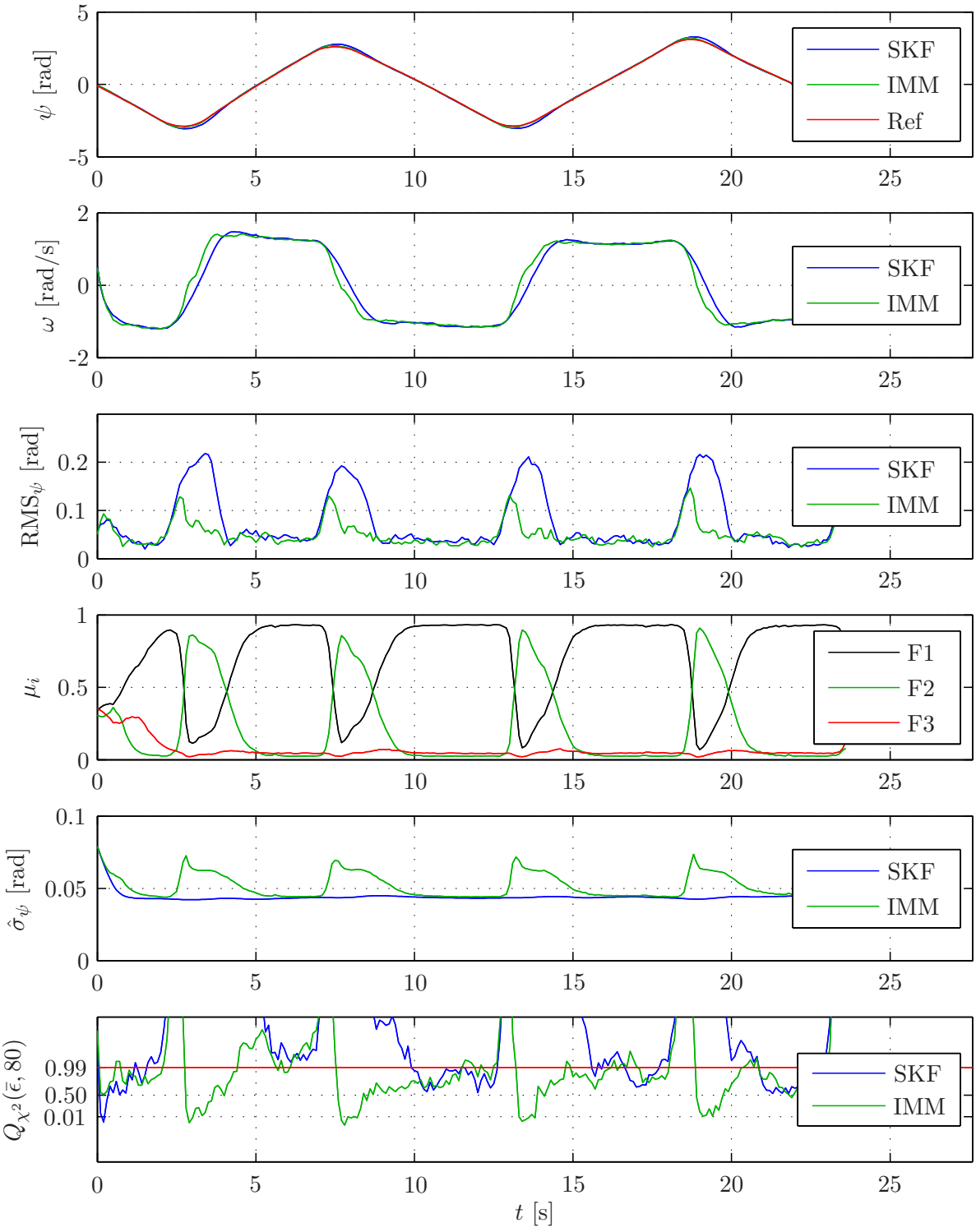


Figure 9.11: Rapid steering maneuvers. Comparison between single Kalman filter (SKF) and IMM. Averaged yaw angle ψ and yaw rate estimates ω, reference yaw angle (Ref), RMS of the yaw angle, model probabilities μ_i, estimated standard deviation of the yaw angle $\hat{\sigma}_\psi$ and cumulative probability of the averaged NEES $Q_{\chi^2}(\bar{\epsilon}, 80)$ which should stay below 99 %.

weight of the model with the higher process noise increases. Compared to a single Kalman filter, the estimation accuracy and filter consistency can thus be significantly increased. The noise reduction in normal driving phases is high and at the same time the estimate follows the trajectory well in highly dynamic situations.

For turning situations at intersections, a special model set with coordinated turn models exhibited in a hybrid simulation a superior performance compared to a single Kalman filter. The estimation accuracy and filter consistency are significantly improved by the IMM in situations of strong acceleration changes and rapid steering maneuvers. Additionally, the latency of the acceleration and yaw rate estimates are strongly reduced by the IMM approach. Thus, the IMM estimates follow the true values earlier compared to the estimates of the single Kalman filter. The coupling of the orientation with the direction of motion is a feature of the coordinate turn models. It enables a significant noise reduction of the orientation error with respect to the orientation measurement, in particular for higher velocities.

It should be mentioned that IMM is computationally more demanding than a single Kalman filter. Compared to a single filter the increase is in the order of $2n + 0.5$, where n is the number of models. In order to keep the computational burden low, a strategy could be to apply the IMM only to objects of interest. However, it is up to the system designer to decide upon the trade-off between the computational factor and the significant benefit of IMM.

10 Situation Assessment for an Emergency Brake

An emergency brake application as defined by Kopischke [96] brakes with maximal deceleration if an accident can not be avoided by any driving maneuver. The driver therefore keeps the full control over the vehicle in all driving phases. Only if an accident becomes unavoidable even for the best driver, the system triggers. During the emergency brake the driver has still the possibility to steer and will be supported in his steering maneuver by the ABS and ESP systems.

Legal aspects play a major role in the discussion about emergency brake systems. If the system falsely engages and then causes an accident in the subsequent traffic, the car manufacturer or the supplier is most certainly liable. Lages therefore underlines in [99] the necessity, that a small accident must occur at the end of the automatically triggered emergency brake. Otherwise it becomes difficult to differentiate retrospectively between situations, in which the system falsely engaged and critical situations which would have led to a collision without the emergency brake.

The system described by Kopischke imposes several assumptions upon the object's trajectory. The system works only for stationary vehicles, standing vehicles which accelerate orthogonally to the host vehicle's driving direction and vehicles which drive orthogonally to the host vehicle's driving direction with a constant velocity.

In this chapter a novel approach for the collision prediction will be presented. The algorithm can be used for any arbitrary object position and driving direction. All physically possible trajectories of the host vehicle and the observed object are considered for a collision avoidance maneuver. Only if all possible trajectory combinations of the opponent and the host vehicle lead to an accident, a collision is predicted and the emergency brake would be engaged. This approach is therefore applicable to any frontal crash vehicle-vehicle scenario. While most systems described so far address only rear end collisions with stationary vehicles in traffic jam situations, the proposed algorithm is additionally applicable to intersections scenarios.

After a short introduction to the physics of vehicles, Kopischke's algorithm will be outlined. The main part of the chapter details the novel collision prediction approach. Finally, it will be evaluated with respect to Kopischke's system.

10.1 Kamm's Circle and Resulting Trajectories

The trajectories a vehicle can drive are constrained by physical limitations. There is a maximal force, which the wheels can transmit to the road surface, which is called the limiting friction force F_f. The limiting friction force is proportional to the normal force

$F_n = mg$, where m is the vehicle's weight and g the standard gravity

$$F_f = \mu_0 mg \tag{10.1}$$

The parameter μ_0 denotes the static friction coefficient. For the acceleration or deceleration of an object a force F_a must be applied against the mass inertia

$$F_a = ma \tag{10.2}$$

where a is the acceleration. From Equations (10.1) and (10.2) the maximal acceleration can be calculated

$$a_{max} = \mu_0 g \tag{10.3}$$

The static friction coefficient $\mu_0 \in [0; 1]$ depends on the road surface. It is low for a wet or icy road and tends towards one for dry surfaces. The Equation (10.3) is however only an approximation. Empirical measurements show that the maximal deceleration of modern cars can be up to 11 m/s^2 [76]. For high velocities the normal velocity increases especially for sport vehicles, as aerodynamic measures enlarge the contact pressure between the vehicle and the road surface.

In contrast, for trucks the maximal cross acceleration is strongly limited by its susceptibility to roll over [126]. For the deceleration the limitations are not as strong. However, the braking distance of a truck is longer than that of a car.

The acceleration a is usually split into two acceleration components. The longitudinal acceleration a_l denotes the acceleration in the driving direction. The cross acceleration a_c represents the acceleration orthogonal to the driving direction. The overall acceleration a is the vector sum of both components. Given the overall acceleration a, different driving maneuvers can be defined by the introduction of an angle γ with

$$a_l = a\cos(\gamma) \tag{10.4}$$
$$a_c = a\sin(\gamma) \tag{10.5}$$

A driving maneuver is therefore defined by the acceleration a and the angle γ as shown in Kamm's circle in Figure 10.1(a). Figure 10.1(b) depicts the possible trajectories for a given acceleration a and an initial velocity v_0. The blue lines indicate the trajectories for maximal acceleration to the left and right ($|a_c| = a$). Interestingly, there are trajectories resulting from a combined steering and braking maneuver, which enable a more effective collision avoidance maneuver assuming a stationary obstacle [125].

The underlying motion model which results in the shown trajectories is given by the differential equations, which are similar to the coordinated turn model equations given in (9.12)

$$\dot{v}(t) = a_{max}\cos(\gamma) \tag{10.6}$$
$$\dot{\psi}(t) = \frac{a_{max}\sin(\gamma)}{v(t)} \tag{10.7}$$
$$\dot{x}(t) = v(t)\cos(\psi(t)) \tag{10.8}$$
$$\dot{y}(t) = v(t)\sin(\psi(t)) \tag{10.9}$$

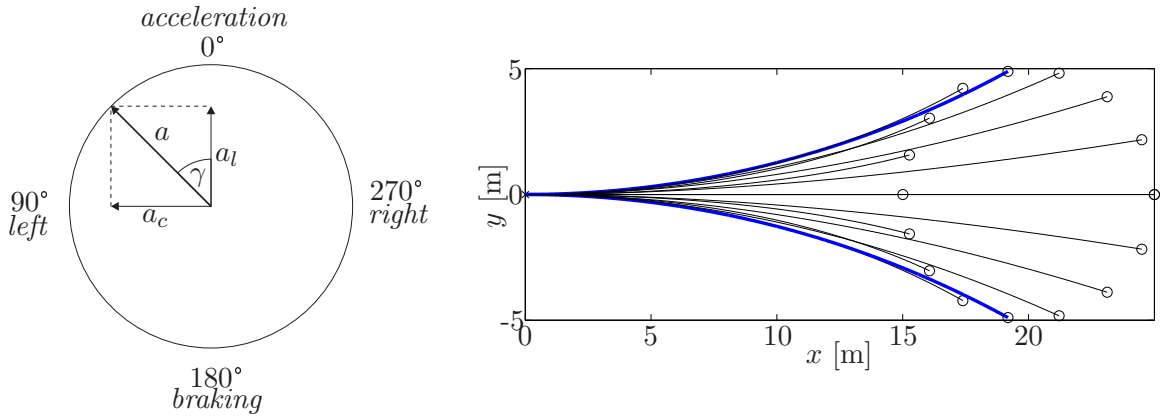


Figure 10.1: (a) Kamm's circle. (b) Possible trajectories for $a = 10$ m/s^2, $v_0 = 20$ m/s for a prediction of one second.

As the analytical solution leads to integral equations an approximate iterative solution is suggested similar to Equation (9.19). The actual state estimates of the object are chosen as the initial estimates for the trajectory prediction. The initial velocity is therefore denoted by v_0, the initial yaw angle by ψ_0 and the position by $[x_0, y_0]$

$$v_{n+1} = v_n + a_{max}\cos(\gamma)\Delta t \tag{10.10}$$

$$\psi_{n+1} = \psi_n + \min\left(\frac{a_{max}\sin(\gamma)}{v_{n+1}}, \frac{v_n}{r_{min}}\right)\Delta t \tag{10.11}$$

$$x_{n+1} = x_n + v_n\cos(\psi_{n+1})\Delta t + \frac{1}{2}a_{max}\cos(\gamma)\cos(\psi_{n+1})\Delta t^2 \tag{10.12}$$

$$y_{n+1} = y_n + v_n\sin(\psi_{n+1})\Delta t + \frac{1}{2}a_{max}\cos(\gamma)\sin(\psi_{n+1})\Delta t^2 \tag{10.13}$$

The predicted trajectory given by the velocity, yaw angle and position is then iteratively calculated in intervals of $\Delta t = 1$ ms.

For small velocities the maximal yaw rate is limited by the turn radius r_{min} of the vehicle. The utilization of v_{n+1} and ψ_{n+1} on the right hand side of Equations (10.11)–(10.13) instead of v_n and ψ_n enables an overestimation. Thus, the emergency brake is triggered rather slightly later and not prematurely.

The approximation assumes a constant yaw angle for the time intervals of the prediction step. Alternatively the prediction could assume a constant yaw rate as given in Equation (9.15). For efficiency reasons however the computationally less demanding solution from Equations (10.10)–(10.13) is chosen.

10.2 Collision Prediction by Kopischke

Before the new approach for collision prediction is introduced, Kopischke's algorithm will be shortly described, as it is used in the evaluation as a comparison with the novel approach presented in this work.

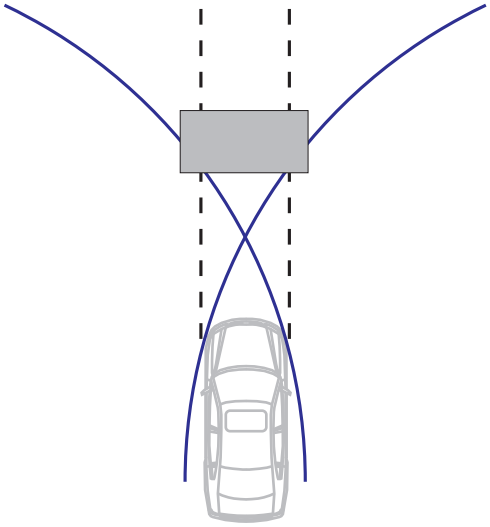


Figure 10.2: Conditions for an emergency brake for stationary obstacles.

Kopischke's emergency brake condition for a stationary vehicle is depicted in Figure 10.2. Three maneuvers of the host vehicle are considered for a collision avoidance: a full brake, which is in Kamm's circle represented by $\gamma = 180°$, a maximal cross acceleration to the left ($\gamma = 90°$) and a maximal cross acceleration to the right ($\gamma = 270°$). If the host vehicle's trajectories of all three maneuvers intersect with the stationary object and therefore lead to a collision with the obstacle, an emergency brake is triggered.

10.3 Situation Assessment

As shown in Figure 10.3 the situation assessment for the emergency brake utilizes the vehicle environment description produced by the feature-level fusion system. Additionally the motion estimate of the host vehicle is necessary for the collision prediction. The situation assessment is composed of two parts. The first part identifies dangerous situations. Collision hypotheses are generated with fast algorithms. They are refined in the collision prediction module using a complex algorithm. This coarse to fine strategy serves as a reduction of the computational burden. Details about the collision hypothesis generation can be found in [124]. The collision prediction algorithm is described in the following.

10.3.1 Collision Prediction Algorithm

In order to eliminate the possibility of a false triggering of the emergency brake, the collision prediction algorithm should account for all possible driving maneuvers. It can be assumed, that a collision is unavoidable, if it can not be avoided by any maneuver at the physical limits. Therefore only trajectories of maximal overall acceleration a_{max} are considered.

The algorithm should check all different combinations of trajectories from the host vehicle and the opponent vehicle. The trajectory combinations of the host vehicle and

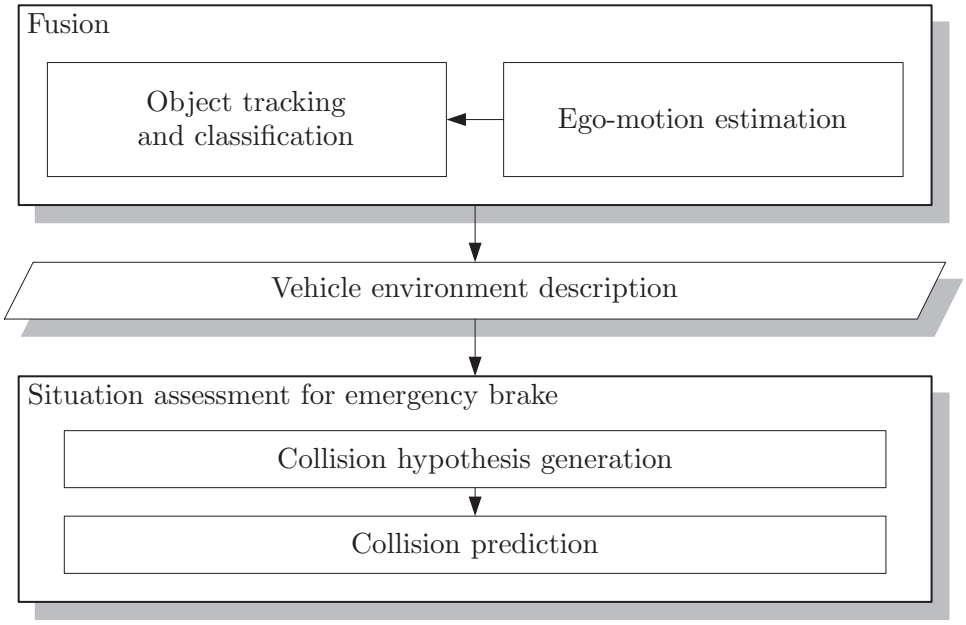


Figure 10.3: System architecture.

the opponent are defined by a tuple of angles $[\gamma_{host}, \gamma_{obj}]$ in Kamm's circle.

However, the number of trajectory combinations is unlimited. One solution could discretize the parameter space. A too rough discretization would however increase the probability, that the trajectory combination which avoids the accident is missed out.

A better alternative performs an intelligent search through the parameter space. For a search to be guided, there must exist a quality measure. Then an optimization, such as a gradient descent algorithms, can be used to search efficiently for the trajectory combination avoiding the accident. The following section describes the collision index, which is used as the quality function for the search algorithm.

10.3.2 Collision Index

For a given trajectory combination $[\gamma_{host}, \gamma_{obj}]$ the time instance is taken, in which the predicted collision occurs. A relative velocity $\mathbf{v}_r$ can be calculated from the velocity of the host vehicle and the opponent, with

$$\begin{aligned} v_{r,x} &= v_{x,host} - v_{x,obj} \\ &= v_{host}\cos(\psi_{host}) - v_{obj}\cos(\psi_{obj}) \qquad (10.14) \\ v_{r,y} &= v_{y,host} - v_{y,obj} \\ &= v_{host}\sin(\psi_{host}) - v_{obj}\sin(\psi_{obj}) \qquad (10.15) \end{aligned}$$

As shown in Figure 10.4 two lines are constructed from the direction of the relative velocity vector and the two corners of the host vehicle's front. These lines intersect with the object and define an area. The collision index is defined by the size of this area.

The collision index is not normalized, which is not necessary for a search of a minimum. The collision index is small, if the two vehicles barely touch with their corners and is

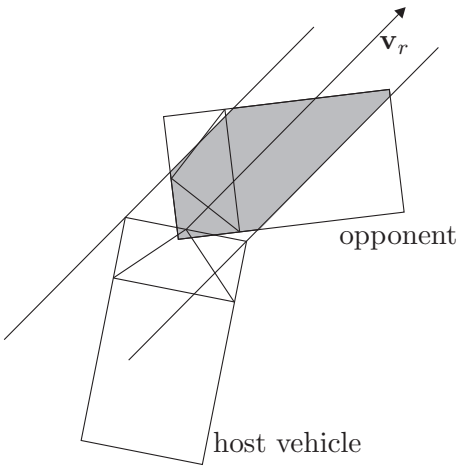


Figure 10.4: Determination of the collision index. The size of the grey area denotes the collision index.

large if the host vehicle crashes centrally into the opponent. Small indices are especially interesting for the collision prediction algorithm, as small variations of the parameters γ_{host} and γ_{obj} could lead to an avoidance of the collision.

10.3.3 Matrix of Collision Indices

Considering all different combinations of γ_{host} and γ_{obj} a two dimensional matrix can be composed. Figure 10.5(a) shows a simulated rear end collision scenario. The host vehicle is driving at a speed of 30 m/s, the opponent vehicle at 10 m/s. The host vehicle has an lateral offset of 0.5 meter to the left with respect to the opponent. The situation leads after 0.5 seconds to a rear end collision.

The matrix of collision indices in Figure 10.5(b) shows the collision index for each trajectory combination $[\gamma_{host}, \gamma_{obj}]$. Bright regions denote a central crash. The trajectories corresponding to dark region almost avoid the collision. The orange and green regions represent the trajectory combinations, which can avoid the accident[1].

Consider the red point in the center of the orange region in the matrix of collision indices. It corresponds to the angle γ_{host} in the Kamm's circle of the host vehicle and to the angle γ_{obj} in the Kamm's circle of the opponent shown in Figure 10.6. If the host vehicle performs a combined right steering and braking maneuver and the opponent a combined left turn and acceleration, the rear end collision can therefore be avoided (orange maneuver in Figure 10.5(a)).

The other possibility of maneuver combinations is given by the green region. For visualization purposes the regions are projected into the Kamm's circles. The collision can be avoided, if the host vehicle steers to the left and the opponent to the right. The green region is much larger then the orange region and therefore leaves more choice for the maneuver combinations. This asymmetry can be explained by the lateral offset between the two vehicles.

[1]For visualization purposes the black regions which denote a collision index of zero, are color coded.

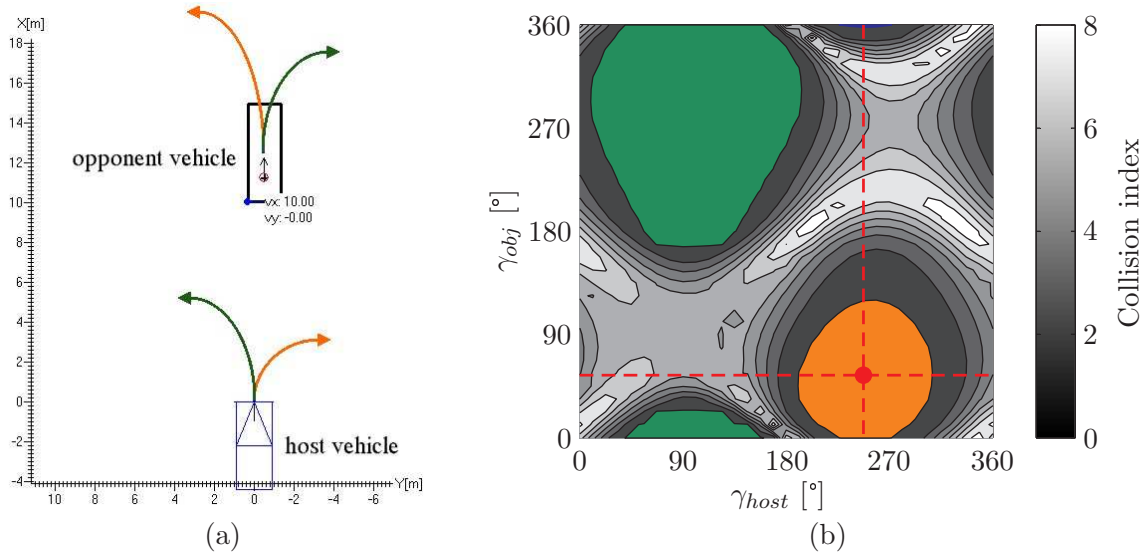


Figure 10.5: Rear end collision scenario with a lateral offset and the corresponding matrix of collision indices (b). The green and orange regions in the matrix in (b) denote the possible collision avoidance maneuver combinations $[\gamma_{host}, \gamma_{obj}]$.

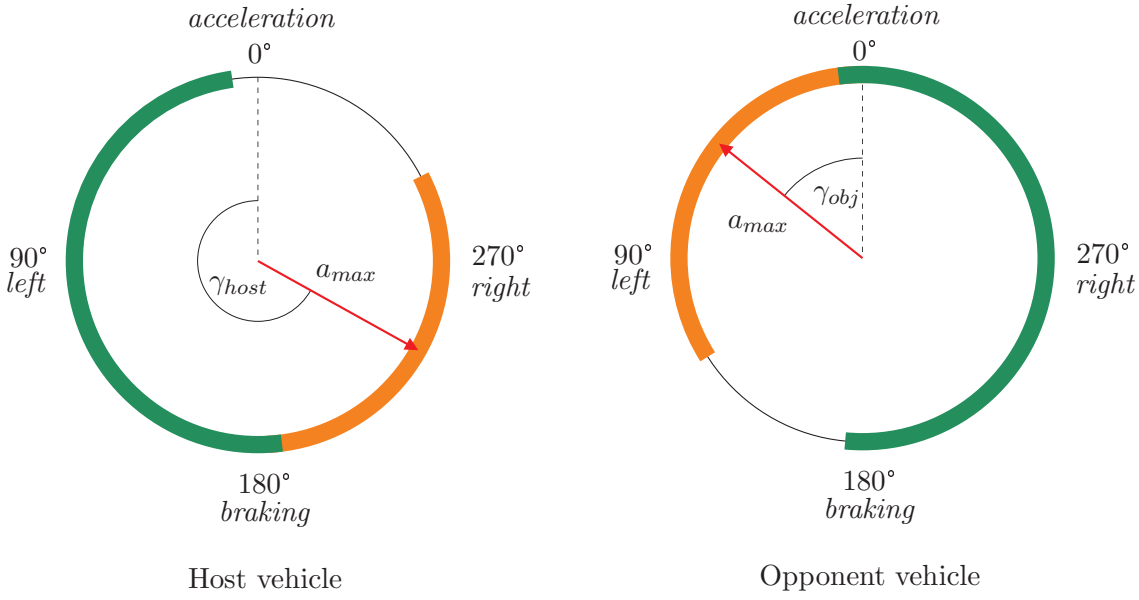


Figure 10.6: Maneuvers which could avoid a collision for the situation depicted in Figure 10.5.

As there are maneuver combinations, represented by the green and orange region, which avoid the accident the emergency brake should not trigger.

10.3.4 Example of a Rear End Collision

Figure 10.7 shows two more time instances of the rear end collision scenario described in the previous section. Over time the host vehicle approaches the opponent and one

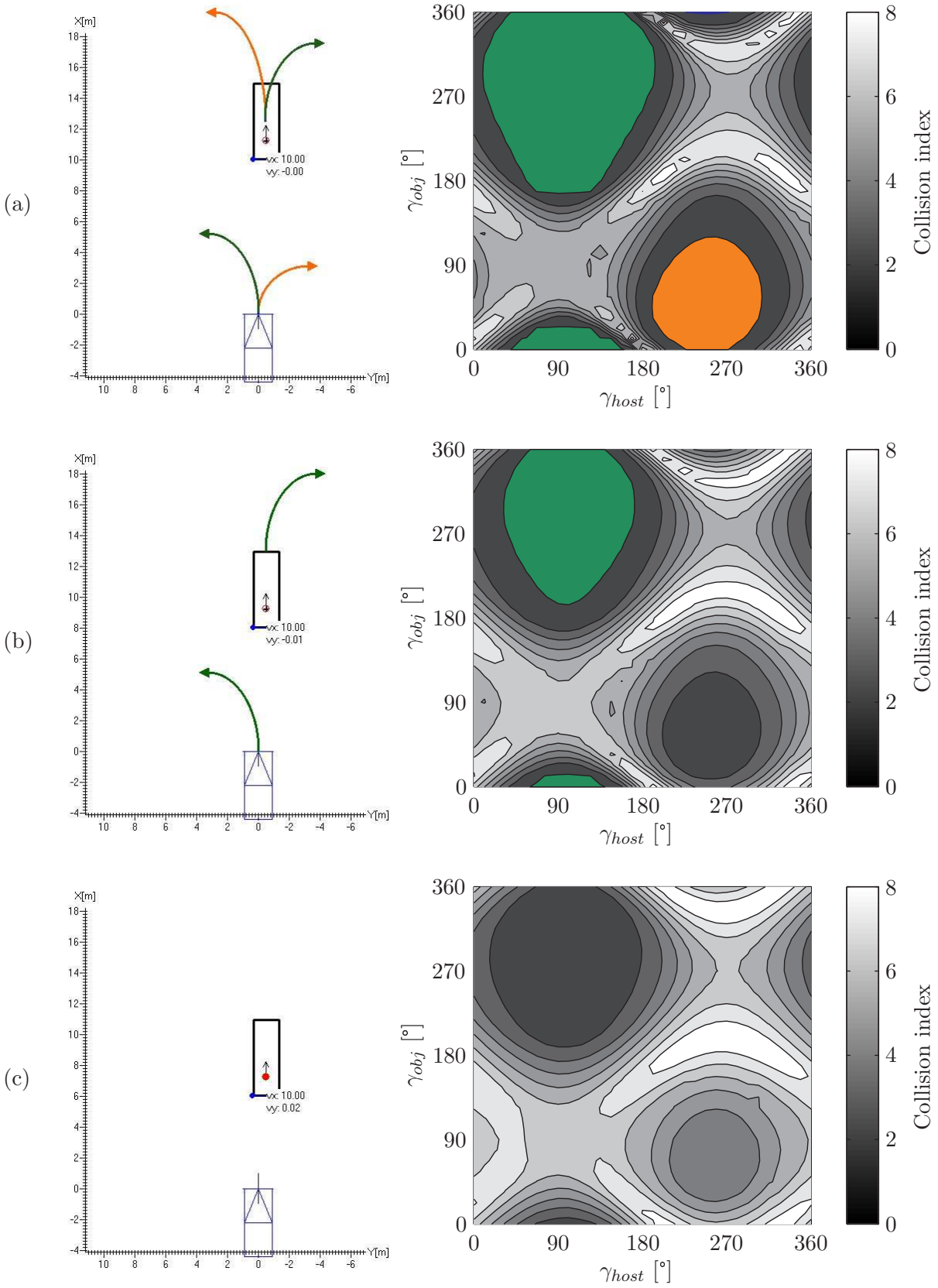


Figure 10.7: Rear end collision scenario with a lateral offset and the corresponding matrix of collision indices. Time intervals (a) 500 ms, (b) 400 ms, (c) 300 ms before the impact. The possible collision avoidance maneuvers are indicated in green and orange.

possibility of an avoidance maneuver disappears. The last change is a left steering of the host vehicle and a right steering maneuver of the opponent vehicle. 100 ms later the collision is unavoidable and the emergency brake would trigger.

10.3.5 Example of a Collision at an Intersection

The second example represents a collision at an intersection presented in Figure 10.8. Both vehicles drive at a speed of 20 m/s. At the end of this simulation the host vehicle crashed centrally into the opponent. From the matrices of collision indices the best collision avoidance maneuvers can be derived. The host vehicle should strongly brake and steer to the right. The opponent should perform a right turn. The collision becomes unavoidable 400 ms before the crash impact.

10.3.6 Search for the Minimal Collision Index

The matrices of collision indices were interpolated from 4096 ($64 \cdot 64$) maneuver combinations. In practise the calculation in real time is impossible for all of these combinations. Therefore, a search along the function given by the collision indices is performed. Using gradient descent algorithms the search is very efficient and needs only few iterations. In order to prohibit that only a local minimum is found, the search is conducted in four areas of the matrix independently. The areas are defined by the sets A_i of trajectory combinations $[\gamma_{host}, \gamma_{obj}]$

$$A_1 = \{[\gamma_{host}, \gamma_{obj}] \mid 0^\circ \leq \gamma_{host} < 180^\circ, 0^\circ \leq \gamma_{obj} < 180^\circ\} \quad (10.16)$$

$$A_2 = \{[\gamma_{host}, \gamma_{obj}] \mid 0^\circ \leq \gamma_{host} < 180^\circ, 180^\circ \leq \gamma_{obj} < 360^\circ\} \quad (10.17)$$

$$A_3 = \{[\gamma_{host}, \gamma_{obj}] \mid 180^\circ \leq \gamma_{host} < 360^\circ, 0^\circ \leq \gamma_{obj} < 180^\circ\} \quad (10.18)$$

$$A_4 = \{[\gamma_{host}, \gamma_{obj}] \mid 180^\circ \leq \gamma_{host} < 360^\circ, 180^\circ \leq \gamma_{obj} < 360^\circ\} \quad (10.19)$$

This procedure accounts for alternative collision avoidance maneuvers. If none of the searches leads to a minimal collision index of zero, the emergency brake is triggered.

10.3.7 Stationary Obstacles

For stationary objects, the object tracking delivers the position, width, orientation and if observable the length. The orientation is however ambiguous. It is not known in which of the four possible direction the object is oriented. Therefore, all combinations are considered and tested by the collision prediction algorithm.

For stationary vehicles seen from the rear, the classification using the reflectors described in Chapter 5.4.2 solves the problem. If the rear of a vehicle is detected, the orientation is known unambiguously. This is especially useful for the prominent case of rear end collisions. For moving vehicles or those which have been observed in motion, the unambiguous orientation is known from the velocity vector.

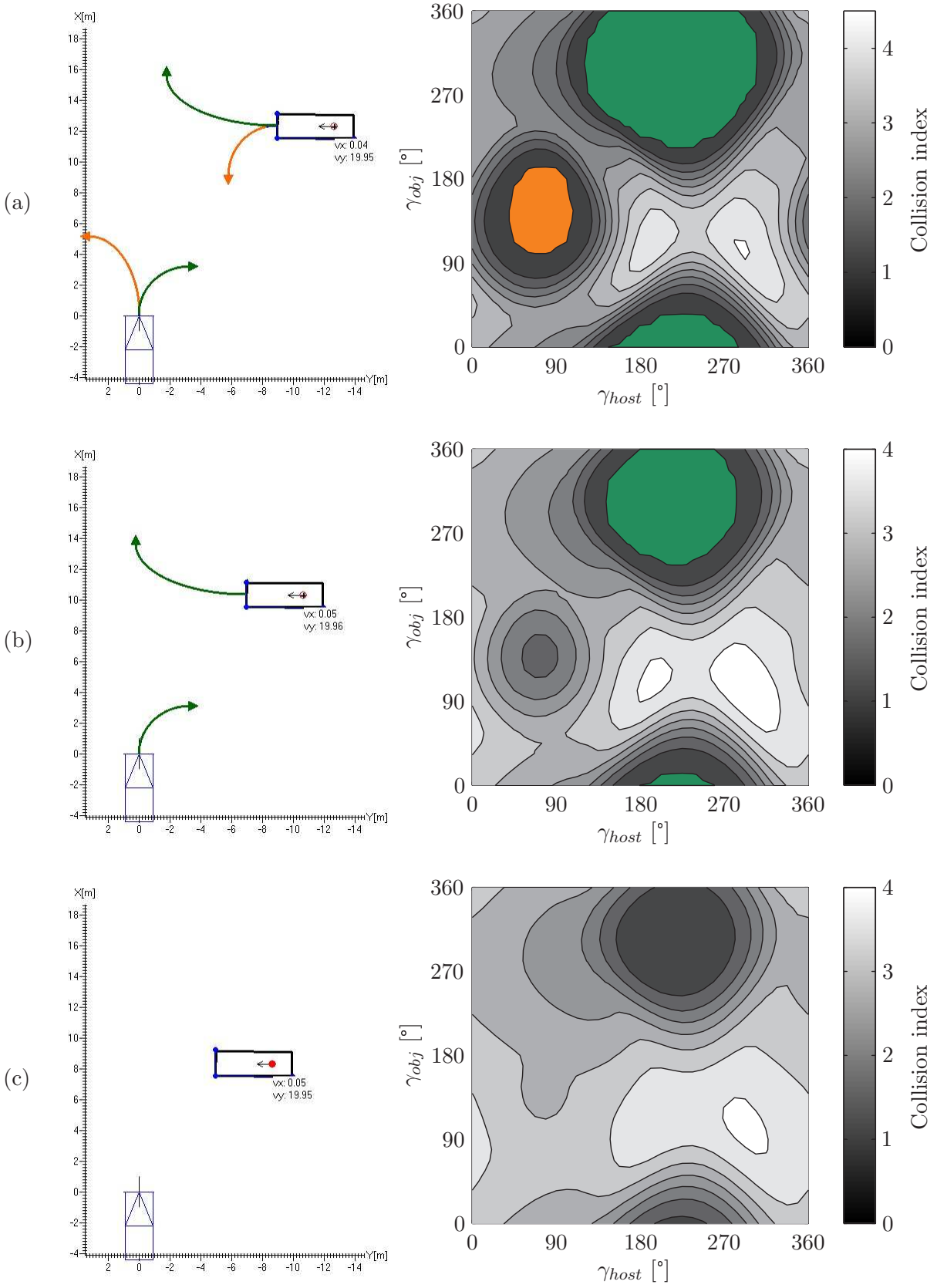


Figure 10.8: Intersection scenario and the corresponding matrix of collision indices. Time instances (a) 600 ms, (b) 500 ms, (c) 400 ms before the impact. The possible collision avoidance maneuvers are indicated in green and orange.

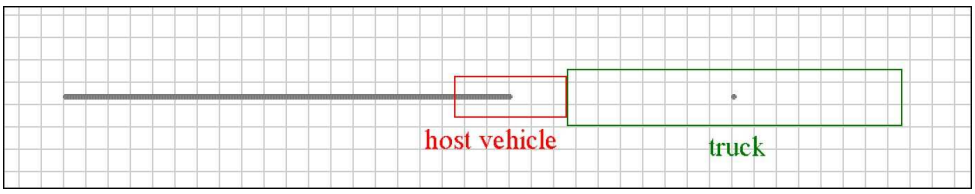


Figure 10.9: Rear end collision as presented in the simulation tool.

10.4 Evaluation

The verification of the functionality is performed in simulation. A catalogue of test cases with stationary and moving objects defines the situations. The sensor measurements are calculated from the resulting trajectories. They are the input to the vehicle environment perception system described in the previous chapters. The output of the fusion module serves as the input to the situation assessment for the emergency brake. For the detection of situations in which the emergency brake falsely triggers, real measurements from a broad spectrum of traffic scenarios were utilized additionally.

10.4.1 Rear End Collision

Many different scenarios were considered for the evaluation including rear end collisions and intersections scenarios with different crash angels. The presentation of the complete results would go beyond the scope of this work. Therefore, the performance of the novel approach with respect to previous work will be demonstrated exemplarily on rear end collisions. Further results can be found in [124].

Experimental Setup

A rear end collision as depicted in Figure 10.9 is considered for the evaluation. The host vehicle is a car with a width of 1.8 meter and a length of 5 meter. The host vehicle crashes centrally into the rear of a stationary truck. The truck is 2.5 meter wide and has a length of 15 meter. As the laser scan can only observe the width of the truck, the length is assumed to be 0.5 meter. A maximal acceleration of $a_{max} = 10$ m/s^2 is assumed and a turn radius of $r_{min} = 5$ m. The reaction time of the system is neglected and the situation is sampled at 1 ms intervals.

The relative impact energy reduction ΔE_{rel} is a good performance measure for the effectiveness of the emergency brake [96]. It is calculated from the initial velocity at the time of the release of the emergency brake v_0 and the velocity reduction Δv achieved by the emergency brake

$$\Delta E_{rel} = \frac{\frac{1}{2}mv_0^2 - \frac{1}{2}m(v_0 - \Delta v)^2}{\frac{1}{2}mv_0^2} = \frac{v_0^2 - (v_0 - \Delta v)^2}{v_0^2} \tag{10.20}$$

where m is the weight of the host vehicle.

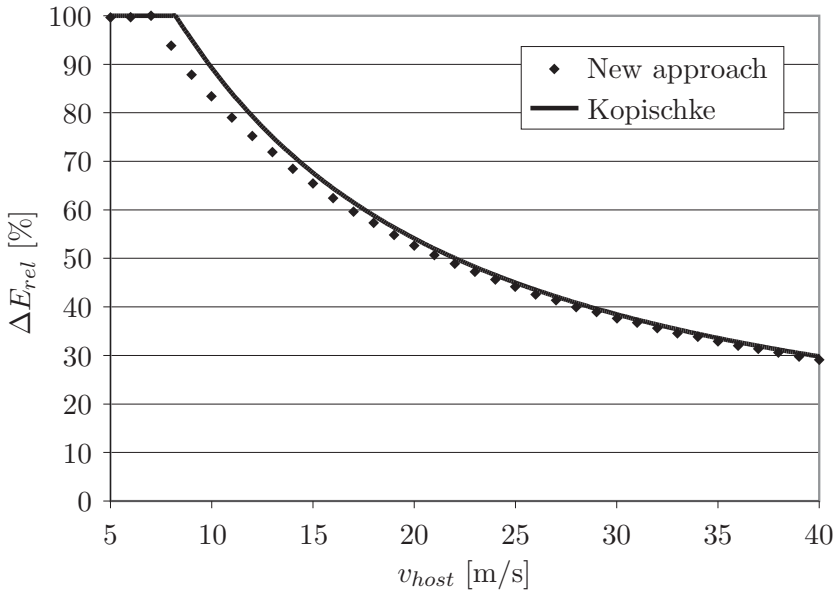


Figure 10.10: Energy reduction for a rear end collision with a stationary obstacle as a function of the host vehicle's velocity. Kopischke's approach is compared to the algorithm presented in this work.

Stationary Object

The first scenario represent the ideal conditions for the emergency brake. The obstacle is assumed to be stationary. All possible maneuvers by the object which would avoid the collision are neglected. The result is an early triggering of the emergency brake compared to other assumptions. Figure 10.10 shows the energy reduction as a function of the host vehicle's velocity. For small velocities the energy can be reduced almost completely.

Compared to Kopischke's analytical solution shown in Figure 10.10, the new approach reduces the energy slightly less. This can be explained by the difference in the two approaches. While Kopischke considers only a full brake, a left and a right turn maneuver, the novel approach accounts for combined maneuvers of braking and steering. As mentioned in Chapter 10.1 a combined steering and braking maneuver is more efficient in collision avoidance than only steering. Therefore the emergency brake triggers later and less energy is reduced. This might seem to be a disadvantage. However, as the novel approach accounts for these combined collision avoidance maneuvers, it is hypothesized, that the algorithms produces fewer false alarms.

Mobile Object with Ambiguous Orientation

In general the obstacle can not be assumed to be stationary. The collision prediction must account for objects, which start moving, as the object could be a vehicle. Different directions must be considered for the motion, if the orientation is only ambiguously known. The 2.5 m wide object could for instance be a motor-bike which stands across the lane.

The time when the emergency brake triggers is thus delayed significantly. Figure 10.11 shows the decreased energy reduction compared to the stationary obstacle assumption from Figure 10.10.

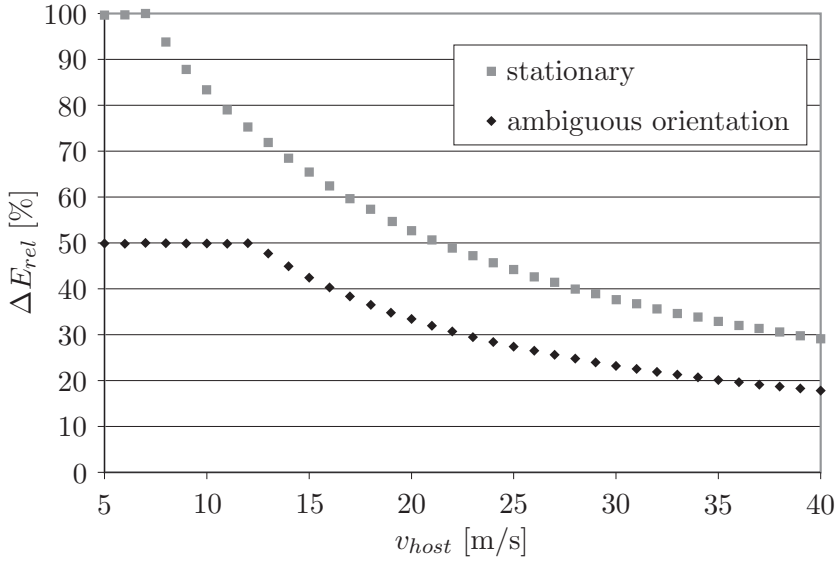


Figure 10.11: Energy reduction for a rear end collision with a mobile, ambiguously oriented obstacle as a function of the host vehicle's velocity, compared to the stationary obstacle assumption.

Mobile Object with Unambiguous Orientation

If an unambiguous object orientation is known, the collision prediction must only account for that direction, in which the vehicle is oriented. The knowledge of the orientation could for instance be provided by the track history, if the vehicle has been observed in motion. Alternatively, the obstacle could be identified as a vehicle seen from the rear, because of the rear reflectors and the number plate.

The object has less degrees of freedom for a collision avoidance maneuver compared to the case, in which the object's orientation is ambiguous. Consequently Figure 10.12 exhibits a higher energy reduction.

Object Classified as Truck

An additional advantage can be obtained by a classification of the obstacle. In the previous assumptions the collision prediction was completely unaware, that the obstacle is a truck. If a truck is identified, the maximal possible acceleration can be reduced. The maximal cross acceleration of a truck is below 5 m/s^2 [126]. The truck tips over if this limit is exceeded. A maximal acceleration of $a_{max} = 5$ m/s^2 and a turn radius of $r_{min} = 15$ m is therefore assumed. Figure 10.13 reveals a significant improvement especially for lower velocities.

Comparison

The results demonstrate, that the efficiency of the emergency brake system can be significantly improved with a detailed knowledge of the obstacle. A comparison of the different approaches is shown in Figure 10.14. While the knowledge of an unambiguous orientation improves the energy reduction for higher host vehicle velocities, the classification

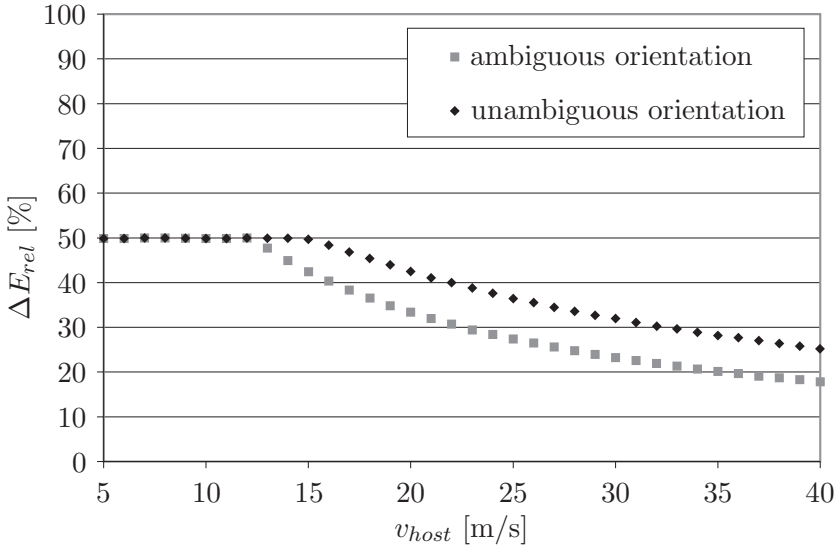


Figure 10.12: Energy reduction for a rear end collision with a mobile, unambiguously oriented obstacle as a function of the host vehicle's velocity, compared to the mobile, ambiguously oriented obstacle assumption.

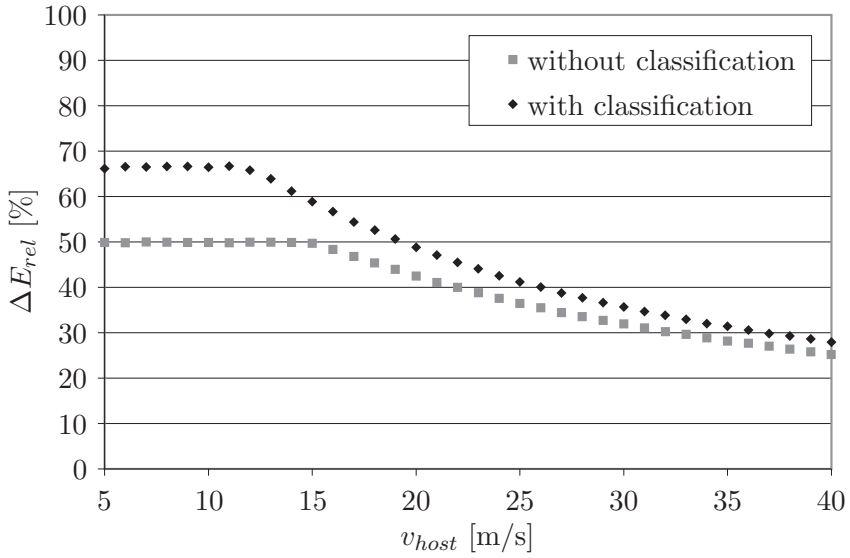


Figure 10.13: Energy reduction for a rear end collision with a mobile, unambiguously oriented obstacle which is classified as a truck, as a function of the host vehicle's velocity. This is compared to the mobile, unambiguously oriented obstacle assumption without classification.

is especially useful in scenarios of lower host vehicle speed. A multi-class classification as described by Wender et al. in [169] which differentiates between pedestrians, bikes, cars, trucks and other objects is therefore important for an efficient emergency brake application.

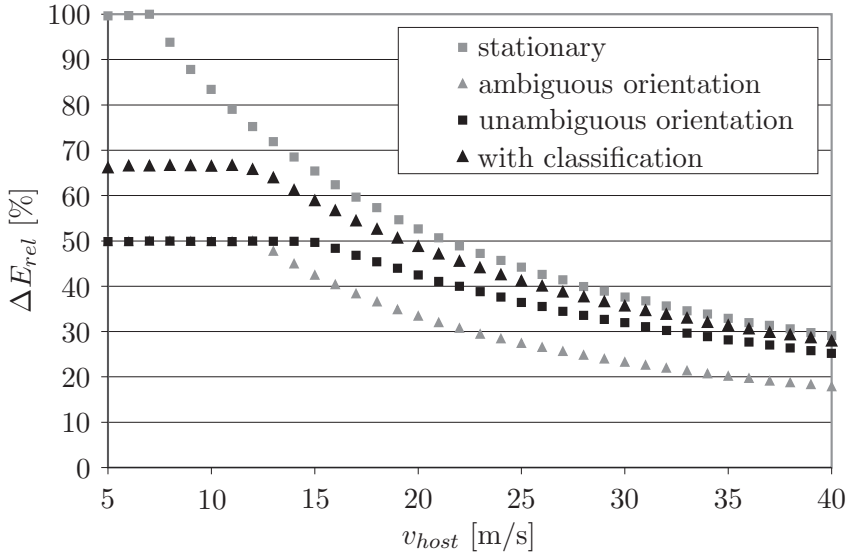


Figure 10.14: Energy reduction for all four cases of a rear end collision.

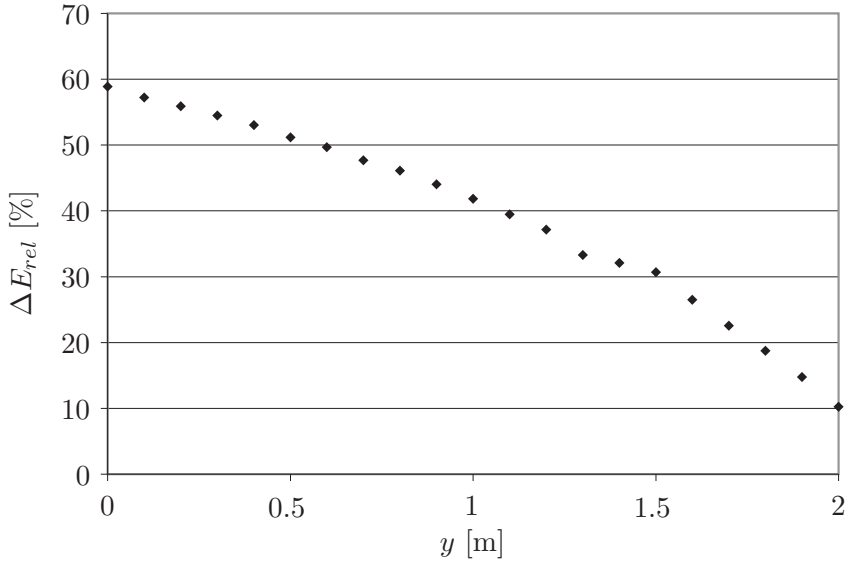


Figure 10.15: Energy reduction for a rear end collision with a mobile, unambiguously oriented obstacle which is classified as a truck, as a function of the offset between the host vehicle and the truck.

Influence of Lateral Object Position

In general there is a lateral offset between the two vehicles in a rear end collision. The larger the offset is, the longer the collision can be avoided by a steering maneuver. Figure 10.15 shows the decreased energy reduction for large offsets. In this situation the host vehicle is driving at a speed of $v_{host} = 15$ m/s. The object is classified as a truck.

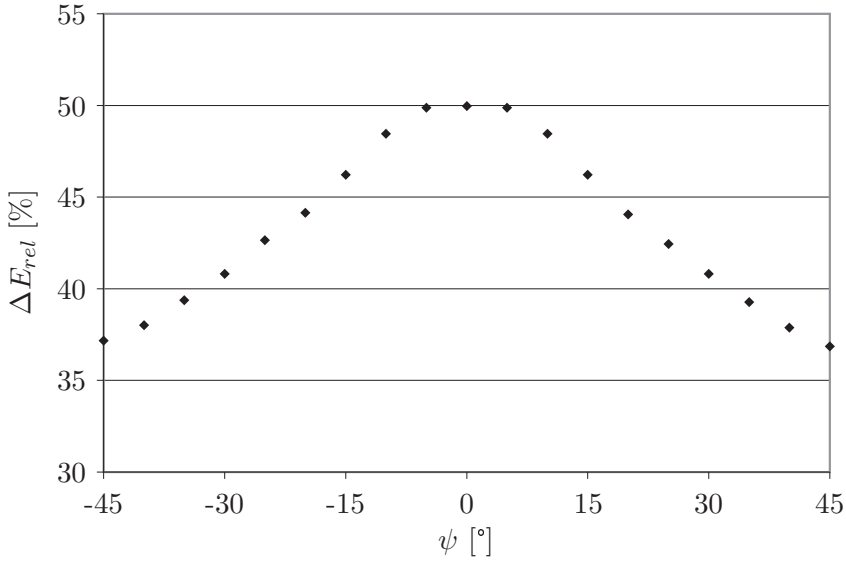


Figure 10.16: Energy reduction for a rear end collision with a mobile, unambiguously oriented obstacle, as a function of its orientation.

Influence of Object Orientation

The last scenario investigates the influence of the obstacle's orientation. The host vehicle approaches a quadratic obstacle with a speed of $v_{host} = 15$ m/s. The edge length of the object is 2.5 meter. The energy reduction decreases with increasing angle as shown in Figure 10.16. For very small variations of the orientation, the change in energy reduction is not very significant. These results underline the importance of the orientation estimation, which is enabled by the laser scanner. Other sensors such as a radar sensor must rely on the estimated orientation of the motion vector, which is impossible for stationary objects.

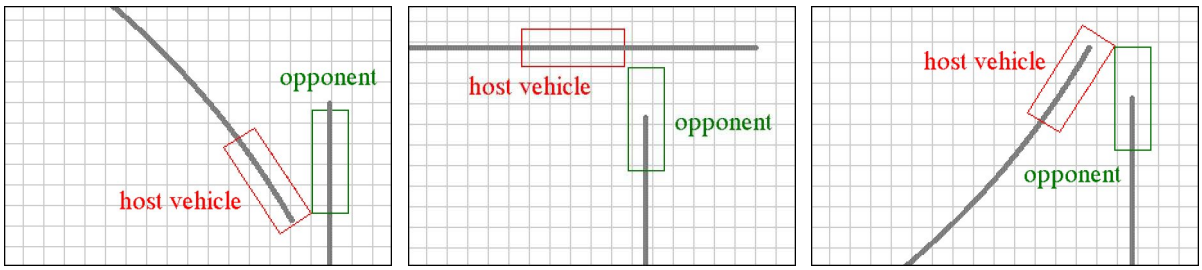
10.4.2 Negative Tests

The previous sections investigate scenarios in which the emergency brake triggers. The analysis exhibits the performance of the collision prediction algorithm in crash scenarios. The correct performance of the emergency brake in situations in which a collision can be avoided is however of equally high importance. These test were performed in simulation and with real measurement data.

Simulation

Three scenarios were developed in which the collision can be narrowly avoided by driving maneuvers at the physical limit. Figure 10.17 depicts these situations.

In the first case shown in Figure 10.17(a) the host vehicle approaches another car which is coming from the right. The host vehicle's speed is $v_{host} = 20$ m/s and the opponent's velocity $v_{obj} = 10$ m/s. The host vehicle avoids the imminent collision by a right turn with maximal cross acceleration. Even if the vehicles almost touch each other, the emergency brake does not trigger in this case.



(a) (b) (c)

Figure 10.17: Three test cases in whch the collision can be narrowly avoided as presented in the simulation tool.

In the second scenario the host vehicle does not perform any maneuver. The opponent, however, performs a full brake and comes to a stand still just when the host vehicle passes in front. Figure 10.17(b) shows the time instant when the opponent stops. In this situation the emergency brake triggers unfortunately, even if the collision could be avoided. The reason lies in the uncertainty of the object state estimation. Because of the filter latency, the estimated velocity diverges from zero although the vehicle has already stopped. The situation assessment therefore predicts the opponent into the driving corridor of the host vehicle and triggers the emergency brake.

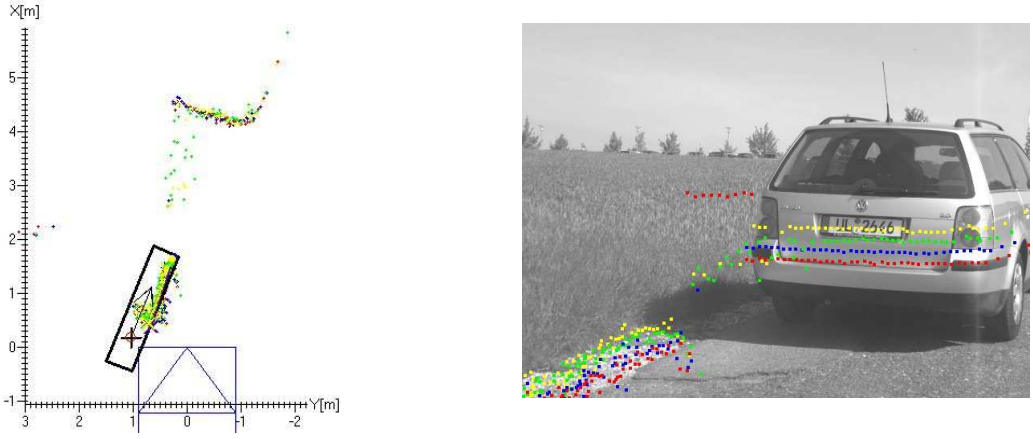
In the third case a car is approaching the host vehicle from the right. The host vehicle's speed is $v_{host} = 20$ m/s and the opponent's velocity $v_{obj} = 10$ m/s. This time the host vehicle avoids the crash with a maximal cross acceleration to the left as shown in Figure 10.17(c). At the same time, the opponent performs an emergency brake at maximal deceleration. The opponent vehicle comes finally to a stand still, while the host vehicle passes in front of the opponent. In this situation, in which both vehicle perform an emergency maneuver in order to avoid the crash and still come very close, the emergency brake performs correctly and does not trigger.

The presented collision prediction does not account for the uncertainty of the opponent state estimates. The negative tests show the necessity to incorporate the uncertainty for an emergency brake and to reduce the uncertainty with sophisticated state estimation algorithms, such as the Interacting Multiple Model Algorithm presented in the previous chapter.

Real World Test

The scenarios described in the previous section represent potentially harmful situations in which a release of an emergency brake is likely to be accepted by the driver and the jurisdiction, even if the collision could be avoided. However, false collision predictions in normal driving situations are unacceptable.

Measurement sequences from test drives on freeways, rural road in in urban areas were recorded. The emergency brake analyzed more than 86,000 measurement frames and triggered only once. In this scenario a preceding vehicle produces a cloud of dust which is detected by the laser scanner as shown in Figure 10.18. A ghost opponent is created and the emergency brake triggers.



(a) (b)

Figure 10.18: A ghost oobject caused by laser scanner measurements on dust behind the preceding vehicle results in a release of the emergency brake signal.

10.5 Discussion

A situation assessment algorithm for an emergency brake application has been presented. The collision prediction detects imminent accidents with other vehicles and analyzes if the collision is avoidable considering each observed vehicle separately. Unlike previous approaches, the situation analysis is not restricted to special scenarios, such as rear end collisions. The emergency brake is applicable to any possible scenario. It is therefore especially useful as a collision mitigation systems for accidents at intersections.

The performance of the emergency brake has been demonstrated in simulated accident scenarios. The analysis exhibited, that the achievable energy reduction strongly depends on the available information about the obstacle. In general, opponents must be assumed mobile. If the orientation of the opponent is ambiguous, the energy reduction is rather small. With the knowledge of the unambiguous orientation the energy reduction can be increased. An opponent classification can further improve the performance. The highest energy reduction is obtained, if the opponent can be assumed to be stationary.

The performance of the emergency brake depends on many parameters. The influence of the opponent's orientation and the offset between the host vehicle and the obstacle were presented. In [124] the necessity of a low system latency is discussed. The system latency is influenced by the measurement update rate, the processing time of the sensor, fusion and collision prediction modules and the response time of the brakes.

The evaluation exhibited an extreme situation which leads to an erroneous engagement of the emergency brake. If the vehicles perform collision avoidance maneuvers close to the physical limit, the system might predict a crash, even if the vehicles' trajectories do not intersect. The situation assessment is mislead by an inaccurate velocity estimate of the observed opponent. As the velocity estimates are always prone to errors due to noisy measurements and the latency of the estimation process, the emergency brake should account for the uncertainties in the state estimates.

Future work could incorporate the infrastructure into the situation assessment. The

knowledge of the road boarder would help to limit the possibilities of collision avoidance maneuvers and the emergency brake could trigger earlier.

11 Conclusions

Advanced driver assistance systems aim at an improved traffic safety, enhanced comfort and driving pleasure. Sensors perceive the objects surrounding the vehicle and produce an environment description. The assistance systems support the driver by assessing the situation recognized by this vehicle environment description.

Current research in the area of advanced driver assistance systems aims at increased functionality. Comfort systems, such as the ACC, are expected to support the driver not only in normal driving phases, but also in more complex situations such as traffic jams. Safety systems will trigger collision avoidance or mitigation measures in a number of potential crash configurations and not only in well defined rear end collision situations. These future advanced applications impose strong requirements on the sensors and demand novel and complex signal processing algorithms for vehicle detection, tracking and situation assessment.

The presented work addresses challenging traffic scenarios. Novel sensor signal processing, senor data fusion and tracking algorithms were developed which provide precise and consistent motion estimates for complex intersection scenarios, lane change maneuvers of vehicles, trucks on neighboring lanes and highly dynamic situations. A novel emergency brake application offers a general solution to the situation assessment and aims in particular at so far unresolved intersection scenarios.

It has been widely recognized that the high standards demanded by advanced driver assistance systems can only be satisfied by exploiting several sensors simultanously. Sensor data fusion algorithms provide methods for an effective combination of the information originating from the diverse sensors. An efficient sensor data fusion architecture has been presented which is widely sensor independent, readily scalable and which requires a limited communication bandwidth. The proposed feature-level fusion aims at an optimal tracking and classification performance as the data is fused on a low abstraction level.

The environment perceiving sensors utilized in this work are a laser scanner and a gray scale camera. The laser scanner offers with its wide angular field of view an excellent sensor specification for the tracking of vehicles in intersection scenarios. The precise distance measurements and good angular resolution enable the observation of the dimensions and orientation of other traffic participants. The novel contour matching algorithm extracts the position and orientation precisely from the laser scanner data. If the rear of a vehicle is observed, the orientation estimate is further improved by exploiting the measurements of reflectors. The evaluation of intersection scenarios showed that the novel cuboid tracking approach can improve the accuracy of velocity estimates by a factor of more than three compared to the standard center of gravity tracking. This is of major importance especially for safety applications, such as the developed emergency brake.

The fusion of the cuboid tracking with the video data based template tracking exhibits its advantages mainly for lane change maneuvers of preceding vehicles. The template

tracking improves the azimuthal position estimate by a factor of more than two. In combination with lane recognition this novel feature-level fusion approach would support an early lane change detection of preceding vehicles.

The tracking of trucks on neighboring lanes at close range addresses another so far unresolved challenge. A precise position and motion estimation is important for potentially harmful side crash scenarios. The novel displacement tracking provides velocity estimates which are calculated from the motion of the structure formed by the laser scanner measurements. The approach is independent of a geometric model and applies for a wide range of object shapes. Ambiguous structures, such as line patterns produced by crash barriers, are reliably identified.

The evaluations show that the novel tracking methods provide precise position, dimensions, orientation and motion estimates of vehicles for a wide range of previously unresolved and challenging scenarios. Statistics from extended test drives exhibit under good weather conditions high detection rates of preceding vehicles of 99.4 % at velocity estimate consistency rates of 100.0 %. Additionally, the estimated velocity of stationary roadside objects has been investigated. Consistent estimates are necessary for ACC applications in order to suppress stationary obstacles for the target selection. The statistics revealed that the developed cuboid tracking approaches are more than twice as accurate as the standard center of gravity tracking. The number of consistent velocity estimates could be increased from 79 % to 92 %.

The analysis of highly dynamic scenarios, such as strong acceleration changes or rapid turns at intersections exhibited the deficiencies of the standard single Kalman filter approach. A state estimation based on multiple models which each represent a special driving phase performs significantly better. The interacting multiple model approach was chosen as an efficient method. The model set was systematically parameterized using statistics from vehicle trajectories in real traffic scenarios. The evaluation of the two model sets for traffic jam and intersection scenarios exhibited a significant performance improvement for the interacting multiple model algorithm with respect to the accuracy and consistency of the estimates.

The benefit of the developed algorithms has been demonstrated exemplarily using an advanced emergency brake algorithm. The novel situation assessment algorithm detects imminent accidents with other vehicles and analyzes whether the collision is avoidable. Unlike previous approaches, the situation analysis is not restricted to special scenarios, such as rear end collisions. The emergency brake is applicable to any possible scenario. It is therefore especially useful as a collision mitigation systems for accidents at intersections. The performance analysis of the emergency brake exhibited that the achievable impact energy reduction strongly depends on the available information about the obstacle. The knowledge of the unambiguous orientation or the object's class can enhance the performance. The unambiguous orientation of stationary vehicles can be obtained for instance from the newly developed classification of the vehicles' rear using the laser scanner measurements of reflectors.

The developed algorithms have been thoroughly evaluated using real traffic scenarios and simulations. A performance degradation is to be expected in adverse weather conditions such as rain, fog or snow as the sensors are sensitive to perturbations in the visible light and near infrared spectrum. The performance of the template tracking using

the data of the passive camera sensor should additionally be tested in challenging light conditions such as at dawn, dusk or at night.

A Proofs and Derivations

A.1 Existence of Maximally One Solution

In theory there are two solutions to the problem stated in Equation (3.16)

$$\alpha_{c,1} = \arctan\left(\frac{bc+\sqrt{1-c^2+b^2}}{1+b^2}, -\frac{b\left(bc+\sqrt{1-c^2+b^2}\right)}{1+b^2}+c\right) \tag{A.1}$$

$$\alpha_{c,2} = \arctan\left(\frac{bc-\sqrt{1-c^2+b^2}}{1+b^2}, -\frac{b\left(bc-\sqrt{1-c^2+b^2}\right)}{1+b^2}+c\right) \tag{A.2}$$

The first proof shows that the second solution $\alpha_{c,2}$ is always located in the third quadrant and is therefore not the correct solution. It has to be shown that $x_2 < 0$ and $y_2 < 0$ with $\alpha_{c,2} = \arctan(y_2, x_2)$. As $b < 0$ and $c > 0$ for $h_c > t_{z,ls} \geq 0$ it can easily be shown that

$$y_2 = \frac{bc-\sqrt{1-c^2+b^2}}{1+b^2} < 0 \tag{A.3}$$

The same can be shown for x_2

$$x_2 = -\frac{b\left(bc-\sqrt{1-c^2+b^2}\right)}{1+b^2}+c \overset{!}{<} 0 \tag{A.4}$$

$$\frac{b\left(bc-\sqrt{1-c^2+b^2}\right)}{c\left(1+b^2\right)} \overset{!}{<} 1 \tag{A.5}$$

$$\frac{b^2+\sqrt{\frac{b^2}{c^2}-b^2+\frac{b^4}{c^2}}}{1+b^2} \overset{!}{<} 1 \tag{A.6}$$

This is true if

$$\frac{b^2}{c^2}-b^2+\frac{b^4}{c^2} \overset{!}{>} 1 \tag{A.7}$$

Using the inequation (A.9) it follows

$$\left(1+b^2\right)\frac{b^2}{c^2}-b^2 > 1+b^2-b^2 = 1 \tag{A.8}$$

This proofs that $x_2 < 0$ and that $\alpha_{c,2}$ is located in the third quadrant. This is therefore not the correct solution.

A.2 Argument under the Root

The second proof shows that the argument under the square root in Equations (3.16), (A.1) and (A.2) is always positive. As $b < 0$ and $c > 0$ for $h_c > t_{z,ls} \geq 0$ it can be shown that

$$\frac{b}{c} = -\frac{\left\|\mathbf{p}_a^{ls}\right\| + \left\|\mathbf{p}_b^{ls}\right\|}{2\left(h_c - t_{z,ls}\right)} < -1 \tag{A.9}$$

by the following proof under the assumption that $w_c > h_c$, which is a sensible restriction on the construction of the calibration object and $\left\|\mathbf{p}_a^{ls}\right\| > w_c$ which is a sensible restriction on the position of the calibration object:

$$\begin{aligned}
\left\|\mathbf{p}_a^{ls}\right\| &> w_c && \text{(A.10)}\\
\left\|\mathbf{p}_a^{ls}\right\| + \left\|\mathbf{p}_b^{ls}\right\| > 2\left\|\mathbf{p}_a^{ls}\right\| &> 2w_c > 2h_c > 2\left(h_c - t_{z,ls}\right) && \text{(A.11)}\\
\frac{\left\|\mathbf{p}_a^{ls}\right\| + \left\|\mathbf{p}_b^{ls}\right\|}{2\left(h_c - t_{z,ls}\right)} &> 1 && \text{(A.12)}\\
\frac{b}{c} = -\frac{\left\|\mathbf{p}_a^{ls}\right\| + \left\|\mathbf{p}_b^{ls}\right\|}{2\left(h_c - t_{z,ls}\right)} &< -1 && \text{(A.13)}
\end{aligned}$$

Equation (A.9) leads to the proof, that the argument under the square root in (A.1) and (A.2) is always positive, with

$$\begin{aligned}
\frac{b}{c} &< -1 && \text{(A.14)}\\
\frac{b^2}{c^2} &> 1 && \text{(A.15)}\\
b^2 &> c^2 > c^2 - 1 && \text{(A.16)}\\
1 - c^2 + b^2 &> 0 && \text{(A.17)}
\end{aligned}$$

A.3 Calculation of the Sub-pixel Correction Vector

The sub-pixel accuracy of the image based template matching is achieved using parabola fitting. The parabola is fitted into the maximal correlation coefficients and the correlation coefficients of the two neighboring pixel positions.

The general equation of a parabola is given by

$$y = m(x - b)^2 + c \tag{A.18}$$

Three points for the parabola fit are given. One point is defined by the pixel position of the maximal correlation coefficient. As a correction vector is searched for and not an absolute image position, the pixel position is $x_2 = 0$. The y-coordinate y_2 is given by the corresponding correlation coefficient $r(\Delta\hat{u}^*, \Delta\hat{v}^*)$. The vectors $[x_1, y_1]'$ and $[x_3, y_3]'$ define the neighboring points. In each coordinate direction a parabola is fitted independently. The points for the parabola fit in u-coordinate direction is given by

$$\begin{aligned}
x_1 = &\ -1, \quad && y_1 = r(\Delta\hat{u}^* - 1, \Delta\hat{v}^*) && \text{(A.19)}\\
x_2 = &\ 0, \quad && y_2 = r(\Delta\hat{u}^*, \Delta\hat{v}^*) && \text{(A.20)}\\
x_3 = &\ 1, \quad && y_3 = r(\Delta\hat{u}^* + 1, \Delta\hat{v}^*) && \text{(A.21)}
\end{aligned}$$

Inserting these three points into Equation (A.18) an equation system can be constructed

$$y_1 = m(-1-b)^2 + c = mb^2 + 2mb + m + c \tag{A.22}$$
$$y_2 = m(-b)^2 + c = mb^2 + c \tag{A.23}$$
$$y_3 = m(1-b)^2 + c = mb^2 - 2mb + m + c \tag{A.24}$$

Subtracting Equation (A.24) from Equation (A.22) delivers

$$\begin{aligned} y_1 - y_3 &= 4mb \\ b &= \frac{y_1 - y_3}{4m} \end{aligned} \tag{A.25}$$

Adding Equation (A.24) and Equation (A.22) provides

$$y_3 + y_1 = 2mb^2 + 2m + 2c \tag{A.26}$$

The parameter m follows from Equation (A.26) and Equation (A.23) with

$$m = \frac{1}{2}(y_3 + y_1) - y_2 \tag{A.27}$$

Inserting Equation (A.27) into Equation (A.25) results in the parameter b

$$\begin{aligned} b &= \frac{y_1 - y_3}{4m} \\ &= \frac{y_1 - y_3}{4[\frac{1}{2}(y_3 + y_1) - y_2]} = \frac{y_1 - y_3}{2(y_3 + y_1) - 4y_2} \end{aligned} \tag{A.28}$$

$$\tag{A.29}$$

Rearranging Equation (A.23) results in

$$c = y_2 - mb^2 \tag{A.30}$$

The parameter c is then calculated as a function of b from Equation (A.30) by insertion of Equation (A.27)

$$\begin{aligned} c &= y_2 - [\frac{1}{2}(y_3 + y_1) - y_2]b^2 \\ &= y_2 + y_2 b^2 - \frac{1}{2}(y_3 + y_1)b^2 \\ &= y_2(1 + b^2) - \frac{1}{2}(y_3 + y_1)b^2 \end{aligned} \tag{A.31}$$

The parameter b from Equation (A.28) defines the u-coordinate $\Delta\hat{u}_{sb}$ of the sub-pixel correction vector. With Equations (A.19) - (A.21) follows

$$\begin{aligned} \Delta\hat{u}_{sb} &= b = \frac{y_1 - y_3}{2(y_3 + y_1) - 4y_2} \\ &= \frac{r(\Delta\hat{u}^* - 1, \Delta\hat{v}^*) - r(\Delta\hat{u}^* + 1, \Delta\hat{v}^*)}{2[r(\Delta\hat{u}^* + 1, \Delta\hat{v}^*) + r(\Delta\hat{u}^* - 1, \Delta\hat{v}^*)] - 4r(\Delta\hat{u}^*, \Delta\hat{v}^*)} \end{aligned} \tag{A.32}$$

The parameter c from Equation (A.31) defines the correlation coefficient $r(\Delta u, \Delta\hat{v}^*)$ of the u-coordinate parabola fit. With Equations (A.19) - (A.21) and Equation (A.28) follows

$$\begin{aligned} r(\Delta u, \Delta\hat{v}^*) &= c = y_2(1+b^2) - \tfrac{1}{2}(y_3+y_1)b^2 \\ &= r(\Delta\hat{u}^*, \Delta\hat{v}^*)(1+\Delta\hat{u}_{sb}^2) \\ &\quad -\tfrac{1}{2}[r(\Delta\hat{u}^*+1, \Delta\hat{v}^*) + r(\Delta\hat{u}^*-1, \Delta\hat{v}^*)]\Delta\hat{u}_{sb}^2 \end{aligned} \tag{A.33}$$

Analogously the points for the calculation of the v-coordinate of the sub-pixel correction vector are given by

$$x_1 = -1, \qquad y_1 = r(\Delta\hat{u}^*, \Delta\hat{v}^* - 1) \tag{A.34}$$

$$x_2 = 0, \qquad y_2 = r(\Delta\hat{u}^*, \Delta\hat{v}^*) \tag{A.35}$$

$$x_3 = 1, \qquad y_3 = r(\Delta\hat{u}^*, \Delta\hat{v}^* + 1) \tag{A.36}$$

The v-coordinate $\Delta\hat{c}_{sb}$ of the sub-pixel correction vector is then given by

$$\begin{aligned} \Delta\hat{v}_{sb} &= \frac{y_1 - y_3}{4[\frac{1}{2}(y_3+y_1) - y_2]} \\ &= \frac{r(\Delta\hat{u}^*, \Delta\hat{v}^* - 1) - r(\Delta\hat{u}^*, \Delta\hat{v}^* + 1)}{2[r(\Delta\hat{u}^*, \Delta\hat{v}^* + 1) + r(\Delta\hat{u}^*, \Delta\hat{v}^* - 1)] - 4r(\Delta\hat{u}^*, \Delta\hat{v}^*)} \end{aligned} \tag{A.37}$$

and the correlation coefficient $r(\Delta u, \Delta\hat{v}^*)$ of the v-coordinate parabola fit

$$\begin{aligned} r(\Delta\hat{u}^*, \Delta v) &= y_2(1+b^2) - \tfrac{1}{2}(y_3+y_1)b^2 \\ &= r(\Delta\hat{u}^*, \Delta\hat{v}^*)(1+\Delta\hat{v}_{sb}^2) \\ &\quad -\tfrac{1}{2}[r(\Delta\hat{u}^*, \Delta\hat{v}^*+1) + r(\Delta\hat{u}^*, \Delta\hat{v}^*-1)]\Delta\hat{v}_{sb}^2 \end{aligned} \tag{A.38}$$

The overall correlation coefficient of the sub-pixel fit is given by the average of the two correlation coefficients provided by Equation (A.33) and Equation (A.38)

$$\begin{aligned} r(\Delta u, \Delta v) &= \tfrac{1}{2}[r(\Delta u, \Delta\hat{v}^*) + r(\Delta\hat{u}^*, \Delta v)] \\ &= r(\Delta\hat{u}^*, \Delta\hat{v}^*)[1 + \tfrac{1}{2}(\Delta\hat{u}_{sb}^2 + \Delta\hat{v}_{sb}^2)] \\ &\quad -\tfrac{1}{4}[r(\Delta\hat{u}^*+1, \Delta\hat{v}^*) + r(\Delta\hat{u}^*-1, \Delta\hat{v}^*)]\Delta\hat{u}_{sb}^2 \\ &\quad -\tfrac{1}{4}[r(\Delta\hat{u}^*, \Delta\hat{v}^*+1) + r(\Delta\hat{u}^*, \Delta\hat{v}^*-1)]\Delta\hat{v}_{sb}^2 \end{aligned} \tag{A.39}$$

A.4 Calculation of the Reference Position

The reference position $\mathbf{p}^{ls}$ results from the intersection of the ray $\mathbf{g}$ and a circle of radius r with its center in the origin of the laser scanner coordinate system (Figure A.1).

The ray $\mathbf{g}$ is defined in the xy-plane of the laser scanner coordinate system and is given by the projection of the ray $\mathbf{g}^*$. The ray $\mathbf{g}^*$ given in the three dimensional laser scanner coordinate system is constructed from the reference image position in the image $\mathbf{p}^{img}$ of the vehicle's rear (Equation (7.8)). The ray is defined by a parametric linear equation in the laser scanner coordinate system

$$\mathbf{g}^* : \quad \begin{bmatrix} x^{ls} \\ y^{ls} \\ z^{ls} \end{bmatrix} = \mathbf{a}^* + \mathbf{b}^* s = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} + \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} s$$

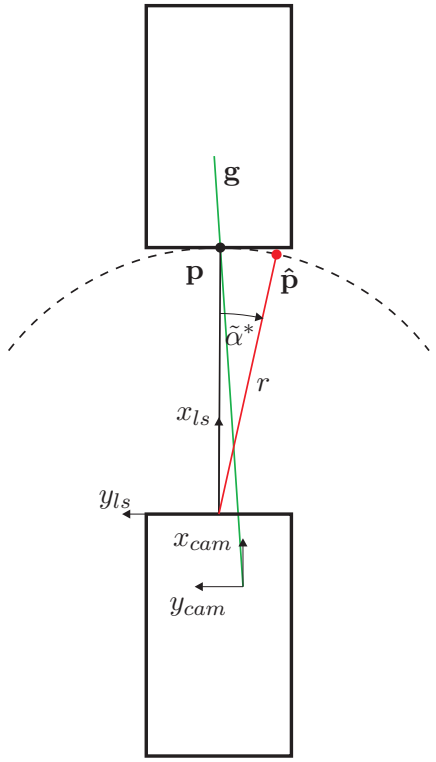


Figure A.1: Azimuth measurement error $\tilde{\alpha}^*$ in the xy-plane of the laser scanner coordinate system. The reference position is denoted by $\mathbf{p}$ and $\hat{\mathbf{p}}$ is the estimate.

The parameters $\mathbf{a}^*$ and $\mathbf{b}^*$ can be derived from the projection matrix $\mathbf{H}_{ls\rightarrow img}$ given in Equation (3.45). This matrix is decomposed into

$$\mathbf{H}_{ls\rightarrow img} = \begin{bmatrix} \mathbf{H}^* & \mathbf{h}^* \end{bmatrix} \tag{A.40}$$

where

$$\mathbf{H}^* \in \mathbb{R}^{3\times 3}, \quad \mathbf{h}^* \in \mathbb{R}^3 \tag{A.41}$$

According to the projection from Equation (3.26) a point in the laser scanner coordinate system $\mathbf{p}^{ls} = \begin{bmatrix} x^{ls} & y^{ls} & z^{ls} & 1 \end{bmatrix}'$ is given in the image $\tilde{\mathbf{p}}^{img}$ with Equation (3.45)

$$\tilde{\mathbf{p}}^{img} = s \begin{bmatrix} u & v & 1 \end{bmatrix}' = \mathbf{H}_{ls\rightarrow img}\mathbf{p}^{ls} \tag{A.42}$$

Using (A.40) the equation can be inverted, resulting in

$$\begin{aligned} s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= \begin{bmatrix} \mathbf{H}^* & \mathbf{h}^* \end{bmatrix} \mathbf{p}^{ls} = \mathbf{H}^* \begin{bmatrix} x^{ls} \\ y^{ls} \\ z^{ls} \end{bmatrix} + \mathbf{h}^* \\ \begin{bmatrix} x^{ls} \\ y^{ls} \\ z^{ls} \end{bmatrix} &= (\mathbf{H}^*)^{-1} \left(s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - \mathbf{h}^* \right) \\ &= s\,(\mathbf{H}^*)^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - (\mathbf{H}^*)^{-1}\,\mathbf{h}^* \end{aligned} \tag{A.43}$$

The parameters $\mathbf{a}^*$ and $\mathbf{b}^*$ of the linear equation are therefore given by

$$\mathbf{a}^* = -\left(\mathbf{H}^*\right)^{-1}\mathbf{h}^* \tag{A.44}$$

$$\mathbf{b}^* = \left(\mathbf{H}^*\right)^{-1}\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \left(\mathbf{H}^*\right)^{-1}\begin{bmatrix} \mathbf{p}^{img} \\ 1 \end{bmatrix}' \tag{A.45}$$

where $\mathbf{p}^{img}$ denotes the reference image position from Equation (7.8).

From the ray equation in three dimensions in Equation (A.40) follows the ray equation in the xy-plane

$$\mathbf{g}: \quad \begin{bmatrix} x^{ls} \\ y^{ls} \end{bmatrix} = \mathbf{a} + \mathbf{b}s = \begin{bmatrix} a_x \\ a_y \end{bmatrix} + \begin{bmatrix} b_x \\ b_y \end{bmatrix} s \tag{A.46}$$

For the determination of the reference position of the vehicle's rear $\mathbf{p}^{ls}$ the intersection between the ray $\mathbf{g}$ and a circle with radius r

$$\left(x^{ls}\right)^2 + \left(y^{ls}\right)^2 = r^2 = \|\hat{\mathbf{p}}^{ls}\| \tag{A.47}$$

is calculated by substitution of x^{ls} and y^{ls} from Equation (A.46)

$$\begin{aligned} \left(a_x + b_x s\right)^2 + \left(a_y + b_y s\right)^2 &= r^2 \\ s^2 + \frac{2\left(a_x b_x + a_y b_y\right)}{b_x^2 + b_y^2}s + \frac{a_x^2 + a_y^2 - r^2}{b_x^2 + b_y^2} &= 0 \end{aligned} \tag{A.48}$$

There are two results for this quadratic equation

$$s_{1/2} = -\frac{\left(a_x b_x + a_y b_y\right)}{b_x^2 + b_y^2} \pm \sqrt{\frac{\left(a_x b_x + a_y b_y\right)^2}{\left(b_x^2 + b_y^2\right)^2} - \frac{a_x^2 + a_y^2 - r^2}{b_x^2 + b_y^2}} \tag{A.49}$$

Inserting these results into Equation (A.46) delivers two solutions for $\begin{bmatrix} x^{ls} & y^{ls} \end{bmatrix}'$, of which only the solution with $x^{ls} > 0$ is valid. This solution denotes therefore the reference position $\mathbf{p}^{ls}$ which is used in Equation (7.13) for the calculation of the azimuth measurement error $\tilde{\alpha}^*$.

It should be noted that the argument under the root is alway positive, as in the experiment the radius r is always larger than the distance to the origin of the camera coordinate system $a_x^2 + a_y^2$, with

$$\begin{aligned} r^2 &> a_x^2 + a_y^2 \\ -\left(a_x^2 + a_y^2 - r^2\right) &> 0 \end{aligned}$$

Bibliography

[1] Aglaia GmbH, http://www.aglaia-gmbh.de, February 2007.

[2] A. Amditis, A. Polychronopoulos, I. Karaseitanidis, G. Katsoulis, and E. Bekiaris, "Multiple-sensor-collision avoidance system for automotive applications using an IMM approach for obstacle detection," in *Proceedings of ISIF/IEEE* 5^{th} *International Conference on Information Fusion*. ISIF, 2002, pp. 812–817.

[3] D. Baehring, S. Simon, W. Niehsen, and C. Stiller, "Detection of close Cut-In and Overtaking Vehicles for Driver Assistance based on Planar Parallax," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

[4] Y. Bar-Shalom, *Estimation and Tracking: principles, techniques, and software*. Boston, London: Artech House, 1993.

[5] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation*. New York: John Wiley & Sons Inc., 2001.

[6] Y. Bar-Shalom and X.-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS, 1995, ISBN: 0-9648312-0-1.

[7] Y. Bar-Shalom and F. T.E., *Tracking and Data Association*. Orlando, FL: Academic Press, 1988.

[8] J. Barron, D. Fleet, and S. Beauchemin, "Performance of Optical Flow Techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, 1994.

[9] J.-C. Becker, "Fusion der Daten der objekterkennenden Sensoren eines autonomen Straßenfahrzeugs," Ph.D. dissertation, Universität Braunschweig, March 2002.

[10] P. J. Besl and N. D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, February 1992.

[11] M. Betke, E. Haritaoglu, and L. Davis, "Multiple Vehicle Detection and Tracking in Hard Real-time," in *IEEE Intelligent Vehicles Symposium*, Tokyo, Japan, 1996, pp. 351–356.

[12] S. S. Blackman, *Multiple-Target Tracking with Radar Applicaitons*. Artech House, 1986.

[13] S. S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999, ISBN: 1-58053-006-0.

[14] H. A. P. Blom and Y. Bar-Shalom, "The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients," *IEEE Transactions on Automatic Control*, vol. 33, no. 8, pp. 780–783, August 1988.

[15] J.-Y. Bouguet, "Camera Calibration Toolbox for Matlab," http://www.vision.caltech.edu/bouguetj/calib_doc/index.html, 2005.

[16] H.-H. Braess and G. Reichart, "Prometheus: Vision des intelligenten Automobils auf intelligenter Straße? Versuch einer kritischen Würdigung — Teil 1," *Automobiltechnische Zeitschrift*, vol. 97, no. 4, pp. 200–205, 1995.

[17] ——, "Prometheus: Vision des intelligenten Automobils auf intelligenter Straße? Versuch einer kritischen Würdigung — Teil 2," *Automobiltechnische Zeitschrift*, vol. 97, no. 6, pp. 330–343, 1995.

[18] A. Broadhurst, S. Baker, and T. Kanade, "A prediction and planning framework for road safety analysis, obstacle avoidance and driver information," Robotics Institute, Carnegie Mellon University, Pittsburgh, USA, Tech. Rep. CMU-RI-TR-04-11, February 2004.

[19] A. E. Broadhurst, S. Baker, and T. Kanade, "A Prediction and Planning Framework for Road Safety Analysis, Obstacle Avoidance and Driver Information," in *Proceedings of* 11^{th} *World Congress on Intelligent Transportation Systems*, Nagoya, Japan, October 2004.

[20] ——, "Monte Carlo Road Safety Reasoning," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

[21] A. Broggi, M. Bertozzi, A. Fascioli, C. L. Bianco, and A. Piazzi, "Visual Perception of Obstacles and Vehicles for Platooning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 3, pp. 164–176, September 2000.

[22] M. Bühler, "Verfahren zur automatisierten Kalibrierung von Laserscanner und Videokamera," Student research project, Universität Ulm, September 2003, supervisor: N. Kämpchen.

[23] ——, "Simultane Schätzung des dynamischen Zustandes und der Form von Objekten mit einem mehrzeiligen Laserscanner," Diploma thesis, Universität Ulm, April 2005, supervisor: N. Kämpchen.

[24] T. Bühler, "Objektverfolgung unter Nutzung der Echopulsbreiten eines mehrzeiligen Laserscanners," Diploma thesis, Universität Ulm, 2005, supervisor: N. Kämpchen.

[25] J. W. Carl, *Handbook of Multisensor Data Fusion*. CRC Press, 2001, ch. 7, Contrasting Apporaches to Combine Evidence, pp. 7–1–7–32.

[26] D. S. Caveney, "Multiple Model Techniques in Automotive Estimation and Control," Ph.D. dissertation, University of California, Berkeley, USA, 2004.

[27] D. S. Caveney, D. Dickmanns, and J. K. Hedrick, "Before the Controller: A Multiple Target Tracking Routine for Driver Assistance Systems," *at — Automatisierungstechnik*, vol. 51, no. 5, pp. 230–238, 2003.

[28] D. S. Caveney and J. K. Hedrick, "Multiple Target Tracking in the Adaptive Cruise Control Environment Using Multiple Models and Probabilistic Data Associtation," in *Proceedings of the ASME International Mechanical Engineering Conference*, New York, NY, USA, November 2001.

[29] J. Chambers, W. Cleveland, B. Kleiner, and P. Tukey, *Graphical Methods for Data Analysis*. Wadsworth, 1983.

[30] Z. Chen, "Local observability and its application to multiple measurement estimation," *IEEE Transactions on Industrial Electronics*, vol. 38, no. 6, pp. 491–496, December 1991.

[31] H. Cramer, U. Scheunert, and G. Wanielik, "Multi Sensor Fusion for Object Detection Using Generalized Feature Models," in *Proceedings of ISIF/IEEE* 6^{th} *International Conference on Information Fusion*, Cairns, Australia, July 2003, pp. 2–10.

[32] H. Cramer, "Modelle zur multisensoriellen Erfassung des Fahrzeugumfeldes mit Hilfe von Schätzverfahren," Ph.D. dissertation, Technische Universität Chemnitz, 2006.

[33] H. Cramer and G. Wanielik, "Road border detection and tracking in non cooperative areas with a laser radar system," in *Proceedings of the German Radar Symposium*, 2002.

[34] M. Darms and H. Winner, "A Modular System Architecture for Sensor Data Processing of ADAS Applications," in *IEEE Intelligent Vehicles Symposium*, June 2005, pp. 729–734.

[35] F. Dellaert, D. Pomerleau, and C. Thorpe, "Model-Based Car Tracking Integrated with a Road-Follower," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Leuven, Belgium, May 1998.

[36] E. D. Dickmanns, R. Behringer, D. Dickmanns, T. Hildebrandt, M. Maurer, F. Thomanek, and J. Schiehlen, "The Seeing Passenger Car VaMoRs-P," in *IEEE Intelligent Vehicles Symposium*, Paris, France, October 1994.

[37] K. Dietmayer, A. Kirchner, and N. Kaempchen, *Fahrerassistenzsysteme*. Springer Verlag, 2004, ch. Fusionsarchitekturen zur Umfeldwahrnehmung für zukünftige Fahrerassistenzsysteme.

[38] K. Dietmayer, N. Kaempchen, K. Fuerstenberg, and J. Kibbel, "Roadway Detection and Lane Detection using Multilayer Laserscanner," in *Advanced Microsystems for Automotive Applications*, Berlin, Germany, 2005.

[39] "DIN 70000, Straßenfahrzeuge; Fahrzeugdynamik und Fahrverhalten; Begriffe, (ISO 8855:1991, modifiziert)," 1994.

[40] D. Ehmanns and H. Spannheimer, "ADASE technology road map," Deliverable D2D Roadmap Development, 2004. [Online]. Available: http://www.adase2.net/

[41] European Commission, "White paper — european transport policy for 2010: time to decide," Office for Official Publications of the European Cimmunities, 2001. [Online]. Available: http://ec.europa.eu/transport/white_paper/index_en.htm

[42] O. Faugeras, *Three-Dimensional Computer Vision*. Cambridge Massachusetts; London, England: Three-Dimensional Computer Vision, 1993.

[43] T. Focke, W. Niehsen, J. Mustaklem, F. Suhling, and M. Mueller, "Method and device for calibrating sensors in a motor vehicle," Patent application, WO2004/003586, 2002.

[44] U. Franke and C. Rabe, "Kalman Filter based Depth from Motion with Fast Convergence," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

[45] R. Freymann, "Möglichkeiten und Grenzen von Fahrerassistenz- und Aktiven Sicherheitssystemen," in *Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, March 2004.

[46] K. Fuerstenberg and K. Dietmayer, "Fahrzeugumfelderfassung mit mehrzeiligen Laserscannern," *Journal Technisches Messen*, vol. 71, no. 3, pp. 164–172, March 2004.

[47] K. C. Fuerstenberg and K. Dietmayer, "Fußgängerschutz unter Nutzung mehrzeiliger Laserscanner," in *FAS Workshop Fahrerassistenzsysteme*, April 2005, pp. 111–120.

[48] K. C. Fuerstenberg and K. C. Dietmayer, "Object Tracking and Classification for Multiple Active Safety and Comfort Applications using Multilayer Laserscanners," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004.

[49] K. C. Fuerstenberg, D. T. Linzmeier, and K. C. Dietmayer, "Pedestrian Recognition and Tracking of Vehicles using a vehicle based Multilayer Laserscanner," in *Proceedings of* 10^{th} *World Congress on Intelligent Transportation Systems*, Madrid, Spain, November 2003.

[50] K. C. Fuerstenberg and V. Willhoeft, "Pedestrian Recognition in Urban Traffic using Laserscanners," in *Proceedings of* 8^{th} *World Congress on Intelligent Transportation Systems*, Syndney, Australia, October 2001.

[51] W. A. Fuller, *Measurement Error Models*. John Wiley & Sohns, 1987.

[52] K. C. Fürstenberg, "Transformation der in rotierenden Polarkoordinaten vorliegenden Messdaten in kartesische Raum-Koordinaten," Universität Ulm, Tech. Rep., September 2002.

[53] ——, "Messgenauigkeit des Laserscanners LD ML," IBEO Automobile Sensor GmbH, Tech. Rep., Juli 2004.

[54] A. Gepperth, J. Edelbrunner, and T. Bücher, "Real-time detection of cars in video sequences," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005, pp. 625–631.

[55] A. Gern, *Multisensorielle Spurerkennung für Fahrerassistenzsysteme*. Berlin, Germany: Logos, 2005, ISBN: 3-8325-1103-2.

[56] H. Goellinger and P. Schulenberg, "Methods for automatic adjustment of laser scanner sensor whose alignment may be altered using adjustable suspension," Patent application, DE19902287 A1, 1999.

[57] M. Grewal and A. Andrews, *Kalman Filtering Theory and Practise*. Prentice Hall, 1993.

[58] D. Gruyer, C. Royere, and V. Berge-Cherfaoui, "Credibilist multi-sensor fusion for the mapping of dynamic environment," in *Proceedings of ISIF/IEEE* 3^{th} *International Conference on Information Fusion*, Paris, France, July 2000.

[59] Y. L. Guilloux, J. Lonnoy, R. Moreira, C. Blanc, J. Gallice, L. Trassoudaine, H. Tattegrain, M.-P. Bruyas, and A. Chapon, "Paroto Project: Infrared and Radar Data Fusion for Obstacle Avoidance," in *Advanced Microsystems for Automotive Applications*, Berlin, Germany, March 2004, pp. 181–200.

[60] A. H. Gunatilaka and B. A. Baertlein, "Feature-Level and Decision-Level Fusion of Noncoincidently Sampled Sensors for Land Mine Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 577–589, June 2001.

[61] D. L. Hall and S. A. H. McMullen, *Mathematical Techniques in Multisensor Data Fusion*. Artech House, 2004.

[62] D. L. Hall and A. N. Steinberg, *Handbook of Multisensor Data Fusion*. CRC Press, 2001, ch. 21, Dirty Secrets in Multisensor Data Fusion, pp. 21–1–21–12.

[63] F. M. Ham and R. G. Brown, "Observability, Eigenvalues, and Kalman Filtering," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-19, no. 2, pp. 269–273, 1983.

[64] H. Hart, *Einführung in die Messtechnik*. VEB Verlag Technik Berlin, 1989.

[65] C. Hoffmann, "Fusing Multiple 2D Visual Features for Vehicle Detection," in *IEEE Intelligent Vehicles Symposium*, Tokyo, Japan, June 2006.

[66] C. Hoffmann and T. Dang, "Cheap Joint Probabilistic Data Association Filters in an Interacting Multiple Model Design," in *Proceedings of ISIF/IEEE* 9^{th} *International Conference on Information Fusion*, Florence, Italy, September 2006, pp. 197–202.

[67] C. Hoffmann, T. Dang, and C. Stiller, "Vehicle detection fusing 2D visual features," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004.

[68] D. Huang and H. Leung, "An Expectation-Maximization-Based Interacting Multiple Model Approach for Cooperative Driving Systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, pp. 206–228, 2005.

[69] IBEO Automobile Sensor GmbH, http://www.ibeo-as.de, June 2006.

[70] Insurance Institute for Highway Safety, "update on electronic stability control," *Status Report*, vol. 41, no. 5, pp. 1–8, June 2006.

[71] M. A. Isard, "Visual motion analysis by probabilistic propagation of conditional density," Ph.D. dissertation, Oxford University, Oxford, UK, 1998.

[72] M. Isard and A. Blake, "Contour Tracking By Stochastic Propagation of Conditional Density ," in *Proceedings of the European Conference on Computer Vision*, Cambridge, UK, 1996, pp. 343–356.

[73] ——, "Condensation - conditional density propagation for visual tracking," *International Journal of Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998. [Online]. Available: citeseer.ist.psu.edu/isard98condensation.html

[74] A. Jahn, "Automatic alignment of a sensor," Patent application, EP1355128, 2002.

[75] H. Janssen and W. Niehsen, "Vehicle Surround Sensing based on Information Fusion of Monocular Video and Digital Map," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004.

[76] J. Jansson, "Dealing with Uncertainty in Automotive Collision Avoidance," in *Advanced Microsystems for Automotive Applications*, 2004, pp. 165–180.

[77] J. Jansson, J. Johansson, and F. Gustafsson, "Decision making for collision avoidance systems," in *Society of Automotive Engineering SAE*, no. 2002-01-0403, March 2002.

[78] O. Jokinen, "Self-Calibration of a Light Striping System by Matching Multiple 3-D Profile Maps," in *2nd International Conference on 3D Digital Imaging and Modeling (3DIM '99)*, Ottawa, Canada, October 1999, pp. 180–190.

[79] S. J. Julier, "Process Models for the Navigation of High-Speed Land Vehicles," Ph.D. dissertation, University of Oxford, Department of Engineering Science, 1997.

[80] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A New Approach for Filtering Nonlinear Systems," in *Proceedings of the American Control Conference*, Seattle WA, USA, 1995, pp. 1628–1632.

[81] N. Kaempchen, M. Clauss, Y. Guenter, R. M. Schreier, M. Stiegeler, K. Tischler, K. Dietmayer, H. P. Grossmann, H. Kabza, H. Neumann, A. Rothermel, and C. Stiller, "Vernetzte Fahrzeug-Umfelderfassung für zukünftige Fahrerassistenzsysteme," in *FAS Workshop Fahrerassistenzsysteme*, Walting, Germany, 2005.

[82] N. Kaempchen and K. Dietmayer, "Fusion of Laserscanner and Video for Advanced Driver Assistance Systems," in *Proceedings of* 11^{th} *World Congress on Intelligent Transportation Systems*, Nagoya, Japan, Oktober 2004.

[83] ——, "Fusion von Laserscanner und Video für aktive Sicherheits- und Komforanwendungen im Automobil," in *21. Internationale VDI/VW-Gemeinschaftstagung: Integrierte Sicherheit und Fahrerassistenzsysteme*, ser. VDI-Berichte, vol. 1864, Okotber 2004.

[84] ——, "IMM vehicle tracking for traffic jam situations on highways," in *Proceedings of ISIF/IEEE* 7^{th} *International Conference on Information Fusion*, Stockholm, Schweden, June 2004.

[85] N. Kaempchen, K. Weiss, M. Schaefer, and K. Dietmayer, "IMM Object Tracking for High Dynamic Driving Maneuvers," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004.

[86] N. Kaempchen, M. Zocholl, and K. Dietmayer, "Spatio-temporal Segmentation using Laserscanner and Video Sequences," in *DAGM 2004, Pattern Recognition, 26th DAGM Symposium*, ser. Lecture Notes in Computer Science, vol. LNCS 3175, Tuebingen, Germany, 2004, pp. 367–374.

[87] N. Kaempchen, M. Buehler, and K. Dietmayer, "Feature-Level Fusion for Free-Form Object Tracking using Laserscanner and Video," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

[88] N. Kaempchen and K. C. J. Dietmayer, "Data synchronization strategies for multi-sensor fusion," in *Proceedings of* 10^{th} *World Congress on Intelligent Transportation Systems*, Madrid, Spain, November 2003.

[89] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transaction of the ASME — Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.

[90] N. Kämpchen, M. Bühler, K. Dietmayer, and U. Lages, "Verfahren zur Kalibrierung eines Abstandsbildsensors," Patent application, DE102004033114A1, EP1615047, Juli 2004.

[91] R. Karlsson, J. Jansson, and F. Gustafsson, "Model-based Statistical Tracking and Decision Making for Collision Avoidance Application," Department of Electrical Engineering, Linkoping University, Tech. Rep. LiTH-ISY-R-2599, March 2004.

[92] T. Kato, Y. Ninomiya, and I. Masaki, "An Obstacle Detection Method by Fusion of Radar and Motion Stereo," in *IEEE Intelligent Vehicles Symposium*, Tokyo, Japan, 2001, pp. 37–42.

[93] A. Kirchner, *Sensordatenverarbeitung eines Laserscanners für autonome Fahrfunktionen von Kraftfahrzeugen*, ser. Fortschritt-Berichte VDI : Reihe 12, Verkehrstechnik, Fahrzeugtechnik. VDI-Verlag, 2000, no. 438, ISBN: 3-18-343812-7.

[94] A. Kirchner and C. Ameling, "Integrated Obstacle and Road Tracking using a Laser Scanner," in *IEEE Intelligent Vehicles Symposium*, Dearborn, USA, October 2000.

[95] L. A. Klein, *Data and Sensor Fusion: A Tool for Information Assessment and Decision Making*. SPIE Press, 2004, ISBN 0-8194-5435-4.

[96] S. Kopischke, "Entwicklung einer Notbremsfunktion mit Rapid Prototying Methoden," Ph.D. dissertation, Technische Universität Braunschweig, Braunschweig, Germany, 2000.

[97] B. Krebs, "Probabilistische Erkennung von 3d Freiformobjekten mit Bayesschen Netzen," Ph.D. dissertation, Technische Universität Braunschweig, Shaker Verlag, 2000.

[98] U. Lages and V. Willhoeft, "Determination of the orientation of an opto-electronic sensor," Patent application, WO2003/087873, 2003.

[99] U. S. Lages, *Untersuchung zur aktiven Unfallvermeidung von Kraftfahrzeugen*, ser. Fortschritt-Berichte VDI : Reihe 12, Verkehrstechnik, Fahrzeugtechnik. VDI-Verlag, 2001, no. 446, ISBN: 3-18-3444612-X.

[100] K. Langwieder, J. Gwehenberger, and T. Hummel, "Sicherheitsgewinn durch ESP — Internationale Erfahrung aus Realunfällen," in *Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, March 2004.

[101] K. Levenberg, "A method for the solution of certain non-linear problem in least squares," *Quarterly Of Applied Mathematics*, vol. 2, pp. 164–168, 1944.

[102] X. R. Li, *Multitarget-Multisensor Tracking: Applications and Advances*. Artech House, 2000, vol. 3, ch. 10, pp. 499–567.

[103] D. T. Linzmeier, D. Vogt, R. Prasanna, M. Mekhaiel, and K. C. Dietmayer, "Probabilistic Signal Interpretation Methods for a Thermopile Pedestrian Detection System," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

[104] M. Maehlisch, R. Schweiger, W. Ritter, and K. Dietmayer, "Multisensor Vehicle Tracking with the Probability Hypothesis Density Filter," in *Proceedings of ISIF/IEEE* 9^{th} *International Conference on Information Fusion*, Florence, Italy, 2006.

[105] M. Maehlisch, R. Schweiger, W. Ritter, and K. C. J. Dietmayer, "Sensorfusion Using Spatio-Temporal Aligned Video and Lidar for Improved Vehicle Detection," in *IEEE Intelligent Vehicles Symposium*, Tokyo, Japan, June 2006.

[106] M. Maehlisch, W. Ritter, and K. C. J. Dietmayer, "ACC Vehicle Tracking with Joint Multisensor Multitarget Filtering of State and Existence," *PReVENT Fusion Forum e-Journal*, vol. 1, pp. 37–43, 2006. [Online]. Available: www.prevent-ip.org

[107] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM Journal on Applied Mathematics*, vol. 11, pp. 431–441, 1963.

[108] T. MathWorks, *'normplot' in the manual of Matlab, release 14 with service pack 3.*, 2005.

[109] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: a survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 103–123, 1998.

[110] E. B. Meier and F. Ade, "Tracking Cars in Range Images Using the Condensation Algorithm," in *International IEEE Conference on Intelligent Transportation Systems*, Tokyo, Japan, 1999, pp. 129–134.

[111] R. Möbus and U. Kolbe, "Multi-Target Multi-Object Tracking, Sensor Fusion of Radar and Infrared," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004.

[112] K. Naab, "Sensorik- und Signalverarbeitungsarchitekturen für Fahrerassistenz und Aktive Sicherheit," in *Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, March 2004.

[113] W. Niehsen, R. Garnitz, M. Weilkes, and M. Stämpfle, *Fahrerassistenzsysteme.* Springer Verlag, 2004, ch. Informationsfusion für Fahrerassistenzsysteme.

[114] W. Niehsen and M. Müller, "IMM-Tracking-Filter für Fahrerassistenzsysteme," in *FAS Workshop Fahrerassistenzsysteme*, September 2003.

[115] PCO, http://www.pco.de, December 2006.

[116] A. Polychronopoulos, Ed., *PReVENT Fusion Forum e-Journal*, September 2006. [Online]. Available: www.prevent-ip.org

[117] R. Putzinger, "Auswirkungen von Fahrerassistenzsystemen auf die Automobilproduktion," in *Elektronik im Kraftfahrzeug 2005 — 12. Internationaler Kongress*, ser. VDI-Berichte/VDI-Tagungsbände, no. 1907, Baden-Baden, Germany, 2005.

[118] T. Röhrig-Gericke, "Objektfusion Radarsensorik," in *1. Fachforum Kraftfahrzeugtechnik Sensorik für Fahrerassistenz*, Heilbronn, Germany, September 2006.

[119] B. Rössler, M. A. Obojski, M. Stanzel, D. Ehmanns, T. Kosch, M. Hopstock, C. Shooter, A. Heenan, P. Woolley, C. Benson, J.-F. Boissou, Y. Martail, L. Serezant, Y. Page, C. Chauvel, and J. M. Kay Fürstenberg, "PReVENT, INTERSAFE SP Deliverable D40.4. Requirements for intersection safety applications." 2005. [Online]. Available: http://www.prevent-ip.org/

[120] C. Royere, D. Gruyer, and V. Berge-Cherfaoui, "Data association with believe theory," in *Proceedings of ISIF/IEEE* 3^{th} *International Conference on Information Fusion*, Paris, France, July 2000.

[121] M. Schäfer, "Untersuchung zur Nutzung der Echopulsbreiten eines Laserscanners," Student research project, Universität Ulm, Ulm, Germany, 2002.

[122] M. Schamberger, "Fahrerassistenzsysteme — wo stehen wir heute, was sind die Potentiale von morgen," in *FAS Workshop Fahrerassistenzsysteme*, Walting, Germany, 2005.

[123] J. Schick, "Gleichzeitige Erkennung von Form und Bewegung durch Rechnersehen," Ph.D. dissertation, Universität der Bundeswehr, Munich, Germany, 1992.

[124] B. Schiele, "Bestimmung des Auslösezeitpunkts einer Automatischen Notbremse," Student research project, Universität Ulm, December 2005, supervisor: N. Kämpchen.

[125] C. Schmidt, F. Oechsle, and W. Branz, "Untersuchungen zu letztmöglichen ausweichmanövern für stehende und bewegte hindernisse," in *FAS Workshop Fahrerassistenzsysteme*, April 2005, pp. 129–138.

[126] K. H. Schönfeld, "Statische und dynamische Kippgrenzen von Sattelkraftfahrzeugen," *Automobiltechnische Zeitschrift*, vol. 82, pp. 119–124, March 1980.

[127] R. Schweiger, H. Neumann, and W. Ritter, "Multiple-cue data fusion with particle filters for vehicle detection in night view automotive applications," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

[128] N. Shimomura, K. Fujimoto, T. Oki, and H. Muro, "An Algorithm for Distinguishing the Types of Objects on the Road Using Laser Radar and Vision," in *IEEE Intelligent Vehicles Symposium*, Tokyo, Japan, 2001, pp. 31–36.

[129] S. Simon, W. Niehsen, A. Klotz, and B. Lucas, "Video-basierte Objekt-Detektion und Verifikation von Radar-Objekt-Hypothesen für Komfort- und Sicherheitsfunktionen," in *FAS Workshop Fahrerassistenzsysteme*, Walting, Germany, 2005.

[130] R. Sittler, "An Optimal Data Association Problem in Surveillance Theory," *IEEE Trans. on Military Electronics*, vol. MIL-8, pp. 125–139, 1964.

[131] A. G. Skibicki, "Positions- und Breitenbestimmung von Fahrzeugen in Laserscanner- und Videobildsequenzen," Diploma thesis, Universität Ulm, March 2004, supervisor: N. Kämpchen.

[132] smart microwave sensors GmbH, http://www.smartmicro.de, June 2006.

[133] J. Sparbert, K. Dietmayer, and D. Streller, "Lane Detection and Street Type Classi.cation using Laser Range Images," in *International IEEE Conference on Intelligent Transportation Systems*, Oakland, USA, August 2001.

[134] Statistisches Bundesamt, DESTATIS, "Verkehrsunfälle 2005," October 2006.

[135] F. Stein, "Efficient Computation of Optical Flow Using the Census Transform," in *DAGM 2004, Pattern Recognition, 26th DAGM Symposium*, ser. Lecture Notes in Computer Science, vol. LNCS 3175, Tuebingen, Germany, 2004, pp. 79–86.

[136] G. P. Stein, O. Mano, and A. Shashua, "Vision-based ACC with a Single Camera: Bounds on Range and Range Rate Accuracy," in *IEEE Intelligent Vehicles Symposium*, Columbus, USA, June 2003.

[137] B. Steux, C. Laurgeau, L. Salesse, and D. Wautier, "Fade : A Vehicle Detection and Tracking System Featuring Monocular Color Vision and Radar Data Fusion," in *IEEE Intelligent Vehicles Symposium*, Versaille, France, 2002.

[138] D. Streller, "Multi-Hypothesen-Ansatz zur Erkennung und Verfolgung von Objekten in Verkehrsszenen mit Laserscannern," Ph.D. dissertation, Universität Ulm, 2005.

[139] D. Streller and K. Dietmayer, "Multiple Hypothesis Classification with Laser Range Finders," in *International IEEE Conference on Intelligent Transportation Systems*, Washington DC, USA, October 2004.

[140] D. Streller, K. Dietmayer, and J. Sparbert, "Object Tracking in Traffic Scenes with Multi-Hypothesis Approach using Laser Range Images," in *Proceedings of* 8^{th} *World Congress on Intelligent Transportation Systems*, Syndney, Australia, October 2001.

[141] T. Strobel and A. Servel, "PReVENT, ProFusion IP Deliverable. Sensor Data Sheets — State-of-the-art of Sensors and Sensor Data Fusion for Automotive Preventive Safety Applications." 2004. [Online]. Available: http://www.prevent-ip.org/

[142] T. Strobel, A. Servel, C. Coue, and T. Tatschke, "PReVENT, ProFusion IP Deliverable. Compendium on Sensors — State-of-the-art of Sensors and Sensor Data Fusion for Automotive Preventive Safety Applications." 2004. [Online]. Available: http://www.prevent-ip.org/

[143] D. Stüker, "Heterogene Sensordatenfusion zur robusten Objektverfolgung im automobilen Straßenverkehr," Ph.D. dissertation, Carl von Ossietzki-Universität Oldenburg, Oldenburg, Germany, 2004.

[144] Z. Sun, G. Bebis, and R. Miller, "On-Road Vehicle Detection Using Optical Sensors: A Review," in *International IEEE Conference on Intelligent Transportation Systems*, Washington, D.C., USA, October 2004.

[145] T. Takahama, T. Kimura, E. Iwasaki, and G. Naitou, "Robust Vehicle Detection for Collision Avoidance Systems by Using a Lidar Sensor and a Vision Sensor," in *Proceedings of* 10^{th} *World Congress on Intelligent Transportation Systems*, Madrid, Spain, November 2003.

[146] H. Takizawa, K. Yamada, and T. Ito, "Vehicles Detection Using Sensor Fusion," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004.

[147] J. Thiem and M. Mühlenberg, "Datafusion of two Driver Assistance System Sensors," in *Advanced Microsystems for Automotive Applications*, Berlin, Germany, 2005.

[148] R. van der Merwe and E. Wan, "Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models," in *Proceedings of the Workshop on Advances in Machine Learning*, Montreal, Canada, June 2003.

[149] R. van der Merwe, "Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models," Ph.D. dissertation, OGI School of Science & Engineering, Oregon Health & Science University, Portland, OR, USA, April 2004.

[150] V. von Holt and M. Maurer, "Aktive Sicherheitssysteme mit maschineller Wahrnehmung - Anforderungen, Potentiale und Einführungshemmnisse," in *Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, March 2004.

[151] A. von Reyher, M. Raab, and H. Winner, "Erweiterte Umgebungserkennung auf der Basis von Lidar-Sensorik," in *21. Internationale VDI/VW-Gemeinschaftstagung: Integrierte Sicherheit und Fahrerassistenzsysteme*, ser. VDI-Berichte, vol. 1864, Okotber 2004.

[152] A. von Reyher, A. Joos, and H. Winner, "A Lidar-Based Approach for Near Range Lane Detection," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

[153] W. von Seelen, U. Handmann, T. Kalinke, C. Tzomakas, and M. Werner, "Image Processing for Driver Assistance," in *Proceedings of the 20. DAGM-Symposium Mustererkennung*, 1998, pp. 11–22.

[154] A. Vukotich and A. Kirchner, "Sensor fusion for driver-assistance-systems," in *Elektronik im Kraftfahrzeug*, Baden-Baden, Germany, 2001.

[155] E. A. Wan and R. van der Merwe, *Kalman Filtering and Neural Networks*. Wiley Publishing, 2001, ch. Chapter 7 : The Unscented Kalman Filter.

[156] C. Wang, C. Thorpe, and S. Thrun, "Online Simultaneous Localization and Mapping with Detection and Tracking of Moving Objects: Theory and Results from a Ground Vehicle in Crowded Urban Areas," in *IEEE Int. Conf. on Robotics and Automation*, May 2003.

[157] C.-C. Wang, "Simultaneous localization, mapping and moving object tracking," Ph.D. dissertation, Carnegie Mellon University, Pittsburgh, USA, April 2004.

[158] C.-C. Wang and C. Thorpe, "A Hierarchical Object Based Representation for Simultaneous Localization and Mapping," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, September 2004.

[159] K. Weiss, N. Kaempchen, and A. Kirchner, "Multiple-Model Tracking for the Detection of Lane Change Maneuvers," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004.

[160] K. Weiss, H. Philipps, and M. Mauthner, "Gütekriterien und Testprozeduren für die Sensordatenfusion in Fahrerassistenzsystemen," in *Elektronik im Kraftfahrzeug*

2005 — 12. Internationaler Kongress, ser. VDI-Berichte/VDI-Tagungsbände, no. 1907, Baden-Baden, Germany, 2005.

[161] T. Weiss, J. Huchler, and K. Dietmayer, "Generierung präziser digitaler Verkehrsinfrastrukturkarten mit Laserscannern," in *22. Internationale VDI/VW-Gemeinschaftstagung: Integrierte Sicherheit und Fahrerassistenzsysteme*, ser. VDI-Berichte, Wolfsburg, Germany, 2006.

[162] T. Weiss, N. Kaempchen, and K. Dietmayer, "Precise Ego-Localization in Urban Areas using Laserscanner and High Accuracy Feature Maps," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

[163] T. Weiss, J. Spruck, and K. Dietmayer, "A Scalable Sensor Fusion Framework for the Localization of a Vehicle on Detailed Digital Maps Using Laserscanners," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Heidelberg, Germany, September 2006.

[164] H. Weisser and R. Bergholz, "Adjustment and alignment method of separation or distance sensors on any motor vehicle using an array of reference markers arranged around the vehicle at varying angular positions," Patent application, DE10116278 A1, 2001.

[165] G. Welch and G. Bishop, "An Introduction to the Kalman Filter," University of North Carolina at Chapel Hill, Department of Computer Science, Chapel Hill, NC, USA, Tech. Rep. TR95-041, 1995.

[166] S. Wender, S. Clemen, N. Kaempchen, and K. C. J. Dietmayer, "Vehicle Detection with Three Dimensional Object Models," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Heidelberg, Germany, September 2006.

[167] S. Wender and K. C. J. Dietmayer, "An Adaptable Object Classification Framework," in *IEEE Intelligent Vehicles Symposium*, Tokyo, Japan, June 2006.

[168] S. Wender, T. Weiss, K. Fuerstenberg, and K. C. J. Dietmayer, "Object Classification exploiting High Level Maps of Intersections," in *Advanced Microsystems for Automotive Applications*, Berlin, Germany, April 2006.

[169] S. Wender, M. Schoenherr, N. Kaempchen, and K. Dietmayer, "Classification of Laserscanner Measurements at Intersection Scenarios with Automatic Parameter Optimization," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, Nevada, USA, June 2005.

[170] V. Willhoeft, K. Fuerstenberg, and R. Krzikalla, "Calibration method," Patent application, WO2002/091095 A1, 2002.

[171] D. Wisselmann, K. Gresser, H. Spannheimer, K. Bengler, and A. Huesmann, "ConnectedDrive - ein methodischer Ansatz für die Entwicklung," in *Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, March 2004.

[172] Q. Zhang and R. Pless, "Extrinsic Calibration of a Camera and Laser Range Finder," in *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2004.

[173] Z. Zhang, "Iterative Point Matching for Registration of Free-Form Curves," Institut National de Recherche en Informatique et en Automatique, Tech. Rep., March 1992.

[174] ——, "Iterative point matching for registration of free-form curves and surfaces," *International Journal of Computer Vision*, vol. 13, no. 2, pp. 119–152, October 1994.

[175] L. Zhao and C. Thorpe, "Qualitative and Quantitative Car Tracking from a Range Image Sequence," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1998*, Santa Barbara, CA, June 1998, pp. 496–501.

[176] D. Zittlau, "Modulares Konzept zur Realisierung zukünftiger videobasierter Fahrerassistenzsysteme," in *Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, March 2004.

Supervised Student Research Projects and Diploma Theses

B. Schiele, “Bestimmung des Auslösezeitpunkts einer Automatischen Notbremse,” Student research project, Universität Ulm, December 2005.

T. Bühler, “Objektverfolgung unter Nutzung der Echopulsbreiten eines mehrzeiligen Laserscanners,” Diploma thesis, Universität Ulm, 2005.

M. Bühler, “Simultane Schätzung des dynamischen Zustandes und der Form von Objekten mit einem mehrzeiligen Laserscanner,” Diploma thesis, Universität Ulm, April 2005.

J. Kibbel, “Fahrspurerkennung in Videobildsequenzen und Entfernungsprofilen eines mehrzeiligen Laserscanners,” Diploma thesis, Universität Ulm, March 2004, supervisors: K. Fürstenberg and N. Kämpchen.

U. Plöckl, “Fahrspurerkennung in Videobildsequenzen,” Diploma thesis, Universität Ulm, May 2004.

T. Weiss, “Globale Positionsbestimmung eines Fahrzeugs durch Fusion von Fahrzeug- und GPS-Daten zur Erstellung einer digitalen Karte,” Diploma thesis, Universität Ulm, March 2004, supervisors: K. Fürstenberg and N. Kämpchen.

A. G. Skibicki, “Positions- und Breitenbestimmung von Fahrzeugen in Laserscanner- und Videobildsequenzen,” Diploma thesis, Universität Ulm, March 2004.

M. Schönherr, “Objektverfolgung im Nahbereich anhand von Videobildsequenzen für aktive Sicherheitssysteme,” Student research project, Universität Ulm, June 2004.

M. Zocholl, “Räumlich-zeitliche Segmentierung in Laserscanner- und Videobildsequenzen,” Diploma thesis, Universität Ulm, February 2004.

M. Schäfer, “Hochdynamische Objektverfolgung mit Multi-Modell-Ansätzen,” Diploma thesis, Universität Ulm, Ulm, Germany, November 2003.

M. Bühler, “Verfahren zur automatisierten Kalibrierung von Laserscanner und Videokamera,” Student research project, Universität Ulm, September 2003.

M. Zocholl, “Synchronisiertes Abspeichern der Daten mehrerer Fahrzeug-Sensoren in Echtzeit,” Student research project, Universität Ulm, October 2002

Publications and Patent Applications

Publications

S. Wender, S. Clemen, N. Kaempchen, and K. C. J. Dietmayer, "Vehicle Detection with Three Dimensional Object Models," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Heidelberg, Germany, September 2006.

N. Kaempchen, M. Clauss, Y. Guenter, R. M. Schreier, M. Stiegeler, K. Tischler, K. Dietmayer, H. P. Grossmann, H. Kabza, H. Neumann, A. Rothermel, and C. Stiller, "Vernetzte Fahrzeug-Umfelderfassung für zukünftige Fahrerassistenzsysteme," in *FAS Workshop Fahrerassistenzsysteme*, Walting, Germany, 2005.

N. Kaempchen, M. Buehler, and K. Dietmayer, "Feature-Level Fusion for Free-Form Object Tracking using Laserscanner and Video," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

K. Dietmayer, N. Kaempchen, K. Fuerstenberg, and J. Kibbel, "Roadway Detection and Lane Detection using Multilayer Laserscanner," in *Advanced Microsystems for Automotive Applications*, Berlin, Germany, 2005.

S. Wender, M. Schoenherr, N. Kaempchen, and K. Dietmayer, "Classification of Laserscanner Measurements at Intersection Scenarios with Automatic Parameter Optimization," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, Nevada, USA, June 2005.

T. Weiss, N. Kaempchen, and K. Dietmayer, "Precise Ego-Localization in Urban Areas using Laserscanner and High Accuracy Feature Maps," in *IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.

T. Weiss, N. Kaempchen, and K. Dietmayer, "High Accuracy Ego-Localization in Complex Dynamic Environments using Laserscanners," in *Proceedings of* 12^{th} *World Congress on Intelligent Transportation Systems*, San Francisco, USA, November 2005.

K. Tischler, M. Clauss, Y. Guenter, N. Kaempchen, R. M. Schreier, and M. M. Stiegeler, "Networked Environment Description for Advanced Driver Assistance Systems," in *International IEEE Conference on Intelligent Transportation Systems*, 2005.

N. Kaempchen, K. Fuerstenberg, A. Skibicki, and K. Dietmayer, "Sensor Fusion for Multiple Automotive Active Safety and Comfort Applications," in *Advanced Microsystems for Automotive Applications*, Berlin, Germany, March 2004, pp. 137–163.

N. Kaempchen, K. Fuerstenberg, and K. Dietmayer, "Ein Sensorfusionssystem für Automotive Sicherheits- und Komfortapplikationen," in *Aktive Sicherheit durch Fahrerassistenz*, Garching, Germany, March 2004.

N. Kaempchen, K. Weiss, M. Schaefer, and K. Dietmayer, "IMM Object Tracking for High Dynamic Driving Maneuvers," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004.

N. Kaempchen and K. Dietmayer, "IMM vehicle tracking for traffic jam situations on highways," in *Proceedings of ISIF/IEEE* 7^{th} *International Conference on Information Fusion*, Stockholm, Schweden, June 2004.

N. Kaempchen, M. Zocholl, and K. Dietmayer, "Spatio-temporal Segmentation using Laserscanner and Video Sequences," in *DAGM 2004, Pattern Recognition, 26th DAGM Symposium*, ser. Lecture Notes in Computer Science, vol. LNCS 3175, Tuebingen, Germany, 2004, pp. 367–374.

N. Kaempchen and K. Dietmayer, "Fusion of Laserscanner and Video for Advanced Driver Assistance Systems," in *Proceedings of* 11^{th} *World Congress on Intelligent Transportation Systems*, Nagoya, Japan, Oktober 2004.

N. Kaempchen and K. Dietmayer, "Fusion von Laserscanner und Video für aktive Sicherheits- und Komforanwendungen im Automobil," in *21. Internationale VDI/VW-Gemeinschaftstagung: Integrierte Sicherheit und Fahrerassistenzsysteme*, ser. VDI-Berichte, vol. 1864, Okotber 2004.

K. Dietmayer, A. Kirchner, and N. Kaempchen, *Fahrerassistenzsysteme*. Springer Verlag, 2004, ch. Fusionsarchitekturen zur Umfeldwahrnehmung für zukünftige Fahrerassistenzsysteme.

K. Weiss, N. Kaempchen, and A. Kirchner, "Multiple-Model Tracking for the Detection of Lane Change Maneuvers," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 2004.

N. Kaempchen and K. C. J. Dietmayer, "Data synchronization strategies for multi-sensor fusion," in *Proceedings of* 10^{th} *World Congress on Intelligent Transportation Systems*, Madrid, Spain, November 2003.

N. Kaempchen, U. Franke, and R. Ott, "Stereo vision based pose estimation of parking lots using 3D vehicle models," in *IEEE Intelligent Vehicles Symposium*, Versaille, France, 2002.

J. Carmena, N. Kaempchen, D. Kim, and J. Hallam, "Artificial ears for a biomimetic sonarhead: from multiple reflectors to surfaces," *Artificial Life*, vol. 7, no. 2, pp. 147–169, 2001.

Patent Applications

N. Kämpchen, M. Bühler, K. Dietmayer, and U. Lages, "Verfahren zur Kalibrierung eines Abstandsbildsensors," DE102004033114A1, EP1615047, Juli 2004.

N. Kämpchen, U. Lages, K. Fürstenberg, and K. Dietmayer, "Verfahren zur Erkennung und/oder Verfolgung von Objekten," DE102004018813A1, EP1589484, April 2004.

N. Kämpchen, K. Fürstenberg, K. Dietmayer, and U. Lages, "Verfahren zur Erkennung von gekennzeichneten Gefahr- und/oder Baustellen im Bereich von Fahrbahnen," DE102004003848A1, EP1557692, January 2004.

K. Fürstenberg, N. Kämpchen, U. Lages, and K. Dietmayer, "Verfahren zur Klassifizierung von Objekten," DE102004003870A1, EP1557694, January 2004.

K. Fürstenberg, N. Kämpchen, U. Lages, and K. Dietmayer, "Verfahren zur Verfolgung von Objekten," DE102004003868A1, EP1557693, January 2004.

K. Fürstenberg, N. Kämpchen, U. Lages, K. Dietmayer, and V. Willhoeft, "Verfahren zur Erkennung von Markierungen auf einer Fahrbahn," DE102004003850A1, EP1557691, January 2004.

N. Kämpchen, U. Lages, and K. Dietmayer, "Verfahren zur gemeinsamen Verarbeitung von tiefenaufgelösten Bildern und Videobildern," DE10312249A1, EP1460454, March 2003.