

# **Vehicle localization with enhanced robustness for urban automated driving**

zur Erlangung des akademischen Grades eines

**Doktors der Ingenieurwissenschaften**

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

**genehmigte**

**DISSERTATION**

von

**Jan Rohde**

aus Hamburg

Tag der mündlichen Prüfung: 17.01.2018  
Erster Gutachter: Prof. Dr.-Ing. J. Marius Zöllner  
Zweiter Gutachter: Prof. Dr.-Ing. Rüdiger Dillmann



## Danksagung

Diese Dissertation wurde während meiner Forschungsarbeit in der Abteilung Fahrzeugsicherheit und Assistenzsysteme im Zentralbereich Forschung (CR/AEV) der Robert Bosch GmbH verfasst. Die nachfolgenden Absätze widme ich denjenigen Menschen die mich durch den damit verbundenen Entstehungsprozess begleitet haben.

Professor Dr.-Ing. J. Marius Zöllner, Direktor für Technisch-Kognitive Assistenzsysteme (TKS) am karlsruher Forschungszentrum für Informatik (FZI), danke ich herzlichst für die wissenschaftliche Betreuung und die Übernahme des Hauptreferats. Dir, Marius, bin ich dankbar für die richtungsweisende und couragierte Begleitung meiner wissenschaftlichen Arbeit von der ersten Vision bis hin zu einem erfolgreichen Abschluss der Promotionsprüfung. Deinen direkten und dabei respektvollen Umgang habe ich sehr zu schätzen gelernt. Professor Dr.-Ing. Rüdiger Dillmann, Leiter des Humanoids and Intelligence Systems Lab (HIS) am Karlsruher Institut für Technologie (KIT) und Direktor am FZI, danke ich zutiefst für die Übernahme des Ko-referats. Zusätzlich zu dem wissenschaftlichen Austausch während der regelmäßigen Klausurtagungen, waren unsere Diskussionen zur gesellschaftlichen Bedeutung der Wissenschaft eine große Bereicherung und Motivation für mich.

Weiterer großer Dank gilt Dr. Holger Mielenz. Als Projektleiter hast Du, Holger, Erhebliches zur Gestaltung und Umsetzung eines ausgezeichneten Arbeitsumfelds beigetragen. Ohne dieses wäre ein Erfolg der Arbeit um einiges unwahrscheinlicher gewesen. Als wissenschaftlicher Betreuer und Mentor hast Du mir in der wissenschaftlichen Arbeit und darüber hinaus geholfen meine Ziele zu erkennen, sowie diese zielgerichtet und fokussiert zu verfolgen. Ich denke gerne an die intensiven fachlichen Diskussionen zurück.

Für die enge und freundschaftliche Zusammenarbeit danke ich meinen Doktorandenkollegen Holger Banzhaf, Ulrich Baumann, Jan Stellet und Benjamin Völz. Es freut mich, dass unser Zusammenhalt die Doktorandenzeit überdauert hat. Besonderer Dank gilt auch Thomas Schamm und Steffen Knoop für unzählige fachliche Diskussionen und moralische Unterstützung. Während der Entstehung dieser Dissertation hast Du, Thomas, mit erheblichem Aufwand korrekturgelesen und bei der Erarbeitung des roten Fadens beraten. Zudem danke ich Peter Biber für zahlreiche Nachmittage mit leidenschaftlichen Diskussionen über Themen der Robotik und ergebnisreiche Programmierleinlagen. Für die menschliche und produktive Arbeitsatmosphäre in der CR/AEV und beim Bosch danke ich meinen Kollegen. Axel Stamm möchte ich für die nicht selbstverständlichen Freiheiten während des Entstehungsprozesses dieser Dissertation danken. Für ihre tatkräftige Unterstützung im Rahmen ihres Studiums danke ich Marc Oliver Höfling, Inga Jatzkowski, Markus Jesenski, Tanja Kempfert, Liana Lattin, Benedikt Mehler, Marc Scott Meißner und Laura Volkenborn. Bei dem gesamten TKS Team bedanke ich mich für die wie selbstverständlich erscheinende freundliche Aufnahme und die interessanten wissenschaftlichen Diskussionen.

Weit über den Rahmen dieser Doktorarbeit hinaus bin ich meinen Eltern, meiner Verlobten und meinem baldigen Schwager unermesslich dankbar. Ihr gabt mir schon immer die nötige Freiheit und Orientierung im Leben. Euer bedingungsloser Rückhalt gibt mir die Gewissheit und Kraft für den nächsten Schritt. Ohne euch gäbe es diese Dissertation vermutlich nicht.

Stuttgart, im Februar 2018

*Jan Rohde*

## Zusammenfassung

Automatisiertes Fahren wird weithin als ein wichtiger Beitrag zur Lösung heutiger und zukünftiger Herausforderungen im Bereich des Personenverkehrs sowie des Waren- und Gütertransports angesehen. Insbesondere im Kontext zunehmend beschleunigter Urbanisierung bietet die Automatisierung Ansatzpunkte zur Verkehrsflussoptimierung und Verwirklichung innovativer Mobilitätskonzepte. Aktuelle Entwicklungen in automatisierten Fahrzeugsystemen für städtische Umgebungen zeichnen sich durch einen steigenden Automatisierungsgrad und sinkende Einschränkungen an die Betriebsumgebung aus. Diese Trends führen zu erhöhten Anforderungen an das maschinelle Situationsverständnis, welches zumeist durch die Anreicherung mit zusätzlichen Karteninformationen, wie z.B. Verkehrsregeln, ergänzt wird. Eine kartenrelative Fahrzeuglokalisierung verbindet diese räumlich referenzierten Karteninformationen mit der fahrzeugrelativen Umfeldwahrnehmung auf Basis der im Fahrzeug verbauten Umfeldsensorik. Folglich ist eine hinreichend genaue und zuverlässige Fahrzeuglokalisierung von fundamentaler Bedeutung für den sicheren Betrieb automatisierter Fahrzeugsysteme. Herkömmliche, rein GPS-basierte Lokalisierungssysteme wie sie in Navigationssystemen verwendet werden, erfüllen die gestellten Anforderungen nicht mit der erforderlichen Zuverlässigkeit. Aufgrund von Störgrößen wie Abschirmungseffekten und Mehrwegeausbreitung ist die Signalqualität häufig einer starken Degradation unterworfen. Vor diesem Hintergrund wird in der vorliegenden Arbeit ein System zur Fahrzeuglokalisierung auf der Grundlage einer Assoziations- und Transformationsschätzung zwischen Messungen einer laserbasierten Sensorik (Light Detection And Ranging - LiDAR) zur Umfelderkennung und hochgenauen Lokalisierungskarten entwickelt. Der Vorgang zur Assoziations- und Transformationsschätzung wird nachfolgend als Registrierung bezeichnet. Degradationseffekte GPS-basierter Systeme werden durch die Bebauung in urbanen Bereichen begünstigt. Aus diesem Grund wird ein Schwerpunkt auf die Unabhängigkeit von GPS Messungen gelegt. Zur zielgerichteten Entwicklung von Lokalisierungssystemen auf Basis von Genauigkeitsanforderungen und Umfeldeigenschaften werden statistische Modelle erarbeitet. Beide Kernbeiträge werden im Folgenden näher ausgeführt.

Der erste Kernbeitrag dieser Arbeit besteht in der Entwicklung eines robusten Verfahrens für die Registrierung von Sensorbeobachtungen und potentiell veralteten Lokalisierungskarten. Robustheit ist hier definiert als die Fähigkeit eines Lokalisierungssystems, spezifizierte Leistungsanforderungen, wie z.B. Genauigkeitsanforderungen, auch unter widrigen Umfeldbedingungen zu erfüllen. Letztere treten in städtischen Umgebungen besonders häufig aufgrund semi-statischer Objekte auf. Zu dieser Klasse von Objekten zählen insbesondere geparkte Fahrzeuge entlang des Fahrbahnrandes. Sobald diese in der Lokalisierungskarte oder der Sensormessung verbleiben, kommt es zu Inkonsistenzen, welche spezieller Berücksichtigung bedürfen. Städtische Umgebung ist weiterhin durch ein hohes Maß an Hetero-

genität, z.B. verursacht durch Variationen zwischen breiten Haupt- und schmalen Wohnstraßen mit hohen Vegetationsmengen, gekennzeichnet. Das zu diesem Zwecke entwickelte Registrierungsverfahren basiert auf einer im Bereich der Bildverarbeitung etablierten Methode auf Basis der Fourier-Mellin Transformation. Dieses beruht auf der Fast Fourier Transform (FFT) für welche eine Vielzahl effizienter und deterministischer Hard- und Softwareimplementierungen existiert. Damit ist deren Einsatz auch für echtzeitfähige Fahrzeugsysteme von großem Interesse. In mehreren Schritten wird diese Methode auf den beschriebenen Anwendungsfall angepasst und erweitert. Die vorgenommenen Erweiterungen führen im Vergleich zu bekannten Methoden zu einer signifikant erhöhten Robustheit gegenüber Inkonsistenzen und Unsicherheiten in der initialen Transformationsschätzung. Das entwickelte Lokalisierungssystem verfügt über analytische Redundanz durch die Verwendung des zuvor beschriebenen spektralen Registrierungsverfahrens und eines pfeilerbasierten Registrierungsverfahrens. Für letzteres werden pfeilerförmige Objekte, z.B. von Straßenlaternen und Lichtsignalanlagen, aus LiDAR Messungen extrahiert und mit der Lokalisierungskarte assoziiert. Konsistenzprüfungen werden durchgeführt, um Mengen konsistenter Fahrzeugposeninformationen auszuwählen. Konsistente Informationen von der Registrierung und der odometriebasierten Eigenbewegungsschätzung des Fahrzeugsystems werden anschließend in einem erweiterten Kalman Filter fusioniert. Eine zusätzlich entwickelte LiDAR-basierte Odometrie erhöht die Präzision der Eigenbewegungsschätzung. Der Betrieb des Fahrzeugsystems in nicht kartierten Bereichen, wie z.B. privaten Parkplätzen, wird mit einem vorgeschlagenen spektralen Simultaneous Mapping and Localization (SLAM) Verfahren ermöglicht. Dieses Verfahren arbeitet auf der Basis von LiDAR Scans und benötigt auch in anspruchsvollen Umgebungen keine weiteren Messdaten. Als weiterer wichtiger Aspekt der Fahrzeuglokalisierung wird die Initialisierung behandelt. Dazu wird ein partikelfilterbasierter Ansatz mit Konsistenzprüfung entwickelt. Basierend auf der verfügbaren Rechenleistung des Fahrzeugsystems wird ein Schwellwert für die Unsicherheit der initialen Posenschätzung definiert. Der spektrale Registrierungsansatz wird bei dessen Unterschreitung verwendet. Dieser ermöglicht eine hochgenaue Posenschätzung unter signifikanter initialer Unsicherheit der Transformationsparameter. Nach erfolgreicher Registrierung wird in den regulären Betrieb des Lokalisierungssystems gewechselt.

Als zweiter zentraler Beitrag wird ein anforderungsorientierter und modellbasierter Ansatz für die Entwicklung von Systemen zur Fahrzeuglokalisierung erarbeitet. In den meisten Anwendungsfällen automatisierter Fahrzeugsysteme müssen spezifische Anforderungen an die Lokalisierungsgenauigkeit eingehalten werden. Diese stehen in einem komplexen Zusammenhang mit den Eigenschaften des gegebenen Betriebsumfeldes. Zu den Eigenschaften zählen auch die Dichte von Umfeldmerkmalen welche zur Lokalisierung verwendet werden können, Ressourcenbeschränkung und Parameter des zur Umfelderkennung verwendeten Perzeptionssystems. Zur Abbildung dieser Zusammenhänge wird ein geschlossen analytisches Modell hergeleitet, welches es erlaubt Parameter von Lokalisierungssystemen für eine gegebene Lokalisierungs-

genauigkeit und Betriebsumfeld zu bestimmen. Zu diesem Zweck werden Lokalisierungssysteme auf Basis von Bayes Filtern wie dem Kalman Filter betrachtet. Als Ausgangspunkt der Fehlerfortpflanzung wird daher eine modifizierte algebraische Riccati Gleichung ausgewählt, welche Fehlerwahrscheinlichkeiten bei der Registrierung berücksichtigt. Eine Schätzung des Registrierungsfehlers des spektralen Registrierungsverfahrens wird aus der verwendeten Gitterauflösung abgeleitet. Für den Pfeilerbasierten Ansatz wird dazu eine Fehlerfortpflanzung über die Lösung des Prokrustes Problems durchgeführt. Die Lösung des Prokrustes Problems ermöglicht bei korrekter Datenassoziation die Berechnung der räumlichen Transformation zwischen zwei Eingangsmengen von Pfeilerkoordinaten. Für die Beschreibung der Unsicherheit in der LiDAR-basierten Messung der Pfeilerkoordinaten wird ein geschwindigkeitsabhängiges Modell hergeleitet. Die Vereinigung dieser Modelle ermöglicht unter den gegebenen Annahmen eine statistische Beschreibung einer Klasse relevanter Lokalisierungssysteme. Eine effiziente Auswertung des Modells wird durch die geschlossenen analytische Formulierung ermöglicht, welche zu einer Senkung der Simulationszeiten gegenüber vergleichbaren numerischen Methoden führt. Die Ergebnisse aus Simulationsstudien werden in einem iterativen Vorgehen für die Entwicklung des zuvor genannten Lokalisierungssystems verwendet.

Eine ausführliche Auswertung der entwickelten Lokalisierungsverfahren wird auf der Basis eines herausfordernden und speziell ausgewählten Satzes von LiDAR- und Odometriedaten aus 14.8 km Testfahrten in städtischem Umfeld durchgeführt. In diesem sind Messungen aus einer Vielzahl unterschiedlicher städtischen Bereichen enthalten, so z.B. dicht beparkte Wohnstraßen und ausgedehnte Kreuzungen. Dieser Datensatz wird um Messdaten von 4 km einer Campus Umgebung ergänzt. Die Ergebnisse zeigen, dass durch die beschriebenen Ansätze das Lokalisierungssystem im Rahmen der untersuchten Anwendungsfälle erhöhten Robustheitsanforderungen gerecht wird. Für das vorgeschlagene Verfahren wird zusätzlich gezeigt, dass die Abhängigkeit von gründlich aktualisierten Lokalisierungskarten und Sensordatenverarbeitung in vielen relevanten Anwendungen reduziert werden kann. Die Korrektheit und Genauigkeit der statistischen Modelle wird an mehreren Stellen durch reale Ergebnissen des Perzeptions- und Lokalisierungssystems und Monte Carlo Simulationen bestätigt. Der zentrale Beitrag dieser Arbeit besteht damit in der Entwicklung eines modularen Lokalisierungssystems mit integrierten Konsistenzprüfungen. Durch den modellbasierten Ansatz können Entwicklungsentscheidungen unterstützt werden, welche zu Lokalisierungssystemen mit erweiterter Robustheit für automatisiertes Fahren in der Stadt führen.



## Abstract

Automated driving contributes to the solution of current and future challenges in personal mobility and transportation. In the context of increasing urbanization, it provides new means for traffic flow optimization and enables innovative mobility concepts. Current developments in automated vehicle systems for urban environments are characterized by progressively increasing levels of automation and decreasing restrictions on operation environments. These trends lead to higher requirements for situational awareness that is commonly fostered by the incorporation of supplementary map information like right of way rules. Vehicle localization links these spatially referenced map information to the vehicle-relative on-board environment perception. Consequently, the safe operation of automated vehicle systems depends on the underlying localization system.

Given this motivation, a light detection and ranging (LiDAR) based localization algorithm with enhanced robustness is developed in this thesis. Robustness is here defined as the ability of a localization system to meet specified performance requirements even under adverse environmental conditions. The latter often occur in urban environments, e.g., due to semi-static objects like parked cars which can cause localization maps to be outdated. The environment is also characterized by high heterogeneity as exemplarily seen from variations between broad main and narrow domestic roads with high amounts of vegetation. Contributions are made by the development of a robust map matching procedure for dealing with outdated localization maps and localization system initialization. As another effect, the dependence on thoroughly updated localization maps and sensor data filtering is generally relaxed. Analytical redundancy is introduced to the modular localization system by utilization of a pole-based map matching modality for intersections and robust spectral matching algorithm for the remaining areas. Consistency checking capability is introduced in order to select sets of consistent vehicle pose information. A developed scan-based odometry permits precise motion estimation and a simultaneous localization and mapping (SLAM) algorithm enable the operation in unmapped areas.

In this work, a requirement-driven approach is taken for localization system design. In most applications, specific localization accuracy requirements have to be met while complex interrelations between environment characteristics, localization requirements, resource constraints and perception system parameters complicate the design process. To this point, a model-based framework for the design of localization systems is developed as another key contribution. Therefore, a statistical localization system model is derived which allows to determine localization system parameters for given localization requirements. An efficient evaluation of the model is fostered by its analytical formulation, leading to a significantly reduced simulation time in comparison to numerical methods. The results from a simulation study based on the derived model are considered within the iterative design of the proposed localization system with enhanced robustness for urban automated driving.



# Contents

<b>Nomenclature</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Problem statement and research questions . . . . .	1
1.2 Key contributions . . . . .	5
1.3 Thesis outline . . . . .	7
<b>I Preliminaries</b>	<b>9</b>
<b>2 Common concepts</b>	<b>11</b>
2.1 Localization . . . . .	11
2.1.1 Localization maps . . . . .	12
2.1.2 Registration . . . . .	13
2.1.3 Bayesian state estimation . . . . .	14
2.1.4 Consistency checking and change detection . . . . .	17
2.1.5 Cramér-Rao bound . . . . .	19
2.2 Simultaneous localization and mapping (SLAM) . . . . .	20
2.3 Dependability, robustness and automation levels . . . . .	22
<b>3 Related work</b>	<b>25</b>
3.1 Registration . . . . .	25
3.1.1 Point-based registration . . . . .	26
3.1.2 Dense registration . . . . .	27
3.2 Localization and SLAM . . . . .	28
3.3 Localization system modeling . . . . .	33
3.4 Conclusion and open research questions . . . . .	35
<b>II Components of vehicle localization systems for urban automated driving (UAD)</b>	<b>37</b>
<b>4 Platform and operation environment</b>	<b>39</b>
4.1 Test vehicle . . . . .	39
4.2 Operation environment . . . . .	41
4.3 Data sets . . . . .	43
<b>5 LiDAR data processing</b>	<b>45</b>
5.1 Processing of 3D LiDAR scans . . . . .	45
5.1.1 Generation of 2D LiDAR scans . . . . .	47
5.1.2 Detection of poles . . . . .	48

5.2	Processing of the localization map . . . . .	49
5.3	Generation of a vehicle pose reference . . . . .	50
<b>6</b>	<b>Registration of cluttered point clouds</b>	<b>52</b>
6.1	Point cloud registration for UAD . . . . .	52
6.2	Spectral registration . . . . .	56
6.2.1	Accuracy improvement and estimation . . . . .	58
6.2.2	3D spectral registration . . . . .	61
6.2.3	Experimental evaluation: runtime and accuracy . . . . .	62
6.3	Robust spectral registration . . . . .	65
6.3.1	Limitations of single-hypothesis spectral registration . . . . .	65
6.3.2	Multiple hypotheses evaluation . . . . .	67
6.3.3	Confidence measures . . . . .	70
6.4	Conclusion . . . . .	71
<b>III</b>	<b>Model-based localization system design</b>	<b>75</b>
<b>7</b>	<b>Measurement and environment models</b>	<b>77</b>
7.1	Environment formalization . . . . .	78
7.1.1	Stochastic distribution models . . . . .	80
7.1.2	Estimation of model parameters . . . . .	81
7.2	Model of LiDAR-based pole detection . . . . .	82
7.2.1	Experimental evaluation . . . . .	84
7.3	Stereo video camera perception model . . . . .	85
<b>8</b>	<b>Localization system models</b>	<b>86</b>
8.1	Map matching model . . . . .	86
8.1.1	Solution of the orthogonal Procrustes problem . . . . .	86
8.1.2	Error propagation . . . . .	89
8.1.3	Experimental evaluation . . . . .	90
8.2	Vehicle pose filtering model . . . . .	92
8.2.1	A closed-form upper bound for Kalman Filtering with inter- mittent observation . . . . .	92
8.2.2	Simulative evaluation . . . . .	96
8.3	Occlusions and missed detections . . . . .	97
<b>9</b>	<b>Model-based design</b>	<b>99</b>
9.1	Localization requirements . . . . .	101
9.2	Matching rate and accuracy requirements . . . . .	102
9.2.1	Interrelation between matching rate and accuracy . . . . .	103
9.2.2	Odometry accuracy . . . . .	104
9.3	Landmark distribution requirements . . . . .	106
9.3.1	Detector selection . . . . .	108

9.3.2	Influence of the sensor field of view . . . . .	109
9.4	Derivation of sensor parameters . . . . .	112
9.4.1	Landmark detection accuracy . . . . .	113
9.4.2	Intrinsic sensor parameters for robust localization . . . . .	114
9.5	Conclusion . . . . .	117
<b>IV Vehicle localization with enhanced robustness for UAD</b>		<b>119</b>
<b>10 LiDAR-based odometry and SLAM</b>		<b>121</b>
10.1	Scan-based odometry . . . . .	121
10.1.1	Experimental evaluation . . . . .	122
10.2	Spectral SLAM . . . . .	127
10.2.1	Experimental evaluation . . . . .	128
10.3	Conclusion . . . . .	132
<b>11 LiDAR-based vehicle localization</b>		<b>133</b>
11.1	Vehicle pose prior generation . . . . .	134
11.2	Precise vehicle localization . . . . .	140
11.2.1	Scan-based map matching . . . . .	140
11.2.2	Landmark-based map matching . . . . .	141
11.2.3	Hybrid vehicle localization . . . . .	142
11.3	Experimental evaluation . . . . .	144
11.3.1	Vehicle pose initialization and recovery . . . . .	145
11.3.2	Multilayer adaptive Monte Carlo localization . . . . .	147
11.3.3	Hybrid vehicle localization . . . . .	149
11.4	Conclusion . . . . .	153
<b>Conclusion and future work</b>		<b>155</b>
<b>Literature</b>		<b>159</b>



---

## Nomenclature

### Acronyms

---

Acronym	Meaning
<b>AD</b>	automated driving
<b>AMCL</b>	adaptive Monte Carlo localization
<b>CPU</b>	central processing unit
<b>CRB</b>	Cramér-Rao bound
<b>CUSUM</b>	cumulative sum
<b>DFT</b>	discrete Fourier transformation
<b>dGPS</b>	differential global positioning system
<b>EKF</b>	extended Kalman filter
<b>FFT</b>	fast Fourier transformation
<b>FMT</b>	Fourier-Mellin transformation
<b>FoV</b>	field of view
<b>GPS</b>	global positioning system
<b>ICP</b>	iterative closest point
<b>IMU</b>	inertial measurement unit
<b>KFIO</b>	Kalman filtering with intermittent observations
<b>KLD</b>	Kullback-Leibler divergence
<b>LiDAR</b>	light detection and ranging
<b>MARE</b>	modified algebraic Riccati equation
<b>MCL</b>	Monte Carlo localization
<b>ML-AMCL</b>	multilayer adaptive Monte Carlo localization
<b>NDT</b>	normal distribution transform
<b>SLAM</b>	simultaneous localization and mapping
<b>OPP</b>	orthogonal Procrustes problem
<b>PF</b>	particle filter
<b>PPP</b>	Poisson point process
<b>RAIM</b>	receiver autonomous integrity monitoring
<b>RANSAC</b>	random sample consensus
<b>ROS</b>	robot operating system
<b>SCGP</b>	single cluster graph partitioning
<b>SNR</b>	signal-to-noise ratio
<b>SPP</b>	stochastic point process
<b>SPRT</b>	sequential probability ratio testing
<b>UAD</b>	urban automated driving

---

## Mathematical notations

Notation	Meaning
$x, y, \dots$	scalar values
$\mathbf{x}, \mathbf{y}, \dots$	column vectors
$\mathbf{A}, \mathbf{B}, \dots$	matrices
$\mathbf{x}^\top, \mathbf{A}^\top$	transpose of a vector and a matrix
$f(\cdot)$	function returning a scalar
$\mathbf{f}(\cdot)$	function returning a vector
$\mathbf{F}(\cdot)$	function returning a matrix
$\mathbf{I}_{3 \times 3}$	$3 \times 3$ identity matrix
$\mathbb{E}[\cdot]$	expected value
$cov(\cdot)$	covariance
$eig(\mathbf{A})$	eigenvalues of matrix $\mathbf{A}$
$max(\cdot)$	maximum value
$trace(\mathbf{A})$	trace of matrix $\mathbf{A}$
$ \cdot $	absolute value
$\ \mathbf{a}\ _2$	$L_2$ norm of $\mathbf{a}$
$(\cdot)$	tuple
$\mathcal{N}(\mathbf{a}; \mu, \Sigma)$	multivariate Gaussian of random variable $\mathbf{a}$ with mean $\mu$ and covariance $\Sigma$
$argmin$	minimizing argument
$\{\cdot\}_{i=i_0}^I$	set with $I - i_0$ members

---

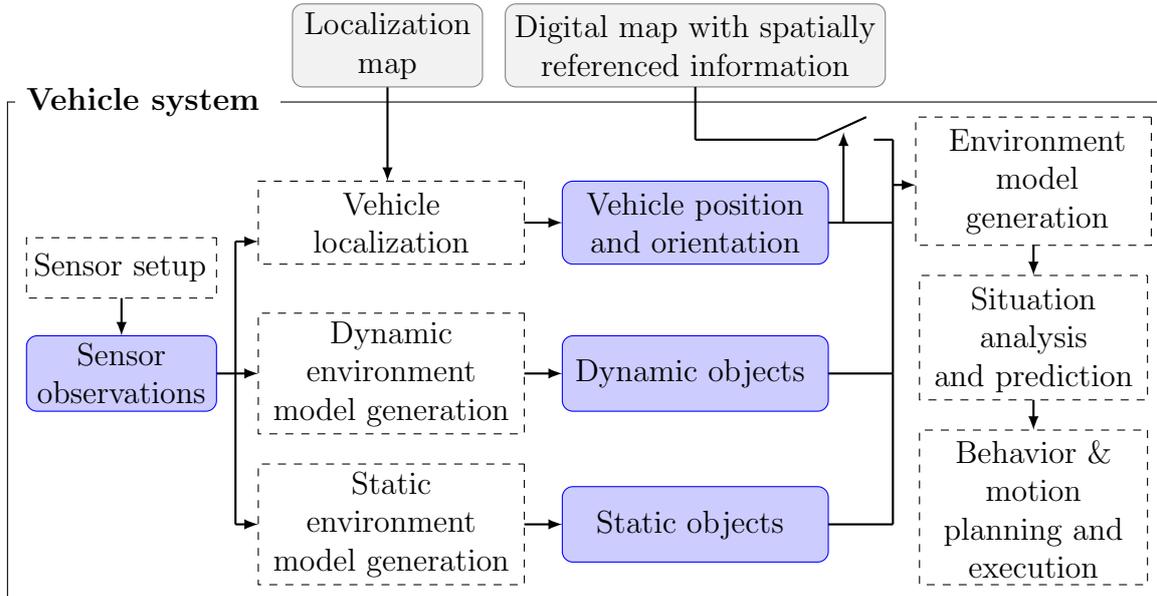
# 1 Introduction

Automated driving (AD) is widely considered a key technology for the solution of current as well as future mobility and transportation challenges. Advanced driver assistance systems like electronic stability control (ESC) and adaptive cruise control (ACC) already nowadays contribute to the reduction of traffic accidents. AD has the potential to further increase driving safety by providing enhanced functional capabilities. An important cause of future challenges is the global trend of urbanization. According to the United Nations' World Urbanization Prospects [oEA14], the percentage of European people living in urban areas is going to increase from 73% in 2014 to 82% in 2050. This development leads to increased traffic density and environmental stress in and around urban areas. At this point, AD enables macroscopic traffic flow optimization and future mobility concepts which could significantly reduce the number of vehicles on the roads. These concepts include mobility on demand services where a customer can get picked up by an automated taxi at a given place at a given time. On the path towards urban automated driving, conceptual and methodical challenges in the development of automated vehicle systems remain to be solved. This thesis is concerned with the solution of some of these challenges.

## 1.1 Problem statement and research questions

An exemplary functional architecture of an automated vehicle system is shown in Figure 1.1. The vehicle system is equipped with a sensor setup which allows to observe its surroundings. Commonly, the sensor measurements are complemented by additional information from digital maps, providing traffic rules or an electronic horizon, covering areas and quantities the sensors cannot observe. These information sources are combined to the environment model, the vehicle system internal representation of its surrounding which includes static objects and other dynamic traffic participants. Before a decision about the vehicle system behavior and motion can be made, the future actions taken by other traffic participants have to be predicted. This step requires situational awareness which can be gained by interpretation of the environment model. For example, an accelerating vehicle driving on the neighboring lane influences the decision of initiating a lane change. Current developments in urban automated driving (UAD) are characterized by progressively increasing levels of automation and decreasing restrictions on operation environments. These trends lead to higher requirements for situational awareness and more frequently necessitate the incorporation of supplementary map information into the environment model. Map-free automated driving would require high perception capabilities of the automated vehicle system which might be infeasible with current affordable sensor systems.

Since, the environment information is typically spatially referenced, i.e. they are only valid for a given location in the map coordinate system, the map-relative pose (position and orientation) of the vehicle has to be known. Determination of this



**Figure 1.1** A simplified version of the automated vehicle system functional architecture outlined in [TKZS16]. Utilization of additional map information in the environment model requires a minimum accuracy of the vehicle position and orientation estimate. Corrupted vehicle pose estimates can lead to a functional degradation of the vehicle system and can have catastrophic impact if localization failures remain undetected.

vehicle pose is denoted map-relative vehicle localization. In a last step, the behavior and motion of the automated vehicle system is planned and executed on the basis of the environment model and vehicle control. Throughout this highly complex procedure from sensor observations to motion execution, vehicle localization is of fundamental importance. This is due to its function of linking the spatially referenced map information and the static as well as dynamic objects that are extracted from the sensor observations. Utilization of additional map information enables well-grounded behavior planning which is paid by the dependence on the underlying localization system for a safe operation of the automated vehicle systems.

Today's navigation systems rely on global positioning system (GPS) signals for vehicle localization during tour guiding. Standard GPS typically provides position information with a typical accuracy between 5 m and 10 m. Centimeter-precise position measurements can be obtained from sophisticated differential global positioning system (dGPS) systems. GPS-based localization has the advantage of directly measurable vehicle positions at nearly every outdoor location. However, GPS-based solutions are considered unsuitable for vehicle localization in urban environments due to signal deteriorating multipath propagation and shadowing. To this point, current localization approaches for environments with unmodified infrastructure commonly rely on highly accurate and up-to-date localization maps (compare

Figure 1.1). Thereby, problems imposed by GPS-based localization can be omitted and no infrastructure modifications are required. Thus, the main research question to be answered throughout this thesis

### **Main research question**

How can GPS-less vehicle localization for urban automated driving be achieved on the basis of an in-vehicle sensor setup and a localization map?

Beginning from this main research question, further relevant research questions arise when looking at the role of localization systems and their components in an automated vehicle system. A fundamental input to the localization system is the localization map to which sensor observations are aligned for vehicle pose estimation. The recording and maintenance of localization maps is complicated due to frequent changes in urban environment and high costs for map updating from measurement collected by specialized mapping vehicle systems. Consequently, a decreased dependence on well-updated localization maps is beneficial. Hence, a promising approach is to decrease the dependence on well maintained localization maps. An outdated localization map and a cluttered LiDAR sensor observation is shown in Figure 1.2 which show the impact changes can have on the level of inconsistencies. This leads to a further research question treated in this thesis:

*rq1.* How can the dependence of localization systems on thoroughly updated localization maps and sensor data processing be relaxed?

A centimeter-precise vehicle position estimate as provided by dGPS systems are commonly only required as a reference. Localization accuracy requirements are highly scenario-dependent and vary within a broad range. In intersection areas, a highly accurate vehicle pose estimate is widely considered mandatory and localization failures might have a catastrophic impact. Infeasible perception and localization system designs increase the likelihood of localization failures which, e.g., can cause an insufficiently low extraction rate of environment information, and have to be avoided. Furthermore, the efficiency of the localization system depends on a plethora of design decisions whose outcomes are strongly coupled to the target operation environment. In this context, this thesis deals with the following research question:

*rq2.* Starting from localization accuracy requirements and operation environment characteristics, how can the localization system design be supported and made repeatable?

For initialization and once a failure state occurred, an initialization procedure has to be executed in order to obtain a precise vehicle pose estimate or maintain a safe vehicle state.



**Figure 1.2** Semi-static objects like building sites and parked cars frequently occur in urban environments. If these remain inside the localization map (grey) or the sensor observation (blue), inconsistencies arise. In combination with high uncertainty in the vehicle pose estimate, as encountered during system initialization and recovery, the alignment of outdated localization maps and sensor observations becomes a challenging task.

*rq3.* How to initialize or recover a vehicle localization system from imprecise priors in challenging urban environments?

False optimism about the current localization accuracy during vehicle operation can have catastrophic impact due to the introduction of false information to the environment model. In this context, the detection of false alignments between sensor observations and the localization map, i.e. map matching failures, is one prerequisite for the avoidance of critical vehicle states

*rq4.* How can map matching failures be detected?

*rq5.* How can adaptations to the localization system architecture lead to increased insensitivity against inconsistent sensor observations?

In parking garages, the high accuracy requirement from the intersection example can most likely be relaxed and an in-vehicle solution for simultaneous localization and mapping (SLAM) might suffice for navigation. Ideally, the additional SLAM capability would not lead to a considerable increase in the complexity of the localization and navigation system.

*rq6.* How to add scan-based odometry and SLAM capability to the vehicle system without significantly increasing the overall system complexity?

The key contributions in this thesis are made in order to answer these research questions. In the following section, an overview over the key contributions is provided.

## 1.2 Key contributions

This thesis contributes to the field of map-relative localization and mapping in the context of urban automated driving. The main contributions are made to LiDAR-based localization systems with enhanced robustness against outdated localization maps and inaccurate priors which are often encountered during system initialization and recovery. The key contributions were presented at international peer-reviewed conferences and in a book chapter. All publications are listed under *Publications* in the *Literature* section.

Real-time capability, determinism, high tolerance against inconsistent sensor observations and imprecise transformation priors are favorable characteristics of map matching algorithms for urban vehicle localization. A thorough evaluation of state of the art map matching algorithms in the context of UAD is conducted and reasonable extensions to the Fourier-Mellin transformation (FMT) based matching algorithm are developed. Integration into a LiDAR-based odometry and SLAM framework leads to high relative motion estimation accuracy. Accurate vehicle localization is achieved by a combination of the developed algorithms in a localization system with consistency checking capability. A further contribution deals with the question of how robustness can be introduced at a higher level of the system architecture. To this point, a particle filter base localization framework with global localization capability is considered and it is shown, how adaptations to the system architecture lead to increased robustness. In this context the following key contribution is made for answering *rq1* and *rq3-rq5*:

*kc1.* A **modular localization system with consistency checking capability** is developed ([RMZ17], [RSMZ16], [RSMZ15]).

- Ad *rq1*: At a lower level of the system architecture, insensitivity against outdated localization maps is achieved by FMT based map matching. Here, a suitable existing algorithm is adapted to the requirements of vehicle systems for UAD.

- Ad *rq3*: Initialization and recovery of vehicle localization systems is achieved by a newly developed robust variant of the FMT based matching algorithm. The evaluation shows a significantly increased insensitivity against inconsistencies between sensor observations and localization maps than the standard algorithm. Due to manifold parameterization possibilities, the algorithm is successfully employed within a variety of localization system components.
- Ad *rq4*: Map matching failures are detected by introduction of analytical redundancy by adding a pole-based map matching procedure to the system architecture and consistency checking.
- Ad *5*: Increased insensitivity of a localization system against inconsistent sensor observations is achieved by incorporation of a consistency checking mechanism into the system architecture. In an example implementation, 3D LiDAR scans are divided into three layers for each of which a separate localizer is instantiated. Consistency checking against odometry measurements is then used to select layers which lead to consistent localization outcomes.
- Ad *rq6*: A spectral SLAM algorithm on the basis of FMT and rFMT based matching is developed. Furthermore, accurate and reliable relative motion estimation is enabled by the developed LiDAR-based odometry.

Furthermore, the model-based design of perception systems for localization is addressed by the development of statistical models and their evaluation in Monte Carlo simulations. It is argued in this thesis, that the design of localization systems should be targeted at the required and not the best achievable performance. This goal shall be achieved in a systematic and repeatable manner. To this point, the proposed approach differs from the vast majority of approaches found in the state of the art. Complex interrelations between environment characteristics, localization requirements, resource constraints and perception system parameters, like sensor setup parameters and feature detectors, complicate the design of localization systems. It is argued, that knowledge of these interrelations can help to identify an advantageous trade-off between robustness against stressful environmental conditions and efficiency of the localization system implementation. It is argued, that statistical models are well suited for the expression of the aforementioned interrelations in an efficient manner compared to numerical models. Hence, the key contribution for answering research question *rq2*:

*kc2*. The design of localization systems is supported and made repeatable on the basis of newly developed statistical models which are evaluated for **model-based localization system design** ([RVMZ16], [RMH<sup>+</sup>16], [RJMZ16]). Error propagation is performed from uncertain sensor observations and over feature extraction, map matching and Bayesian ego vehicle pose tracking. The model framework captures the environment characteristics, i.e. feature densities and distributions, by utilization of stochastic point processes. This enables the determination of minimum feature detection rates, feature detector selection, sensor parameters and map matching rates for a given localization accuracy requirement and environment.

### 1.3 Thesis outline

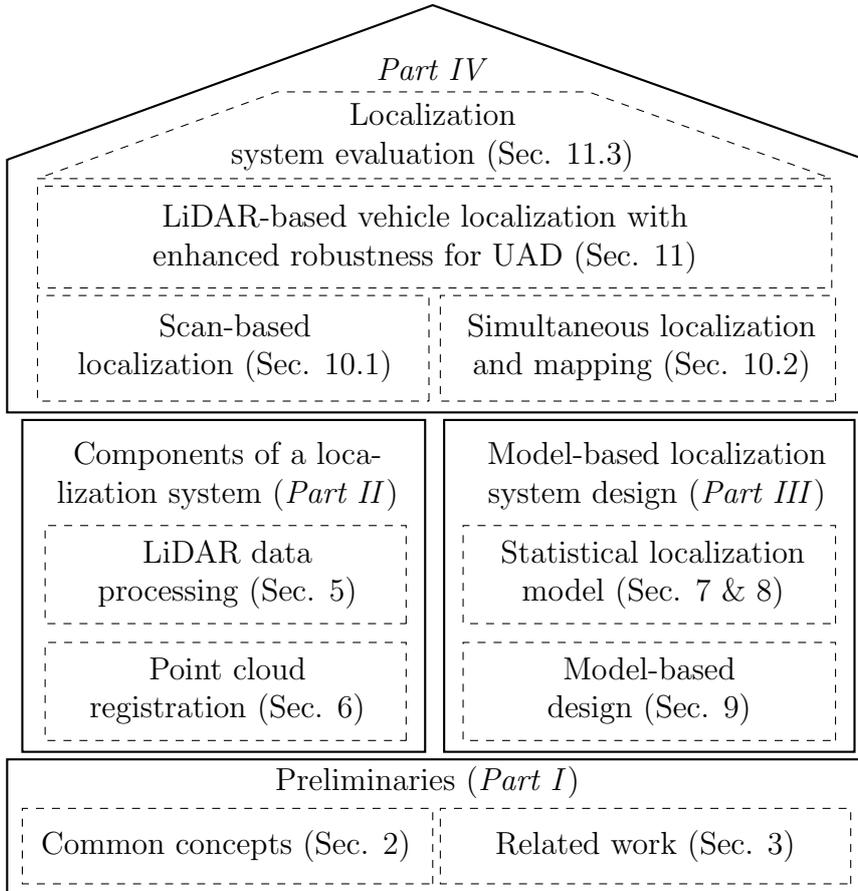
The thesis at hand is organized in four parts. A graphical overview is given in Figure 1.3. *Part I* provides the knowledge base required throughout this thesis. It contains definitions of relevant terms and the mathematical foundations (Section 2) of following sections. The terms dependability and robustness are introduced and considered for vehicle localization in connection with increasing automation levels. A comprehensive summary of related work in the treated fields of research is developed in Section 3. A discussion of open research challenges arising from the state of the art ends this part and serves as the motivation for the key contributions developed in this thesis.

The components of the proposed localization system are introduced in *Part II* (*rq1*, *rq3*, *rq4*, *kc1*). This part starts with the introduction of the test vehicle system, the sensor setup and a short overview over characteristics of the urban operation environment. Consecutively, the concept of spectral registration is presented. Its applicability to vehicle localization in urban environments is discussed based on a thorough experimental comparison to other state of the art matching algorithms (Section 6). A robust variant of the spectral matching algorithm is developed in order to overcome the limitations of the standard algorithm (Section 6.3). This development enables higher matching rates and initialization from inaccurate priors.

Statistical models of the localization system and their interrelations with environment characteristics are derived and experimentally evaluated in *Part III* (*rq2*, *kc2*). The statistical models are developed in Section 7 & 8. Comprehensive simulation studies and evaluations based on real measurement data are discussed and principles for robustness-oriented localization system design are formulated (Section 9).

The localization system is developed in *Part IV* (*rq1* - *rq6*, *kc1*). This part begins with the development of a LiDAR-based odometry which makes use of spectral registration. The latter is instrumented in a SLAM framework where the matching approach is successfully utilized for scan matching and loop closing (Section 10). Insensitivity to inconsistent sensor observations is added at architectural level to a particle filter (PF) based localization approach, denoted adaptive Monte Carlo localization (AMCL), and results in the implementation of multilayer adaptive Monte

Carlo localization (ML-AMCL) (Section 11.1). In the overall localization system, ML-AMCL is utilized as a robust global prior (Section 11). Under the assumption of inaccurate priors for the vehicle pose and highly inconsistent sensor observations, a suitable extended Kalman filter (EKF) based localization framework is developed (Section 11.2). Analytical redundancy is introduced by utilization of scan- and pole-based map matching algorithms with broad basins of convergence. False matching results are detected by consistency checking mechanisms. A concluding discussion of the results presented in this thesis and future research perspectives are given in Section 11.4.



**Figure 1.3** Outline of the thesis at hand.

# **Part I**

## **Preliminaries**

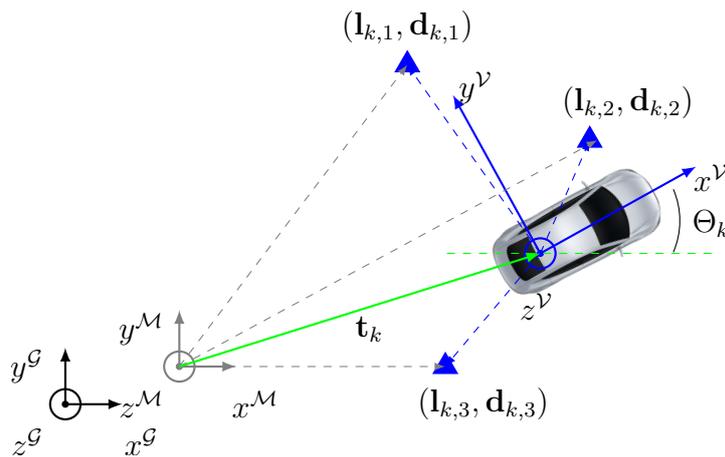


## 2 Common concepts

### 2.1 Localization

A sufficiently accurate estimate of the current map-relative vehicle pose is considered mandatory for urban automated driving. The vehicle pose  $\mathbf{x}_k$  at time  $k$  contains a position and orientation of the vehicle frame  $\mathcal{V}$  relative to the map reference frame  $\mathcal{M}$  (Figure 2.1). Commonly,  $\mathcal{M}$  is aligned with the global (e.g. WGS48) frame  $\mathcal{G}$ . Vehicle localization is then defined as the task of estimating  $\mathbf{x}_k$  from uncertain sensor observations.

Current navigation systems provide global pose information on the basis of global positioning system (GPS) measurements. However, this approach is widely considered insufficiently accurate and reliable for vehicle localization, especially in urban environments where multi-path propagation and shadowing effects frequently occur. Thus, vehicle localization is commonly based on exteroceptive, e.g. light detection and ranging (LiDAR) and video, and proprioceptive, e.g. wheel speed and steering angle, measurements. Therefore, localization systems entail a perception module, a localization map (Section 2.1.1) and estimation algorithms. Localization approaches can be categorized by the level of uncertainty in the utilized prior. In global localization the initial pose is unknown. This approach is particularly challenging due to ambiguities and computational complexity. If not stated differently, the availability of a rough pose estimate, also denoted prior, is assumed. Map-relative pose estimates are calculated from the registration (Section 2.1.2) of uncertain, i.e. noisy and cluttered, exteroceptive sensor measurements and the localization map. Relative motion



**Figure 2.1** Vehicle reference frame  $\mathcal{V}$  (blue) in accordance to DIN ISO 8855 [ISO12]. Localization names the process of aligning  $\mathcal{V}$  with the map frame  $\mathcal{M}$  (gray) based on the matching of sensor observation and a localization map. It is commonly assumed, that the transformation between  $\mathcal{M}$  and the global (e.g. WGS48) frame  $\mathcal{G}$  is known which enables the incorporation of GPS measurements if desired.

estimates can be utilized for further refinement of absolute pose estimates and for temporary unavailability of registration results. Relative pose estimation is subject to drift and can therefore be used for limited distances.

Fusion of absolute and relative pose information is often conducted by Bayesian filtering which is described in Section 2.1.3 and used throughout this thesis. Inconsistent vehicle states can be detected by consistency checking methods (Section 2.1.4). The accuracy in dynamic state estimation is bounded by the Cramér-Rao bound introduced in Section 2.1.5 which forms an important basis for the derivation of statistical localization system models in later sections.

### 2.1.1 Localization maps

A localization map  $\mathbf{m}$  is a model of the operation environment. Environment information is typically represented as sensor-specific sets of objects like points and lines or in form of a grid map. A feature-based map contains a set of  $M$  features  $\{\mathbf{m}_i\}_{i=1}^M$  which can be observed by an appropriate perception system.

The utilized feature types have significant influence on the computational efficiency and hardware requirements. Furthermore, it affects the selection of suitable registration and localization algorithms. On the lowest level of abstraction, accumulated sensor measurements like point clouds serve as localization maps, yielding dense environment representations. As pointed out in [PCS15], this map representation can increase the localization reliability. In order to enable a sparser representation, geometric primitives like cylinders or plains can be extracted from sensor measurements and used as features. A feature point might then represent a reference point on a pole-shaped object such as a traffic sign post. In contrast to LiDAR measurement points, the location of the pole can be repeatably measured. Repeatably detectable features are hereafter denoted as landmarks. Some definitions in literature are narrower in the sense that some additional uniqueness requirement is imposed upon the repeatably detectable feature. The extraction of sparsely distributed geometric primitives also leads to an increased uniqueness of the features. In the case of pole landmarks this means that a reference point of the pole is repeatably detectable but different landmarks are not distinguishable. Further meta-information about the point feature type and possibly about its surrounding increases this uniqueness. Descriptors  $d_i$  are commonly utilized to describe the local neighborhood of a feature point, leading to an increased uniqueness. Examples for descriptors are ORB [RRKB11] for video and FLIRT [TBA14] for LiDAR measurements. For artificial landmarks, QR codes can be used for achieving globally uniqueness which simplifies the correspondence estimation step during registration (Section 2.1.2).

A common assumption about localization maps is that these exclusively contain representations of static environment entities. Due to suboptimal mapping procedures, semi-static objects like parked cars and building sites, as well as artifacts from dynamic objects or processing steps can be contained in the map. This clutter leads to inconsistent sensor readings and hence threatens correct functioning of vehicle

localization systems. The goal of localization systems with enhanced robustness is the consideration of such inconsistencies. Registration algorithms (Section 2.1.2) enable the alignment of sensor observations and localization maps.

### 2.1.2 Registration

Registration, also denoted as matching, is utilized in order to align a set of  $N_k$  noisy point measurements  $\mathbf{L}_k = \{\mathbf{l}_i\}_{i=1}^{N_k}$  with covariance matrix  $\Sigma_{\mathbf{l},i}$  and a second data set. Without loss of generality, it is assumed that  $\mathbf{L}_k$  is given relative to  $\mathcal{V}$ . Absolute vehicle poses are estimated by map matching which is the task of matching a local sensor observation with a localization map. The result of map matching is the map-relative vehicle pose estimate  $\mathbf{z}_k = [x_k, y_k, \theta_k]$  that forms the transformation between  $\mathcal{V}$  and  $\mathcal{M}$  (Figure 2.1):

$$\begin{aligned} \mathbf{L}_k^{\mathcal{M}} &= \mathbf{R}_k \mathbf{L}_k^{\mathcal{V}} + \mathbf{t}_k = \mathbf{R}_k \mathbf{L}_k + \mathbf{t}_k, \\ \mathbf{R}_k &= \begin{bmatrix} \cos(\theta_k) & -\sin(\theta_k) \\ \sin(\theta_k) & \cos(\theta_k) \end{bmatrix}, \mathbf{t}_k = \begin{bmatrix} x_k & y_k \end{bmatrix}. \end{aligned} \quad (2.1)$$

Registration of consecutive sensor observations  $\mathbf{L}_{k-1}$  and  $\mathbf{L}_k$  can be used to describe the movement of the ego vehicle. The transformation parameters are obtained from a two step procedure which consists of correspondence estimation (data association) and transformation calculation (see Figure 2.1). A correspondence between a map feature  $m_i$  and sensor observation  $l_j$  is modeled as a correspondence variable  $c_j = i$ , leading to a set of correspondence pairs  $\{\{\mathbf{l}_j, \mathbf{m}_{c_j}\}\}_{j=1}^{n_k}$ ,  $n_k \leq N_k$  [TBF05]. The implementation of this step has significant influence on the robustness and computational efficiency of the registration algorithm (compare Section 3.1). For an estimated set of correspondences, the transformation parameters are obtained by optimization of a cost function  $J$

$$(\mathbf{R}, \mathbf{t}) = \arg \min_{\mathbf{R}, \mathbf{t}} J(\{\{\mathbf{l}_i, \mathbf{m}_i\}\}_{i=1}^n, (\mathbf{R}, \mathbf{t})) . \quad (2.2)$$

For known correspondences, the problem of minimizing quadratic cost function  $J$  is known as the orthogonal Procrustes problem (OPP). Correspondence pairs can be weighted in order to account for measurement noise. The weighted cost function is given as

$$J = \sum_{i=1}^n \|\mathbf{W}_i (\mathbf{R}\mathbf{l}_i + \mathbf{t} - \mathbf{m}_i)\|^2 . \quad (2.3)$$

In literature, different variants and solutions for the minimization of  $J$ , i.e. the solution of the OPP, can be found. A non-exhaustive overview is given in Table 2.1. The choice of the weighting matrix  $\mathbf{W}_i$  is used to weight the point correspondences. Anisotropic measurement noise can be considered by defining  $\mathbf{W}_i$  as a densely populated or diagonal matrix. By choosing  $\mathbf{W}_i^T \mathbf{W}_i = \Sigma_{\mathbf{l},i}^{-1} + \mathbf{R} \Sigma_{\mathbf{m}_i}^{-1} \mathbf{R}^T$ , the most accurate

**Table 2.1** Different formulations and solutions of the orthogonal Procrustes problem.

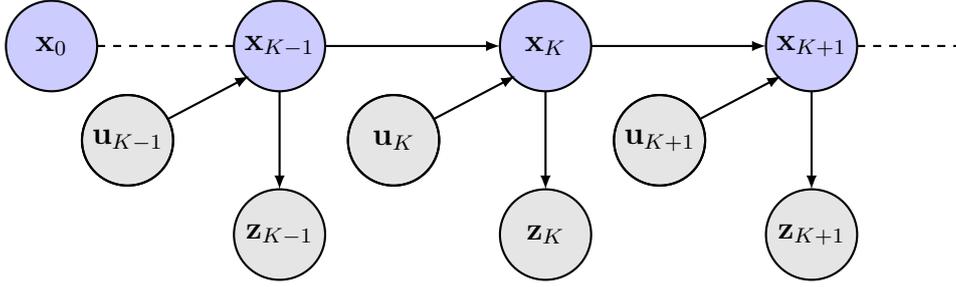
Weighting strategy	Solution
<ul style="list-style-type: none"> <li>• <math>\mathbf{W}_i = \frac{1}{n}\mathbf{I}</math> (unweighted)</li> <li>• <math>\mathbf{W}_i = w_i\mathbf{I}</math> (scalar-weighted (isotropic noise))</li> </ul>	closed-form <ul style="list-style-type: none"> <li>• singular value decomposition (SVD) [AHB87]</li> <li>• unit quaternions [Hor87]</li> <li>• dual quaternions [WSV91]</li> </ul>
<ul style="list-style-type: none"> <li>• <math>\mathbf{W}_i = \Sigma_{\mathbf{l},i}^{-1} + \mathbf{R}\Sigma_{\mathbf{m},i}^{-1}\mathbf{R}^\top</math> (matrix-weighted (anisotropic noise))</li> </ul>	closed-form (approximate) and iterative <ul style="list-style-type: none"> <li>• approximate [WCR92]</li> <li>• iterative [CT98], [KS91]</li> </ul>

registration results can be obtained. For  $\mathbf{W}_i = w_i\mathbf{I} = \max(\text{eig}(\Sigma_{\mathbf{l},i}))^{-\frac{1}{2}}$  the scalar-weighted and for  $\mathbf{W}_i = n^{-1}\mathbf{I}$  the unweighted Procrustean problem is obtained. Both formulations can be solved efficiently in closed-form. For high numbers of landmarks, the optimization results of all three cost functions are similar.

The notation  $\mathbf{L}_k$  will be used for sets of pole-shaped features extracted from LiDAR data, whereas scans are denoted  $\mathbf{s}_k$  instead. The covariance of the registration result is denoted  $\Sigma_{\mathbf{z}}$ . The correctness of the correspondences is a fundamental prerequisite for the validity of  $\Sigma_{\mathbf{z}}$ . Accurate covariance estimation is an important prerequisite for a successful deployment of matching algorithms in localization and SLAM frameworks. Together with the estimates vehicle pose, the covariance estimate is used for vehicle state estimation as described in Section 2.1.3.

### 2.1.3 Bayesian state estimation

Localization is an example of state estimation in the context of automated driving. In this thesis, dynamic state estimation is considered in a probabilistic manner by application of Bayesian filtering. The contents in this section are based on *Thrun et al.* [TBF05]. The vehicle state is modeled as a random variable  $\mathbf{X}$  which allows to consider state uncertainties. The probability of the value  $\mathbf{X} = \mathbf{x}$  is described by the probability density function (PDF)  $p(\mathbf{X} = \mathbf{x})$ . In the following, the shorter common notation  $p(\mathbf{x})$  is used instead. Bayesian filtering uses this representation in order to propagate vehicle state estimates over time. The Markov assumption is fundamental to Bayesian filtering. Accordingly, all future states only depend on the current state and, hence, are independent of all past states. Consequently, the vehicle state at time step  $k$  is completely represented by the belief  $\text{bel}(\mathbf{x}_k)$ . Based on the Markov assumption, Bayes filters are utilized to recursively update the belief function, also denotes as posterior distribution or posterior, given a set of uncertain sensor measurements and the posterior from the preceding time step. The evolution



**Figure 2.2** Under the Markov assumption, the dynamic Bayesian network [TBF05] models the evolution of the vehicle states  $\mathbf{x}_k$  (blue) in dependence on noisy odometry (or control inputs)  $\mathbf{u}_k$  and map matching results  $\mathbf{z}_k$ .

and causality are modeled as a dynamic Bayesian network as shown in Figure 2.2. A detailed derivation of the Bayes filter under utilization of the algorithm theorem of total probability and the Bayes rule is given in [TBF05].

The Bayes filtering algorithm consists of a prediction (Equation 2.5) and a measurement update (Equation 2.7) step [TBF05].

$$\overline{\mathbf{bel}}(\mathbf{x}_k) = p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}, \mathbf{u}_{1:k}) \quad (2.4)$$

$$= \int p(\mathbf{x}_k \mid \mathbf{u}_k, \mathbf{x}_{k-1}) \mathbf{bel}(\mathbf{x}_{k-1}) d\mathbf{x}_{k-1} \quad (2.5)$$

$$\mathbf{bel}(\mathbf{x}_k) = p(\mathbf{x}_k \mid \mathbf{z}_{1:k}, \mathbf{u}_{1:k}) \quad (2.6)$$

$$= \eta p(\mathbf{z}_k \mid \mathbf{x}_k) \overline{\mathbf{bel}}(\mathbf{x}_k) . \quad (2.7)$$

Vehicle motion is modeled by the motion model  $p(\mathbf{x}_k \mid \mathbf{u}_k, \mathbf{x}_{k-1})$ . The measurement model  $p(\mathbf{z}_k \mid \mathbf{x}_k)$  captures the interrelation between sensor observations and the absolute vehicle state. In order to obtain a probability density distribution, the left side of Equation 2.7 is multiplied by the normalization constant  $\eta$ . The Bayes filter equations can only be implemented for finite state spaces which is not the case for vehicle localization in open environments. Nonetheless, approximate implementations can be derived and have gained exceptional popularity. The *Kalman filter* (KF) is an efficient Bayes filter implementation which models the belief function as a multivariate normal distribution (Gaussian) with mean  $\boldsymbol{\mu}$  and covariance  $\mathbf{P}$  [TBF05]

$$\mathbf{bel}_{KF}(\mathbf{x}_k) = p(\mathbf{x}_k) \det(2\pi\boldsymbol{\Sigma}_k)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{x}_k - \boldsymbol{\mu}_k)^\top \mathbf{P}_k^{-1}(\mathbf{x}_k - \boldsymbol{\mu}_k)\right) . \quad (2.8)$$

The major disadvantage of this representation is the fact that noise is modeled as zero-mean Gaussian and only one hypothesis about the current vehicle state can be represented. This unimodality can be problematic in the presence of ambiguous sensor observations. Several extensions to the KF equations exist, including the

extended Kalman filter (EKF) for nonlinear system and measurement models which is considered in this thesis. The following nonlinear system model is considered throughout this thesis.

$$\mathbf{x}_{k+1} = g(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k \quad (2.9)$$

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k \quad (2.10)$$

$$\mathbf{u}_k = [\Delta x, \Delta y, \Delta \theta] \quad (2.11)$$

$$\mathbf{x}_{k+1|k} = g(\mathbf{x}_{k|k}, \mathbf{u}_k, \mathbf{0}) \quad (2.12)$$

$$= \begin{pmatrix} x_k \\ y_k \\ \theta_k \end{pmatrix} + \begin{pmatrix} \mathbf{R} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta \theta \end{pmatrix} \end{pmatrix} \quad (2.13)$$

$$= \begin{pmatrix} x_k + \Delta x \cos(\theta) - \Delta y \sin(\theta) \\ y_k + \Delta x \sin(\theta) + \Delta y \cos(\theta) \\ \theta_k \end{pmatrix} \quad (2.14)$$

From Equation 2.9 and Equation 2.8 the following EKF equations can be derived (a detailed derivation can be found in [TBF05]).

$$\mathbf{P}_{k+1|k} = \mathbf{A}_{k+1} \mathbf{P}_k \mathbf{A}_{k+1}^\top + \mathbf{Q}_{k+1} \quad (2.15)$$

$$\mathbf{S}_k = \mathbf{P}_{k+1|k} \mathbf{C}_k^\top (\mathbf{C}_{k+1} \mathbf{P}_{k+1|k} \mathbf{C}_k^\top + \Sigma_{\mathbf{z},k}) \quad (2.16)$$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{S}_k (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_{k|k-1})) \quad (2.17)$$

$$= \mathbf{A} \mathbf{x}_{k-1|k-1} + \mathbf{K}_k \epsilon_k \quad (2.18)$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{S}_k \mathbf{C}_k) \mathbf{P}_{k|k-1} \quad (2.19)$$

with the system and measurement matrices  $\mathbf{A}$  and  $\mathbf{C}$  given as the Jacobians

$$\mathbf{A} = \frac{\partial g}{\partial \mathbf{x}} (\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k, \mathbf{0}) \quad (2.20)$$

$$= \begin{pmatrix} 1 & 0 & -\Delta x \sin(\theta) - \Delta y \cos(\theta) \\ 0 & 1 & \Delta x \cos(\theta) - \Delta y \sin(\theta) \\ 0 & 0 & 1 \end{pmatrix} \quad (2.21)$$

$$\mathbf{C} = \frac{\partial h}{\partial \mathbf{x}} (\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k, \mathbf{0}) . \quad (2.22)$$

Information fusion at the level of map matching results lead to the simplification  $\mathbf{C} = \mathbf{I}$ .

Multimodality can be achieved by utilization of particle filter (PF) which are a special case of nonparametric filters and represent the belief as a set of  $L$  state samples

$$\mathbf{bel}_{PF}(\mathbf{x}_k) = \mathbf{X}_k \quad \text{with} \quad (2.23)$$

$$\mathbf{X}_k := \left\{ \mathbf{x}_k^{[1]}, \dots, \mathbf{x}_k^{[L]} \right\} . \quad (2.24)$$

The samples are called particles and represent different hypotheses about the current system state. The belief propagation is conducted in a three step procedure:

1. The prediction step for time  $k$  is done for each particle  $\mathbf{x}_{k-1}^{[i]} \in \mathbf{X}_{k-1}$  separately by sampling from the motion model  $p(\mathbf{x}_k | \mathbf{x}_{k-1}^{[i]}, \mathbf{u}_k)$ . The result of this step is a set  $\bar{\mathbf{X}}_k$  of particles which represents  $\overline{\mathbf{bel}}(\mathbf{x}_k)$ .
2. Integration of measurements  $\mathbf{z}_k$  is done by weighting in accordance to the fit between particle and measurement such that  $w_k^{[i]} = p(\mathbf{z}_k | \mathbf{x}_k^{[i]})$ .
3. Resampling is conducted by drawing with replacement of  $L$  particles with a probability given by the particle weight. The result is a set  $\mathbf{X}_k$  of  $L$  unweighted particles which represents  $\mathbf{bel}(\mathbf{x}_k)$ . This step is also denoted as importance sampling [TBF05]. This step can lead to a convergence of all particles to one hypothesis, resulting in an insufficient particle spread. This can be avoided by deployment of adaptive resampling strategy like Kullback-Leibler divergence (KLD) resampling [Fox01].

#### 2.1.4 Consistency checking and change detection

Correctness of the pose estimate  $\hat{\mathbf{x}}_k$  and the covariance  $\mathbf{P}_k$  is commonly expressed as the level of consistency between redundant measurement sources with a stated test strength. *Gustafsson* [Gus00] distinguishes between whiteness-, likelihood ratio, multiple model and algebraical consistency test based change detection. The system model in Equation 2.9 is extended by a deterministic parameter  $\mathbf{f}_k$  which is subject to changes in case of failures

$$\mathbf{x}_k = g(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{f}_k) + \mathbf{w}_{k-1} \quad \mathbf{y}_k = h(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{f}_k) + \mathbf{v}_{k-1} . \quad (2.25)$$

The signal model  $\mathbf{Y}_k$  is given as

$$\mathbf{Y}_k = \mathbf{CZ}_k + \mathbf{V}_k + \mathbf{C}_f \mathbf{F}_k \quad (2.26)$$

where  $\mathbf{Z}_k$  denotes the available pose estimates,  $\mathbf{V}_k$  the measurement noise and  $\mathbf{F}_k$  the faults. Consequently, the residuals can be defined as

$$\epsilon_k = \mathbf{W}(\mathbf{Y}_k - \mathbf{CZ}_k) \quad (2.27)$$

$$= \mathbf{W}\boldsymbol{\nu}_k . \quad (2.28)$$

In the context of Bayesian filtering, consistency is given if [BSLK04]

- state errors are zero-mean Gaussian and the error magnitudes fit the covariance given by the filter,
- innovations have the same characteristics as the state errors,

- innovations have a zero-mean Gaussian noise characteristic.

For online application, the last two conditions can be tested. The second condition can be tested based on the time-average normalized innovation squared statistic [BSLK04]

$$\bar{\epsilon}_{\nu}^2 = \frac{1}{K} \sum_{k=1}^K \boldsymbol{\nu}_k^{\top} \mathbf{S}_k^{-1} \boldsymbol{\nu}_k . \quad (2.29)$$

Under  $\mathcal{H}_0$  the filter innovations are zero-mean Gaussian which corresponds to  $\bar{\epsilon}_{\nu}^2$  being  $\chi^2$ -distributed with  $K = n_{\mathbf{z}}$  degrees of freedom. By performing a  $\chi^2$ -test, it can be evaluated whether  $\mathcal{H}_0$  can be accepted with a given probability.  $\mathcal{H}_0$  is accepted if  $\boldsymbol{\nu}_k \in [r_1, r_2]$  with acceptance interval  $P\{\boldsymbol{\nu}_k \in [r_1, r_2] | \mathcal{H}_0\} = 1 - \alpha$  [BSLK04]. For  $\boldsymbol{\nu}_k < r_1$  the covariance estimates are pessimistic or conservative, and optimistic for  $\boldsymbol{\nu}_k > r_2$ . The first case can lead to suboptimal estimation performance, but the latter one could potentially cause catastrophic failures. If  $\mathcal{H}_0$  is rejected, integrity alarm is triggered. Depending on the width of the time window for averaging, a fast jump detection is enabled. The limitation of this approach is the lack of detectability of drifts. This limitation is often overcome by utilization of multiple model approaches [WHLS15]. sequential probability ratio testing (SPRT) and its special case, the cumulative sum (CUSUM) test [Pag54]. The one-sided SPRT is given by [Gus00]

$$g_k = g_{k-1} + s_k - d \quad (2.30)$$

$$g_k = 0 , \text{ and } \hat{k} = k \text{ if } g_k < a < 0 \quad (2.31)$$

$$g_k = 0 , \text{ and } k_a = \hat{k} \text{ and alarm if } g_k > h > 0 . \quad (2.32)$$

The CUSUM test is then obtained for  $a = 0$ . The test statistic  $g_k$  is obtained by summing up the distance measure  $s_k$  and a drift compensation variable  $d$ . An alarm is triggered if  $g_k$  exceeds the threshold  $h$  and change time estimate is denoted  $k_a$ . The design of change detection methods requires to trade-off the false alarm rate and the fast detection of faults. For faults detection in hybrid localization procedures, the detection of a fault can be used as a triggering signal for the reinitialization of a localization modality. This step underlines the importance of this trade-off, since the initialization procedure is time consuming and might simultaneously lead to decreased localization performance. An integrity measure is then given by the result of the consistency check and the protection level ([WHLS15], chapter 26).

Alternatively, a set of consistent hypotheses can be found from a pairwise consistency matrix  $\mathbf{R}_k^c$  by means of spectral clustering [VL07].  $\mathbf{R}_k^c$  is built from the set of residuals  $\{r_k^j\}_{j=1}^{n_r}$  which can be calculated from the Mahalanobis or any other distance between two hypotheses

$$\mathbf{R}_k^c = \begin{Bmatrix} 1 & r_k^1 & r_k^2 \\ r_k^1 & 1 & r_k^3 \\ r_k^2 & r_k^3 & 1 \end{Bmatrix} . \quad (2.33)$$

Consecutively, a spectral clustering algorithm like single cluster graph partitioning [Ols08] is applied to  $\mathbf{R}_k^c$  by calculation of its singular value decomposition. The dominant eigenvector  $\mathbf{v}_k$ , belonging to the largest eigenvalue  $\lambda_{\max,k} = \max(\lambda_{i,k})$ , is considered in order to determine the set of consistent hypotheses. This set is described by the indicator vector  $\mathbf{i}_k$ . It is determined by applying a threshold  $c$  to  $\mathbf{v}_k$  [Ols08]

$$\mathbf{i}_{j,k} = \begin{cases} 1 & , \text{if } v_j > c \\ 0 & , \text{else} \end{cases} . \quad (2.34)$$

Pairs of hypotheses corresponding to residual  $r_{j,k}$  are in the consistency set if  $i_{j,k} = 1$ .

### 2.1.5 Cramér-Rao bound

The state estimation accuracy for unbiased estimators like the previously introduced Bayesian filtering framework is limited by the Cramér-Rao bound (CRB). For an unbiased estimator  $\hat{\mathbf{x}}_k(\mathbf{z}_k)$  of the vehicle state  $\mathbf{x}_k$  it holds

$$\mathbb{E}[\hat{\mathbf{x}}_k(\mathbf{z}_k) - \mathbf{x}_k] = \mathbf{0} . \quad (2.35)$$

The bound is then given by the calculation of the filtering Fisher information matrix [HRFT04]  $\mathbf{J}_k$  such that

$$\text{cov}(\hat{\mathbf{x}}_k(\mathbf{z}_k) - \mathbf{x}_k) \geq \mathbf{J}_k^{-1} . \quad (2.36)$$

The CRB, also denoted Cramér-Rao lower bound (CRLB), is then given by  $\mathbf{J}_k = \mathbf{P}_k^{-1}$ . The state  $\mathbf{x}_k$  is assumed to be unknown and random. Consequently, the posterior CRB is considered in the context of vehicle localization. For known state sequences  $\mathbf{x}_{1:k}$ , the parametric CRB is used instead. The posterior CRB is given by the solution of

$$\mathbf{J}_k = -\mathbb{E}[\Delta \ln p(\mathbf{Z}_k, \mathbf{X}_k)] . \quad (2.37)$$

For the considered case of non-linear systems with zero-mean Gaussian noise this expression can be solved in closed-form. *Tichavsky* et al. [TMN98] propose the following solution for the iterative calculation of  $\mathbf{J}_k$ . Following the formulation provided in [HRFT04], the solution is given as:

$$\begin{aligned} \mathbf{J}_{k+1} &= \mathbf{D}_{33,k} - (\mathbf{D}_{12,k})^\top [\mathbf{J}_k + \mathbf{D}_{11,k}]^{-1} \mathbf{D}_{12,k} + \mathbf{J}_Z(k+1) , & (2.38) \\ \mathbf{D}_{11,k} &= \mathbb{E} \left[ [\nabla_{\mathbf{x}_k} \ln p(\mathbf{x}_{k+1} | \mathbf{x}_k)] [\nabla_{\mathbf{x}_k} \ln p(\mathbf{x}_{k+1} | \mathbf{x}_k)]^\top \right] \\ &= \mathbf{A}_k^\top \mathbf{Q}_k^{-1} \mathbf{A}_k , \\ \mathbf{D}_{12,k} &= \mathbb{E} \left[ [\nabla_{\mathbf{x}_k} \ln p(\mathbf{x}_{k+1} | \mathbf{x}_k)] [\nabla_{\mathbf{x}_{k+1}} \ln p(\mathbf{x}_{k+1} | \mathbf{x}_k)]^\top \right] \\ &= \mathbf{A}_k^\top \mathbf{Q}_k^{-1} , \end{aligned}$$

$$\begin{aligned}
\mathbf{D}_{33,k} &= \mathbb{E} \left[ \left[ \nabla_{\mathbf{x}_{k+1}} \ln p(\mathbf{x}_{k+1} | \mathbf{x}_k) \right] \left[ \nabla_{\mathbf{x}_{k+1}} \ln p(\mathbf{x}_{k+1} | \mathbf{x}_k) \right]^\top \right] \\
&= \mathbf{Q}_k^{-1}, \\
\mathbf{J}_z(k+1) &= \mathbb{E} \left[ \left[ \nabla_{\mathbf{x}_{k+1}} \ln p(\mathbf{z}_{k+1} | \mathbf{x}_k) \right] \left[ \nabla_{\mathbf{x}_{k+1}} \ln p(\mathbf{z}_{k+1} | \mathbf{x}_k) \right]^\top \right] \\
&= \mathbb{E} \left[ \mathbf{C}_k^\top \boldsymbol{\Sigma}_{z,k}^{-1} \mathbf{C}_k \right] = \mathbb{E} \left[ \boldsymbol{\Sigma}_{z,k}^{-1} \right].
\end{aligned}$$

By inverting this recursion, the Riccati equations for the EKF are obtained.

In model-based design, an upper bound for state estimation uncertainty is of high interest. For a set of worst case assumptions  $\mathbf{Q}_u \geq \mathbf{Q}_k$ ,  $\forall k$  and  $\boldsymbol{\Sigma}_u \geq \boldsymbol{\Sigma}_{z,k}$ ,  $\forall k$  as well as constant system matrix  $\mathbf{A}$  such a bound can be derived. *Mirzaei et al.* [MMR07, MR11] proved that in this case, the state uncertainty is always bounded by the resulting CRB for  $k \rightarrow \infty$ .

## 2.2 Simultaneous localization and mapping (SLAM)

Simultaneous localization and mapping requires a robotic system to create a map of a previously unknown environment and to perform localization in this map at the same time. The following introduction to simultaneous localization and mapping (SLAM) is mainly based on the work presented by *Levinson et al.* [LMT07]. Based on the system model introduced in Equation 2.9, the following motion constraint can be formulated

$$(\mathbf{x}_k - g(\mathbf{u}_k, \mathbf{x}_{k-1}))^\top \mathbf{Q}_k^{-1} (\mathbf{x}_k - g(\mathbf{u}_k, \mathbf{x}_{k-1})) . \quad (2.39)$$

This constraint is obtained by odometry and scan matching. These relative motion constraints are interpreted as edges in a sparse Markov graph [LMT07]. Absolute pose measurements  $\mathbf{z}_{G,k}$  from GPS can be incorporated by the following constraint [LMT07]

$$(\mathbf{x}_k - (\mathbf{z}_{G,k} + b_k))^\top \boldsymbol{\Gamma}_k^{-1} (\mathbf{x}_k - (\mathbf{z}_{G,k} + b_k)) . \quad (2.40)$$

$b_k$  is the latent variable in the Markov chain based model of the systematic GPS measurement uncertainty [LMT07]. It is modeled by a random walk of the form  $b_k = \gamma b_k + \beta_k$  with Gaussian noise variable  $\beta_k$  which has a covariance  $\mathbf{B}_k$  [LMT07]. This model adds an additional constraint to the optimization problem

$$(\mathbf{b}_k - \gamma \mathbf{b}_{k-1})^\top \mathbf{B}_k^{-1} (\mathbf{b}_k - \gamma \mathbf{b}_{k-1}) . \quad (2.41)$$

Loop closures form additional constraints of the form [LMT07]

$$(\mathbf{x}_k + \delta_{sk} - \mathbf{x}_s)^\top \mathbf{L}_{sk}^{-1} (\mathbf{x}_k + \delta_{sk} - \mathbf{x}_s) . \quad (2.42)$$

The offset  $\delta_{sk}$  is obtained from registration of the overlapping submaps around the poses  $\mathbf{x}_k$  and  $\mathbf{x}_s$  with covariance  $\mathbf{L}_{sk}$ .

The poses  $\mathbf{x}_k$  are obtained from optimization of the overall cost function  $J$  [LMT07]

$$\begin{aligned}
J = & \sum_k (\mathbf{x}_k - g(\mathbf{u}_k, \mathbf{x}_{k-1}))^\top \mathbf{Q}_k^{-1} (\mathbf{x}_k - g(\mathbf{u}_k, \mathbf{x}_{k-1})) \\
& + \sum_k (\mathbf{x}_k - (\mathbf{z}_{G,k} + b_k))^\top \mathbf{\Gamma}_k^{-1} (\mathbf{x}_k - (\mathbf{z}_{G,k} + b_k)) \\
& + \sum_k (\mathbf{b}_k - \gamma \mathbf{b}_{k-1})^\top \mathbf{B}_k^{-1} (\mathbf{b}_k - \gamma \mathbf{b}_{k-1}) \\
& + \sum_k (\mathbf{x}_k + \delta_{sk} - \mathbf{x}_s)^\top \mathbf{L}_{sk}^{-1} (\mathbf{x}_k + \delta_{sk} - \mathbf{x}_s) .
\end{aligned} \tag{2.43}$$

Accumulation of the sensor measurements in an occupancy grid according to the set of poses results in the map  $\mathbf{m}$ . The solution of  $J$  can be obtained by standard least-squares optimization or robust functions like the Huber kernel. The error terms in Equation 2.43 can be rewritten as

$$\mathbf{e}_k^\top \mathbf{\Sigma}_k \mathbf{e}_k = \rho \left( \sqrt{\mathbf{e}_k^\top \mathbf{\Sigma}_k \mathbf{e}_k} \right) \tag{2.44}$$

$$= \rho(\epsilon_k) . \tag{2.45}$$

For least-squares optimization, the kernel  $\rho_2$  is given as  $\rho_2(\epsilon) = \epsilon^2$ . By replacing  $\rho_2$  by the robust Huber kernel

$$\rho_H(\epsilon_k) = \begin{cases} \epsilon_k^2 & , \text{if } |\epsilon_k| < b \\ 2b|\epsilon_k| - b^2 & , \text{if } |\epsilon_k| \geq b \end{cases} \tag{2.46}$$

the quadratic influence of error terms can be reduced, yielding more robust optimization performance. The Huber kernel can be incorporated into the optimization process by weighting of the error terms [GKSK11]

$$(w_k \mathbf{e}_k)^\top \mathbf{\Sigma}_k^{-1} (w_k \mathbf{e}_k) \tag{2.47}$$

with

$$w_k = \frac{\sqrt{\rho_H(\epsilon_k)}}{\epsilon_k} . \tag{2.48}$$

The utilization of robust kernels has been shown to be equivalent to switchable constraints [LCN14] used to remove inconsistent edges. Consequently, the applicability of robust optimization procedures is limited in the presence of loop closing constraints which would be likely removed. Therefore, loop closing constraints are often added as fixed edges which requires a thorough outlier rejection procedure to avoid false loop closures. The optimization step can be performed offline or online by iterative optimization. Throughout this work, the  $g^2o$  optimization library [KGS<sup>+</sup>11] is used for optimization.

## 2.3 Dependability, robustness and automation levels

Increasing levels of automation have significant influence on the requirements concerning the localization system performance. In this section, an overview over different classifications of automation levels is provided. Furthermore, definitions for robustness and related terms are provided and discussed in the context of different automation levels. Consecutively, the focus is narrowed and the role of robustness in vehicle localization for highly automated driving is surveyed in the end of this section.

Well recognized standardized definitions of automation levels for automated driving (AD) are provided by the Society of Automotive Engineers (SAE) and used throughout this thesis. An overview over the automation levels defined by the SAE, the German Federal Highway Research Institute (BAST) and the National Highway Traffic Safety Administration (NHTSA) is given in [Int14] and restated in Table 2.3. Systems with automation levels 0 (no automation), 1 (driver assistance) and 2 (par-

**Table 2.3** A comparison of different common automation level definitions [Int14].

	level 0	level 1	level 2	level 3	level 4	level 5
<b>SAE</b>	no automation	driver assistance	partial automation	conditional automation	high automation	full automation
<b>BAST</b>	driver only	assisted	partly automated	highly automated	fully automated	-
<b>NHTSA</b>	no automation	function specific automation	combined function automation	limited self-driving automation	full self-driving automation	

tial automation) require the driver to continuously monitor the environment. For automation levels 0 to 3 (conditional automation), the human driver serves as a fallback in case the automated vehicle systems fails. Automation levels 4 and 5 do not rely on the human driver as a fallback. A more detailed view on selected automation levels is provided in Table 2.5.

Current technological developments aim at level 4 and 5 automated vehicle systems. This has significant implications on the required system capabilities. The relaxed restrictions on the operation environment require the automated system to deal with an even increasing manifold of use cases and environment characteristics. Furthermore, the capability of failure detection and functional degradation is

**Table 2.5** Progression in automation levels is characterized by decreased restrictions on vehicle system use cases and increasing requirements concerning the capabilities of the vehicle system. This can be seen from the automation level specific minimum capabilities in [Int14].

capability	level 1	level 3	level 4	level 5
transition time until human intervention	no	yes	yes	yes
failure detection and functional degradation	no	no	yes	yes
covered driving modes	some	some	some	all

emphasized. The combination of relaxed use case restrictions and less human intervention forms the main challenge inherent to increasing automation levels. This leads to the notions of dependability in general and robustness in particular which are introduced in the following. Both are required for a safe execution of automated driving tasks. Since the number of these tasks increases and the level of human intervention decreases, both become more and more important with increasing automation levels. As discussed in the introduction of this work, map-relative vehicle localization enables the usage of additional spatially referenced information from digital maps in automated decision processes and is part of a vehicle system with robustness requirements. Consequently, localization systems have to be dependable, once behavior decisions are made and executed based on the provided pose estimates. An overview over important definitions is provided in Table 2.7.

**Table 2.7** Definitions of relevant terms in the context of this work. The definitions are cited from ISO 24765 [ISO10] and [LCI<sup>+</sup>04].

<b>dependability</b>	“trustworthiness of a computer system such that reliance can be justifiably placed on the service it delivers” [ISO10]
<b>availability</b>	“the degree to which a system or component is operational and accessible when required for use” [ISO10]
<b>reliability</b>	“the ability of a system or component to perform its required functions under stated conditions for a specified period of time” [ISO10]
<b>robustness</b>	<ul style="list-style-type: none"> <li>• “the degree to which a system or component can <b>function correctly</b> in the presence of <b>invalid inputs</b> or <b>stressful environmental conditions</b>” [ISO10]</li> <li>• “the delivery of a correct service in <b>implicitly-defined adverse situations</b> arising due to an uncertain system environment.” [LCI<sup>+</sup>04]</li> </ul>

Availability and reliability are important aspects of dependability. Robustness contributes to dependability by advantageously influencing the system reliability and availability. Both definitions of robustness underline a limitation of robustness, namely the availability and feasibility of assumptions concerning adverse situations. In the previous sections, vehicle localization was defined as the task of inferring absolute vehicle pose information from uncertain sensor observations. This process of inference entails a manifold of explicit and implicit assumptions about the nature of the sensor observations and their interrelation with the vehicle pose. Challenges in vehicle localization typically arise from a mixture of these two sources of uncertainty. Examples for uncertain models include the Markov assumption and the belief function type that were mentioned in the previous section. Also map matching algorithms entail assumptions, often expressed by heuristics like in iterative closest point (ICP). It is in the nature of model assumptions, that open environments cannot be exhaustively described. Furthermore, the assumptions can lead to increased requirements concerning sensing and processing resources which might be constrained. The robustness definitions according to [ISO10] and [LCI+04] entail four formulations summarized as

- correct functioning,
- invalid inputs,
- (implicitly-defined) stressful environmental conditions

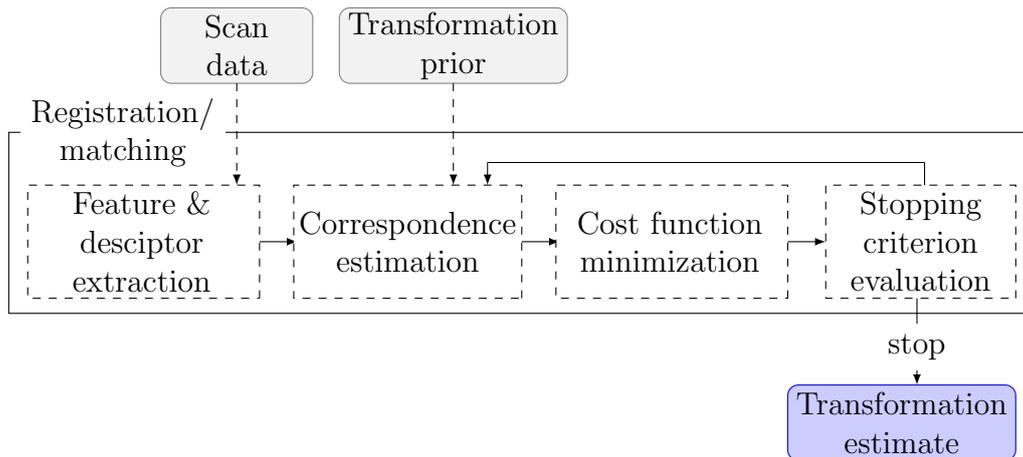
which are interesting for the thesis at hand and are now defined in the context of localization. The explicit meaning of correct functioning is highly dependent on the requirements of the intended use case. For the considered automation levels 4 and 5, the requirements arising from behavior planning are considered. These are expressed as a localization accuracy requirement (see Section 9.1) and by demanding a detection capability for false vehicle pose estimates. Both shall be met during automated operation in urban environments. The introduction of various characteristics of urban environments in Section 4.2 showed the environmental conditions which can be considered as stressful for localization systems. Examples are occlusions, semi-static objects in the map and sensor measurements and heterogeneity. These are the main aspects which are considered in this thesis. Implicitly, these aspects are defined by utilization of a suitable map matching procedure which takes inconsistent point cloud data as valid inputs. Thereby, the probability of invalid map matching results to higher hierarchical layers of the localization system can be reduced. Failures in data association, especially during map matching, can lead to invalid inputs to the localization system, such as inconsistent information about the vehicle pose. Ideally, the occurrence of invalid system inputs can be avoided since outliers cannot always be detected. Hence, in this thesis failure avoidance instead of compensation is fostered where possible. This is achieved during runtime by utilization of the aforementioned matching algorithm (Section 6) and during design time by a model-based design approach (Section 9).

## 3 Related work

The presented related work is structured in categories which reflect the part-structure of this thesis. Section 3.1 reflects contents from *Part II* by providing a survey on matching algorithms and applications of these. Furthermore, a concept for registration algorithm categorization is elaborated. Localization systems are surveyed in Section 3.2, providing the state of the art for *Part IV*. In consideration of *Part III*, an overview over current localization performance models is given in Section 3.3.

### 3.1 Registration

Data association remains a fundamental challenge in automated driving and robotics in general. During registration, this task occurs as correspondence estimation between two sets of input data. Given the correspondences, the transformation parameters can be calculated in closed form [LSL76, HJL<sup>+</sup>89] (unweighted, scalar weighted OPP). Incorporation of additional knowledge about the uncertainty in the measurement points can increase the matching accuracy. Consideration of anisotropic noise as encountered in stereo video depth measurements, requires the utilization of optimization procedures [BF00] or can be solved approximately in closed-form [WCR92] (see Section 8.1.1). In practical applications, correspondences are commonly unknown and must be estimated. The steps of point data registration are shown in Figure 3.1. The proposed concept overview is used throughout this section to categorize state of the art registration algorithms. In the following subsections, state of the art iterative and non-iterative registration algorithms are reviewed, where different approaches for correspondence estimation are employed.



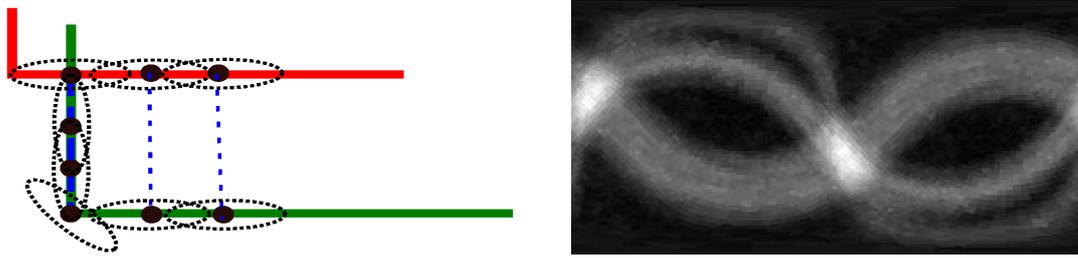
**Figure 3.1** The proposed concept for categorization of iterative point registration procedures (Section 3.1.1) and their distinction from dense or region-based approaches (Section 3.1.2).

### 3.1.1 Point-based registration

Point-based registration approaches are predominant in the area of robotics. In point-based registration, the data association or point correspondence estimation is explicitly addressed. Correspondence and transformation estimation is jointly conducted in an iterative manner and commonly relies on strong assumptions on the input data. The assumptions are typically formulated as distance and similarity measures, cost functions and stopping rules (Figure 3.1). Variations of the assumptions and submodules of standard registration algorithms led to the introduction of a solid body of different contributions.

The ICP algorithm [BM92] and its variants have gained great popularity. The original ICP variant iteratively determines corresponding point pairs from two overlapping scans based on the Euclidean point-to-point distance and minimizes the squared distances between corresponding point pairs. Modifications to the cost function and similarity measures for point correspondence estimation have led to the introduction of ICP variants with increased robustness. Modification of the point-to-point distance include point-to-plane [CM91], measures based on local geometric similarity in generalized ICP [SHT09] (Figure 3.2) and approximate solutions [GY03]. The efficiency and robustness of correspondence estimation can be increased by the introduction of search intervals [LM97]. Robust cost functions for the optimization step have been fostered by introduction of M-estimation in [KKM03] and robust cost function like the Huber kernel [Fit03]. The selection of suitable features like corners or abstract features is of crucial importance for the performance of correspondence-based registration algorithms. The extraction of features from point cloud data has shown good performance in scenarios, where a sufficiently high number of repetitively detectable features is available. This might not be the case for highly heterogeneous operation environments. Under the assumption of a partially planar environment like corridors, minimally uncertain maximum consensus (MUMC) [PBVP10] extends the ICP approach to the usage of planes. According to [Zha94], care should be taken when utilizing geometric primitives since the feature extraction from point data introduces another source of uncertainty which might lead to matching failures. A covariance estimate for point-based ICP is provided in [Cen07a]. Comprehensive surveys on ICP variants can be found in [PCS15] and [BSMW14].

Recently, the normal distribution transform (NDT) [BS03] came into the focus of robotic scan matching applications. NDT like matching approaches address the correspondence search by modeling the input point cloud data as sums of normal distributions. Since, this model is piecewise smooth, standard optimization procedures can then be used for iterative data alignment. Thereby, NDT inherits all advantages and disadvantages of iterative optimization procedures. In comparison to ICP, the correspondence estimation between single points is addressed implicitly. The original algorithm for 2D scan measurements has been extended to the 3D case [MLD07]. Simultaneous matching of several scans using was introduced in [BS06]. *Magnusson et al.* provide a comparison between NDT and ICP in [MNL<sup>+</sup>09] with



**Figure 3.2** Left image: iterative point correspondence estimation on the basis of a plane-to-plane distance measure (image source: [SHT09], *Segal et al.*); Right image: dense matching on the basis of the Hough-Radon transformation (image source: [CC09], *Censi*), where the transformation parameters are found by a maximum search in the correlation domain.

a focus on runtime and reliability. This comparison is expanded to ICP, MUMC and NDT in [MVS<sup>+</sup>15] and shows the broader basin of convergence of NDT and MUMC compared to ICP. The accuracy of NDT has been shown to be comparable to ICP [SMAL11], although the results tend to vary for different data sets. Clustering approaches have been introduced in order to deal with sparse and cluttered sensor measurements. In [RBH<sup>+</sup>15] clustering is used to adapt NDT to the characteristics of radar measurements. The authors of [SASL13] propose to use NDT for localization within a Monte-Carlo localization framework. NDT and ICP perform a local search for correct transformation parameters which leads to a high computational efficiency of this algorithm class. Since, the underlying optimization problem is typically non-convex this local approach can lead to convergence against local optima.

This problem is addressed by sample-based registration approaches. random sample consensus (RANSAC) [FB81] randomly draws samples of correspondence pairs from the data and calculates the transformation parameters. Consecutively, the sample sets are ranked by the score of the transformation model. Commonly the score is expressed as the number of model-supporting points. In case of exhaustive drawing, i.e. all possible samples are drawn, the fittest model can be selected. *Olson et al.* introduced an approach based on single cluster graph partitioning (SCGP) [Ols09, OWTL05] which overcomes some performance issues of RANSAC. Both approaches are less sensitive to prior inaccuracy and noise. Depending on the size of the input data, sample-based approaches are computationally demanding. Thus, in practice a maximum number of iterations or a threshold for acceptable fitness scores is commonly defined.

### 3.1.2 Dense registration

Dense or region-based registration methods solve the data association and error minimization in one step (compare Figure 3.1). Hence, steps with high amounts of assumptions, like feature extraction, correspondence search and error minimization

can be omitted. This class of algorithms typically shows good performance under the presence of severe noise in the input data and in heterogeneous operation environments. This class of algorithms has a long history in image processing and pattern recognition [ZF03, Gos12].

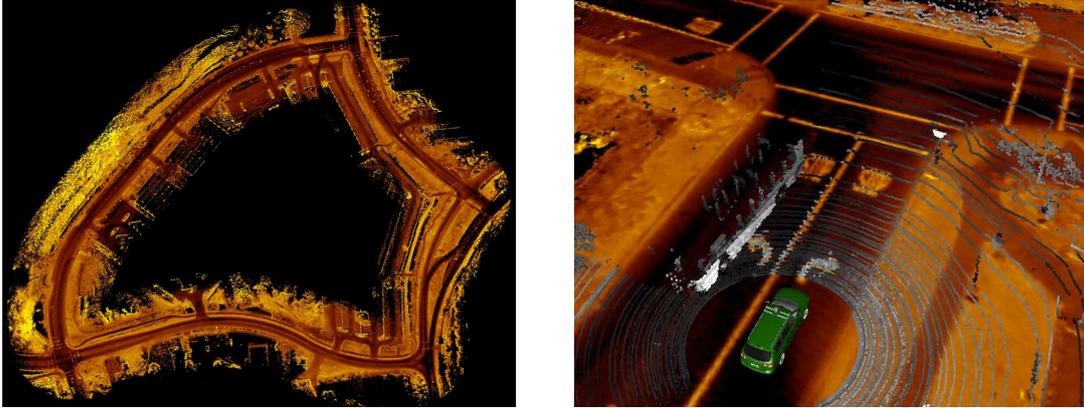
A basic approach to dense registration is the application of correlation-like methods to the unmodified input data. However, according to *Zitova et al.* [ZF03], this method is sensitive to noise in the input data. Hence, the initial application of a transformation, e.g. Fourier transformation, to the input data was found to provide more robust registration results [ZF03]. Additional transformations can be used in order to achieve a decoupling between rotation and translation parameter estimation. This is done by rotation- or translation invariant transformation of the input data. To this point, state of the art dense registration algorithms often employ the Fourier-Mellin transformation (FMT) [CM87] or the Hough-Radon transformation ([CC09], Figure 3.2). The transformation parameters are then found by evaluation of a transfer function which correlates the input data. The transfer function can be designed to maximize the signal-to-noise ratio (SNR) or other characteristics of its output [KH90].

Fourier transformation based registration, also named spectral registration, of translated images has been extended to rotation estimation by introduction of the translation-invariant FMT [CM87]. The estimation of scale changes is considered in [RC96]. An extension to 3D data is presented in [LL02, BB13]. The phase correlation step is of significant importance for the matching quality. An overview over phase correlation approaches is provided in [CDD94, KH90]. Subpixel methods can be utilized in order to increase the matching accuracy [FZB02b]. Masking has been shown to enhance the matching robustness [Pad12, FZB02a]. The computational complexity of dense registration procedures is often stated as a major disadvantage. Hence, several works contribute to more efficient implementations. Efficiency can be gained, e.g., by approximate solutions for phase correlation [KBN04]. Hierarchical registration strategies which start from coarse and progress towards fine resolution are other means which have been applied in the medical domain [LA99].

Beside the FMT, the Hough-Radon transformation has been recently proposed by *Censi et al.* in [CIG05] for registration of 2D range finder data. This approach was extended to 3D range data in [CC09]. A similar method has been applied to localization with radar data in [WRK<sup>+</sup>15, WBK<sup>+</sup>15]. The application of dense registration algorithms is considered advantageous, if the existence of a minimum density of specific features for correspondence-based registration is not given with a sufficient probability.

## 3.2 Localization and SLAM

In this thesis, the focus is put on vehicle localization in urban environments relative to an accurate localization map. To this point, the following review on related work is concerned with map-relative vehicle localization in urban environments under



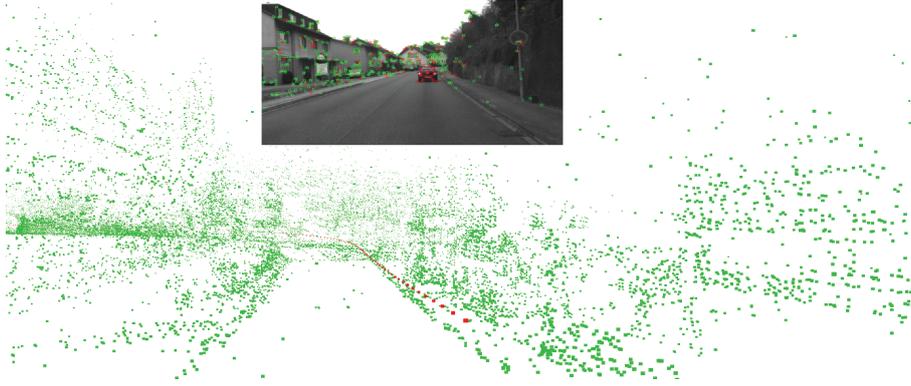
**Figure 3.3** Left image: a LiDAR intensity measurement based ground plane map of an urban test track; Right image: a detailed view of an intersection scene with other traffic participants (source of both images: [LT10], *Levinson et al.*). Due to the sole utilization of ground plane features, the influence of non-static objects is significantly reduced at the cost of the dependence on ground features.

special consideration of mechanisms for increasing the robustness of pose estimation. The focus on urban environments is broadened where it is required.

Many works on vehicle localization originated from indoor robotics. In an early contribution, *Dellaert et al.* introduce the Monte Carlo localization (MCL) approach [DFBT99]. MCL incorporates a particle filter (PF) for propagation of the vehicle pose over time. Additionally, the PF enables global localization within the computational capabilities of the robotic system. This approach has been more recently extended to adaptive Monte Carlo localization (AMCL) which utilizes KLD resampling in the PF [Fox01, TBF05]. As pointed out by *Carlone et al.* in [CB09], the utilization of particle filters and robust variant come at the price of a loss in accuracy while gaining robustness against false pose measurement updates. This observation has also been stated by *Huber et al.* [Hub11] in the far more general context of robust statistics. A thorough survey on earlier state of the art can be found in [GBFK98] with an extension to more recent works in [GF02].

*Levinson et al.* were the first to introduce a probabilistic localization framework explicitly developed for urban environments [LT10] based on 360° LiDAR measurements. The influence of dynamic and semi-static objects is reduced to a negligible level by exclusive utilization of ground plane features for localization (Figure 3.3). Consequently, a static world can be assumed and is met with high probability. A video-based approach can be found in [ZLS<sup>+</sup>14, LSZS13]. The authors extract features from mono camera measurements and use them for estimation of a map-relative vehicle pose (Figure 3.4). These map-relative pose estimates are then refined by the fusion with precise IMU measurements. The density of the feature measurements leads to an increased robustness against outliers and occlusions.

The aforementioned approaches use dense data for localization. The utilization of sparse landmark-based environment representations can lead to decreased computa-



**Figure 3.4** Dense visual feature with high dimensional descriptors used for localization with a monocular camera (image source: [LS12], *Lategahn et al.*). This approach leads to high localization accuracy and storage as well as computational costs.

tional costs. Early works on landmark-based localization approaches are introduced in [LDW91, BG97]. *Pink et al.* use aerial images to extract landmarks for localization [PMB09, Pin10]. In their work, road markings are automatically extracted from high-resolution aerial images for map generation. A stereo-video camera is then used to detect these landmarks during operation of the automated vehicle system. *Schlichting et al.* propose to use poles and planes extracted from LiDAR measurements for localization [SB14]. The level of abstraction can be further increased by using semantic environment information as opposed to metrically accurate environment representation. *Oberländer et al.* propose to extract semantic information from sensor observations and use these for localization in simplified, semantic localization maps which have to be topologically correct but can be metrically approximate. Depending on the metric accuracy and the number of semantic objects (e.g. poles) in the map, parts of the navigation task have to be implemented reactively.

With the increasing interest in challenging operation environments, the research focus shifted to the enhancement of robustness of localization systems. It is the author’s opinion, that there is no sharp transition between robust and non-robust localization algorithms. Especially the introduction of probabilistic approaches [TBF05] led to an increase in robustness in comparison to the formerly known state of the art. Hence, in this thesis, the line is drawn between classical probabilistic approaches and their combination with additional measures for increasing robustness.

As shown in, e.g., [LT10], the selection of features has significant influence on the robustness of the resulting localization system against clutter caused by semi-static and dynamic objects. In urban environments in particular and open environments in general, a single feature type does commonly not suffice for accurate localization. This observation has been recognized by many current state of the art algorithms and led to the trend of hybrid localization frameworks. In [ZCO<sup>+</sup>11] visual landmarks and range data are combined in order to achieve increased localization accuracy.

The problem of high local differences in feature types has also been addressed in [WSG10], where a combination of sparse feature-based and dense grid-based environment representations is proposed for indoor operation and parking lots. The authors of [Cho14] and [YYM<sup>+</sup>15] use different landmark types that are extracted from LiDAR measurements. Matching results obtained from registration of sensor observations and different map types are fused in order to obtain the overall pose estimate. Since, these approaches utilize non-robust fusion procedures, the results provided by the different information sources must be correct. An approach for a safe fusion of information sources with unknown correlation is presented in [NDH16].

In [DRN<sup>+</sup>14] the need for several maps is overcome by utilization of a map in combination with different sensor-modalities. Maximally stable extremal regions are introduced to perform map matching. Another way of increasing the number of sources for vehicle pose information is presented in [BGU13]. Absolute pose estimates are obtained based on runs of visual odometry measurements which are registered with a topological map. All these approaches can be seen as means for increasing the system-immanent level of analytical redundancy by feature diversification. Another redundancy-based strategy aims at decreasing the influence of inaccurate models by utilization of sets of models and an adequate weighting. In an exemplary approach, localization is performed by using GNSS and INS in combination with an interacting multiple model EKF [TMZIUMGS07].

An alternative strategy is the selection of landmarks which lead to desired increases in localization performance. Early approaches instrumented machine learning procedures [Thr98, GI96] and several other approaches followed [SSSD06, LRS07]. The characteristics of the operation environment are averaged and consequently it can not be accounted for significant heterogeneities. This approach has been extended to sensor and landmark selection in [TBCC07]. The selection of landmarks to reduce the size of a localization map has been approached in [HS09]. Selection of sensed landmarks for localization and mapping is closely related to the task of the placement of artificial landmarks. The latter can be another option in some application domains like indoor logistics [BMB11, BMKB13, RL00]. Closely related is the field of measurement selection which is mainly concerned with the detection of outliers.

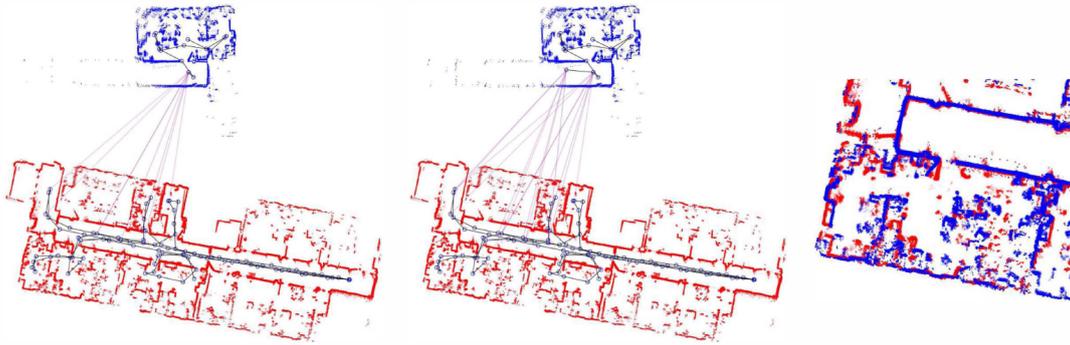
The existence of corrupted measurements is a significant challenge in urban vehicle localization. Corrupted measurements can occur at different stages of the signal processing chain, including map matching results with a wrong covariance estimate which might lead to wrong weighting in consecutive fusion steps and outlier feature detections. To this point, redundancy as introduced in hybrid localization frameworks can be used as basis for outlier detection and rejection.

Examples for the detection of outlier features include a parity space approach [TSG08], outlier detection in the SLAM optimization process [CCD14] and in visual odometry [BW16]. Another approach for the detection of outlying features based on parity space is presented in [TSG08]. Integrity monitoring has a long history in GPS-based navigation for aviation systems and is commonly denoted receiver

autonomous integrity monitoring (RAIM). Recent research attempts are concerned with the application of receiver autonomous integrity monitoring (RAIM) methods to vehicle navigation and localization in urban environments [ZBIG16, DB13, MBIGB09, GB05, AAHAJ10, TENAH<sup>+</sup>14, KG17]. The localization systems mainly use inertial measurement unit (IMU) and GPS pseudo range measurements as inputs and enable lane accurate pose estimation in many cases. Consistency checking by evaluation of the quadratic error between predicted poses (wheel odometry) and differences of consecutive absolute GPS measurements for increasing the localization accuracy is proposed in [WWBN15]. An evaluation of different GPS based localization sources is conducted in [SJ06] and used for mobile robot navigation. It uses an approach similar to [LCJS04]. A diagnosis-based approach for failure detection based on logistic regression has been proposed in [FTY<sup>+</sup>15]. The authors use logistic regression for the detection of failures in a particle filter based localization framework. Other diagnosis procedures use particle filter based approaches for failure detection in an extraterrestrial rover [DWS<sup>+</sup>04] and for classical robotic applications [PSB06].

As opposed to wheel odometry, solution for motion estimation on the basis of LiDAR or other exteroceptive sensor measurements are independent of wheel slippage which likely occurs during operation in urban environments (compare Section 4.2). Furthermore, LiDAR odometry has shown the potential to regularly provide accurate motion estimates. LiDAR based odometry enables accurate estimation of relative sensor movement. The core algorithms and the assumptions made about the operation environment are similar to the ones used for map matching and localization in general. The same holds for measures to enhance robustness. The existence of planes is assumed in [GVI13] and an algorithm for fast plane detection for LiDAR odometry is proposed. Outlier removal is achieved, e.g. by utilization of RANSAC for correspondence estimation [AB13, KGL10]. Another popular odometry framework is LOAM [ZS16]. Other approaches adapt algorithms known from visual odometry [FS12] to LiDAR data [TB13]. Redundancy is increased by combinations of different sensor modalities and signal processing algorithms [ZS15].

SLAM algorithms are of interest for vehicle system operation in unmapped areas. Additionally, methods for consistency checking have been used for SLAM which might as well be of great interest in localization. A comprehensive introduction to SLAM algorithms can be found in [TBF05, DWB06, BDW06] and the references therein. A commonly used optimization framework for SLAM graphs is found in [KGS<sup>+</sup>11] and used in this thesis. SLAM has been used in areas with high density of semi-dynamic objects for building temporary localization maps [MDHGB10]. In [ATS<sup>+</sup>13, AGDT<sup>+</sup>14] covariance scaling is used for SLAM graph optimization. This robust kernel method is capable of dealing with false edges in the graph. SLAM graph optimization with switchable constraints [SP12] or removal of edges based on spectral graph partitioning [OLT06] give similar results. A thorough review on current robust SLAM backends can be found in [LCN14] and shows the equivalence of some of these robustness enhancing methods. Spectral registration has been



**Figure 3.5** Indoor mapping on the basis of spectral registration. The presented results indicate a good performance of spectral registration in the context of changing environments (image source: [ORD13], *Oberländer et al.*).

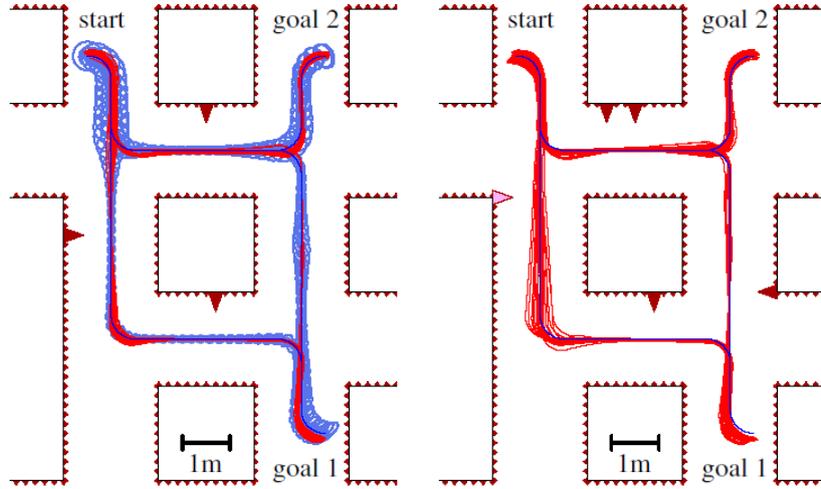
applied to 2D SLAM for indoor environments [ORD13] (Figure 3.5) and SLAM for underwater robotics [PBS<sup>+</sup>10, BB13]. First results presented by *Oberländer et al.* indicate that spectral registration performs well when dealing with outdated localization maps.

### 3.3 Localization system modeling

Localization systems consist of a whole signal processing chain containing sensing, object detection, data association, registration and tracking algorithms. State of the art approaches for modeling parts of these signal processing steps. In addition to the system itself, the operation environment is of crucial importance and has to be abstracted for gaining a comprehensive model of the whole localization system in the context of its operation environment. Different models and model-based design approaches are reviewed in this section. In this context, approaches for localization system modeling, model-based design of localization systems in general and perception systems in particular are surveyed in this section.

*Censi* published a covariance estimation approach for ICP based point cloud registration [Cen07a]. Since, ICP is an iterative algorithm it does not have a closed form solution. Therefore, it is proposed to use the implicit function theorem in order to calculate the covariance of the ICP estimate. In another publication *Censi* introduces a model for the whole localization system based on range finder measurements [Cen07b]. The operation environment is described based on its basic geometry. Operation environment models based on stochastic geometry are popular in the design of cellular and sensor networks. *ElSawy et al.* use stochastic point processes in order to model the spatial distribution of network elements [EHH13] in order to derive performance bounds for cognitive cellular networks.

In guided surgery systems, markers, also denoted as fiducials, are attached to the surgical instruments and are commonly tracked by a vision-based system. This task is closely related to point cloud matching and hence considered related work.



**Figure 3.6** Covariance estimation for landmark-based navigation along a predefined trajectory has been used for the calculation of near-optimal landmark placement (image sources:[BMKB13], *Beinhofer et al.*).

The goal is to provide assistance to the surgeon in order to reach the operation area with the tip of the surgical instrument. Knowledge of the uncertainty in the instrument tip location estimate is hence of fundamental importance. The work of *Fitzpatrick et al.* [FWM98] has found wide attention where the authors consider the case of unweighted point matching. Error propagation from the fiducial measurements to the location estimates of the instrument tip is conducted. In [vdBAG11] the authors utilize an error model to evaluate the quality of candidate driving paths. The model captures uncertainties in sensor measurements and characteristics of the utilized controller to execute a given path. Thus, it enables the calculation of the a-priori probability distributions of the robotic system for a given controller and path. This approach is utilized for near-optimal placement of artificial landmarks for previously defined paths in [Bei14, BMB11, BMKB13]. Uncertainty accumulation in scan matching based relative motion estimation is studied in [SHK<sup>+</sup>14]. Statistical models are derived for the case of matching results calculated by applying the unweighted OPP to corresponding features from consecutive scans.

*Mirzaei et al.* introduce a probabilistic model for localization performance estimation for multi-robot formations [MMR07] which extends previous works in [MR06]. It is assumed that at least one of the robots receives absolute pose measurements, e.g. from GPS. Based on the assumed uncertainty in the global pose measurements and the odometry measurements, upper bound for the state uncertainty are calculated. The results are obtained from calculation of the steady-state solution of the algebraic Riccati equation of the utilized EKF. This model assumes the existence of pose measurements in every time step. A more general perspective on lower state uncertainty bounds is provided by works on general estimation problems. *Hernandez et al.* compare different Cramér bounds for detection probabilities lower than

one. This and comparable works form a possible basis for the derivation of statistical localization models. *Bansal* et al. study in an empirical fashion, how the configuration of a camera system influences the localization performance [BBH14]. A similar approach in the context of visual odometry can be found in [PKB14]. Sensor planning is an important task in the field of industrial robotics. In [TAT95] a thorough survey on this field of research is provided. The placement of sets of cameras is studied in [FSK06] for active object detection. The authors propose to use the Cramér-Rao lower bound as a performance criterion to rank different camera arrangements.

### 3.4 Conclusion and open research questions

From the reviewed related work some commonly used strategies for map-relative localization with enhanced robustness (compare Section 2.3) can be identified. The following observations concerning map matching algorithms, hybrid localization systems and SLAM have led to the development of key contribution *kc1* (Section 1.2). Firstly, robustness of a localization system can be increased by adapting separate algorithmic building blocks like map matching algorithms. Iterative registration algorithms are widely used in robotic applications including urban automated driving. The often utilized heuristics and associated model assumptions, e.g. static world and vertical world, lead to efficient implementations of the respective algorithms. Deviations from the assumptions will, however, likely deteriorate matching performance. Furthermore, accurate prior pose estimates are commonly assumed to be known which might not be the case during system initialization or recovery. To this point, some current SLAM frameworks incorporate non-iterative matching procedures for dealing with challenging measurement data, like sonar measurements for underwater SLAM or laser scans from indoor environments. The utilized data are characterized by high noise levels, but due to the temporal proximity of their collection, inconsistencies between data subsets are limited to a small amount. In this point, the SLAM approaches differ from localization with outdated maps. First respective results in [ORD13] indicate, that spectral registration can have advantages for measurement data with high levels of inconsistencies. Approaches based on the FMT showed accurate and noise insensitive matching performance.

Spectral registration on the basis of the FMT shows great potential for practical application, since efficient software and hardware implementations of the fast Fourier transformation (FFT) are available. To the knowledge of the author, the utilization of spectral registration algorithms in the context of localization and mapping for urban automated driving (UAD) has not been addressed. This includes the use case of localization with partially outdated localization maps and high clutter rates in sensor observations. It has to be shown, that spectral registration can be used in such applications with significant non-Gaussian noise while meeting the requirements for real-time operation in automated vehicle systems. Furthermore, the applicability of different state of the art registration algorithms has to be evaluated for the context

of robust vehicle localization for UAD. Due to the prevalent application domains, algorithmic efficiency especially for registration of 3D data has not been in focus of state of the art implementations. Place recognition procedures enable a rough estimate of the global vehicle pose without necessitating a transformation prior. Some state of the art particle filter based localization system, frameworks enable global localization given the processing power to evaluate high amounts of particles. Challenges arise once the vehicle pose has to be refined given an outdated map and cluttered LiDAR scans. Fast, single-shot initialization without the need for an initialization drive has been addressed by randomized methods like RANSAC. An alternative, non-iterative and exhaustive algorithm for accurate pose initialization is not known to the author. All registration approaches are potentially subjects to failures. Reasons include sparse or ambiguous features and false data association. Thus, frequently used solutions are the selection of features which lead to a sufficient localization performance and the introduction of redundancy in the feature and registration algorithm domain. Hybrid localization algorithms utilize a set of two or more localization modalities, e.g. based on different features, algorithms or sensor technologies, in order to increase the system availability and robustness. Hybrid localization approaches commonly perform well if the number of different localization modalities is high enough for the considered operation environment. This approach to robustness enhancement might lead to challengingly high computational costs. Consistency checking and outlier rejection is commonly not fostered in order to further increase the robustness of the localization system. The results from the map matching modalities are typically fused without the removal of invalid matching results. Consequently, a biased fusion result due to corrupted matching modalities cannot be effectively prevented despite the introduction of further modalities. On the contrary, robust mapping procedures make extensive use of consistency checking procedures which might be of interest for localization as well. The application of methods from RAIM to GPS-based solutions for urban automated driving has been shown to provide high reliability but insufficient accuracy.

The design of localization systems is mainly based on sole expert knowledge which led to highly accurate and reliable localization system implementations. These results are often based on the utilization of high-dimensional descriptors in combination with frequently occurring feature types and sets of features. The localization accuracy requirement for the target applications is often exceeded, neglected or the localization system is adapted in an empirical manner. These design approaches frequently lead to powerful, yet unnecessarily costly system designs and might result in increased efforts during and less repeatability of the design process. For these reasons, in some works single parts of the localization system or the operation environment were designed on the basis of numerical and statistical models, e.g. for determination of minimal numbers of artificial landmarks and their placement. The design and deeper understanding of whole localization systems on the basis of suitable statistical models has not been comprehensively conducted, to the knowledge of the author. Hence, the key contribution *kc2*.

**Part II**  
**Components of vehicle localization**  
**systems for urban automated driving**  
**(UAD)**



## 4 Platform and operation environment

### 4.1 Test vehicle

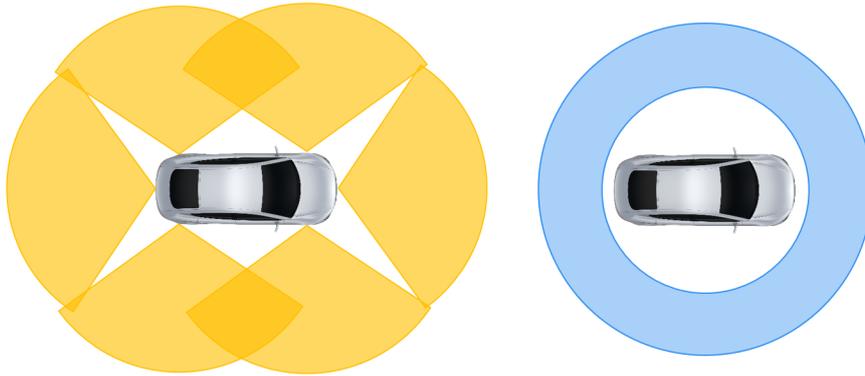
The evaluation results and data sets in this thesis were recorded by utilization of the perception system installed in the test vehicle shown in Figure 4.1. Included in the sensor setup are light detection and ranging (LiDAR) sensors for environment perception (Figure 4.2), a differential global positioning system (dGPS) for global pose and sensors for motion measurement.

A Velodyne HDL-E64 LiDAR sensor [Velb] is mounted at the roof of the vehicle. The opening angles of the Velodyne are  $360^\circ$  horizontally and  $26.9^\circ$  vertically with an angular resolutions of  $0.08^\circ$  and  $0.4^\circ$ , respectively. The vertical opening angle is covered by 64 separate laser beams, resulting in 64 measurement layers. It is configured to provide a full  $360^\circ$  measurement at a frequency of 10 Hz. All LiDAR measurements are transformed into the vehicle frame by application of a known rigid transform. Hence, without loss of generality and for better readability, exteroceptive measurements are stated relative to the vehicle frame (compare Figure 2.1). In addition to the Velodyne sensor, six Ibeo LUX LiDAR sensors are mounted at bumper height. Each sensor has four layers, a horizontal opening angle of  $85^\circ$ , a vertical opening angle of  $3.2^\circ$ ,  $0.125^\circ$  horizontal resolution and operates at a measurement frequency of 12.5 Hz. Relative motion measurements are provided by wheel speed and steering angle sensors. Global pose measurements are collected from a Genesys ADMA Pro+ dGPS system [Gen]. The Adma dGPS system contains an inertial measurement unit (IMU) as an additional source of relative motion estimates. The latter combines acceleration sensors with a fibre optic gyroscope for angular rate



**Figure 4.1** The test vehicle with a Velodyne HDL-64E (blue box) mounted on its top. The six Ibeo LUX sensors (orange boxes) are mounted at bumper-height.

measurements. dGPS and IMU measurements are fused in a filtering step to obtain a pose estimation accuracy of up to 0.01 m.



(a) Sensor fields of view: left: setup of six Ibeo LUX LiDAR sensors mounted at bumper height, right: Velodyne HDL-64E mounted at the top of the test vehicle.



(b) Ibeo LUX LiDAR sensor with four scan layers are arranged around the vehicle contour in order to cover a  $360^\circ$  FoV (image source: [Ibe]).



(c) Velodyne HDL-64E LiDAR sensor with 64 layers (image source: [Vela])

**Figure 4.2** The partially redundant LiDAR sensor setup of the test vehicle covers a  $360^\circ$  field of view around the vehicle.

## 4.2 Operation environment

Urban environments comprise a great variety of different dynamic, semi-static and static objects. Narrow roads with densely parked cars and bounded by tall buildings are as representative to urban environments as extensive intersection areas. These attributes account for a wide variety of challenges for localization systems, including

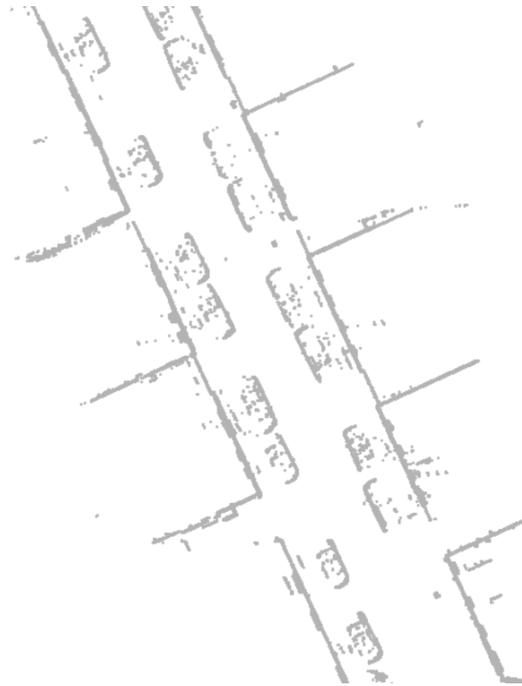
- high clutter rates due to, e.g., dynamic and semi-static objects such as parked cars,
- occlusions,
- high localization accuracy requirements,
- regions of sparse features,
- heterogeneity (compare Figure 4.3),
- high curvatures and
- mixture of high and low speed sections.

Figure 4.3 depicts different sceneries from an urban environment in Stuttgart as these are contained in the localization map (Section 5.2). Narrow domestic roads are often characterized by densely parked cars beside both sides of the road, missing road markings and dense vegetation like hedges. In some section, a good visibility of house walls is given. Due to the heterogeneity of this road type, the applicability of a universal feature type might be strongly limited. Narrow roads alternate with expansive areas such as roundabouts and intersection. The depicted roundabout has several areas with vegetation. Shapes of vegetation areas might serve as features for localization. Additionally, many poles are existent in this scenery. Intersections usually have less distinct vegetation features. Interchangeably, poles and road marking might serve as sole localization features. The localization requirements in intersection areas might be higher than for narrow roads, where an accurate lateral localization often suffices. Hence, localization in intersection areas can be challenging, particularly with high occlusion rates. Parking lots are a common scenery with sparse features and high densities of semi-static objects, i.e. parked cars. Therefore, operation in these areas potentially has lower localization accuracy requirements.

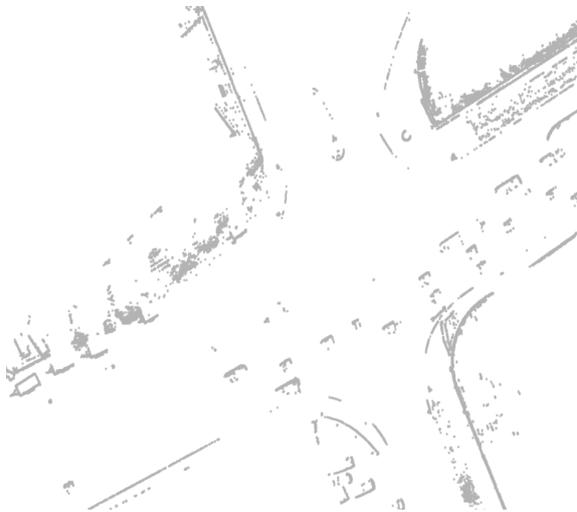
This non-exhaustive set of example sceneries from urban environments underline the inhomogeneity of urban environments with respect to availability of feature types, densities, occlusion rates and occurrence rate of semi-static objects.



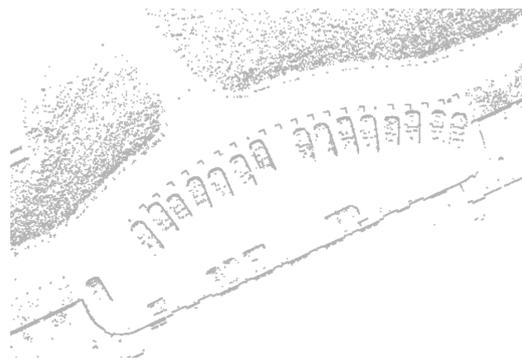
(a) An expansive roundabout with vegetation.



(b) A narrow domestic road with densely parked cars and visible house walls.



(c) A parking lot with dense vegetation and sparse structures otherwise.



(d) An intersection with vegetation and sparse structures otherwise.

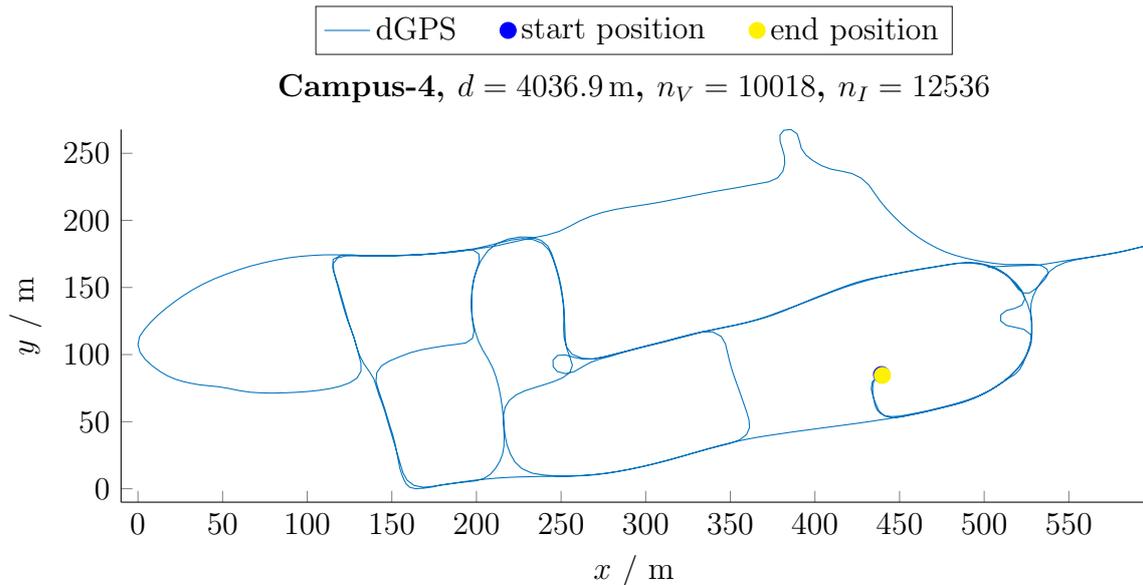
**Figure 4.3** Selected urban sceneries from a localization map for Stuttgart, Germany characterize the heterogeneity of urban environments. These local differences in the operation environment of an automated vehicle system render vehicle localization a challenging task.

### 4.3 Data sets

The evaluation results contained in this thesis are derived on the basis of self-recorded data sets. All data sets for evaluation were collected from test drives with the test vehicle (Section 4.1) and contain dGPS, GPS, LiDAR (Velodyne HDL-64E, Ibeo LUX), IMU and wheel odometry measurements. This makes the contained data comparable to those contained in the popular KITTI data set [GLU12]. Since an outdated map is required for the evaluation of the localization procedure, a suitable data set was recorded. The data sets were collected from an urban and a campus environment.

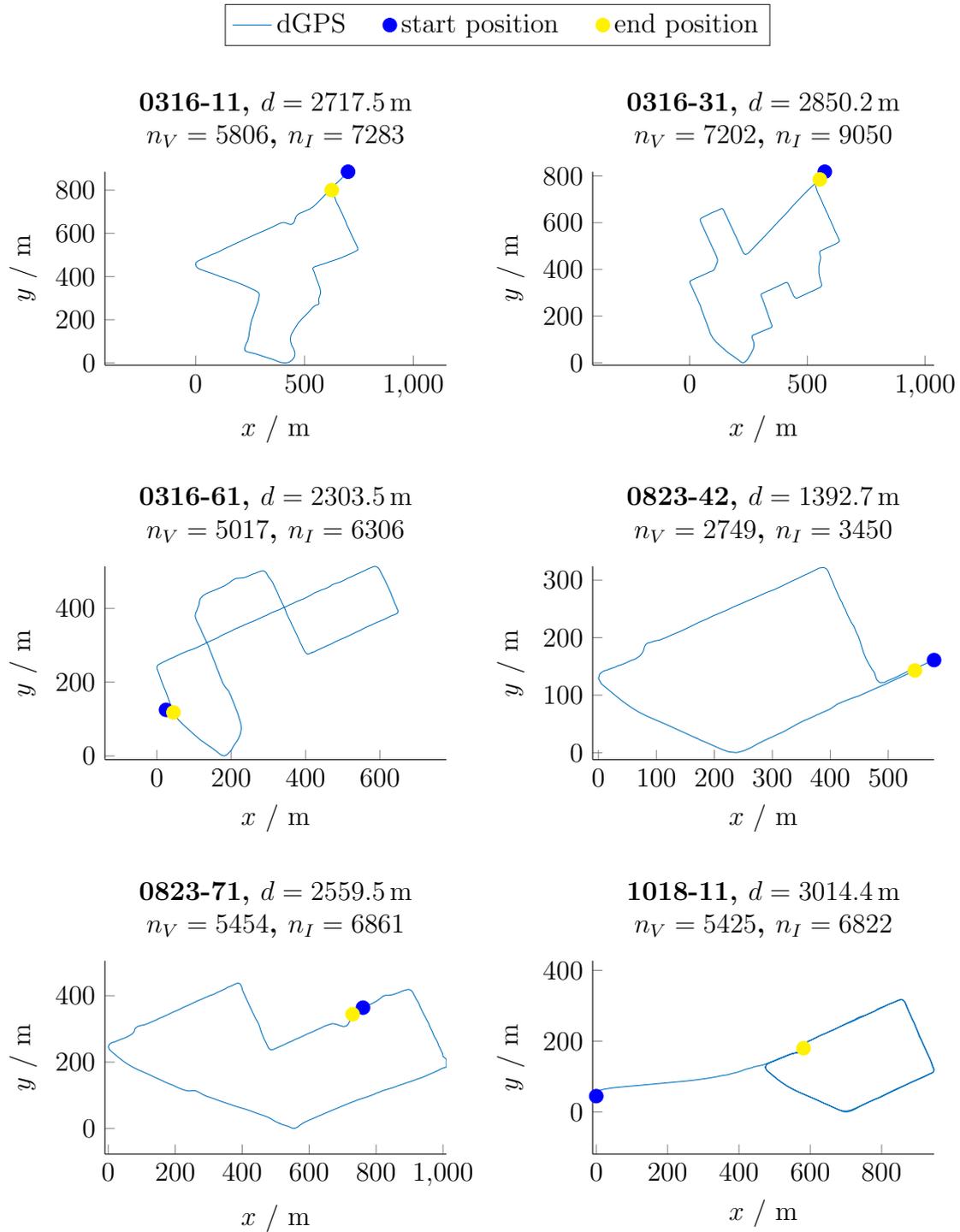
The campus data set is recorded on the Robert Bosch Campus in Renningen, Germany. The width of the roads is comparable to inner city residential roads. Occlusions do not occur as often, since the number of dynamic and semi-static objects is significantly decreased. The feature distribution is equally heterogeneous as in urban environments. The campus data set contains data from a 4.0 km long drive with  $n_V = 10018$  LiDAR scans from a Velodyne HDL-64E and  $n_I = 12536$  scans from the Ibeo LUX setup. An example path from the campus data set is contained in Figure 4.4.

The urban data sets were recorded in Stuttgart-Feuerbach, Germany. This area is characterized by narrow urban canyons and densely parked cars. The whole area is also covered by a localization map which was recorded two years in advance (compare Section 5.2). An overview over the urban data sets is given in Figure 4.5. It contains a total of  $n_{V,sum} = 31653$   $360^\circ$  LiDAR scans from a Velodyne HDL-64E sensor and  $n_{I,sum} = 39772$   $360^\circ$  scans collected from the Ibeo LUX sensor setup (compare Section 4.1) which cover 14.8 km urban roads. The data sets differ in the length



**Figure 4.4** The campus data set.

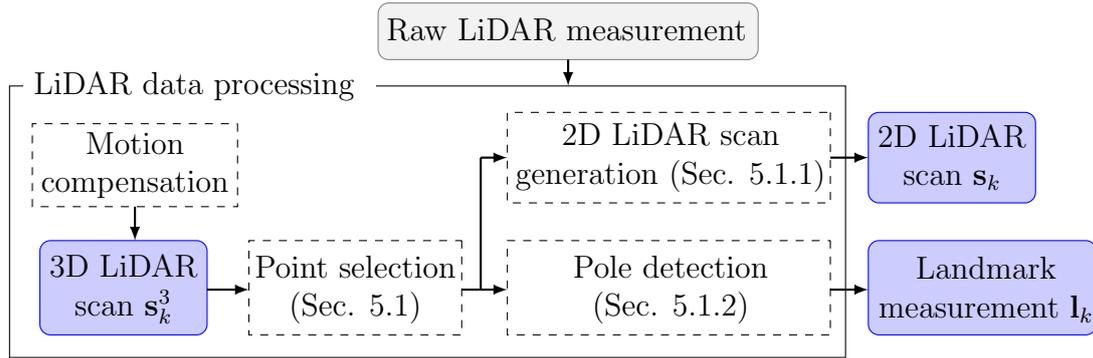
and number of the driven loops as well as the number of turns. Thus, the data sets enable a comprehensive experimental evaluation of localization, matching, odometry and mapping performances with regard to different trajectory characteristics.



**Figure 4.5** The urban data sets cover 14.8 km of inner-city roads.

## 5 LiDAR data processing

In this work, vehicle pose estimation based on 2D LiDAR scans and localization maps is considered. The scans and localization map are recorded in 3D and are projected to the  $xy$ -plane. Therefore, points from vertical surfaces are extracted and used for further processing. The functional architecture of the LiDAR data processing procedure is shown in Figure 5.1.

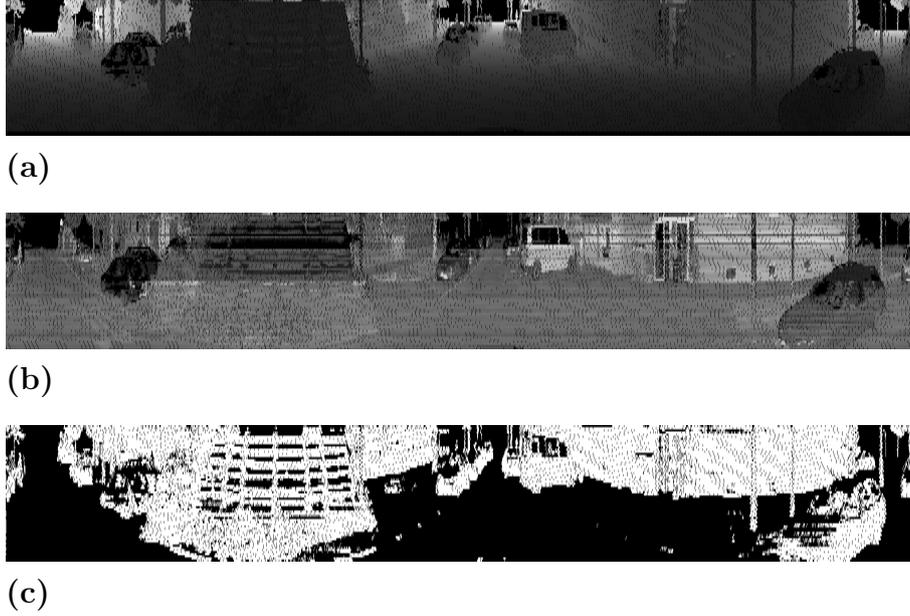


**Figure 5.1** The proposed LiDAR signal processing chain. Points from vertical structures are selected once and used for scan projection and pole detection.

Given a speed limit of 50 km/h for urban highways and a LiDAR revolution frequency of 10 Hz, a vehicle moves 1.4 m between the first and last point measurement. Hence, the LiDAR scans have to be motion compensated before further processing in order to avoid motion induced inaccuracies. This task is also denoted as deskewing. The next step is the selection of a subset of all measurement points which shall be used for localization. In this thesis, the utilization of geometrical instead of intensity-based features is fostered which has the advantage of being less sensitive to challenging weather conditions and presence of high reflectivity features like road markings. Additionally, urban environments are particularly rich of locally unambiguous structures like house walls. Therefore, points from vertical surfaces are extracted from given LiDAR measurements which forms the basis for 3D LiDAR scan processing (Section 5.1).

### 5.1 Processing of 3D LiDAR scans

The main challenge is the high density of the 3D point cloud. To this point, a method for point selection with increased efficiency is developed. A first procedure uses normal vector estimation (Figure 5.3a) as basis for scan projection. This approach was developed in [RJMZ16], [Jat16] and extends existing approaches by an adaptive point neighborhood selection. Based on the local point neighborhoods, a normal vector for every scan point can be estimated. The radius for the neighborhood is adapted to the distance of the point to the sensor frame origin. Additionally, the velocity-dependent variation of the neighborhood radius can lead to increased



**Figure 5.2** All measurement points from a  $360^\circ$  LiDAR measurement are added to a depth image (a). Parked cars and house walls are visible in the corresponding intensity image (b). Segmentation in vertical and horizontal structures is achieved by utilization of Equation 5.1 (c).

detection performance of small structures like curb stones. The measurement points are added to a kd-tree structure to enable efficient local neighborhood search. This approach gives good scan projection results at the cost of decreased runtime performance for non-parallelized implementations.

To this point, a method on the basis of a two beam point selection criterion (Figure 5.3b) is developed. For points on vertical surfaces (e.g. walls) and with the range measurements  $r_1$ ,  $r_2$ , the distance of the surface from the sensor and the beam angles  $\alpha_1$ ,  $\alpha_2$ , it holds

$$d = r_2 \cos(\alpha_2) = r_1 \cos(\alpha_1) \quad (5.1)$$

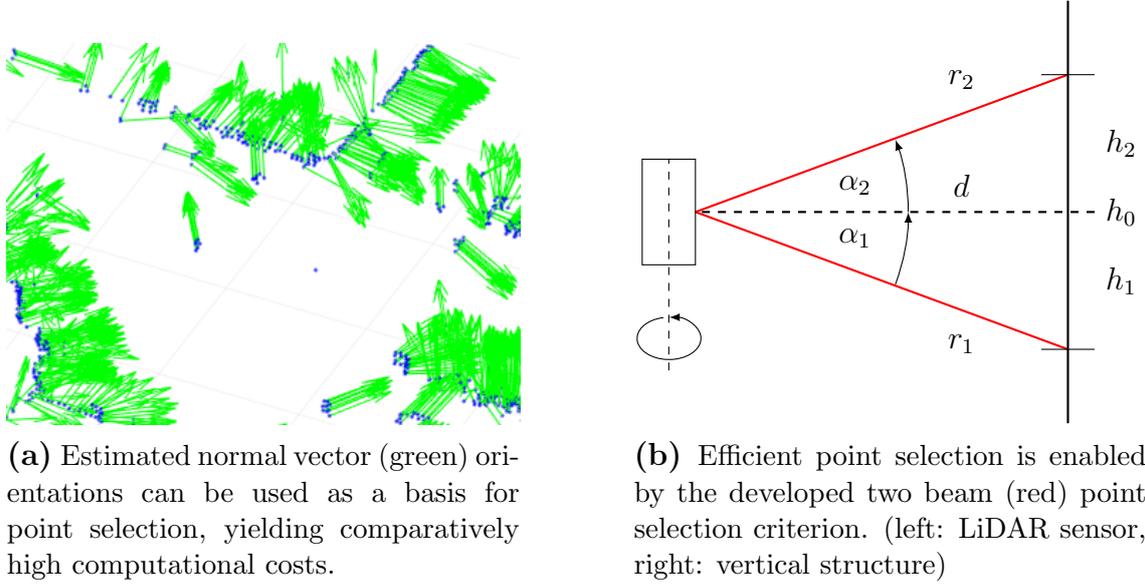
$$0 = r_1 \cos(\alpha_1) - r_2 \cos(\alpha_2) . \quad (5.2)$$

It follows the resulting condition for points in vertical surfaces:

$$r_1 \cos(\alpha_1) - r_2 \cos(\alpha_2) < \Delta_{thr} . \quad (5.3)$$

This criterion can be efficiently applied to depth images which are created from full  $360^\circ$  Velodyne spins as in Figure 5.2. For a better visibility of the objects in the scan, the corresponding intensity image is shown in Figure 5.2b.

Due to measurement noise, the difference in Equation 5.3 differs from 0 which is accounted for by definition of a threshold  $\Delta_{thr}$ . Threshold selection is conducted by analysis of the errors induced by range measurement uncertainty and under the



(a) Estimated normal vector (green) orientations can be used as a basis for point selection, yielding comparatively high computational costs.

(b) Efficient point selection is enabled by the developed two beam (red) point selection criterion. (left: LiDAR sensor, right: vertical structure)

**Figure 5.3** Prior to the projection of the LiDAR scan to the ground plane, point selection is performed. In Figure 5.3a and Figure 5.3b two possible approaches are presented.

assumption of uncertainty in the LiDAR sensor calibration

$$f := \frac{r_1}{r_2} - \frac{\cos(\alpha_2)}{\cos(\alpha_1)} \quad (5.4)$$

$$\Delta_{thr} = \left| \frac{\partial f}{\partial r_1} \right| \Delta r_1 + \left| \frac{\partial f}{\partial r_2} \right| \Delta r_2 + \left| \frac{\partial f}{\partial \alpha_1} \right| \Delta \alpha_1 + \left| \frac{\partial f}{\partial \alpha_2} \right| \Delta \alpha_2 \quad (5.5)$$

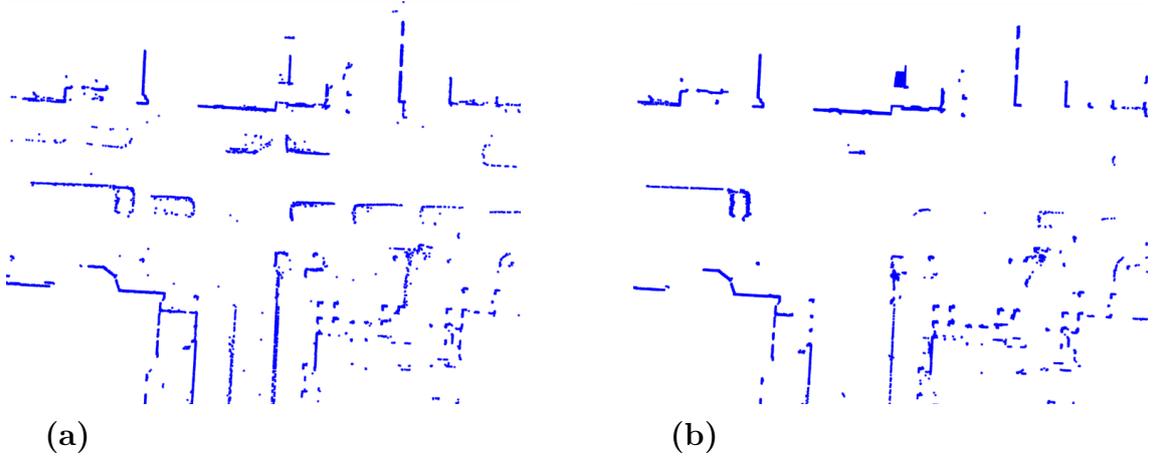
$$= \cos(\alpha_1) \Delta r_1 + \cos(\alpha_2) \Delta r_2 + r_1 \sin(\alpha_1) \Delta \alpha_1 + r_2 \sin(\alpha_2) \Delta \alpha_2 . \quad (5.6)$$

For an uncertainty in Velodyne HDL-64E range measurements of 0.06 m as stated in [GL10] and an angular uncertainty of  $0.01^\circ$ , the distance dependent threshold is calculated. In a distance of 10.0 m the threshold is  $\Delta_{thr,10} = 0.061$  m.

### 5.1.1 Generation of 2D LiDAR scans

2D scans  $\mathbf{s}_k$  are created from selected points in a given  $z$ -range, defined by the upper and lower bounds  $z_u, z_l$ . Point within this height range are projected to the  $xy$ -plane. A high range of  $z_l = -1.5$  m and  $z_u = 0.5$  m in sensor coordinates is chosen in order to maintain as much as possible structural information as possible while removing treetops. An exemplary result from each scan projection approach is shown in Figure 5.4. The normal vector based method yields the highest detection performance of vertical structures. This advantage is partially compensated by the increased computational complexity.

Both projection procedures are implemented in C++ and run in the robot operating system (ROS) under Ubuntu. On an Intel Core i7-4800MQ central processing



**Figure 5.4** The results from the scan projection approaches based on: (a) estimated normal vector orientations, (b) the point selection criterion

unit (CPU) (single core at 2.7 GHz), the average processing times for a data set with 2749 LiDAR scans (0823-42, Section 4.3) were 145.40 ms for the normal vector based procedure and 12.48 ms for the proposed projection approach. For the given hardware and implementations, the two-beam point selection approach runs in real time and is used throughout the following sections.

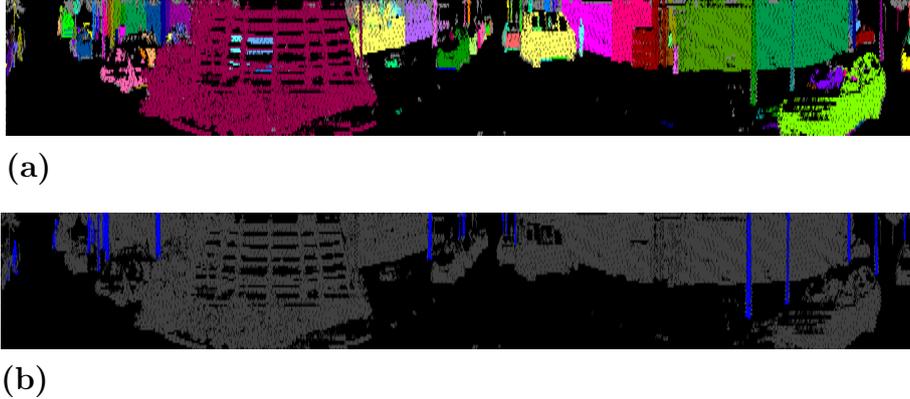
### 5.1.2 Detection of poles

In this subsection, the method used for the extraction of pole-like structures from laser point clouds is summarized. The LiDAR measurement points are added to a depth image (Figure 5.2) in which every row contains the data gathered by one laser diode during a 360° spin. Points on vertical surfaces are selected based on the criterion presented in Section 5.1.1 and all other points are removed. Segmentation is conducted for the remaining set of points based on a connected component approach [GMBN14, SS01] (Figure 5.5a). Segments of interest are selected based on the respective height to width ratios.

For every segment, a random sample consensus (RANSAC) based procedure is utilized in order to fit a circle model to the point  $xy$ -component of cloud data. Therefore,  $n = 3$  sample points are drawn at every iteration and the following cost function is used to estimate the circle model consisting of radius  $r$  and center point  $\mathbf{c}$

$$J_c = \sum_{i=1}^n \|\mathbf{W}_i (|\mathbf{c} - \mathbf{z}_i| - r)\|^2. \quad (5.7)$$

The number of model-supporting points is used to evaluate the model fitness. Additionally, only cylinder models with  $r < 0.2$  m are considered valid which seems feasible for most of the pole structures found in urban environments. Due to this



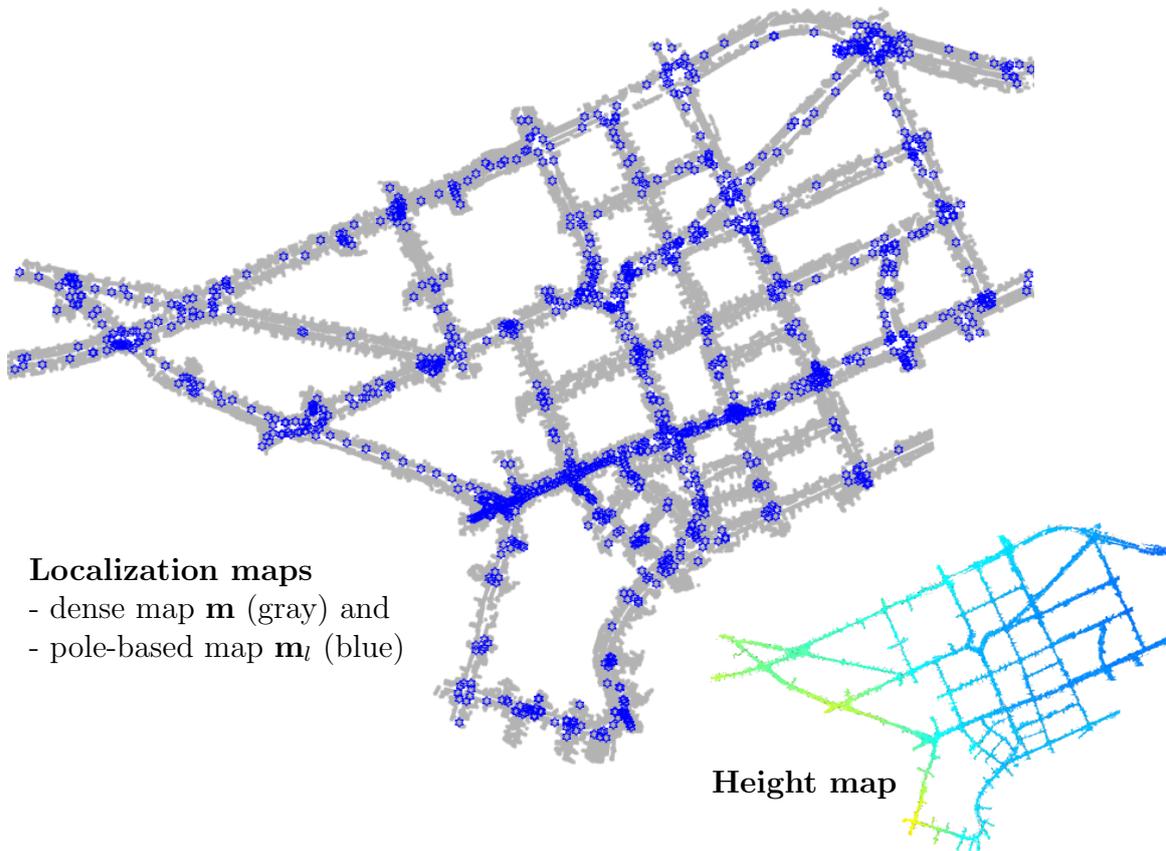
**Figure 5.5** Segmentation is performed for all points on vertical structures based on a connected component approach (a). The set of segments is filtered by the respective height-to-width ratios and a RANSAC-based algorithms is utilized to fit a cylinder model to the data (b).

constraint, pedestrians are not classified as poles. The runtime was measured for scan processing on a single CPU core. The pole detector has an average runtime of 75.3 ms for one Velodyne LiDAR scan for the considered data set 0823-42 with 2749 scans.

## 5.2 Processing of the localization map

Highly accurate localization maps are considered an important prerequisite for vehicle localization. Throughout this work, a prerecorded map of an urban environment in Stuttgart, Germany is used for the presented experiments. Map generation was based on 3D LiDAR scans from two single-layer LiDAR sensors and a high precision IMU. Both sensing planes were orthogonal to the driving direction and vertically oriented in order to capture features with high vertical offsets as well as the ground plane. The resulting map has a higher point density than maps created based on Velodyne HDL-64E measurements from a single run. For further utilization of the map in the 2D scan-based localization process, it has to be projected to the  $xy$ -plane. Thus, points on vertical structures are selected based on estimated normal vector orientations which yields the highest detection performance of vertical structures. The selected points are then projected to the ground plane and added to a grid map. In order to decrease random noise effects, a cell is marked as occupied if more than five points are enclosed at a grid discretization of 0.1 m. All other cells are marked as free or unobserved (Figure 5.6). Height information are stored in a height map (Figure 5.6) and can be accessed by  $xy$ -coordinates. The  $z$  coordinates are created from interpolation of the height values from the 3D map. For the pole-based localization procedure, poles have to be extracted from the dense 3D LiDAR map. To this point, it is resorted to manual extraction of the pole structures. The landmark positions are then given by the center of a set containing five points at the base of

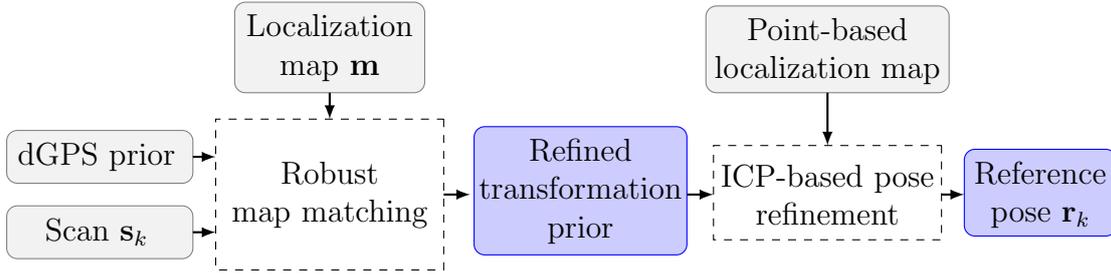
the respective pole. All map information are accessible via a database, developed in [Höf16].



**Figure 5.6** The localization map with a dense 2D grid  $\mathbf{m}$  and sparse pole-landmarks  $\mathbf{m}_l$ . The height map contains  $z$ -coordinates which can be requested from the developed map backend by the corresponding  $xy$ -coordinates.

### 5.3 Generation of a vehicle pose reference

Reference values for the map-relative localization systems have to be sufficiently accurate and insensitive to systematic errors. The latter might be caused by a misalignment between the reference frame, e.g. WGS84, of the reference system and the map frame. Additionally, in urban environments, popular global positioning system (GPS) based reference systems suffer from shadowing and multi-path propagation. These effects can lead to decreased accuracy and false covariance estimates. dGPS based evaluation of localization algorithms might lead to erroneous results as well. Potential errors in temporal synchronization further increase this effect.



**Figure 5.7** The proposed system for the generation of reference vehicle poses for localization system evaluation.

A possible option to mitigate problems imposed by GPS is to check the availability of satellites and weather conditions during the intended time window for test data recording. Additionally, the selection of a suitable test track with low urban canyons can lead to good results. However, in this thesis, the selection of the test track shall only be based on the environment characteristics that are interesting for the generation of test cases and not on the requirements of the reference system. Consequently, the evaluations in this thesis shall be based on a non-real time capable GPS-independent reference system which is developed in this section.

The developed (Figure 5.7) approach is similar to [LT10], where the influence of this misalignment is removed by looking at the differences between a reference graph and the one created with the localization system. Similarly, *Goshtasby* proposes to use a gold standard algorithm for the evaluation of image registration algorithms [Gos12]. In the proposed localization reference generation framework, measurements of a dGPS system are used to initialize the robust map matching algorithm developed in Section 6.3 with a grid resolution of 0.1 m. The resulting transformation is used as an prior for an pose refinement step on the basis of a robust iterative closest point (ICP) variant [Fit03]. For the ICP based matching, the original 2D point cloud of the localization map is utilized in order to gain an increased matching accuracy. The settings of the ICP are chosen in order to provide accurate matching result with high insensitivity to outliers, i.e. the maximum number of iterations is set to 1000, the maximum residual for stopping the matching process is set to  $10^{-16}$  and only correspondence pairs with an Euclidean distance of less than 0.1 m are considered. Correct convergence of the ICP is ensured by visual inspection and corrupted matching results are corrected manually. In order to evaluate arbitrary localization results relative to the reference poses, interpolation of the reference poses is conducted.

## 6 Registration of cluttered point clouds

The registration of sensor observations and localization maps forms the basis for map-relative vehicle localization. Many registration approaches are known from literature and have been successfully adapted to and applied within a manifold of application domains. In Section 6.1, an evaluation of the frequently used iterative registration algorithms ICP and NDT in the context of urban automated driving is conducted. This selection of registration algorithms is extended by a spectral matching approach which is shown to have favorable characteristics compared to well established iterative algorithms.

In Section 6.2, the registration algorithm is adapted to the field of urban vehicle localization. Covariance estimation for the case of subpixel-accurate matching is addressed. The characteristics of the spectral matching algorithm include a separate calculation of the rotation and translation transformation parameters. This fact is instrumented in Section 6.2.2 for the development of 3D spectral matching approach which outperforms other state of the art algorithms with respect to computational efficiency. A robust variant of the well known spectral registration algorithm is developed in Section 6.3 as an important key contribution. Furthermore, confidence measures are evaluated and their applicability to the assessment of the current matching quality is discussed.

### 6.1 Point cloud registration for UAD

This section is concerned with the evaluation and discussion of state-of-the-art matching algorithms in the context of urban automated driving (UAD). To this point, qualitative map matching requirements arising from their utilization in vehicle localization systems are derived first. Thereafter, the basin of convergence of spectral registration, ICP variants and NDT is evaluated and discussed in the context of the requirements imposed by the application domain. The result of this section is the well-grounded selection of a suitable class of registration algorithms which fits these requirements.

In the analysis of urban environment characteristics in Section 4.2, several aspects were identified which complicate the registration process. Outdated localization maps and cluttered LiDAR measurements are dominant sources of noise. Consequently, the following requirement is given:

- Registration algorithms shall be insensitive to noise.

Reusability of algorithmic building blocks has the potential to reduce the number of components of the overall system. Relevant applications include, localization system initialization and recovery from inaccurate priors. Hence, the second requirement:

- Registration algorithms shall be able to deal with priors of varying accuracy.

Vehicle localization systems have to be able to run in real-time which leads to the following requirement:

**Table 6.1** The parameters used for evaluation. For NDT, G-ICP and ICP the implementations from the Point Cloud Library [RC11] are used.

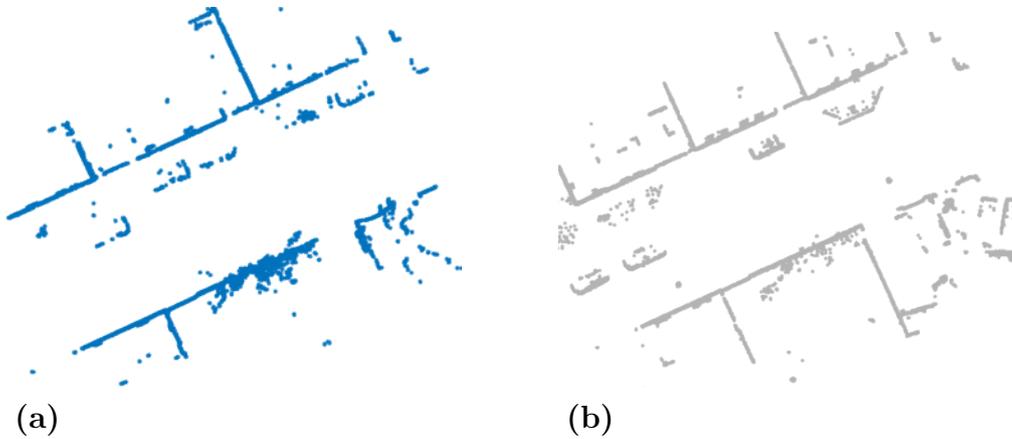
Algorithm	Parameters
spectral registration [CDD94]	<ul style="list-style-type: none"> <li>• grid resolution: 0.1 m</li> <li>• number of grid cells: <math>N = 512 \times 512</math></li> </ul>
Point-to-Point ICP [BM92]	<ul style="list-style-type: none"> <li>• convergence threshold: <math>\ \Delta\epsilon\  &lt; 10^{-6}</math></li> <li>• maximum iterations: 1000</li> </ul>
G-ICP [SHT09]	<ul style="list-style-type: none"> <li>• number of point neighbors for statistic calculation: 25</li> <li>• convergence threshold: <math>\ \Delta\epsilon\  &lt; 10^{-6}</math></li> <li>• maximum iterations: 1000</li> </ul>
NDT [BS03]	<ul style="list-style-type: none"> <li>• convergence threshold: <math>\ \Delta\epsilon\  &lt; 10^{-6}</math></li> <li>• number of iterations: 1000</li> <li>• cell size: 1.5 m</li> <li>• step size: <math>\ \Delta s\  = 0.05</math> for rotation and translation estimation</li> </ul>

- Registration algorithms shall be real-time capable and have a deterministic runtime.

For the evaluation, publicly available implementations of point-to-point [BM92] as well as plane-to-plane (G-) [SHT09] ICP and normal distribution transform (NDT) [BS03] from the Point Cloud Library (PCL) [RC11] are utilized. For spectral registration, it is resorted to an own C++ implementation of the original spectral registration algorithm proposed in [CDD94]. The data used for evaluation were collected from a Velodyne HDL-64E and are described in Section 4.3. The matching procedure parameters for ICP and NDT were selected similarly to [MVS<sup>+</sup>15] and are listed in Table 6.1.

In the first experiment, the basin of convergence of the different matching algorithms is evaluated. To this point, a scan  $\mathbf{s}$  from the urban data set 0823-42 and the corresponding submap from the localization map  $\mathbf{m}$  as shown in Figure 6.1 are selected. For the sake of clarity, only results for one scan-submap set are considered here. However, experiments with additional data were conducted and agreed with the results presented below. The reference pose is obtained based on the reference generation procedure proposed in Section 5.3. Consecutively, the transformation offsets  $\Delta\Theta$ ,  $\Delta x$  and  $\Delta y$  are added to  $\mathbf{s}$ .  $\Delta\Theta$  range from  $0^\circ$  to  $70^\circ$ ,  $\Delta x$  and  $\Delta y$  from  $-5$  m to  $5$  m. The ranges are selected as single-sided since the basin of convergence is almost symmetric for negative angles and such that the basin of convergence of the iterative algorithms is enclosed. This experiment enables the study of the sensitivity to Gaussian and non-Gaussian noise of the different registration algorithms.

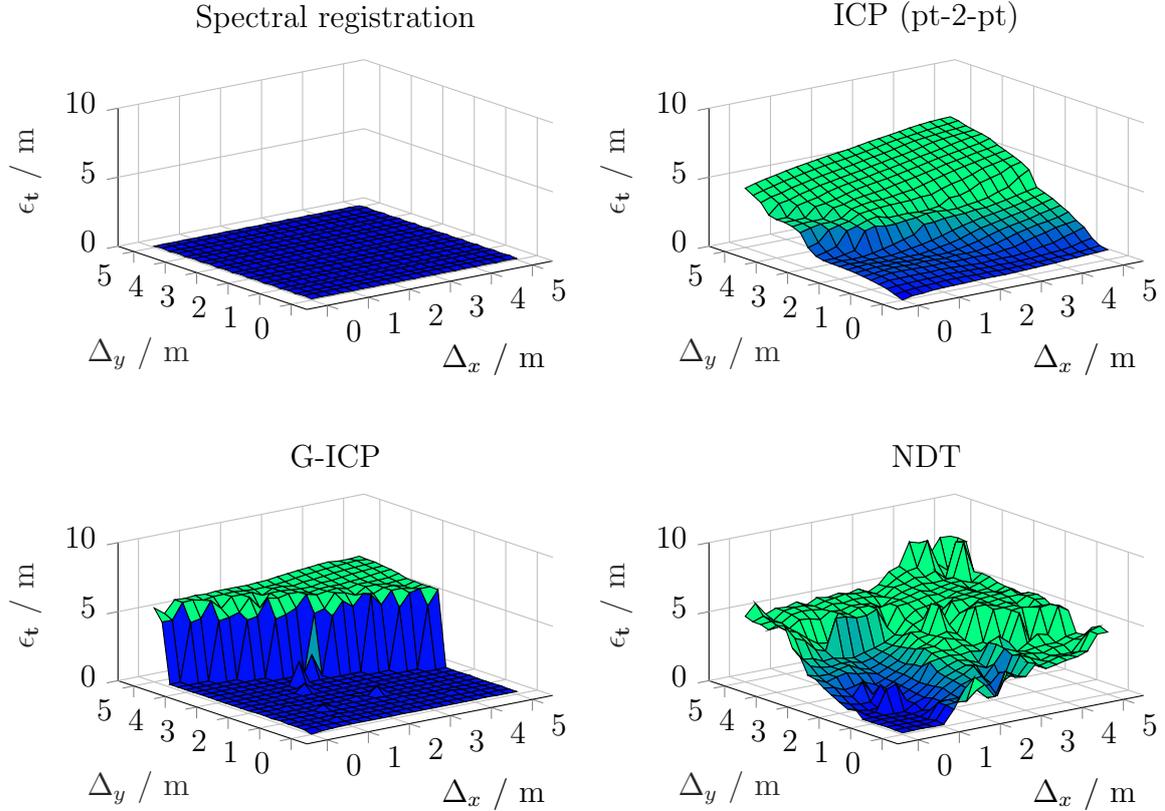
The results presented in Figure 6.2 illustrate the significant differences between the matching procedures. Although, the shape of the error surfaces might depend on the structure of the scan, the results correctly represent the general characteristics of the considered algorithms as observed from other test runs. Low offsets lead to good matching accuracy for all approaches. Due to the discretization of the input point clouds, spectral registration provides slightly less accurate transformation estimates. The differences occur for higher offsets, where NDT and ICP lead to steeply increasing errors in the matching results. Matching results obtained from spectral registration have the same uncertainty, independent from the prior accuracy. Hence, this algorithm is well suited for dealing with localization system initialization. For the considered input data and algorithm parameterizations, generalized ICP has the lowest sensitivity to inaccurate priors.



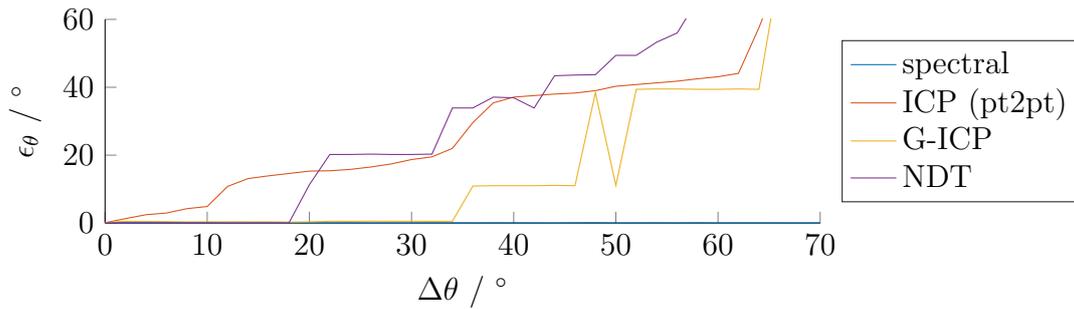
**Figure 6.1** The scan (Figure 6.1a) from data set 0823-42 and the corresponding submap (Figure 6.1b) used for experimental evaluation.

Implications of these results can be observed during motion estimation based on consecutive LiDAR scans without incorporation of prior information about the transformation parameters. In this case, narrow basin of convergence leads to imprecise motion estimates and increased the probability of failures. On the contrary, the spectral registration algorithm enables reliable relative motion estimation in this application. In following sections, this characteristic of the spectral registration procedure will prove as useful for the implementation of a self-contained scan-based odometry and SLAM framework which do not rely on additional sensor measurements. Thereby, redundancy is introduced to the localization system which decreases the sensitivity against failures of subsystems by decreasing the self-diagnosis capability.

Throughout the conducted evaluations, spectral registration had a constant runtime. This is opposed to the prior- and point data dependent runtime of iterative registration algorithms. Parallelization and removal of a subset from the point cloud can be used in order to reduce this effect. The removal of point measurements would lead to a randomization of the matching procedure which could potentially increase



**Figure 6.2** The matching errors  $\epsilon_t$  of the iterative procedures (ICP, G-ICP, NDT) depends on the initial transformation offsets  $\Delta_x$  and  $\Delta_y$ . Application of spectral registration omits this dependency.



**Figure 6.3** As for translation estimation, the rotation error  $\epsilon_\theta$  is independent of the accuracy of the initial transformation for spectral registration.

the susceptibility to registration performance degradation. Consequently, in the set of evaluated registration algorithms, inherently deterministic runtime behavior is exclusive to spectral registration.

In conclusion, it can be stated that spectral registration enables a non-iterative and deterministic transformation parameter estimation. Furthermore, no point cor-

respondence estimation is required, hence, omitting a common pitfall of iterative algorithms which commonly involve heuristics, e.g. for correspondence estimation. Especially for inconsistent sensor measurements, iterative approaches inherit the risk of convergence to local optima. For spectral registration, the amount of assumptions concerning the input data is low in comparison to other state of the art algorithms. The potentially loss of robustness due to high amounts of assumptions has already been addressed in this and other works, including the comprehensive point cloud survey by *Pomerleau et al.* [PCS15].

## 6.2 Spectral registration

In this section, the spectral registration approach according to [CDD94] is described and adapted to the intended application for UAD. Furthermore, a method for covariance estimation is developed. As discussed earlier in Section 6.1, matching based on a complete correlation between the inputs has advantages over classical point cloud registration procedures. A major disadvantage are the comparatively high computational costs for correlation calculation. Hence, significant research on the utilization of the Fourier phase of the input data has been conducted. An important advantage of this approach is the existence of efficient hardware implementations of the fast Fourier transformation (FFT). Due to the utilization of the spectral phase, registration algorithms from this category are denoted spectral registration procedures.

In the context of this work, scans and submaps are represented as binary occupancy grids with the states occupied and free. Additionally, two scans from consecutive time steps are matched for relative motion estimation. It is a well known fact in the image processing community, that the transformation between two translated image functions can be obtained by phase correlation methods. By interpretation of the input data as image function, these methods can be adapted to the described input data.

An overview over the spectral matching algorithm is provided in Figure 6.4. In a first step the discrete Fourier transformation (DFT) of the input images are calculated. The 2-dimensional DFT and its inverse are given as follows:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp\left(-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right) \quad (6.1)$$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \exp\left(j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right) \quad (6.2)$$

Windowing of the input data reduces the impact of occupied grid cells at the edges of the grid which can lead to artifact, potentially resulting in matching failures. To this point, a Hanning window is utilized. Highly efficient hardware and software implementations are available for the FFT which is of significant advantage since, the main computational costs in spectral registration arise from calculation of the

FFTs. The algorithm's characteristics in combination with efficient implementations render it well suited for vehicle localization with real-time requirements.

For a given pair of input data sets  $r(x, y)$  (reference) and  $s(x, y)$  (sensed), the Fourier transformed are denoted  $R(u, v) = \mathcal{F}\{r(x, y)\}$  and  $S(u, v) = \mathcal{F}\{s(x, y)\}$ .

Translational transformations of the input data occur as phase shifts in the frequency domain. Matched filter (MF), phase-only matched filter (POMF) and the symmetric phase-only matched filter (SPOMF) (Equation 6.3) allow the retrieval of the translation parameters from the phase shift. The transfer or filter functions are given as

$$H_{MF}(u, v) = \frac{R^*(u, v)}{|N(u, v)|^2} \quad (6.3)$$

$$H_{POMF}(u, v) = \frac{R^*(u, v)}{|R^*(u, v)|} \quad (6.4)$$

$$H_{SPOMF}(u, v) = \frac{R^*(u, v)}{|S(u, v)||R^*(u, v)|} . \quad (6.5)$$

After application of the filter function to  $S(u, v)$ , the filter output  $C(u, v)$  is obtained. The translation parameter estimates in pixels are then found by a maximum peak search in the inverse Fourier transformed filter response  $c(x_p, y_p)$

$$c(x_p, y_p) = \mathcal{F}^{-1}\{C(u, v)\} \quad (6.6)$$

$$= \mathcal{F}^{-1}\{S(u, v)H(u, v)\} . \quad (6.7)$$

The metric translation vector  $\mathbf{t}$  is given by the product of the grid resolution  $\Delta^p$  and  $\mathbf{t}^p$

$$\mathbf{t}^p = \arg \max_{x, y} (c(x_p, y_p)) \quad (6.8)$$

$$\mathbf{t} = \Delta^p \mathbf{t}^p . \quad (6.9)$$

Since, the phase of the input data is rotation variant, this approach enables the calculation of the translation between two inputs with insignificantly small rotational offsets.

In order to overcome this limitation, the rotation parameter has to be calculated in advance. *Castro* et al. addressed this problem in [CM87] by iteratively stepping through all possible rotation angles and applying the phase-only matched filter. The solution with the highest peak is then selected as the correct solution. Due to the repetitive application of the filtering step, this approach has high computational costs. An elegant extension for the calculation of rotation parameters has originally been introduced in [CDD94]. The central idea is the calculation of a translation invariant descriptor of the inputs and, thereby, decouple rotation and translation estimation. One possible solution is the calculation of the Fourier-Mellin invariant

descriptor (FMID) which is obtained by the Fourier-Mellin transformation (FMT) with polar coordinates  $\rho$ ,  $\Theta$  and the image function  $f$  as follows [CDD94]

$$\mathcal{M}(a, b) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{2\pi}^0 f(\Theta, \exp(\rho)) \exp(-j(a\Theta + b\rho)) d\Theta d\rho. \quad (6.10)$$

The FMID is translation invariant and rotations occur as phase shifts. Thus, the rotational offset between two inputs can be determined by the described filter functions. Due to the periodicity of the FMID, the result is  $\pi$ -periodic. If no prior for the rotation parameter is available, both rotation hypotheses have to be evaluated by copying and rotating the moving scan as shown in Figure 6.4b. The result is then given by the rotation hypothesis which is found by selection of the hypothesis with the highest peak in the translation estimation step together with the respective translation estimate.

The filter output gives a quasi-global matching result as shown for rotation estimation in Figure 6.5, i.e. it can contain several peaks which correspond to different transformation hypotheses. The distinctiveness of the peaks depends on the utilized filter function and noise characteristics of the inputs. Consequently, the selection of a filter function  $H$  which suites the input data is fundamental for the registration procedures correct functioning.  $H_{SPOMF}$  results in sharp and distinct peaks in  $c$ . This can be seen from the comparison of the normalized filter outputs in Figure 6.5. Due to the significant noise in the input data, the SPOMF filter function is selected for utilization in the context of urban vehicle localization.

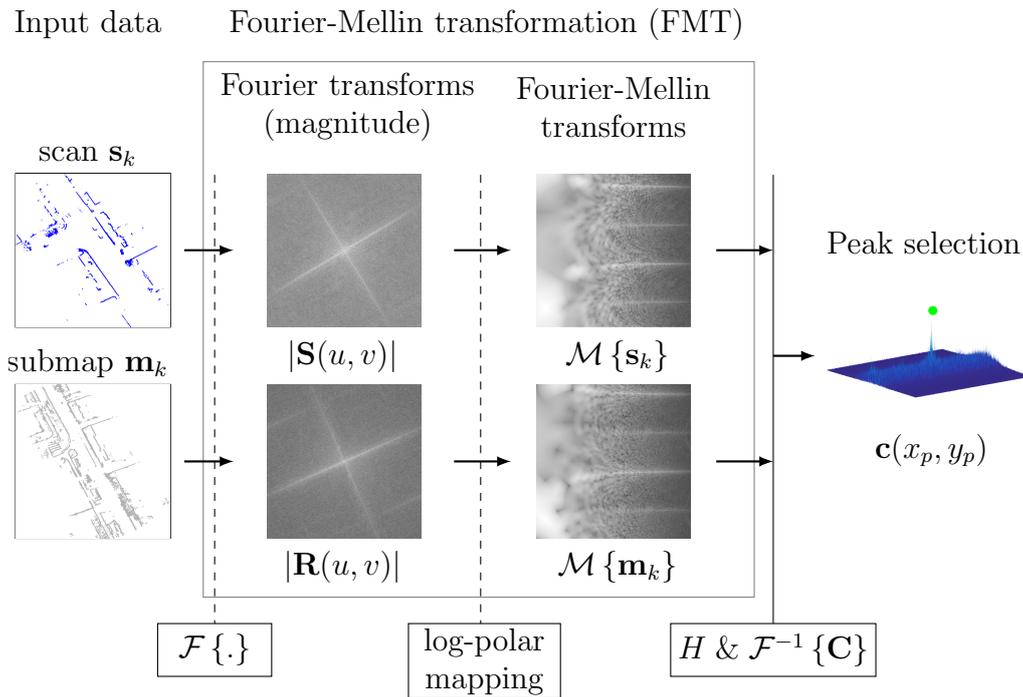
### 6.2.1 Accuracy improvement and estimation

For noise free input data and without discretization artifacts, the uncertainty in the matching result is limited by the grid resolution and the grid dimension

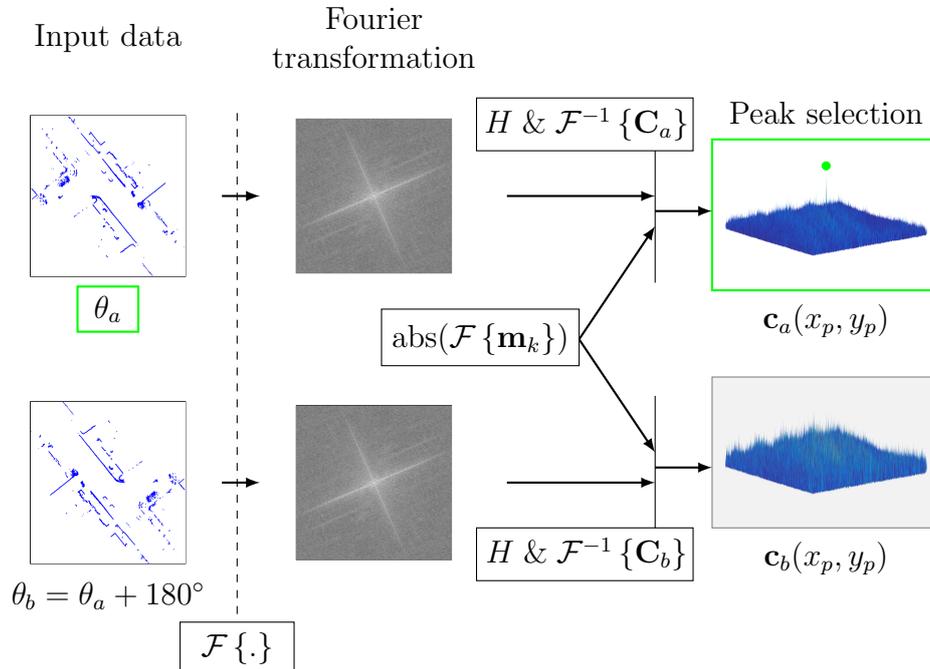
$$\sigma_{x,res} = \sigma_{y,res} = \frac{\Delta_{res}}{2} \quad (6.11)$$

$$\sigma_{\Theta,res} = \frac{360^\circ}{N}. \quad (6.12)$$

This bound is often exceeded due to noisy sensor measurements and discretization effects. By calculation of the subpixel-accurate peak coordinate  $\mathbf{x}_s$ , this limitation can be partly overcome. In the following, the weighted centroid method is used since it enables computationally efficient calculation of  $\mathbf{x}_s$  in comparison to most other interpolation-based methods.  $\mathbf{x}_s$  is calculated from all pixels within the radius  $K$  around the peak position (Equation 6.13). The gain in accuracy is evaluated in

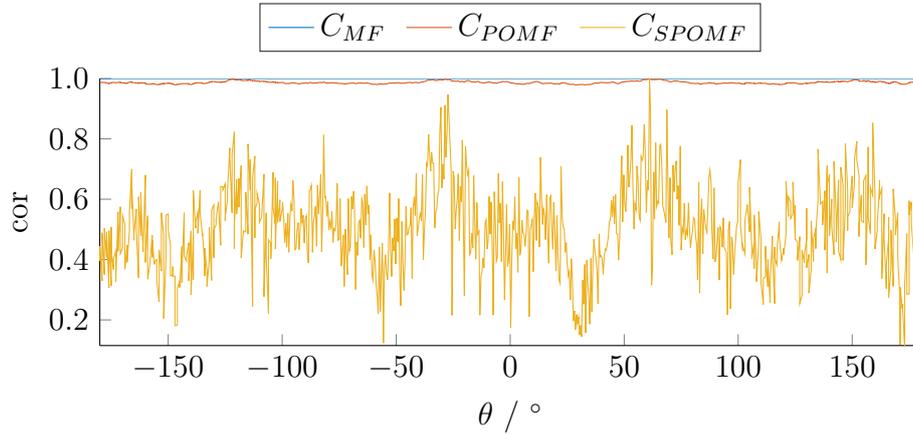
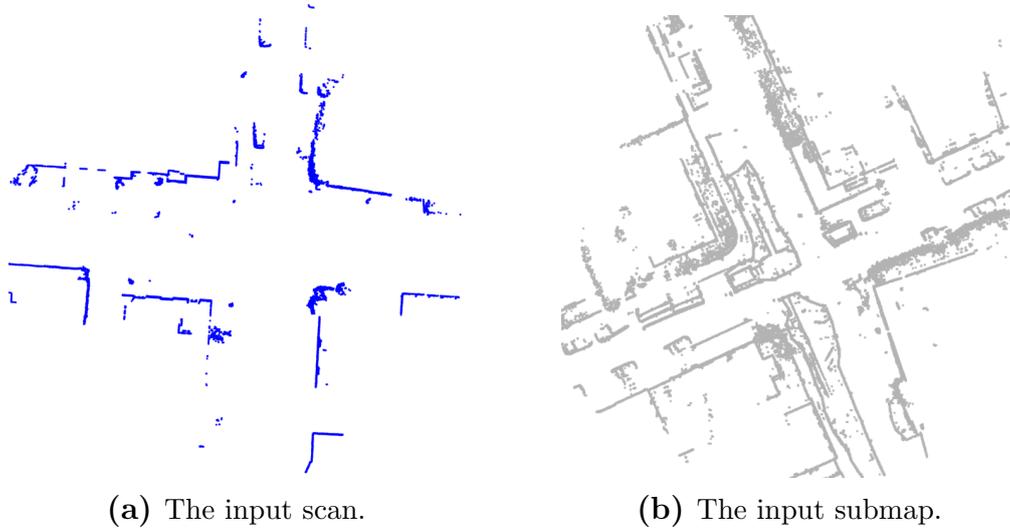


(a) Rotation estimation gives a  $\pi$ -ambiguous angle pair  $(\theta_a, \theta_b)$ .



(b) The best peak from both spectra is selected which gives the rotation and translation estimate.

**Figure 6.4** The Fourier-Mellin transformation is used in spectral registration to decouple rotation and translation estimation. Firstly, a rotation estimate is obtained (a). Consecutively, the input scan is rotated for translation estimation (b).



(c) Comparison of different filter responses for rotation estimation between the input scan and the submap.

**Figure 6.5** The filter function used to calculate the normalized filter responses for the input scan and submap have significant influence on the detectability of peaks in the filter response. Symmetric phase-only matched filtering maximizes the signal to noise ratio, resulting in well separated peaks. The symmetry of the intersection can be seen from the occurrence of several distinct peaks in the filter response.

Section 6.2.3.

$$\mathbf{x}_s = \begin{pmatrix} x_s \\ y_s \end{pmatrix} = \frac{\sum_{x=x_t-K/2}^{x_t+K/2} \sum_{y=y_t-K/2}^{y_t+K/2} c(x, y) \begin{pmatrix} x \\ y \end{pmatrix}}{\sum_{x=x_t-K/2}^{x_t+K/2} \sum_{y=y_t-K/2}^{y_t+K/2} c(x, y)}. \quad (6.13)$$

In mobile robotics and especially for SLAM and localization an accurate *covariance estimate* of the utilized matching algorithms is of fundamental importance. To this point, a covariance estimate is proposed as an extension to [PBB12] with respect to the subpixel-accurate estimate of the peak coordinates. To this point, the rotation and translation estimates are considered uncorrelated and a Gaussian is fitted to  $c(x, y)$  within a given region around  $\mathbf{x}_s$ . For the 2-dimensional case, the estimate can be obtained as follows and its extension to 3D is straight forward.

$$\mathbf{C}_{\mathbf{x}_s} = \frac{1}{N} \sum_{x=x_t-K/2}^{x_t+K/2} \sum_{y=y_t-K/2}^{y_t+K/2} \frac{c(x, y)}{c(x_t, y_t)} [x - x_s \ y - y_s]^\top [x - x_s \ y - y_s] . \quad (6.14)$$

$x_t$  and  $y_t$  denote the pixel coordinates of the peak in the filter output,  $K$  is the window size,  $c(x, y)$  is the filter output at point  $(x, y)$  and  $N$  the number of pixels in the input data.

### 6.2.2 3D spectral registration

The separation of rotation and translation calculation in the spectral registration algorithm enables a re-parameterization of the algorithm between both estimation steps. This characteristic is now utilized to extend the 2D to a 3D spectral matching approach with scalable computational costs. The aspect of computational efficiency has so far been widely neglected by the current state of the art what might be due to the difference in application domains. The main challenge in 3D spectral registration is the computationally demanding calculation of the roll, pitch and yaw angles. The main idea behind this 3D matching approach is the compensation of the global roll and pitch angles which can be measured by utilization of IMUs. Afterwards, the yaw-angle can be calculated in accordance to the 2D case.

Firstly, roll and pitch have to be compensated, e.g. by utilization of IMU measurements or under the assumption of vertical structures in the vehicle surrounding. 3D LiDAR scans are rotated accordingly and projected to 2D using the criterion presented in 5.1.1. The yaw angle is then calculated on the basis of the 2D scan. The 3D translation can now be calculated based on the rotation-compensated 3D scan. Computational efficiency can be gained by decreasing the vertical resolution which, however, results in decreased matching accuracy. The 3D FFTs are decomposed into several 2D (horizontal slices) and 1D (vertical) FFTs. Hence, parallelization can be used to decrease the runtime of the transformation step.

The separation of the estimation steps can be used for the development of various hybrid matching algorithms. The term hybrid is chosen in order to express the combination of different matching strategies for rotation and translation calculation. In the authors opinion, this is an interesting direction for future research in point cloud registration as it would enable the combination of point correspondence based and dense matching.

### 6.2.3 Experimental evaluation: runtime and accuracy

The evaluations in this section focus on the previously developed methods for the improvements and adaptation of 2D and 3D spectral registration. Data set 0823-42 (Section 4.3) is used due to the high outdatedness of the corresponding localization map and the coverage of heterogeneously structured areas. These include a round-about with vegetation and narrow urban canyons. All algorithms are implemented in C++ using methods from the OpenCV Library and run under the Robot operating system (ROS). The computer used for evaluation runs Ubuntu 14.04 (64-Bit) and is equipped with a Intel i5-4690 ( $4 \times 3.5$  GHz) from which one core is used. Execution times are measured based on the system time and, therefore, are approximate. An exact evaluation of the processing time is complicated, e.g. due to dynamic frequency scaling and the parallel execution of other processes. Nonetheless, the evaluation results show the gross-differences between the different approaches and enable a well-grounded discussion of their respective effectiveness. The average runtime and the standard deviations are determined based on the runtime for matching each of the 2749 Velodyne HDL-64E LiDAR scans with a corresponding localization submap. The runtime of the spectral registration procedure is evaluated for a grid size of  $N = 512$  and a resolution of 0.1 m of each grid cell. These setting correspond to a edge length of 51.3 m which is feasible for the utilization with the given sensor setup. The results are summarized in Table 6.3.

**Table 6.3** Results from the runtime evaluation.

Algorithm	$\bar{T}$ / ms	$\sigma_T$ / s
standard	88.30	2.81
storage of FFT (submap)	53.60	3.87
windowing	62.68	3.73
subpixel accuracy ( $K = 3$ )	64.30	3.51
rotation from prior	48.19	3.29

The standard version of the spectral registration algorithm takes 88.3 ms with a standard deviation of 2.81 ms for the calculation of the transformation parameters. Since the LiDAR sensor runs at 10 Hz, the matching and data preprocessing is close to real-time. By storing the FFT of the submap to be used during rotation and translation estimation, the runtime reduces to 61.7 ms which enables map matching in real time. Additional windowing and calculation of the subpixel accurate peak positions, results in a runtime of 64.30 ms. The lowest runtime is achieved if only one rotation hypothesis has to be evaluated, i.e. is a transformation prior is available. Since, all implementations, except the standard case, run in real-time, the version without prior incorporation is favored. This is due to the decreased requirements concerning the transformation prior.

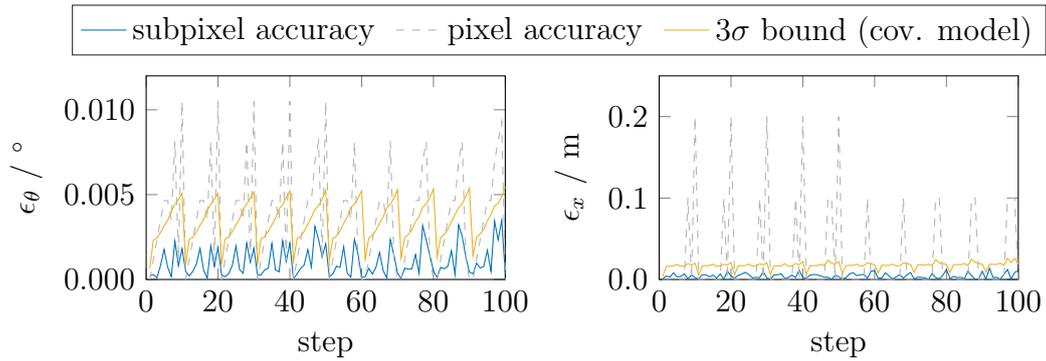
The matching performance of the FMT implementations is measured in the number of inlier pose estimates. For this purpose, outliers are defined as measurements with an error of more than 0.2 m relative to the reference described in Section 5.3. GPS measurements were used as an initial guess for the transformation parameters for each scans in data set 0823-42. The scans were then aligned to a submap of the localization map. For 2317 poses a correct matching result could be calculated based on the optimized spectral registration algorithm which translates to a matching rate of 84.3%. The most frequent error sources are analyzed in Section 6.3.1 and addressed by the introduction of a robust spectral registration variant in Section 6.3. The average matching accuracy for the standard spectral matching procedure is 0.12 m. The introduction of subpixel-accuracy leads to a significantly increased accuracy of 0.06 m.

For evaluation of the proposed 3D spectral registration approach, it is compared to the one presented in [LL02]. The results from Table 6.5 were determined based on a pair of identical 3D LiDAR scans which were transformed. The runtime for the standard algorithm is determined for a grid with an edge length of 128 voxels at a resolution of 0.2 m. It can be seen from the results, that the calculation of the roll and pitch angles is the main reason for the high computational complexity of 3D spectral registration. Consequently, the compensation of the two rotational degrees of freedom seems advisable.

**Table 6.5** Runtime comparison of the proposed and the state-of-the-art 3D spectral matching algorithm [LL02].

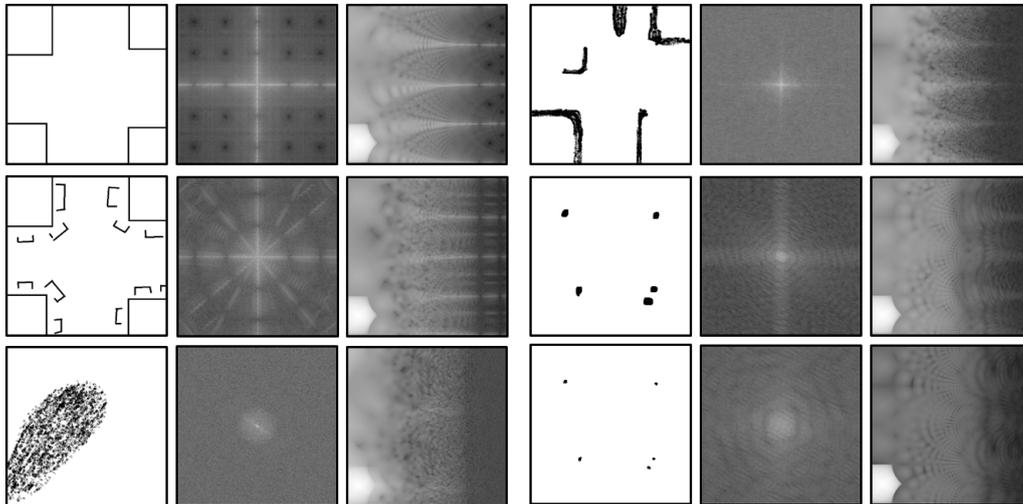
Algorithm	$t_{avg}/s$	$t_{min}/s$	$t_{max}/s$	multi-core processing
2d/3d	0.35	0.34	0.37	enabled
2d/3d	0.90	0.89	0.93	disabled
3d	9.43	9.10	9.73	enabled
3d	12.82	12.55	12.97	disabled

The covariance estimation performance on the basis of Equation 6.14 is conducted for a scan pair which was created by duplication of a scan from data set 0823-42. One scan is transformed so that the real transformation parameters are known and matched to the duplicated version. The evaluation results in Figure 6.6 were derived for  $K = 20$ . It can be seen from the results, that the covariance estimate is accurate for subpixel-accurate peak location calculation. For the standard matching algorithm, the covariance estimate is optimistic and it can be observed that the matching error varies between nearly discrete steps. Some results are very accurate, some have the magnitude of one or two grid cell discretization. In this case, the definition of peak-width intervals for determination of the discrete error level could potentially be preferable to the continuous formulation proposed in [PBB12].



**Figure 6.6** The covariance model gives good estimates for subpixel-accurate scan matching. In each evaluation step, a copied scan is translated and rotated, then registered with the original scan. Utilization of subpixel accuracy results in a significant increase in matching accuracy.

Various example input patterns were considered to determine the dependence of spectral registration performance on the existence of specific feature types and structures. The results are summarized in Figure 6.7. A prerequisite for reliable matching is the existence of distinct line structures in the Fourier-Mellin transformed. These are due to high frequencies in the Fourier spectrum which are mainly caused by line structures. The line structures do not have to be clearly defined but can be noisy as in the case of hedges. Sparse structures are a challenge for spectral registration.



**Figure 6.7** Examples for the Fourier- (right from input pattern) and Fourier-Mellin transformations (FMT) (right) of different input pattern (figures with white background). Lines and corners lead to high frequencies and distinguished lines in the transformed. Sparse structures as shown in the bottom left are challenging.

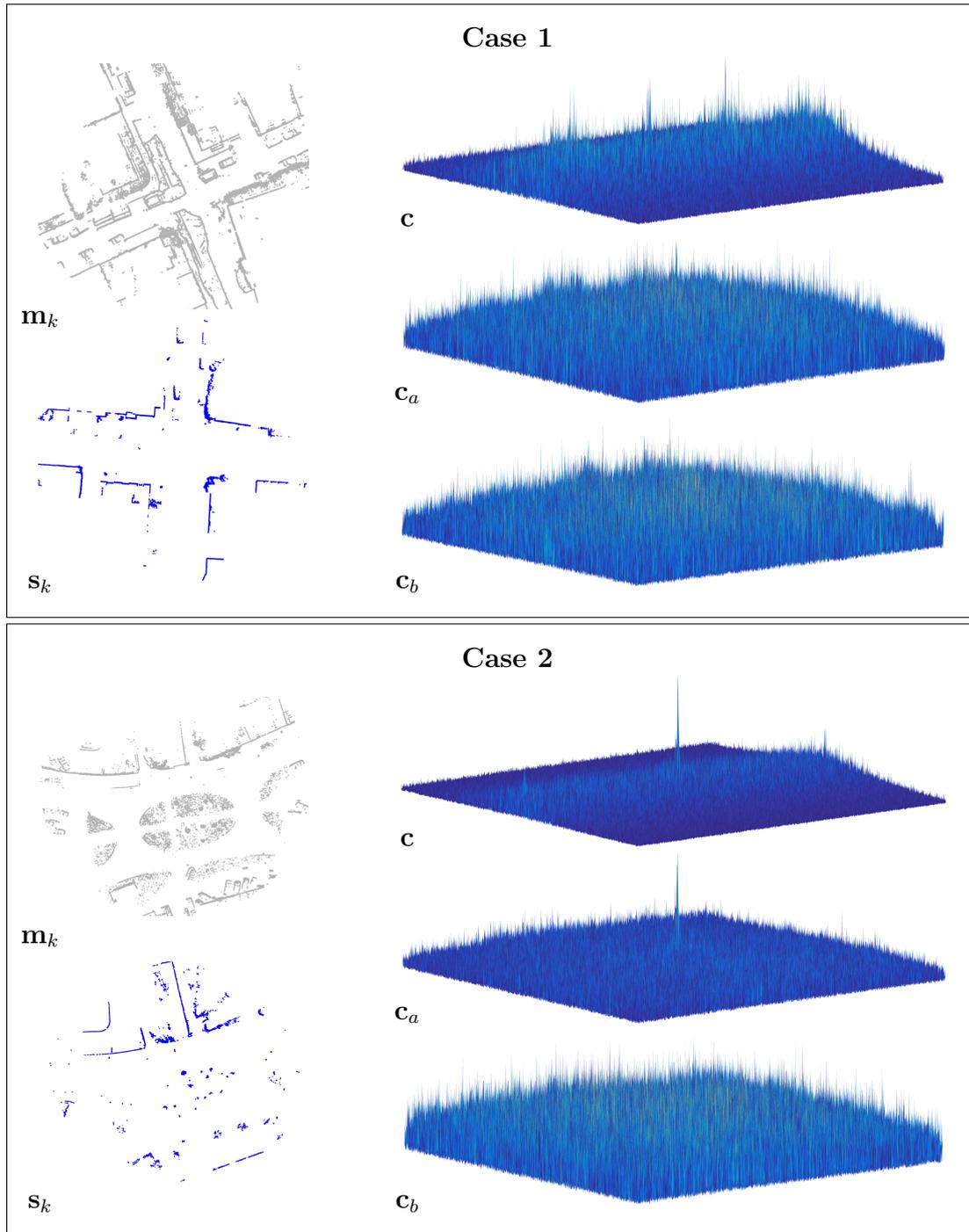
## 6.3 Robust spectral registration

In this section, a spectral registration algorithm is developed which involves the evaluation of multiple transformation hypotheses, resulting in an increased robustness against inaccurate transformation priors and noise. This algorithm is denoted robust FMT (rFMT) and leads to a significantly increased rate of successful matching attempts as compared to the standard single-hypothesis algorithm (Section 6.2). Spectral localization provides quasi-global matching results where hypotheses are represented by peaks in the filter response. Consequently, multiple hypotheses can be obtained by selection of several peaks. Challenges arise, once the highest peaks do not correspond to the correct set of transformation parameters. Two example scan-submap pairs are discussed in Section 6.3.1 that underline the limitations of the single-hypothesis approach to spectral registration. In Section 6.3.2, a procedure for the evaluation of multiple transformation parameter hypotheses is developed which overcomes the discussed limitations.

### 6.3.1 Limitations of single-hypothesis spectral registration

Common causes of matching failures in spectral registration are given in Table 6.7 and exemplified in Figure 6.8. Figure 6.8 illustrates different pairs of scans (data set 0823-42) and submaps. For each pair, the scan and submap are shown according to the prior rotation estimate and together with the set of filter outputs. In cases with sharply defined peaks, the correct translation and rotation parameters are easily obtained by selection of the highest peaks. For the following frequently occurring two examples, this approach fails.

For the first data set, the peak in the SPOMF response for the rotation calculation is less pronounced and ambiguities occur. However, in this case, choosing the highest peak still results in determination of the correct two rotation hypotheses. On the contrary, the highest peak in the translation calculation step corresponds to the false rotation hypothesis. This is due to the fact, that the overall noise in the filter responses is high and the peaks for rotation hypothesis selection are of equal height. Consequently, a limitation of the matching approach is reached if no rotation hypothesis can be selected with a sufficiently high probability which might be related to the distinctiveness of the peaks and the overall noise level. In this example, the obtained pose estimate has an error of several meters. Another limiting case is given by the second example in Figure 6.8. The filter response from the SPOMF for rotation estimation has a dense cluster of peaks. This cluster corresponds to several similar rotation hypotheses. Nonetheless, the correct rotation hypothesis is obtained during the translation estimation step. However, ambiguities occur in the SPOMF output corresponding to the correct rotation hypothesis, leading to the selection of an incorrect translation peak. Due to the peak layout, the resulting error is smaller than the one encountered in example two. This failure type is harder to detect due to the small error and consequently has an increased potential for having a deteriorating impact on the pose estimation integrity.



**Figure 6.8** Dealing with multiple hypotheses from the filter outputs is of fundamental importance for robust point cloud matching as shown by the following limiting cases. Case 1: The highest peaks in the translation estimation filter outputs  $c_a$  and  $c_b$  are of similar height. Due to noise, the wrong translation hypothesis is selected. Case 2: Selection of a wrong side peak in the rotation estimation step (c) leads to an increased error in parameter estimation.

**Table 6.7** An overview over common error sources in spectral registration.

Processing step	Error	Causes
rotation hypothesis selection	false peak selected	<ul style="list-style-type: none"> <li>• ambiguities</li> <li>• heavy clutter</li> </ul>
	false hypothesis selection	<ul style="list-style-type: none"> <li>• selection of the false translation hypothesis</li> </ul>
translation hypothesis selection	false peak selected	<ul style="list-style-type: none"> <li>• ambiguities</li> <li>• heavy clutter</li> </ul>

Therefore, the avoidance of this category of matching failures has a high priority. The main causes for matching failures are summarized in Table 6.7. Ambiguities in the SPOMF response often occur due to the corridor- (e.g. narrow roads) or x-structure (e.g. intersections) of the vehicle surrounding. Despite the structural-ambiguities, the correct rotation peak is typically clearly distinctive, but can have small side-peaks. Within these dense clusters of peaks, the highest peak does frequently not correspond to the correct peak. Especially for severe inconsistencies between or small overlap of the input data, small deviations in angle estimation can result in significantly increased noise levels in the translation filter responses (similar to case 2, Figure 6.8). This error type significantly contributes to increased matching failure rates. This observation is underlined by the results presented in Section 11.3.1 and motivates the development of rFMT in this section.

In conclusion, the example cases underline the challenges in peak selection. Before this study, these limiting cases for the spectral registration have by great extend been neglected in literature. Due to their importance for UAD, these are further evaluated in Section 6.3.2. As an important contribution in the next section, methods for overcoming some of the limitations are developed and evaluated.

### 6.3.2 Multiple hypotheses evaluation

The examples from Figure 6.8 underline the significant difficulty introduced by multiple similar hypotheses and unveil the limitations of a peak selection solely based on the information provided by the filter outputs. To this point, further means are developed in this section for reliable peak selection and matching failure detection. Obviously, the introduction of prior knowledge can be used for narrowing the hypotheses search space and thereby decreasing the impact of ambiguities. This approach effectively leads to a loss of the broad basin of convergence of the matching algorithm and prohibits an application for localization system initialization (Section 11.3.1). Additionally, the second example in Figure 6.8 shows that the distance between hypotheses can be small, leading to high prior accuracy requirements.

Each transformation hypothesis can be evaluated based on point-based metrics

which are used in point-based registration algorithms like ICP and RANSAC. The information contained in the filter responses used for spectral registration typically not equivalent. In the following, a general approach for the application of point-based fitness metrics in spectral registration is developed. Transformation parameter hypotheses can be extracted in an iterative manner from the filter outputs calculated during the rotation and translation estimation steps. To this point, the highest peak is searched and the peak is set to zero. This procedure is repeated until the required number of hypotheses is reached. Typically, the peaks are broadened due to noise in the filter input data. Therefore, it is argued that a neighborhood around the peak shall be set to zero as well. The neighborhood size has to be carefully chosen in order to avoid consideration of parts of broad peaks as distinct hypotheses and to avoid rejection of densely clustered hypotheses. Alternatively, a gradient-based algorithm leads to accurate results by setting all adjacent cells with a negative gradient to zero. For efficiency reasons, it is resorted to a fixed neighborhood of  $3 \times 3$  grid cells around the peak.

For an arriving LiDAR scan  $\mathbf{s}_k$  at time step  $k$ , a set  $\mathbf{U}_k$  of  $U_k$  random sample points

$$\mathbf{U}_k = \{\mathbf{u}_j\}_{j=1}^{U_k} \quad (6.15)$$

is extracted (Figure 6.9). The hypothesis set  $\mathbf{H}_k$

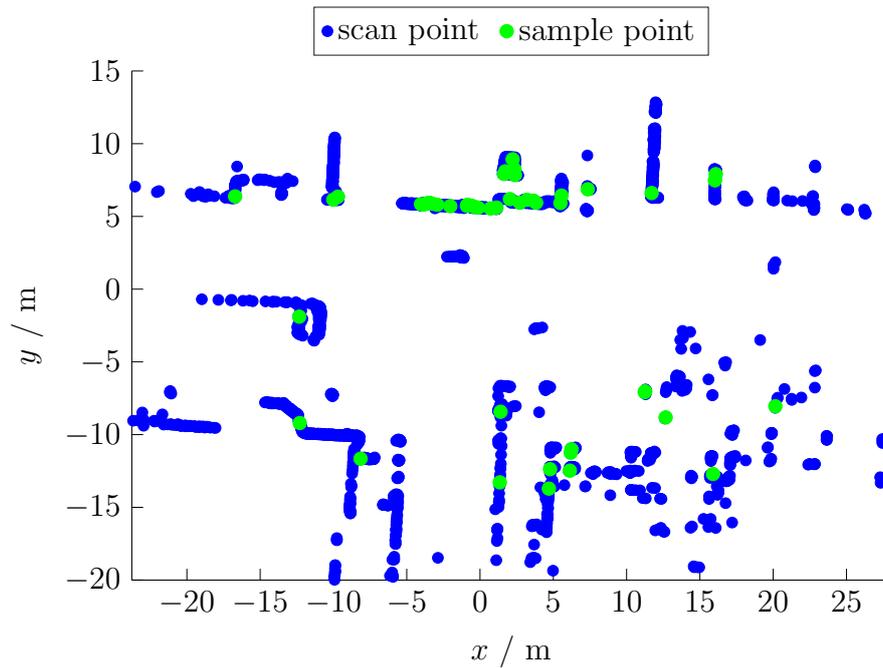
$$\mathbf{H}_k = \{\mathbf{h}_i\}_{i=1}^{H_k} \quad (6.16)$$

is selected.  $\mathbf{U}_k$  is then transformed according to the  $\mathbf{H}_k$  to obtain a set of transformed sample point sets

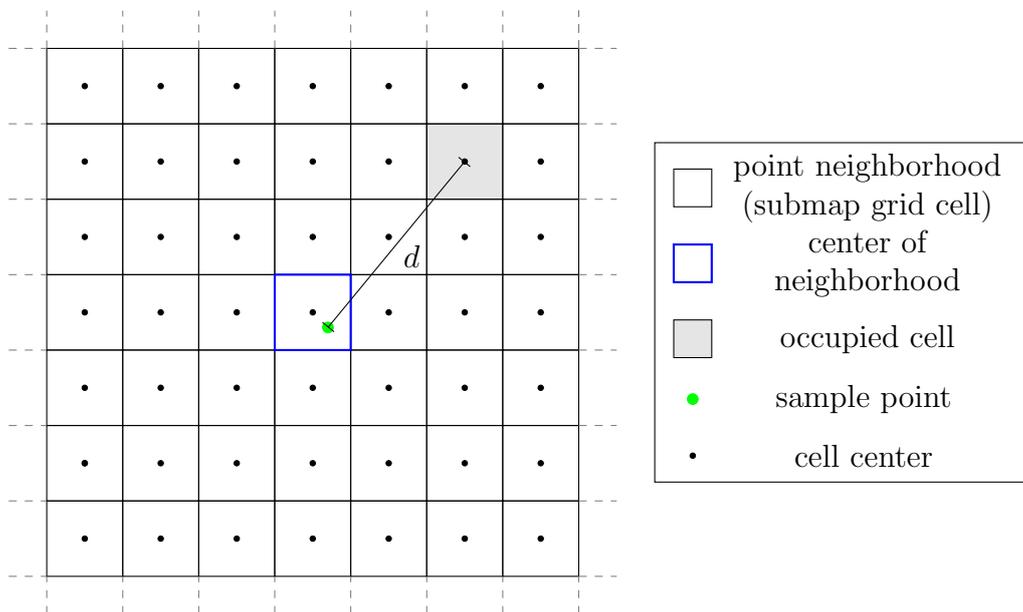
$$\mathbf{P}_k^{\mathbf{H}} = \{\mathbf{U}_{k,i}\}_{i=1}^{H_k} . \quad (6.17)$$

Consecutively, a nearest neighborhood search is used to determine the corresponding points in the reference data (submap  $\mathbf{m}_k$  or scan  $\mathbf{s}_{k-1}$ ). The efficiency of the nearest neighbor search is crucial for the applicability of the hypothesis selection procedure. To this point, a suitable procedure is developed in the following. For the distinction of good and poor registration results, the consideration of a tight neighborhood around the transformed sample point as shown in Figure 6.10 is sufficient. The center of the neighborhood is obtained by a mapping between sample point coordinates and grid cells of the current submap. If no occupied grid cell is part of the neighborhood, the Euclidean distance is set to  $\infty$ . A  $7 \times 7$ -neighborhood is characterized by a good trade-off between computational costs and accuracy.

The correspondence pairs are then utilized to calculate the median distances  $d_i$  for each hypothesis. Each  $\mathbf{h}_i \in \mathbf{H}_k$  is then weighted according to the respective distance  $d_i$ . Accordingly, rFMT is a hybrid matching algorithm which uses elements from the dense registration domain for data association and means from the point-based registration field for hypothesis weighting.



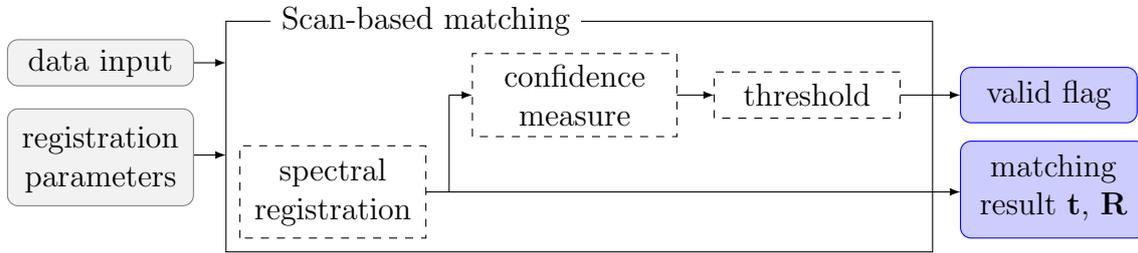
**Figure 6.9** An example scan with  $U_k = 50$  randomly drawn sample points used in rFMT as basis for the weighting of multiple transformation hypotheses.



**Figure 6.10** The sample points from the input scan are transformed in accordance to the transformation hypotheses. The Euclidean distances are determined between the map frame relative sample points from the scan and  $7 \times 7$  grid cell neighborhoods. This approach leads to accurate results and efficient implementations.

### 6.3.3 Confidence measures

In order to increase the localization system robustness, confidence measures are required which can be used to evaluate the current matching performance beyond covariance estimation. Based on the current confidence, a threshold can be enforced as mean for outlier rejection as shown in Figure 6.11. Several confidence metrics have been introduced in the past for spectral image registration, including works in [KH90] and an heuristic approach in [AAS15]. In most case the evaluation of the confidence metrics has been conducted under the assumption of zero-mean Gaussian noise. In the context of this thesis, the latter assumption does not sufficiently reflect the noise characteristics of the utilized data and, therefore, a re-evaluation of existing approaches and the point distance measure from Section 6.3.2 is performed.



**Figure 6.11** The concept of confidence measure based matching result selection.

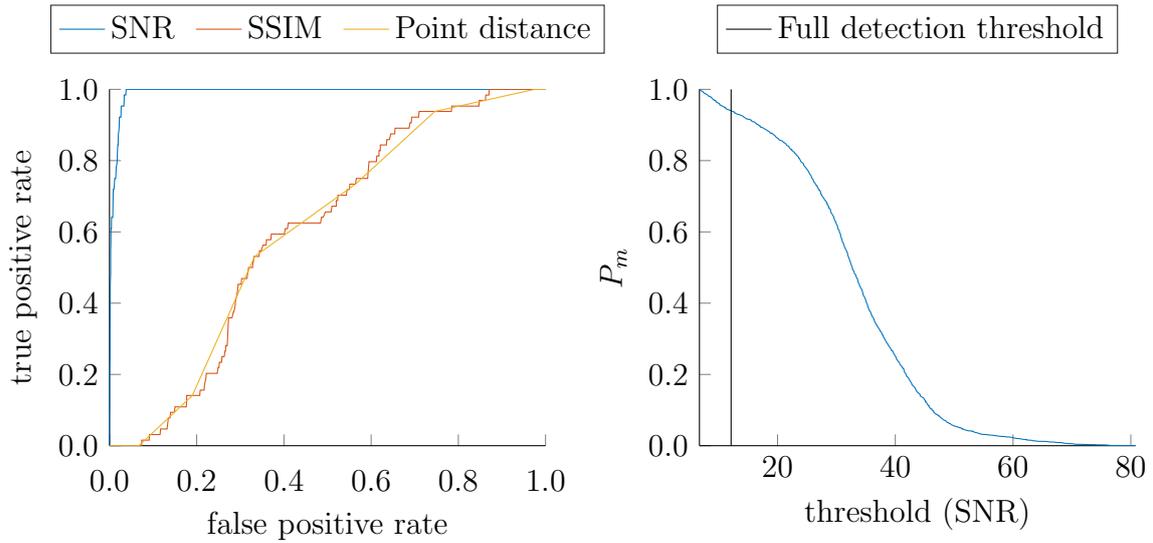
Quantification of the matching quality can be achieved by evaluation of the signal-to-noise ratio (SNR) of the selected peak. The SNR is defined as [CDD94]

$$SNR = \frac{c(x_s, y_s) - \text{mean}(c(x, y))}{\sigma_c} . \quad (6.18)$$

Alternatively, structural similarity (SSIM) [WBSS04] can be used to evaluate the similarity of the transformed scan and the corresponding submap. SNR, SSIM and the point distance measure based approach from Section 6.3.2 are now further evaluated with respect to their applicability to outlier detection. The results for data set 0823-42 in Figure 6.12 indicate a high correlation of SNR with the correctness of the matching result. SSIM and the point distance based approach developed in Section 6.3 showed superior performance for the selection of the best out of multiple hypotheses. In comparison to the SNR, both measures are insensitive to incorrectly selected peaks. Consequently, both are not well suited for outlier detection.

For the SNR-based outlier detection, all false matching results are detected for a threshold of 12.06. By selection of this threshold and under consideration of the corresponding false positive rate, the overall matching rate is 94.0% which is below 97.8% from Table 11.5. The runtime for calculation of the SNR and the point distance is around 0.5 ms on average. The calculation of the structural similarity takes several ms.

In conclusion, the evaluation results from the this section show that the confidence estimation at matching algorithm level is well performing but is subject to limita-



**Figure 6.12** Receiver operating characteristic curve (ROC) for matching outlier detection based on different confidence measures. The SNR gives the highest detection performance. Due to false rejections, the matching rate  $P_m$  is decreased while the true rejection rate is increased. Consequently, robustness of map matching can be achieved and scaled at the cost of a decreased matching rate.

tions. Consequently, the decision about accepting or rejecting matching results shall be supported on a higher hierarchical level under consideration of redundant map matching results. This approach will be elaborated in Section 11 in the context of the development of a hybrid localization framework.

## 6.4 Conclusion

This section started with an evaluation of matching algorithms for challengingly heterogeneous urban environments, outdated localization maps and cluttered LiDAR measurements. Especially the aspect of insensitivity against noise and inaccurate priors motivated the selection of spectral registration as a mean for the enhancement of localization robustness (compare Section 2.3). Adaptation of the matching algorithm to the intended usage in the field of urban automated driving were made. A robust spectral registration algorithm (rFMT) was developed which enabled the consideration of multiple transformation hypotheses. This led to increased matching performance and can be utilized as a basis for localization system initialization from inaccurate transformation priors.

The contributions in this section are part of key contribution 1 (*kc1*, Section 1.2) and comprise the following contributions.

- The selection and adaptation of a spectral registration algorithm as central component in a vehicle localization system for urban automated driving (*rq1*), comprising
  - the adaptation to the expected noise characteristics by selection of a suitable filter function,
  - covariance estimation and
  - evaluation of confidence measures for the detection of matching failures.
- The developed robust spectral registration algorithm (rFMT) based on a hypothesis selection procedure led to a significantly increased matching rate (*rq1*, *rq3*, *rq4*).
- The development of a 3D spectral registration procedure yielded scalable computational costs for point cloud alignment.

The evaluation of the spectral matching approach in Section 6.2.3 on the basis of a challenging urban data set indicates its insensitivity to inconsistencies between sensor measurements and the localization map. Accordingly, the main characteristics of spectral registration can be summarized as follows

- high insensitivity to inconsistencies between the input data (*rq1*, Section 1),
- low requirements on prior accuracy which enables system initialization and recovery (*rq3*, Section 1),
- deterministic runtime behavior and availability of well researched software and hardware implementations for the most important algorithmic building blocks,
- quasi-global matching results and
  - expressive confidence metrics (*rq4*, Section 1)
  - hypotheses selection strategies (*rq1*, *rq3*, Section 1)

These characteristics enable to bridge the gap between place recognition and iterative algorithms used for pose refinement. Furthermore, it covers important tasks previously addressed by the latter procedures. Confidence estimation based on the SNR showed good performance for the identification of valid matching results. High probabilities of outlier removal on the basis of the SNR trades-off with suboptimal matching rates due to the rejection of valid results. The probability of detecting false matching results is smaller than one and hence further means have to be taken in order to answer *rq4* (Section 1). This question will be addressed by analytical redundancy in combination with consistency checking in Section 11.2.

Especially for areas with sparse structures (compare Figure 6.7) and depending on the parameterization of the algorithm, the spectral matching approach can suffer from degraded matching performance. In order to underline this argument, results from a short experiment are presented. Robust spectral registration with a grid resolution of 0.2 m is used in order to align LiDAR scans to the localization map. Figure 6.13 contains the results for the urban data set 0823-42 (Section 4.3) which was selected due to the compact occurrence of diverse challenges, fostering lucidity for this short overview. Based on the SNR values, it is possible to identify areas in which spectral matching delivers reliable matching performance. This is especially given for narrow roads and generally in areas with dense structures. The amount of clutter has no significant influence on the matching performance in these areas. Given the beneficial characteristics of the spectral registration algorithm and the discussed challenges, the following applications are further treated in this thesis:

- relative motion estimation (Section 10.1) contributes to the compensation of matching failure clusters,
- simultaneous localization and mapping (Section 10.2) enables operation in areas without valid localization maps and
- vehicle localization (Section 11) with redundant map matching procedures leads to increased failure detection capability and enhanced robustness.

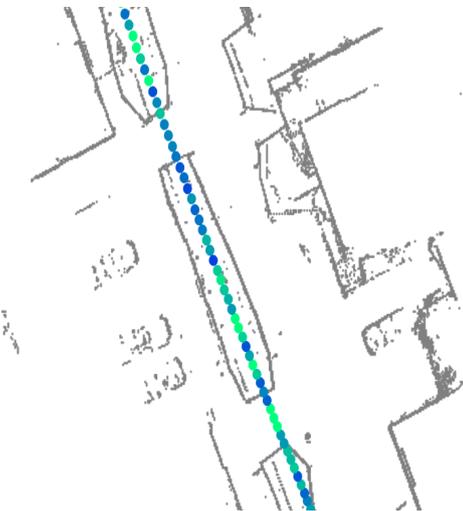
The distribution of the matching failures is of significant importance for the environment and localization system model, developed in Section 9 and following sections. In the evaluated data set, two major categories of matching failure sequences can be distinguished:

1. approximately uniformly distributed failures and
2. clusters of failures.

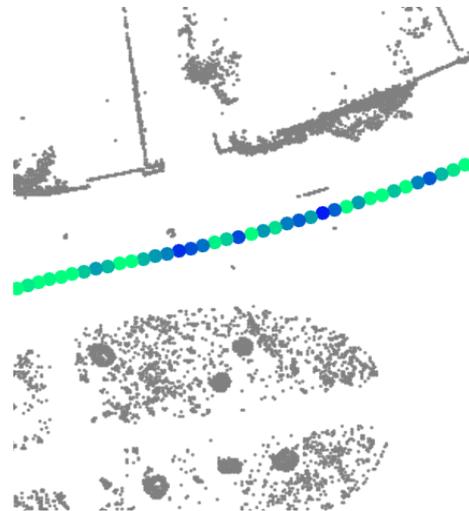
The question of how the observed challenges in map matching affect the localization system and influence the localization system design remains unanswered (*rq2*, Section 1). More generally, it lacks a systematical approach for localization system design which captures the operation environment characteristics, localization subsystems and the actual localization requirements. To this point, the design of localization systems is treated in Part III as another key contribution (*kc2*, Section 1.2) in this thesis. It is argued, that statistical models of the overall localization system can support the design process. Therefore, statistical localization system models are developed in Section 7 and Section 8. The statistical models also capture the discussed matching failure sequence characteristics. Comprehensive simulation studies are conducted in Section 9 to derive important design parameters, like grid resolutions and matching rate for spectral matching as well as odometry accuracy. Additionally, strategies for adaptation of localization systems to the respective target operation environment are discussed.



(a) The SNR values significantly vary along the driven trajectory. Blue dots indicate low SNR values and green dots high values above 40.0. Detailed views are shown in the following figures.



(b) Significant inconsistencies between sensor observation and submaps lead to decreased SNR values.



(c) Weakly structured areas have a comparable influence on the matching performance.

**Figure 6.13** Spectral registration works well in areas with narrow roads and house walls in proximity of the LiDAR sensor (compare Section 6.2.3), whereas in expansive intersections, the matching becomes less reliable. This is indicated by a decreased SNR in these areas which indicate degradations in map matching performance.

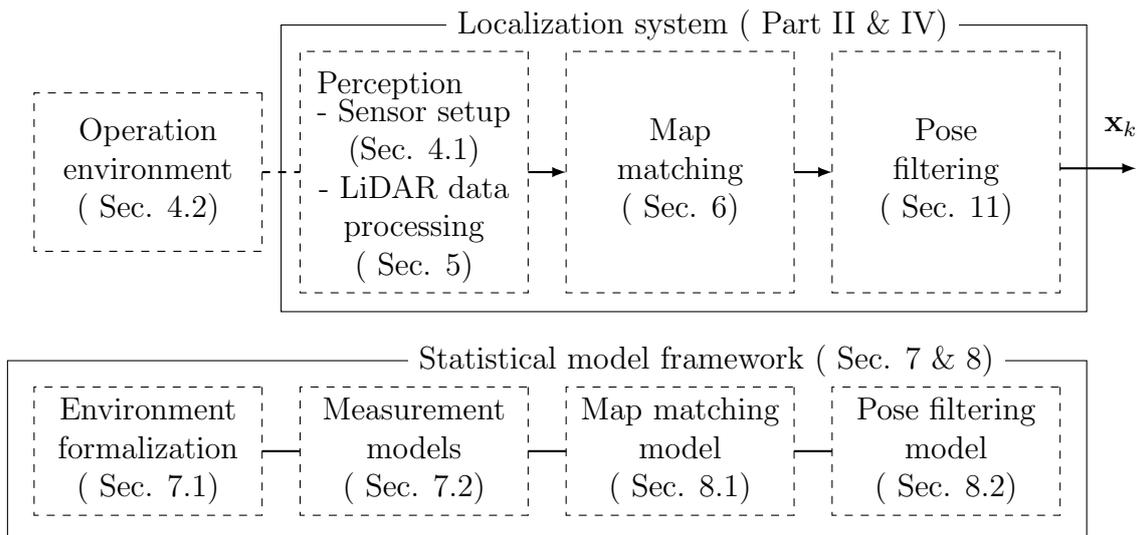
**Part III**  
**Model-based localization system  
design**



## 7 Measurement and environment models

The design of localization systems remains a challenging task and requires high amounts of expert knowledge. Statistical models foster the quantification and harmonization of this valuable expert knowledge. Thereby, a model-based design approach can contribute to a localization requirement oriented system development. Furthermore, the design process can be made more repeatable since the models can be applied to a diverse set of localizer concepts and operation environments. Utilization of statistical models can also lead to a deeper understanding of localization systems and the impact of specific design decisions on system performance parameters. A model-based design approach can thereby provide an additional source of ideas for approaching the remaining challenges in vehicle localization for urban automated driving (UAD). To this point, a statistical model framework for the whole signal processing chain of a localization system (Figure 7.1) is derived in this and the following section. It is resorted to analytical model formulations, facilitating a general discussion of the interrelation between localization system design parameters, operation environment characteristics and localization accuracy requirements (as conducted in Section 9) in comparison to numerical models. The models are given in closed-form which enables an efficient evaluation in simulation studies. The described models have been published in [RSMZ15], [RSMZ16] and [RMZ17] to great extend and offer the basis for answering research question 2 (*rq2*, Section 1.1).

In this section, the detection of pole-features as described in Section 5.1.2 and utilized in the proposed localization framework (Section 11) is considered. Map matching on the basis of scans as described in Section 6 is treated beginning with



**Figure 7.1** The developed statistical models cover the whole signal processing chain of a localization system and capture the operation environment characteristics.

Section 8.2, where the matching result is characterized by its uncertainty.

The measurement models form a direct link between the environment characteristics, sensor properties and the utilized feature extraction algorithm. Exteroceptive sensors are utilized in order to provide a set of landmark observations  $\mathbf{L}_k = \{\mathbf{l}_{i,k}\}_{i=1}^{n_k}$  at every time step  $k$ . Every landmark observation  $\mathbf{l}_{i,k}$  corresponds to a  $\mathbf{m}_{i,k} \in \mathbf{m}_l$  which requires the detection and removal of false positive landmark detections from the sensor observation. This assumption is feasible since it is achieved by a suitable matching algorithm with a sufficiently high probability. The sensor field of view (FoV)  $\mathcal{S}$  is defined by a maximum detection distance  $r_{max}$  and an opening angle  $\alpha$ . The location of a landmark relative to the sensor reference frame is given by the homogeneous transformation  $\mathbf{T}(\mathbf{m}_{i,k}, \mathbf{x}_k)$ . A landmark is observed with the, potentially distance- and angle-dependent, detection probability  $P_D$  if  $\|\mathbf{T}(\mathbf{m}_{i,k}, \mathbf{x}_k)\| < r_{max}$  and  $-\frac{\alpha}{2} \leq \angle \mathbf{T}(\mathbf{m}_{i,k}, \mathbf{x}_k) \leq \frac{\alpha}{2}$ .

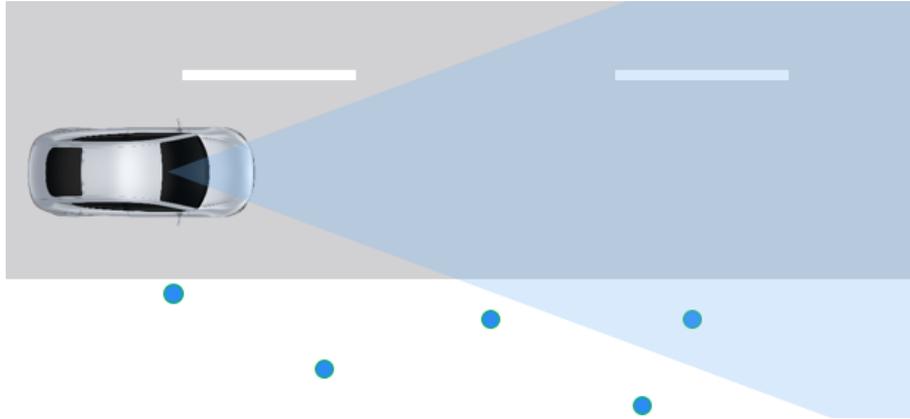
$$\mathbf{l}_{i,k} = \begin{cases} \mathbf{T}(\mathbf{m}_{i,k}, \mathbf{x}_k) + \Delta \mathbf{l}_{i,k} & \mathbf{T}(\mathbf{m}_{i,k}, \mathbf{x}_k) \in \mathcal{S}(\mathbf{m}_l) \\ \emptyset & \mathbf{T}(\mathbf{m}_{i,k}, \mathbf{x}_k) \notin \mathcal{S}(\mathbf{m}_l) \end{cases}. \quad (7.1)$$

The measurement error  $\Delta \mathbf{l}_{i,k}$  is characterized by a measurement noise  $\Sigma_{\mathbf{l}_{i,k}}$  which is considered Gaussian with zero mean. The derivation of statistical models for  $\Sigma_{\mathbf{l}_{i,k}}$  is described in the first half of this section. Each model states the expected covariance of a sensor-relative landmark position measurement in Cartesian coordinates. A measurement model for the light detection and ranging (LiDAR) based detection of poles is developed in Section 7.2 and complemented by a model for stereo video depth measurements in Section 7.3. In this thesis, the focus is put on LiDAR based localization and the case of stereo-video based feature detection is exclusively considered to underline the impact of its distinct measurement noise characteristics on localization system parameterization.

Consideration of operation environment characteristics is of equally fundamental importance for the model-based design of localization systems. A formalization of the environment forms the connection between the modeled localization system characteristics and the intended operation environment. An environment characteristic formalization is developed in Section 7.1 and models the distribution of landmarks relative to the vehicle system. These models in combination with the map matching and sensor data fusion model in Section 8 are capable of describing the interrelations between localization system parameters, operation environment characteristics and localization performance. Thereby, a precise adaptation of a localization system to the target operation environment under consideration of localization accuracy and robustness requirements is achieved.

## 7.1 Environment formalization

The adaptation of localization systems to the intended operation environment is a key aspect of robustness enhancement. Consequently, relevant environment characteristics are to be formalized for consideration in the model framework. The latter



**Figure 7.2** The arrangement of infrastructural features like lane markings can be considered deterministic and described by classical geometric models. Deviations from building guidelines and feature types with less deterministic distributions (indicated by blue dots) are captured by means of stochastic geometry.

allows to account for the interrelation of environment characteristics and the localization performance. Urban environments in specific and road transport infrastructure in general are structured environments. This classification relates to the fact, that building guidelines exist which regulate the arrangement of infrastructure elements like road marking, street lamps and traffic signs. Automated vehicle systems with localization capacity often use measurements of sensor-relative infrastructure object poses or other feature types as inputs. Additionally, the intended vehicle motion is typically narrowly constraint by the road topology. And consequently, the detection often occur in the same section of the sensor field of view. In this thesis, two perspectives on suitable formalizations of feature distributions are developed and incorporated into the Monte Carlo simulations used for model evaluation in Section 9.

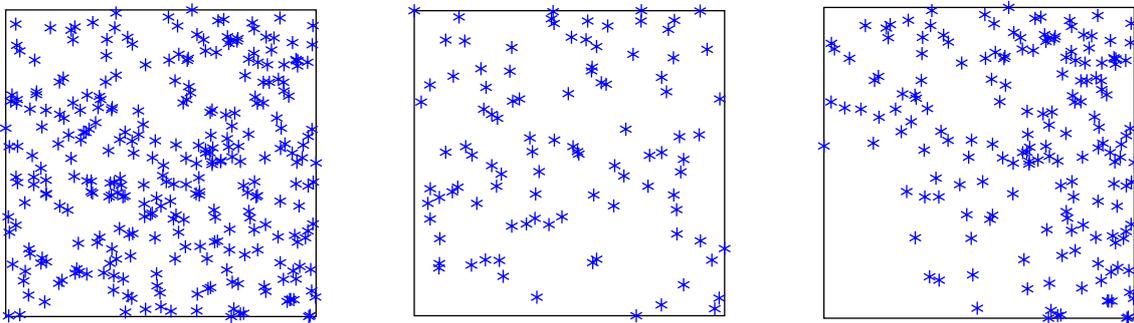
Road marking and other structured features often have well defined geometric arrangements and can hence be modeled based on data from available standardization documents (Figure 7.2). This environment abstraction is well suited for supporting design decisions for parts of the urban operation environment. In order to capture different environments, worst case assumptions about geometric dimensions can be taken. The design of a road marking based localization system might then be based on the assumption, that road marking occur at a minimum rate and the lane width does not exceed a worst case assumption. The design task would be the derivation of detection accuracy and rate requirements given a maximum lane width and road marking density. Thereby, all sceneries providing richer information than the considered worst case can be subsumed. The worst case assumptions have to be made cautiously, since there exists a strong correlation between the latter and perception system requirements like sensor accuracy. By assuming a at least partially stochastic point feature distribution, the environment model can be made more realistic.

Furthermore, other point features like depth or contrast discontinuities might have a more random distribution and can be modeled by means of stochastic geometry which are introduced in Section 7.1.1.

### 7.1.1 Stochastic distribution models

The assumption of a deterministic distribution can be feasible for special types of landmarks. Nonetheless, in real operation domains, landmark distributions are subject to stochastic variations. The latter arise from discrepancies between building guidelines and actual placement of infrastructure elements. Other feature types might be naturally characterized by stochastic distributions. In this thesis, stochastic geometry is utilized for the incorporation of stochastic variations in geometric arrangements from building guidelines and general feature distributions. One example is the modeling of distance distributions between street lamps by an exponential distribution with an average distance corresponding to some assumed average distribution [RSMZ15]. Stochastic geometry offers a sound basis for a more general modeling of the stochastic nature of landmark distributions. Especially in the domain of communication and sensor networks, stochastic point processes (SPPs) have been successfully applied to system parameter design. In this context, SPPs are commonly used to formalize the distribution of transmission towers and sensors. In this section, the applicability of SPPs for formalization of landmark distributions are studied. In the case of point landmarks, this can be achieved by utilization of stochastic point processes like the Poisson point process (PPP) [RSMZ16] (Figure 7.3).

PPP with the distribution are a widely used SPPs and enable the modeling of independent identically distributed landmarks. One realization  $\xi$  of the PPP represents a continuous model of a landmark arrangement in global state space. Realizations  $\xi$  of the PPP are contained in the event space  $\Xi$  and are obtained by a two step procedure. In a first step the number  $N$  of landmarks is drawn from a Poisson



**Figure 7.3** Poisson point processes are used to model stochastic landmark distributions with different densities (left:  $\lambda = 300$ , middle:  $\lambda = 100$ ). Non-homogeneous Poisson point processes can be used to model spatially varying landmark densities as shown on the right.

probability density function with the intensity parameter  $\lambda$ :

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!} . \quad (7.2)$$

Consecutively, the  $N$  continuous landmark positions are considered as independently identically distributed in space. Possible localization maps  $\xi_m$  are then given by a PPP realization in the world state space  $\mathcal{W}$

$$\xi_m = (N_m, \{\mathbf{m}_1^m, \dots, \mathbf{m}_{N_m}^m\}) , \xi_m \in \Xi(\mathcal{W}) . \quad (7.3)$$

Landmarks are observed within the sensor field of view  $\mathcal{S} \in \mathcal{W}$ . In case of homogeneous PPPs, spatial samples are drawn from an uniform distribution, Consequently, the map properties are spatially invariant and the probability of observing a given number of landmarks is modeled as constant. A sensor observation  $\xi_l$  is then given as

$$\xi_l = (N_l, \{\mathbf{l}_1^l, \dots, \mathbf{l}_{N_l}^l\}) , \xi_l \in \Xi(\mathcal{S}) . \quad (7.4)$$

The intensity value of a point process can be estimated from real measurement data as will be shown in Section 7.1.2. These estimated values can then be re-used for the design of different localization systems for the same type of features.

### 7.1.2 Estimation of model parameters

The intensity value  $\lambda$  in the Poisson distribution (Equation 7.2) can be estimated from the landmark data obtained in a map. Since,  $\lambda$  is defined in relation to a unit area, the area for estimation, e.g. a 2m broad stripe beside the road where most of the poles are situated, has to be previously defined. In most environments and for most feature-types the landmark density is subject to significant spatial variation and can only be estimated for small patches or as an average value. In such environments, a non-homogeneous PPP model can be preferable if a general model is needed. In the following, homogeneous PPP are now considered and the maximum likelihood estimate is calculated for a given landmark map (Figure 5.6). For the Poisson distribution in Equation 7.2 it is easily found that the maximum likelihood estimate of  $\lambda$  equals the observation  $x$ . The estimation procedure has to be conducted once and can be reused for the design of different localization and perception systems which use the considered map. For a 1.39km long trajectory (data set 0823-42, Section 4.3), an overall roadway width of 5.1m and a width of the area of landmark occurrences beside the roadway of 5m the intensity  $\lambda = 121(100\text{m})^{-2}$  is obtained.

## 7.2 Model of LiDAR-based pole detection

The detection of poles in LiDAR data is one step in the proposed hybrid localization system (Section 11). To this point, a model for the uncertainty in pole detection is derived in this subsection. Point measurements provided by a stationary LiDAR sensor are characterized by a weakly distance independent measurement noise. The latter is mainly caused by laser beam broadening due to imperfect beam focusing and is further amplified by incorrect sensor calibration. This uncertainty type is quantified in the following.

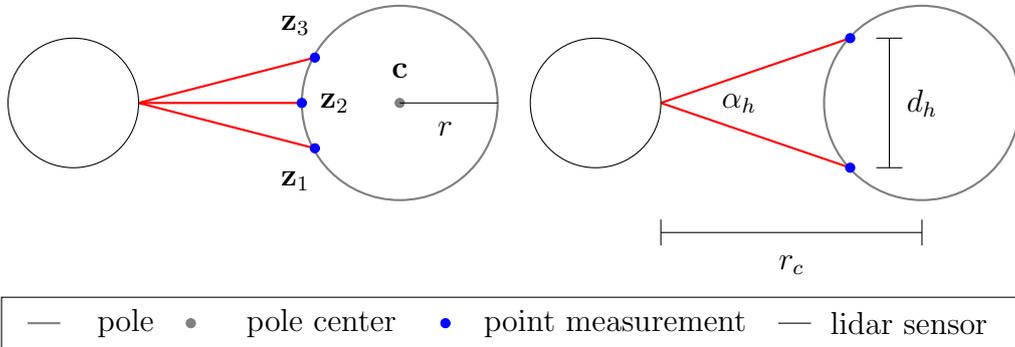
A basic approach for determining the position  $\mathbf{l}_{i,k}$  of a pole is the fitting of a circle to a set of measurement point (Figure 7.4). To this point, the z-component of the overall point cloud is set to zero and samples of three points are drawn from the points in the  $xy$ -plane. For every sample, a circle model  $\mathbf{u}$  with the circle center  $\mathbf{c}$  (compare Figure 7.4) and  $r$  the circle radius is determined.  $\mathbf{u}$  is obtained by minimization of the quadratic cost function  $J_c$

$$\mathbf{u} = [c_x, c_y, r]^\top = \arg \min_{\mathbf{u}} J_c \quad (7.5)$$

$$= \arg \min_{\mathbf{u}} \sum_{i=1}^n \|\mathbf{W}_i (|\mathbf{c} - \mathbf{z}_i| - r)\|^2 \quad (7.6)$$

$$= \arg \min_{\mathbf{u}} \sum_{i=1}^n \left\| \mathbf{W}_i \left( \sqrt{(c_x - z_{x,i})^2 + (c_y - z_{y,i})^2} - r \right) \right\|^2. \quad (7.7)$$

The circle model with the most model-supporting point measurements is selected and  $\mathbf{l}_{i,k}$  is given by the corresponding circle center coordinates. For error propagation from uncertain sensor measurements to  $\Sigma_{\mathbf{l}_{i,k}}$  it is important to observe that  $J_c(\mathbf{z}; \tilde{\mathbf{u}}) = 0$  forms an implicit function connecting the input measurements and the true circle parameters  $\tilde{\mathbf{u}}$ . Hence, a covariance estimate can be found by applying the implicit function theorem to  $J_c$ . Based on the implicit function theorem, the



**Figure 7.4** The setting for the derivation of the error model for LiDAR based pole detection.

covariance estimate  $\text{cov}(\mathbf{u})$  is given as follows:

$$\Sigma_{\mathbf{u}} = \sum_{i=1}^n \left( \frac{\partial^2 J_c}{\partial \mathbf{u}^2} \right)^{-1} \frac{\partial^2 J_c}{\partial \mathbf{u} \partial \mathbf{z}} \Sigma_{\mathbf{z}} \frac{\partial^2 J_c}{\partial \mathbf{u} \partial \mathbf{z}}^\top \left( \frac{\partial^2 J_c}{\partial \mathbf{u}^2} \right)^{-1} \quad (7.8)$$

$$\frac{\partial^2 J_c}{\partial \mathbf{u}^2} = \begin{bmatrix} \frac{\partial^2 J_c}{\partial c_1^2} & \frac{\partial^2 J_c}{\partial c_1 \partial c_2} & \frac{\partial^2 J_c}{\partial c_1 \partial r} \\ \frac{\partial^2 J_c}{\partial c_2 \partial c_1} & \frac{\partial^2 J_c}{\partial c_2^2} & \frac{\partial^2 J_c}{\partial c_2 \partial r} \\ \frac{\partial^2 J_c}{\partial r \partial c_1} & \frac{\partial^2 J_c}{\partial r \partial c_2} & \frac{\partial^2 J_c}{\partial r^2} \end{bmatrix}, \quad (7.9)$$

$$\frac{\partial^2 J_c}{\partial \mathbf{u} \partial \mathbf{z}} = \begin{bmatrix} \frac{\partial^2 J_c}{\partial c_1 \partial x} & \frac{\partial^2 J_c}{\partial c_1 \partial y} \\ \frac{\partial^2 J_c}{\partial c_2 \partial x} & \frac{\partial^2 J_c}{\partial c_2 \partial y} \\ \frac{\partial^2 J_c}{\partial r \partial x} & \frac{\partial^2 J_c}{\partial r \partial y} \end{bmatrix}.$$

For  $\mathbf{W}_i = (1/n)\mathbf{I}$ ,  $\Sigma_{\mathbf{u}}$  basically scales with  $1/n$ , although the point distribution on the circle surface also has an influence on the circle model accuracy.

In the pole detection procedure developed in Section 5.1.2,  $\mathbf{c}$  is always calculated from three measurement points and hence, the derived model has to be evaluated for  $n = 3$ . Consequently, the measurement uncertainty  $\Sigma_{\mathbf{1}}$  of a pole is assumed to be constant and isotropic:

$$\Sigma_{\mathbf{1}} = \Sigma_{\mathbf{u}}^{\mathbf{c}} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{yx}^2 & \sigma_y^2 \end{bmatrix} \quad (7.10)$$

$$\approx \sigma_l^2 \mathbf{I}. \quad (7.11)$$

Obviously, this is an approximation, since the LiDAR beams are not perfectly focused which results in a increase in the cross-range measurement uncertainty. This uncertainty source is considered to be significantly smaller than other sources of uncertainty and is therefore neglected.

An additional significant source of uncertainty is introduced during motion compensation of the LiDAR scans. In an urban environment with a typical speed limitation of  $50 \frac{\text{km}}{\text{h}}$ , the vehicle moves 1.4 m during a  $360^\circ$  rotation in 100 ms. Accordingly, a velocity estimation error of  $1 \frac{\text{km}}{\text{h}}$  could potentially result in an additional pole location measurement error of 0.3 m. This leads to an addition of the following error term to Equation 7.8

$$\epsilon_{\mathbf{v}, \max} = \frac{|\mathbf{v}_k - \hat{\mathbf{v}}_k|}{3.6 \frac{\text{kms}}{\text{hm}}} 100 \text{ ms}. \quad (7.12)$$

In practical application the motion-imposed error typically outweighs the error due to the uncertainty in single LiDAR point measurements. It is worth noticing, that  $\text{cov}(\mathbf{u})$  does not explicitly depend on the distance between the origin of the sensor frame and the pole object. However, the number of measurement points on the pole

surface does depend on the distance

$$n(r_c) = \frac{2r}{d_h(r_c)} \quad (7.13)$$

$$= \frac{2d}{r_c \tan(\alpha_h)}. \quad (7.14)$$

For  $n = 3$ ,  $r = 0.2$  m and the horizontal angular resolution  $\alpha_h = 0.125^\circ$  of the Velodyne LiDAR sensor, the resulting maximum detection range  $r_{c,max} \approx 47.75$  m. However, this theoretical maximum range is typically irrelevant, since the outlier rate significantly increases for higher distances. Consequently, the maximum range is often defined independently from the theoretical range.

### 7.2.1 Experimental evaluation

The pole detector and error model evaluation is based on real pole measurements which are extracted from 3D LiDAR measurements recorded with a Velodyne HDL-64E. All results are obtained for one pole within the sensor field of view. The evaluation is conducted for different distances between the static origin of the sensor frame and the pole. At each distance approximately 2500 pole measurements are recorded. Based on the measurement sets, the statistics of the pole detector performance are derived. The results are listed in Table 7.1.

**Table 7.1** Results from the pole detection accuracy experiment show an approximately distance independent accuracy of LiDAR based pole detection.

$d / \text{m}$	11.84	15.88	20.33	25.79	31.85	36.43
$\sigma_x / \text{m}$	0.025	0.021	0.022	0.039	0.027	0.017
$\sigma_y / \text{m}$	0.018	0.016	0.017	0.021	0.022	0.019
$n_{scans}$	2350	2518	2308	2476	2616	1967
$P_D$	0.999	1.000	0.996	1.000	0.964	0.211

The experimental results indicate that the detection accuracy can be considered approximately distance independent. In down-range direction the standard deviation was  $\sigma_{d,lm} = 0.025$  m which is slightly higher than the observed cross-range value  $\sigma_{c,lm} = 0.018$  m. The model is evaluated for the manually measured 0.15 m radius of the pole and  $\sigma_l = 0.06$  m ([GL10] state a similarly high value). The modeled value for  $\sigma_{lm} = 0.016$  m is slightly optimistic. The detection range decreases significantly for pole distances higher than 32.0 m which is due the finite angular resolution of the LiDAR sensor.

### 7.3 Stereo video camera perception model

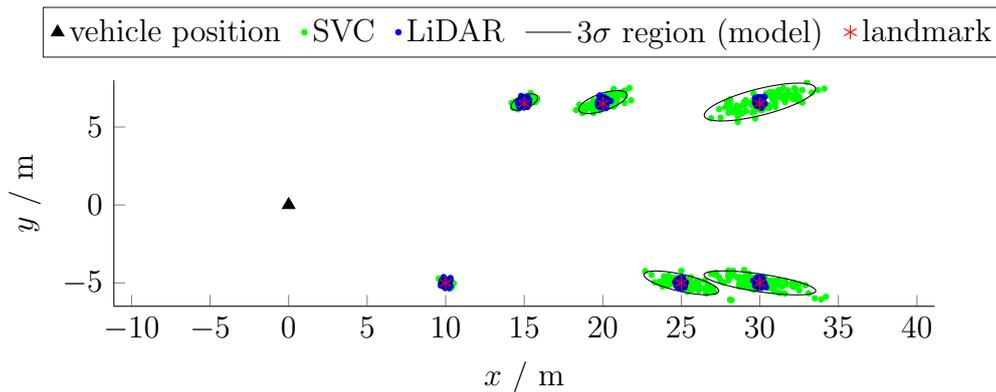
The detection of poles in video images is state of the art and not further treated within this thesis. Once a pole is detected, the position of a predefined reference point of the pole, e.g. its upper end, can be obtained from a depth measurement. A model of the depth measurement is now obtained by modification of the model presented in [Bad09]. Under the assumption assumption of a precise intrinsic calibration, it can be written [RSMZ15]

$$\Sigma_{\mathbf{l}_i} = \frac{1}{(c \cdot b_w)^2} z_{i,x}^2 \begin{bmatrix} z_{i,x}^2 \sigma_d^2 & z_{i,x} z_{i,y} \sigma_d^2 \\ z_{i,x} z_{i,y} \sigma_d^2 & z_{i,y}^2 \sigma_d^2 + b_w^2 \sigma_u^2 \end{bmatrix}. \quad (7.15)$$

Image columns are denoted  $u$ , disparity  $d$ , base-width  $b_w$ , measurement noise variances  $\sigma_d^2$ ,  $\sigma_u^2$  in disparity and image column and the camera constant  $c$ . This model can be further simplified [RSMZ15] (compare Equation 7.16) by assuming that  $\sigma_d \approx \sigma_u$  and  $z_{i,k,y} \ll 0.5b_w$ . The first assumption is justified by the observation that  $d$  is estimated from noisy values of  $u$ . Furthermore, the second assumption is reasonable, since landmarks are typically situated outside the driving tube and hence, the lateral offset is several times larger than  $b_w$ . These assumptions lead to the following simplified model [RSMZ15]

$$\Sigma_{\mathbf{l}_{i,k}} \approx \xi z_{i,k,x}^2 \begin{bmatrix} z_{i,k,x}^2 & z_{i,k,x} z_{i,k,y} \\ z_{i,k,x} z_{i,k,y} & z_{i,k,y}^2 \end{bmatrix}, \quad \xi = \frac{\sigma_d^2}{(c \cdot b_w)^2}. \quad (7.16)$$

Example noise realization for the pole detection and the SVC model are given in Figure 7.5 and show the difference between the approximately isotropic measurement noise in LiDAR based pole detection and the highly anisotropic noise in SVC depth measurement. The models perform well in both cases.



**Figure 7.5** All noise realizations and the  $3\sigma$  region from model evaluation are scaled by a factor of 5 for better visibility. For the SVC model, the parameters  $\sigma_d = 0.12\text{px}$ ,  $\sigma_u = 0.21\text{px}$  (in accordance to [KASS05]),  $c = 10\text{ mm}$  and  $b_w = 0.12\text{ m}$  as well as  $\sigma_1 = 0.03\text{ m}$  for LiDAR based pole detection were selected.

## 8 Localization system models

In the signal processing chain of the considered localization system, the detection of landmarks is followed by a map matching step. The error model for map matching (Section 8.1) is derived by propagation of the landmark measurement uncertainty to the map matching result. To this point, the solution of the orthogonal Procrustes problem (OPP) which minimizes the quadratic cost function  $J$  (Equation 8.1) is considered. This solution to the OPP enables the calculation of the transformation parameters between two data sets  $\{\mathbf{m}_{l,k}\}_{i=1}^n$  from the localization map and the sensor observation  $\{\mathbf{l}_{i,k}\}_{i=1}^n$  under the assumption of known correspondences.

The absolute vehicle pose information from map matching is then fused with motion estimates from odometry measurement in a Bayesian filtering framework. This step is modeled by a Cramér-Rao like performance bound for map matching probabilities smaller than one. The closed-form bound is developed and discussed in Section 8.2. At this model stage, the performance of the spectral map matching algorithm (Section 6) can be considered. This section concludes with the development of a strategy for dealing with occlusions and landmark detection failures in Section 8.3 to make the model realistic and applicable to real world scenarios.

### 8.1 Map matching model

In this subsection, a model for the map matching procedure is developed. The central assumption for the derivation of the map matching model is the existence of a correct correspondence estimates between a set  $\mathbf{L}_k$  of landmark observations and landmarks contained in the localization map  $\mathbf{m}_l$ . Furthermore, it is assumed that the transformation parameters are calculated based on the solution of the Procrustes problem described in Section 2.1.2. The incorporation of additional information about the sensor noise is achieved by weighting the correspondence pairs by a weighting matrix  $\mathbf{W}_i$  which leads to the quadratic cost function

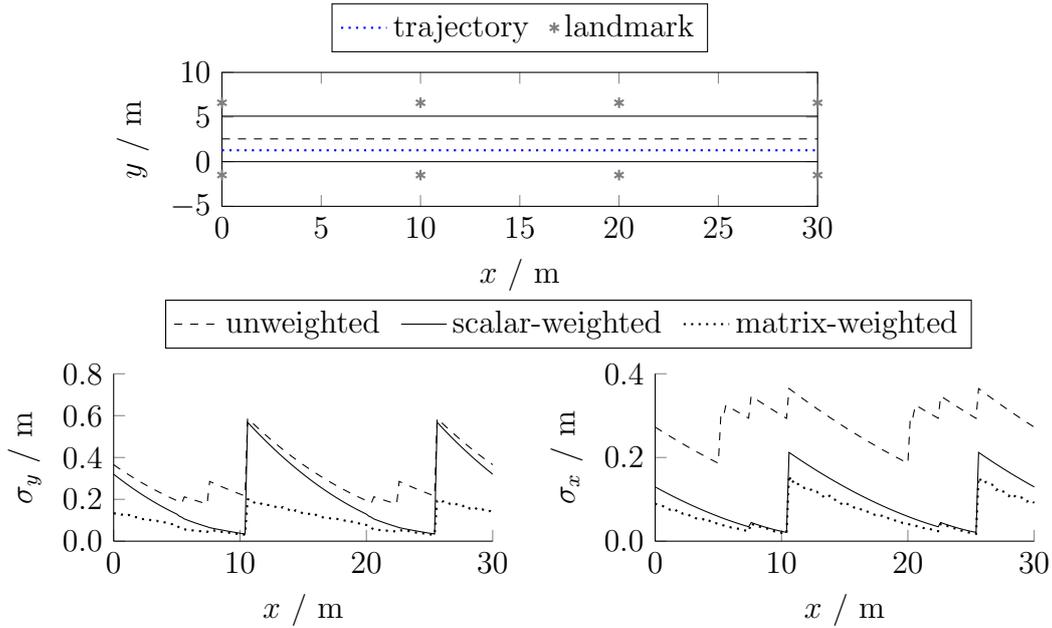
$$J = \sum_{i=1}^n \|\mathbf{W}_i (\mathbf{R}\mathbf{l}_i + \mathbf{t} - \mathbf{m}_i)\|^2 . \quad (8.1)$$

The scalar-weighted OPP is considered here and is obtained by the substitution  $\mathbf{W}_i = w_i \mathbf{I}$  in Equation 8.1. The solution for the scalar-weighted OPP is given in Section 8.1.1 and followed by the derivation of the map matching error model in Section 8.1.2. Based on real pole measurements and map matching results, the model is verified in Section 8.1.3.

#### 8.1.1 Solution of the orthogonal Procrustes problem

In a first step, a suitable formulation of the OPP as basis for the map matching model derivation shall be selected. Therefore, three variants of the OPP formulations are considered, the unweighted scalar- and matrix-weighted OPP. The accuracy of the

different implementations is studied in Monte Carlo simulation for the landmark distribution shown in Figure 8.1 and the stereo-video depth measurement model (Section 7.3). The results in Figure 8.1 show, that the information about the noise characteristic which is contained in the weighting factor leads to an increase in accuracy.



**Figure 8.1** Top: simulated landmark distribution; Bottom: Incorporation of noise information by a weighting factor leads to increased matching accuracy ( $\sigma_x, \sigma_y$ ). The scalar weighted cost function gives conservative estimates of the matrix-weighted case and is therefore considered for the derivation of the map matching model. The data were obtained from Monte Carlo simulations with utilization of the SC model.

The noise of LiDAR based pole measurements is approximately distance independent and isotropic (Section 7.2). Hence, the results from all OPP solutions would result in a similar accuracy. For the stereo video model (Section 7.3), weighting results in an increased matching accuracy. In practical application the performance difference between the solutions depends on the availability of accurate sensor noise models. To this point, it is resorted to the scalar-weighted OPP for model derivation since it gives a conservative estimate of the matching performance which accounts for some imprecision in the sensor noise model.

In the case of anisotropic measurement noise, the choice  $w_i = \max(\text{eig}(\Sigma_{1,i,k}))^{1/2}$  gives precise matching results. In the following, a closed-form solution of the scalar-weighted Procrustes problem is derived by means of the Lagrangian multiplier method. For the calculation of the translation parameters, the centroid coincidence

theorem [WCR92] is applied first:

$$\sum_{i=1}^n w_i^2 \mathbf{m}_i = \sum_{i=1}^n w_i^2 (\mathbf{R}\mathbf{l}_i + \mathbf{t}) \quad (8.2)$$

$$\mathbf{t} = \bar{\mathbf{m}} - \mathbf{R}\bar{\mathbf{l}} \quad (8.3)$$

with the weighted centroids

$$\bar{\mathbf{l}} = \frac{\sum_{i=1}^n w_i^2 \mathbf{l}_i}{\sum_{i=1}^n w_i^2}, \quad \bar{\mathbf{m}} = \frac{\sum_{i=1}^n w_i^2 \mathbf{m}_i}{\sum_{i=1}^n w_i^2}. \quad (8.4)$$

This equation allows to cancel out the translation parameter from the cost function by substitution of  $w_i (\mathbf{R}\mathbf{l}_i + \mathbf{t} - \mathbf{m}_i) = w_i (\mathbf{R}\mathbf{l}_i + \bar{\mathbf{m}} - \mathbf{R}\bar{\mathbf{l}} - \mathbf{m}_i)$  in Equation 8.1 which results in

$$J = \sum_{i=1}^n \left\| \mathbf{R}\tilde{\mathbf{l}}_i - \tilde{\mathbf{m}}_i \right\|^2, \quad \text{with } \tilde{\mathbf{l}}_i := w_i (\mathbf{l}_i - \bar{\mathbf{l}}), \quad \tilde{\mathbf{m}}_i := w_i (\mathbf{m}_i - \bar{\mathbf{m}}). \quad (8.5)$$

The orthogonality constraint on the rotation matrix is now enforced by defining the Lagrangian function  $\Lambda$  with the Lagrangian multiplier  $\lambda$

$$\Lambda(c, s, \lambda) = \sum_{i=1}^n \left\| \mathbf{R}\tilde{\mathbf{l}}_i - \tilde{\mathbf{m}}_i \right\|^2 + \lambda (c^2 + s^2 - 1) \quad (8.6)$$

$$(8.7)$$

And the solution is obtained by solving the following set of equations

$$\nabla_{c,s,\lambda} \Lambda(c, s, \lambda) \stackrel{!}{=} 0. \quad (8.8)$$

Finally, it is solved for  $c$  and  $s$

$$c = \frac{f_1}{\sqrt{f_1^2 + f_2^2}}, \quad (8.9)$$

$$s = \frac{f_2}{\sqrt{f_1^2 + f_2^2}}, \quad (8.10)$$

$$f_1 = \sum_{i=1}^n \tilde{\mathbf{l}}_i^\top \tilde{\mathbf{m}}_i, \quad f_2 = \sum_{i=1}^n \tilde{\mathbf{l}}_i^\top \mathbf{F} \tilde{\mathbf{m}}_i, \quad \mathbf{F} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

The translation parameters are then calculated by inserting Equation 8.9 and Equation 8.10 in Equation 8.3.

### 8.1.2 Error propagation

Propagation of the measurement error  $\Sigma_{1,i}$  to the matching uncertainty  $\Sigma_{\mathbf{z}}$  is now performed for the solution of the scalar-weighted OPP [RSMZ15]. As the model is intended to describe vehicle localization relative to a highly accurate map, the point data in the map are assumed to be noise-free, yielding  $\Delta \mathbf{m}_i = \Delta \tilde{\mathbf{m}}_i = \mathbf{0}$ . Accordingly, the perturbation in  $f_1$  and  $f_2$  are calculated as

$$\Delta f_1 = \sum_{i=1}^n \tilde{\mathbf{m}}_i^\top \Delta \tilde{\mathbf{I}}_i \quad \text{and} \quad \Delta f_2 = \sum_{i=1}^n \tilde{\mathbf{m}}_i^\top \mathbf{F} \Delta \tilde{\mathbf{I}}_i. \quad (8.11)$$

Consequently, the variance of the perturbation  $\Delta \mathbf{f} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{f}})$  is given as

$$\Sigma_{\mathbf{f}} = \text{cov} \left( \begin{bmatrix} \Delta f_1 \\ \Delta f_2 \end{bmatrix} \right) = \sum_{i=1}^n \begin{bmatrix} \tilde{\mathbf{m}}_i \\ \mathbf{F} \tilde{\mathbf{m}}_i \end{bmatrix} \Sigma_{1,i} \begin{bmatrix} \tilde{\mathbf{m}}_i^\top & \tilde{\mathbf{m}}_i^\top \mathbf{F}^\top \end{bmatrix}. \quad (8.12)$$

$\Delta \mathbf{f}$  is then propagated to the translation perturbation  $\Delta \mathbf{t} = (\mathbf{R} + \Delta \mathbf{R})(\bar{\mathbf{I}} + \Delta \bar{\mathbf{I}}) - \mathbf{R} \bar{\mathbf{I}}$ . The dependence of  $\Delta \mathbf{t}$  on all transformation parameters is now expressed by defining the stacked parameter vector  $\mathbf{p} = [\bar{\mathbf{I}}^\top \ c \ s]^\top$  and rewriting  $\Delta \mathbf{t}$  as a function of  $\Delta \mathbf{p}$

$$\Delta \mathbf{t} = \begin{bmatrix} \Delta \mathbf{p}^\top \mathbf{A} \Delta \mathbf{p} \\ \Delta \mathbf{p}^\top \mathbf{B} \Delta \mathbf{p} \end{bmatrix}. \quad (8.13)$$

In order to consider the correlations between  $\Delta \mathbf{t}$  and all transformation parameters, the normal approximation of the quadratic form in Equation 8.13 is used for estimation of  $\Sigma_{\Delta \mathbf{t}}$

$$\Sigma_{\Delta \mathbf{t}} = \begin{bmatrix} 2\text{tr}(\mathbf{A}_1 \Sigma_{\mathbf{p}} \mathbf{A}_1 \Sigma_{\mathbf{p}}) + 4\mathbf{p}^\top \mathbf{A}_1 \Sigma_{\mathbf{p}} \mathbf{A}_1 \mathbf{p} & 2\text{tr}(\mathbf{A}_1 \Sigma_{\mathbf{p}} \mathbf{A}_2 \Sigma_{\mathbf{p}}) + 4\mathbf{p}^\top \mathbf{A}_1 \Sigma_{\mathbf{p}} \mathbf{A}_2 \mathbf{p} \\ 2\text{tr}(\mathbf{A}_1 \Sigma_{\mathbf{p}} \mathbf{A}_2 \Sigma_{\mathbf{p}}) + 4\mathbf{p}^\top \mathbf{A}_1 \Sigma_{\mathbf{p}} \mathbf{A}_2 \mathbf{p} & 2\text{tr}(\mathbf{A}_2 \Sigma_{\mathbf{p}} \mathbf{A}_2 \Sigma_{\mathbf{p}}) + 4\mathbf{p}^\top \mathbf{A}_2 \Sigma_{\mathbf{p}} \mathbf{A}_2 \mathbf{p} \end{bmatrix} \quad (8.14)$$

$$= \begin{bmatrix} \sigma_{t11}^2 & \sigma_{t12} \\ \sigma_{t21} & \sigma_{t22}^2 \end{bmatrix} \quad (8.15)$$

with the matrices

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{0} & \mathbf{F}_1 \\ \mathbf{F}_1^\top & \mathbf{0} \end{bmatrix}, \quad \mathbf{F}_1 = \begin{bmatrix} 0.5 & 0 \\ 0 & -0.5 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} \mathbf{0} & \mathbf{F}_2 \\ \mathbf{F}_2^\top & \mathbf{0} \end{bmatrix}, \quad \mathbf{F}_2 = \begin{bmatrix} 0 & 0.5 \\ 0.5 & 0 \end{bmatrix}.$$

The expressions for  $\sigma_{t12}$  and  $\sigma_{t21}$  are valid for symmetric  $\mathbf{A}_1$  and  $\mathbf{A}_2$  which is always given. Calculation of  $\Sigma_{\bar{\mathbf{I}}}$ ,  $\Sigma_{c,s}$  and the cross-correlation between  $\bar{\mathbf{I}}$  and  $[c \ s]^\top$  is lengthy, but straight forward and is therefore not stated here. Several simplification and transformation steps lead to the final result for the variance of  $\mathbf{p}$

$$\Sigma_{\mathbf{p}} = \text{cov}(\mathbf{p}) = \sum_{i=1}^n \Gamma_i \Sigma_{1,i} \Gamma_i^\top, \quad (8.16)$$

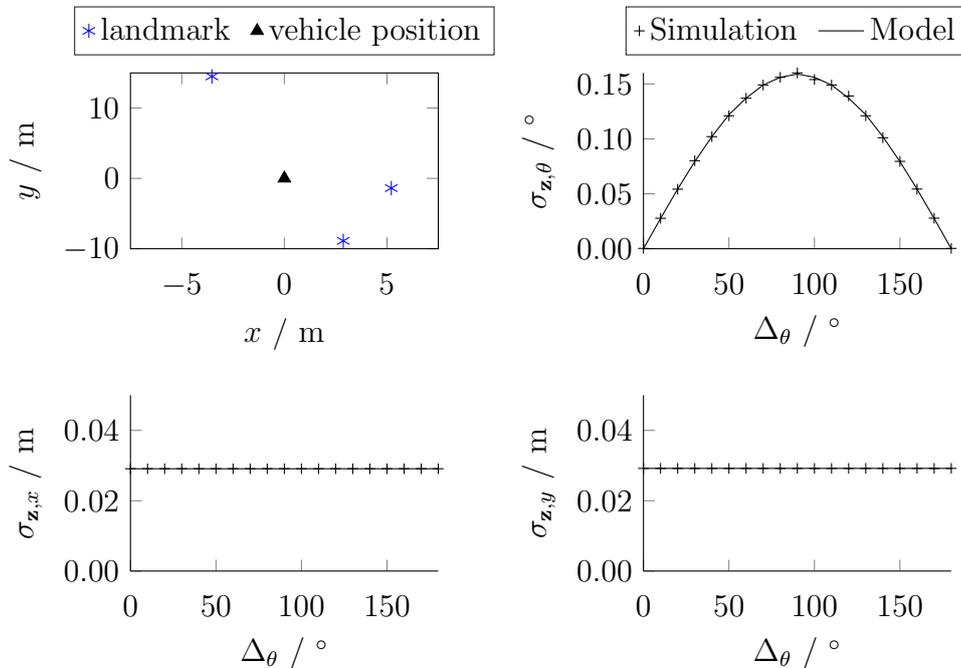
$$\Gamma_i = w_i \begin{bmatrix} \frac{w_i}{\sum_i w_i^2} \mathbf{I}_{2 \times 2} \\ \frac{1}{f_1^2 + f_2^2} \begin{bmatrix} f_2 \\ -f_1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_i^\top & \mathbf{F} \end{bmatrix} \end{bmatrix}. \quad (8.17)$$

An estimate of the rotational map matching error is then obtained from  $\sigma_c$  or  $\sigma_s$ .

### 8.1.3 Experimental evaluation

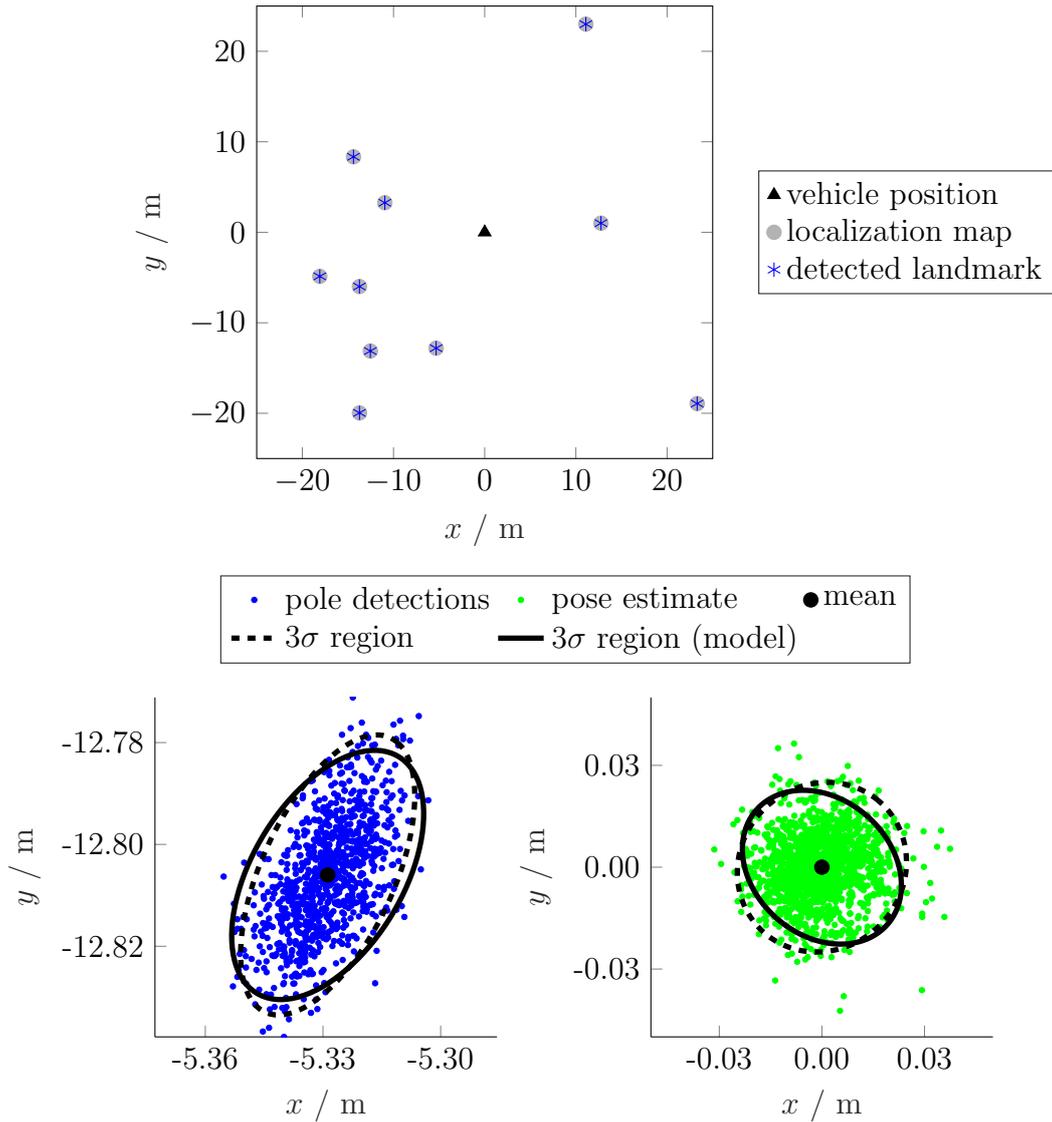
The derived model is evaluated in Monte Carlo simulations and based on real measurement data. For simulative evaluation,  $10^6$  realizations of the zero-mean Gaussian measurement noise were used in the Monte Carlo simulation. A fixed landmark arrangement is defined and the point set is rotated by  $0 \leq \Delta_\theta \leq 180^\circ$  (Figure 8.2). The landmark measurement error is set to  $\Sigma_1 = (0.05 \text{ m})^2 \mathbf{I}$ . The results in Figure 8.2 show a good fit of the model with the simulated standard deviations for the translational  $\sigma_{z,x}$ ,  $\sigma_{z,y}$  and rotational matching result  $\sigma_{z,\theta}$ .

The approach used for real data based evaluation of the map matching model is closely related to the procedure from Section 7.2.1. Firstly, pole measurements from a stationary LiDAR sensor are subdivided into two equally sized sets. The measurements from the first set are used to remove false detections and to calculate the average pole locations. The latter are used as a localization map for the matching procedure as shown in Figure 8.3. Based on this map and the remaining pole measurements, the map matching statistics in Figure 8.3 are calculated. The estimated values for pole detection  $\Sigma_1^m$  and map matching  $\Sigma_z^m$  from the model evaluation fit the empirically determined data. The matching accuracy is 0.01 m which is highly accurate in comparison to the results presented in Section 11.3. The evaluation was conducted for the static case and hence, the results indicate the impact of



**Figure 8.2** Top: landmark distribution for simulation; Remaining figures: simulation results. The Monte Carlo simulation and map matching model results comply.

the motion compensation error term in Equation 7.13. This observation underlines the high significance of accurate motion compensation of LiDAR scans, particularly if sparse features like poles are extracted, and motivated the development of an accurate LiDAR odometry in Section 10.



**Figure 8.3** Top: vehicle-relative landmark distribution; Bottom: pole detections (left) and map matching results (right) with empirical and modeled  $3\sigma$  trust regions. The evaluation of the map matching model was conducted based on real pole detections from the campus environment.

## 8.2 Vehicle pose filtering model

Absolute pose estimates provided by map matching procedures form the basis for map-relative localization. Localization systems typically rely on the fusion of such absolute pose estimates with relative motion estimates provided by an odometry (Figure 7.1). Odometry measurements are typically provided at a fixed rate with a high reliability. On the contrary, the availability of environment perception based absolute pose measurements depends on the characteristic of the operation environment and failures might occur at a higher rate as shown in Section 6. Based on the odometry measurements, the pose estimation accuracy can be increased in comparison to a single-shot map matching based localization procedure. Furthermore, it enables the compensation map matching failure sequences by integration of motion estimates starting from the last valid absolute pose estimate. The fusion step often involves the utilization of widely used Bayesian filtering algorithms which are outlined in Section 2.1.3 and are utilized in the proposed localization system in Section 11.

In this section, a performance bound for Bayesian filtering based localization systems in closed-form is developed which considers fusion of map matching and odometry results (Section 8.2.1). Additionally, it accounts for probabilities  $P_m < 1$  of obtaining a valid map matching results at time  $k$ . Hence, the model is developed for the most frequently occurring type of localization systems and is capable of comprehensively describing the localization system developed in Section 11. The derived statistical model establishes the connection between the odometry and map matching performance and the resulting localization performance. Thereby, important aspects for the design of localization systems are captured by the model and form the prerequisite for model-based design of localization systems (Section 9). The contents from this section were published in [RSMZ16] and [RMZ17].

### 8.2.1 A closed-form upper bound for Kalman Filtering with intermittent observation

The Cramér-Rao bound (CRB) [BSLK04] (Section 2.1.5) is well established for the calculation of upper accuracy bounds in Bayesian state estimation problems as vehicle localization. A central assumption in the derivation of the classical CRB is the availability of a measurement at every time step  $k$  for conducting the measurement update step. This assumption corresponds to the case of  $P_m = 1$  in the context of the developed localization system. As for localization, this assumption is often violated practical applications and therefore a discussion started more recently on how to address this challenge in the context of Kalman filtering. A thorough overview over this discussion can be found in [HRFT04] and a solid body of related literature exists. In [HRFT04], the characteristics of the information reduction factor CRB and the sequential enumerated (EN) CRB are discussed and compared. The information reduction factor (IRF) CRB [NWBS01] considers an average matching rate  $P_m$  up to a time  $k$ . The bound  $\mathbf{P}_k$  is given by the inverse Fisher information matrix

$\mathbf{P}_k = \mathbf{J}_k^{-1}$  which is stated in iterative form [HRFT04]

$$\mathbf{J}_{k+1} = P_m \mathbf{C}^\top \Sigma_{\mathbf{z},k} \mathbf{C} - \mathbf{Q}^{-\top} \mathbf{A} (\mathbf{J}_k + \mathbf{A}^\top \mathbf{Q}^{-1} \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{Q}^{-1} + \mathbf{Q}^{-1} . \quad (8.18)$$

For  $P_m = 1$  and by inverting this expression, the Riccati equation for the standard Kalman filter is obtained. The utilized matrices correspond to the ones used in the standard extended Kalman filter (EKF) (Section 2.1.3). In [HRFT04], the IRF CRB is compared to the EN CRB which is calculated on the basis of explicitly defined sequences  $\mathbf{S}$  of map matching successes and failures. Evaluation is computational demanding since it requires the calculation of  $\mathbf{J}_k(\mathbf{S})$  for all possible  $\mathbf{S}$  up to time step  $k$ , resulting in an exponentially increasing number of  $2^k$  evaluation steps. The EN CRB is then given by the expected value of all results for  $\mathbf{J}_k(\mathbf{S})$ . Due to the computational costs, the accurate performance estimation based on the EN CRB is not considered here. The results from Section 8.2.2 and [HRFT04] indicate that the IRF CRB potentially provides optimistic performance estimates. To this point, an alternative with a computational efficiency similar to the IRF CRB is elaborated in Section 8.2.1.

A parallel development to the CRB for  $P_m < 1$  has been taken in the research area of dynamic system control with intermittent observations. *Sinopoli* et al. were the first to introduce a set of equations for Kalman filtering with intermittent observations (KFIO) in [SSF<sup>+</sup>04] (see also [MS12], [SRA15], [HD07], [YX11]). Accordingly, the existence of a map matching result at time  $k$  can be modeled as a binary random variable  $\gamma_k$  following a Bernoulli process. This assumption reflects the observations made in Section 6.4 and is of great significance for the applicability of the model for treating occlusions as will be discussed in Section 8.3. Following [SSF<sup>+</sup>04], the covariance of the matching result noise is described by the probability distribution function

$$p(\mathbf{v}_k | \gamma_k) = \begin{cases} \Sigma_{\mathbf{z}} & , \text{if } \gamma_k = 1 \\ \sigma^2 \mathbf{I} & , \text{if } \gamma_k = 0 \end{cases} .$$

The matching probability  $P_{m,k}$  is now defined as  $P_{m,k} = p(\gamma_k = 1)$ . In the following, a homogeneous Bernoulli process such that  $P_{m,k} = P_m, \forall k$  is considered. The KFIO equations are now obtained for  $\sigma \rightarrow \infty$  [SSF<sup>+</sup>04]. Applied to the problem of localization and with the variable names from Section 2.1.3, the KFIO equations [SSF<sup>+</sup>04] are defined as

$$\begin{aligned} \hat{\mathbf{x}}_{k+1|k} &= \mathbf{A} \hat{\mathbf{x}}_{k|k} \\ \mathbf{P}_{k+1|k} &= \mathbf{A} \mathbf{P}_{k|k} \mathbf{A}^\top + \mathbf{Q} \\ \hat{\mathbf{x}}_{k+1|k+1} &= \hat{\mathbf{x}}_{k+1|k} + \gamma_{k+1} \mathbf{K}_{k+1} (\mathbf{z}_{k+1} - \mathbf{C} \hat{\mathbf{x}}_{k+1|k}) \\ \mathbf{P}_{k+1|k+1} &= \mathbf{P}_{k+1|k} - \gamma_{k+1} \mathbf{K}_{k+1} \mathbf{C} \mathbf{P}_{k+1|k} \\ \mathbf{K}_{k+1} &= \mathbf{P}_{k+1|k} \mathbf{C}^\top (\mathbf{C} \mathbf{P}_{k+1|k} \mathbf{C}^\top + \Sigma_{\mathbf{z},k})^{-1} . \end{aligned}$$

In this formulation, the state covariance becomes stochastic since the arrival sequence of valid map matching results is random. For the derivation of the performance bound, we are interested in the expected value of the state covariance denoted as  $\mathbf{P}_k^e$  which leads to the following modified algebraic Riccati equation (MARE) [SSF<sup>+</sup>04]

$$\mathbf{P}_{k+1}^e = \mathbf{A}\mathbf{P}_k^e\mathbf{A}^\top + \mathbf{Q} - P_m\mathbf{A}\mathbf{P}_k^e\mathbf{C}^\top (\mathbf{C}\mathbf{P}_k^e\mathbf{C}^\top - \Sigma_{\mathbf{z},k})^{-1} \mathbf{C}\mathbf{P}_k^e\mathbf{A}^\top. \quad (8.19)$$

This MARE does not have any stochastic components which is the central prerequisite for the following derivation of a performance bound.

The model formulation is now compared to the state of the art in CRBs for  $P_m < 1$ . *Niu* et al. introduce the scaling factor CRB in [NWBS01] which considers false detection probabilities smaller one by introducing a constant scaling factor for the measurement uncertainty. In [BD09] it is proven, that the following relation between the MARE, IRF and EN CRB holds

$$\mathbf{P}_k^{IRF} \leq \mathbf{P}_k^{EN} \leq \mathbf{P}_k^{MARE}.$$

This result is in agreement with [HRFT04] where the IRF CRB is found to be optimistic. The EN CRB requires the enumeration of all possible detection sequences which leads to an exponential increase in computation for increasing  $k$ . To this point, the bound provided by the MARE is chosen here as a basis for system modeling and a closed-form formulation is derived in consecutive paragraphs.

Its results are close to the ones provided by the EN CRB but can be obtained significantly more efficient. The authors in [SSF<sup>+</sup>04] propose to obtain an upper bound for the state uncertainty by iteration of the MARE or by the solution of a semidefinite programming problem. In order to derive a closed-form solution for the MARE for efficient evaluation in simulations, we now derive its steady-state solution. The MARE (Equation 8.19) is equivalent to

$$\mathbf{P}_{k+1} = \mathbf{A} \left( \lambda \mathbf{P}_k + (1 - \lambda) \mathbf{P}_k - \lambda \mathbf{P}_k \mathbf{C}^\top (\mathbf{C} \mathbf{P}_k \mathbf{C}^\top + \Sigma_{\mathbf{z},k})^{-1} \mathbf{C} \mathbf{P}_k \right) \mathbf{A}^\top + \mathbf{Q}.$$

From this point on, the transformations are widely equivalent to the ones used for obtaining the steady-state solution of the Kalman filter Riccati equation [MMR07]. Application of the Kailath variant of the Woodbury identity [PP12] gives the following expression

$$\mathbf{P}_{k+1} = \mathbf{A} \left( \lambda \mathbf{P}_k (\mathbf{I}_{3 \times 3} + \mathbf{C}^\top \Sigma_{\mathbf{z},k}^{-1} \mathbf{C} \mathbf{P}_k)^{-1} + (1 - \lambda) \mathbf{P}_k \right) \mathbf{A}^\top + \mathbf{Q}. \quad (8.20)$$

Without consideration of the correlation between position and rotation uncertainty  $\mathbf{A} := \mathbf{I}_{3 \times 3}$  which leads to a simplification of Equation 8.20:

$$\begin{aligned} \mathbf{P}_{k+1}^Q &= \lambda \mathbf{P}_k^Q \left( \mathbf{I}_{3 \times 3} + \mathbf{X} \mathbf{P}_k^Q \right)^{-1} + (1 - \lambda) \mathbf{P}_k^Q + \mathbf{I}_{3 \times 3}, \\ \mathbf{X} &:= \mathbf{Q}^{1/2} \mathbf{C}^\top \Sigma_{\mathbf{z},k}^{-1} \mathbf{C} \mathbf{Q}^{1/2}, \quad \mathbf{P}_k^Q := \mathbf{Q}^{-1/2} \mathbf{P}_k \mathbf{Q}^{-1/2}. \end{aligned}$$

Application of a singular value decomposition to  $\mathbf{X}$  gives

$$\begin{aligned}\mathbf{X} &= \mathbf{V}\mathbf{diag}(e_n)\mathbf{V}^\top \\ &= \mathbf{V}\mathbf{E}\mathbf{V}^\top\end{aligned}$$

with eigenvalues  $e_n$ . Based on this definition, it can be written

$$\begin{aligned}\mathbf{P}_{k+1}^{\mathbf{Q}} &= \lambda\mathbf{P}_k^{\mathbf{Q}}\left(\mathbf{I}_{3\times 3} + \mathbf{V}\mathbf{E}\mathbf{V}^\top\mathbf{P}_k^{\mathbf{Q}}\right)^{-1} + (1-\lambda)\mathbf{P}_k^{\mathbf{Q}} + \mathbf{I}_{3\times 3} \\ \mathbf{V}^\top\mathbf{P}_{k+1}^{\mathbf{Q}}\mathbf{V} &= \lambda\mathbf{V}^\top\mathbf{P}_k^{\mathbf{Q}}\mathbf{V}\left(\mathbf{I}_{3\times 3} + \mathbf{E}\mathbf{V}^\top\mathbf{P}_k^{\mathbf{Q}}\mathbf{V}\right)^{-1} + (1-\lambda)\mathbf{V}^\top\mathbf{P}_k^{\mathbf{Q}}\mathbf{V} + \mathbf{I}_{3\times 3} \\ \mathbf{P}_{k+1}^{\mathbf{V}} &= \lambda\mathbf{P}_k^{\mathbf{V}}\left(\mathbf{I}_{3\times 3} + \mathbf{E}\mathbf{P}_k^{\mathbf{V}}\right)^{-1} + (1-\lambda)\mathbf{P}_k^{\mathbf{V}} + \mathbf{I}_{3\times 3}.\end{aligned}$$

For  $k \rightarrow \infty$  and time invariant  $\mathbf{Q}$ ,  $\mathbf{C}$  and  $\Sigma_{\mathbf{z},k}$  it holds  $\mathbf{P}_{k+1} = \mathbf{P}_k = \mathbf{P}_{ss}$  and  $\mathbf{P}_{k+1}^{\mathbf{V}} = \mathbf{P}_k^{\mathbf{V}} = \mathbf{P}_{ss}^{\mathbf{V}}$ . Consequently, the following expression can be derived

$$\mathbf{P}_{ss}^{\mathbf{V}} = \lambda\mathbf{P}_{ss}^{\mathbf{V}}\left(\mathbf{I}_{3\times 3} + \mathbf{E}\mathbf{P}_{ss}^{\mathbf{V}}\right)^{-1} + (1-\lambda)\mathbf{P}_{ss}^{\mathbf{V}} + \mathbf{I}_{3\times 3}. \quad (8.21)$$

Equation 8.21, it can be solved for the single entries of the  $n \times m$  matrix  $\mathbf{P}_{ss}^{\mathbf{V}}$

$$\begin{aligned}\mathbf{P}_{ss}^{\mathbf{V}}(n, m) &= \frac{\lambda\mathbf{P}_{ss}^{\mathbf{V}}(n, m)}{1 + e_n\mathbf{P}_{ss}^{\mathbf{V}}(n, m)} + (1-\lambda)\mathbf{P}_{ss}^{\mathbf{V}}(n, m) + \mathbf{I}_{3\times 3}(n, m) \\ 0 &= \mathbf{P}_{ss}^{\mathbf{V}}(n, m)^2 - \frac{\mathbf{P}_{ss}^{\mathbf{V}}(n, m)}{\lambda} - \frac{\mathbf{I}_{3\times 3}(n, m)}{\lambda e_n} \\ \mathbf{P}_{ss}^{\mathbf{V}}(n, m) &= \frac{1}{2\lambda} \pm \sqrt{\frac{1}{4\lambda^2} + \frac{\mathbf{I}(n, m)}{\lambda e_n}}.\end{aligned}$$

This solution is now inserted into  $\mathbf{P}_{ss}$  to obtain the steady-state solution of the MARE

$$\mathbf{P}_{ss} = \mathbf{Q}^{1/2}\mathbf{V}\mathbf{P}_{ss}^{\mathbf{V}}\mathbf{V}^\top\mathbf{Q}^{1/2}. \quad (8.22)$$

For the case of  $A \neq \mathbf{I}_{3\times 3}$ , Equation 8.22 provides optimistic covariance estimates. The exact solution for the state covariance  $\mathbf{P}_{ss}$  is given as

$$\mathbf{P}_{ss}^{\mathbf{A}} = \begin{bmatrix} \mathbf{P}_{ss}(1, 1) - c_{1,1} & \mathbf{P}_{ss}(1, 2) - c_{1,2} & -\mathbf{A}(1, 3)\mathbf{P}_{ss}(3, 3) \\ \mathbf{P}_{ss}(2, 1) - c_{2,1} & \mathbf{P}_{ss}(2, 2) - c_{2,2} & -\mathbf{A}(2, 3)\mathbf{P}_{ss}(3, 3) \\ \mathbf{A}(1, 3)\mathbf{P}_{ss}(3, 3) & \mathbf{A}(2, 3)\mathbf{P}_{ss}(3, 3) & \mathbf{P}_{ss}(3, 3) \end{bmatrix} \quad (8.23)$$

$$= \lambda\mathbf{P}_{ss}\left(\mathbf{I}_{3\times 3} + \mathbf{C}^\top\Sigma_{\mathbf{z}}^{-1}\mathbf{C}\mathbf{P}_{ss}\right)^{-1} + (1-\lambda)\mathbf{P}_{ss} + \mathbf{Q}^{\mathbf{A}} \quad (8.24)$$

with the following relations

$$\mathbf{P}_{ss}^{\mathbf{A}} = \mathbf{A}^{-1}\mathbf{P}_{ss}\mathbf{A}^{-T} \quad (8.25)$$

$$\mathbf{Q}^{\mathbf{A}} = \mathbf{A}^{-1}\mathbf{Q}\mathbf{A}^{-T} \quad (8.26)$$

$$c_{1,1} = \mathbf{P}_{ss}(3, 3)\mathbf{A}(1, 3)^2 \quad (8.27)$$

$$c_{2,2} = \mathbf{P}_{ss}(3, 3)\mathbf{A}(2, 3)^2 \quad (8.28)$$

$$c_{1,2} = \mathbf{P}_{ss}(3, 3)\mathbf{A}(1, 3)\mathbf{A}(2, 3) \quad (8.29)$$

$$c_{2,1} = c_{1,2}. \quad (8.30)$$

For  $\mathbf{C} = \mathbf{I}_{3 \times 3}$  (see Equation 2.22),  $\mathbf{A}$  from Equation 2.21 and  $\Sigma_{\mathbf{z},t} = \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_\Theta^2 \end{bmatrix}$ ,

this expression can be easily solved in closed-form. This is done by considering the components of  $\mathbf{P}_{ss}$  separately, starting with  $\mathbf{P}_{ss}(3,3)$ .

By choosing  $\mathbf{Q}_k = \mathbf{Q}^u$ ,  $\Sigma_{\mathbf{z},k} = \Sigma_{\mathbf{z}}^u$  such that  $\mathbf{Q}_k \leq \mathbf{Q}^u$ ,  $\forall k$  and  $\Sigma_{\mathbf{z},k} \leq \Sigma_{\mathbf{z},k}^u$ ,  $\forall k$  it can be shown that  $\mathbf{P}_{ss}^e \leq \mathbf{P}_{ss}^{e,u}$ ,  $\forall k$ . The correctness of this bound is verified in Section 8.2.2 by means of simulation studies.

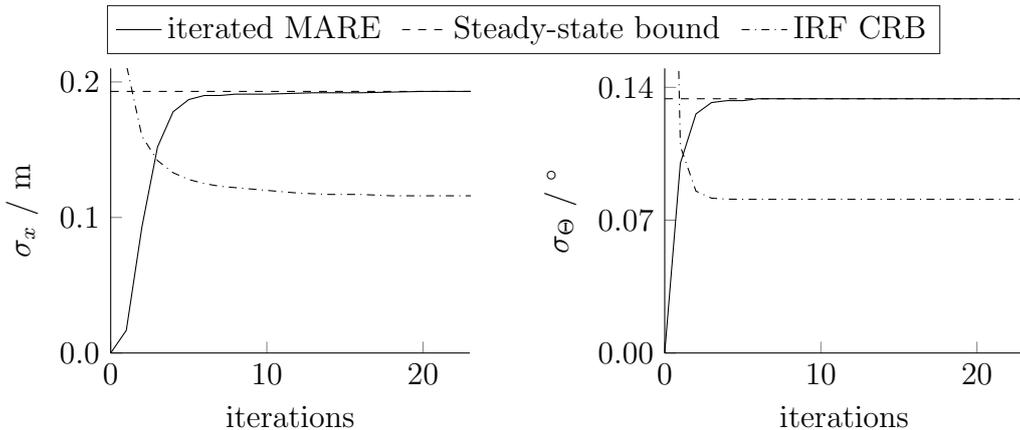
### 8.2.2 Simulative evaluation

In this evaluation section, the proposed bound is evaluated in Monte Carlo simulations and compared to two variants of the CRB. The results in Figure 8.4 were calculated for the following parameters:

$$P_m = 0.84, \quad \Sigma_{\mathbf{z}}^u = \begin{bmatrix} (0.2 \text{ m})^2 & 0 & 0 \\ 0 & (0.2 \text{ m})^2 & 0 \\ 0 & 0 & (0.1^\circ)^2 \end{bmatrix}, \quad (8.31)$$

$$\mathbf{Q}^u = \begin{bmatrix} (0.0167 \text{ m})^2 & 0 & 0 \\ 0 & (0.0167 \text{ m})^2 & 0 \\ 0 & 0 & (0.2^\circ)^2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1.67 \\ 0 & 0 & 1 \end{bmatrix}. \quad (8.32)$$

$P_m$  is chosen in accordance to the map matching rates for standard spectral registration in Section 11.3.1 and upper translational and rotational map matching errors of 0.2 m and  $0.1^\circ$  are assumed. Odometry measurements are assumed to have a standard deviation of 1% of the driven distance which gives 0.0167 m for a velocity of 60 km/h and a LiDAR measurement frequency of 10 Hz.



**Figure 8.4** The simulation results of the proposed closed-form bound. Its closed-form formulation leads to a runtime performance increase of an order of magnitude during model evaluation in Section 9.

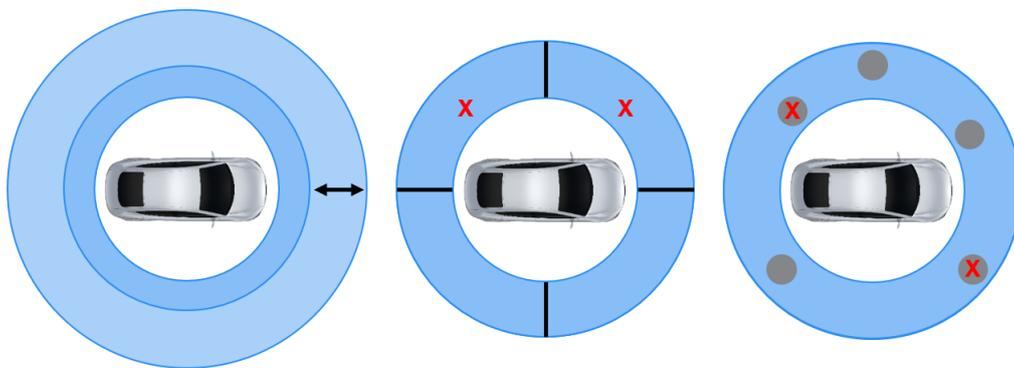
The values obtained by iterating the MARE converge against the proposed closed-form bound as seen in Figure 8.4. By iteration of the MARE, the bound is reached after approximately 20 iterations for translational performance estimation. Hence, utilization of the closed-form solution leads to a significantly decreased number of iterations during model evaluation. Considering the commonly used number of  $10^5$  and more measurement noise and landmark distribution realizations during Monte Carlo simulation, this leads to a significantly increased efficiency of the evaluation. Evaluation of the IRF CRB leads to state covariance estimates which are approximately a factor two smaller than the proposed bound. While the MARE based bound seems feasible in the context of the localization results presented in Section 11.3, the IRF CRB seems to be optimistic. This result is in agreement with [HRFT04] and [BD09] who present similar simulation outcomes. The evaluation results further motivate the usage of the proposed close-form performance bound for model-based design in Section 9.

### 8.3 Occlusions and missed detections

Challenges arise where temporal or spatial correlation between consecutive measurements exist. These are frequently caused by occlusions. An example is the partial occlusion of the sensor FoV by a truck driving next to the ego vehicle, potentially leading to low feature observation rates which result in sequences of map matching results with degraded accuracy. Nonetheless, the consideration of occlusions and missed landmark detections is required in order to ensure the applicability of the statistical models to real world scenarios. Especially the localization model (Section 8.2.1) assumes an independently and identically distributed occurrence of degradation events. The limited expressiveness for bundled appearance of false or inaccurate estimates is not specific to the considered application and the previously developed models of vehicle localization systems. In fact, *Huber* [Hub11] and other authors described it as significantly important and mainly unresolved in the broad context of (robust) statistics. To this point, a comprehensive overview over degradation scenarios and respective modeling strategies is elaborated in this section.

An overview over frequently occurring perception degradation scenarios is provided in Figure 8.5. Prolonged partial occlusions of the sensor FoV and challenging weather conditions like rain and fog effectively lead to a decreased FoV. In this thesis, these effects are considered by adaptation of the sensor FoV during the design process. Transition effects occurring during variation of the occlusion scenario are considered small and are hence neglected. The system designer has then to separately specify the localization performance requirements for the different levels of FoV deterioration. A possible strategy is the definition of a worst case scenario until which the localization and, hence, the automated vehicle system is intended to provide full functionality. Starting from this basic design, degraded system operation modes could be designed. Prolonged periods of map matching failures can be treated separately by specification of odometry accuracy requirements. For sen-

sors with high mounting positions, such as the utilized Velodyne LiDAR sensor (Section 4.1), the probability of missed detections is significantly higher than the occlusion probability if high features like house walls or poles are considered. In this case it is assumed that the missed detections are uniformly distributed in time and space and are uncorrelated. This type of missed detections is considered at the sensor model during numerical integration. If required, missed detection probabilities can be modeled by other probability density functions which can be easily incorporated into the numerical integration step. Adaptation of the FoV and uniformly distributed missed detections are considered in Section 9, for model-based localization system design on basis of the previously developed statistical models.



**Figure 8.5** Decreased detection ranges due to bad weather conditions (left) and long sequences of partially occluded sensor field of view (middle) are addressed by an adaptation of the sensor field of view (blue areas). Random missed detection (right) are captured by the sensor model.

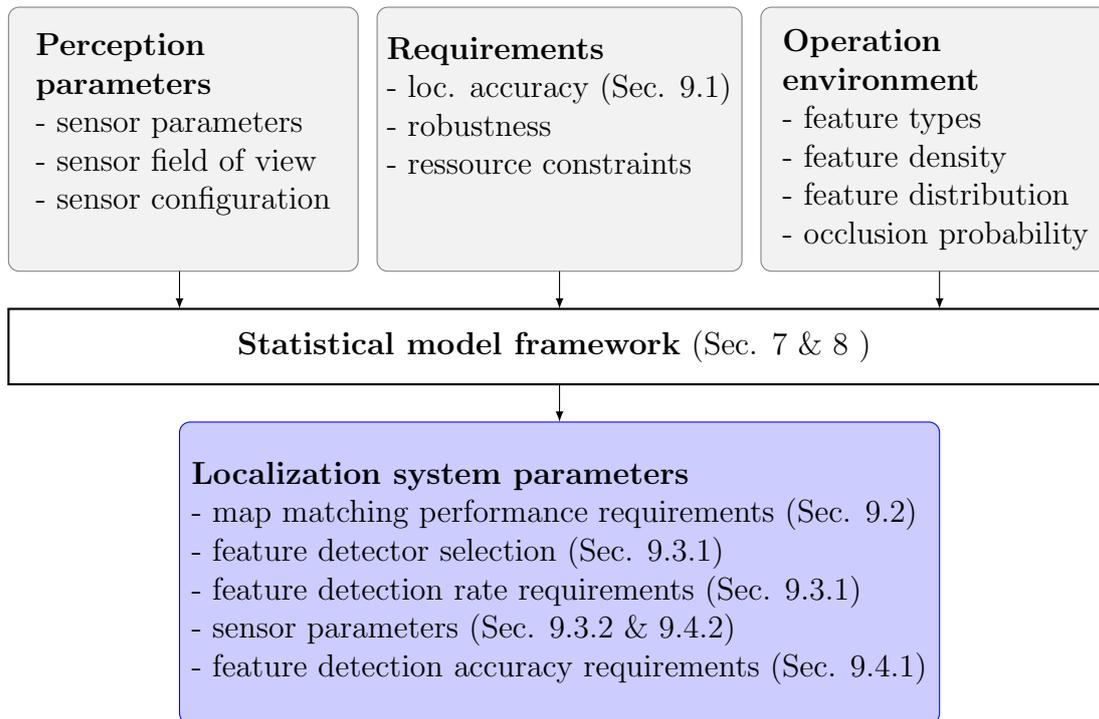
## 9 Model-based design

This section is concerned with the derivation of localization system design parameters like the selection of feature types, detection ranges and accuracy by utilization of the developed statistical models. Vehicle localization in stressful open environments as well as the design of suitable localization systems remain challenging tasks, even after decades of research. The incorporation of additional spatially referenced information from digital maps into the behavior planning process of an automated vehicle system necessitates a sufficiently accurate and robust vehicle pose estimate. Localization accuracy requirements typically arise from the intended use case for which the additional information source shall be used. The derivation of localization requirements is exemplified for a right turn scenario in Section 9.1.

In this section, a model-based approach (Figure 9.1) for the requirement-driven design of localization systems is elaborated which is conceptually different from commonly utilized experience- and test-based development procedures. Most localization systems found in literature were designed purely based on the knowledge and experience of the expert developers. The resulting implementations are evaluated and, if possible or desired, parameterized under consideration of the localization performance requirements. It is argued here, that starting a system design process from the actual performance requirements, can yield advantageous system designs, e.g. with respect to efficiency and robustness. Furthermore, this requirement-driven approach contributes to a repeatable and efficient design process.

A great challenge in this approach is the manifold of influencing factors for vehicle localization system design decisions and their impact on localization performance. To this point, the interrelations between operation environment characteristics, perception parameters and localization performance were modeled in the two previous sections (7 and 8). The statistical models are used with different sets of fixed input parameters in order to derive a feasible set of remaining localization system design parameters. Therefore, Monte Carlo simulations are used if required and otherwise, the solutions are directly calculated. The results from this section indicate, that many aspects of localization system design that are known by expert developers can be qualitatively and quantitatively described by the proposed models. Hence, the results underline that the examination of statistical models can lead to a deeper theoretical understanding of the whole system and thereby provide ideas and approaches for further improvements and optimizations, e.g. data rates and computational performance. The general concept is summarized in Figure 9.1 and an overview over the contents in the following sections is given hereafter.

Hybrid localization algorithms can lead to system implementations with enhanced robustness. Failures of single redundant subsystems can be detected and compensated. Beside other beneficial impacts, the utilization of different feature types and matching approaches leads to a higher probability of obtaining a correct and accurate absolute vehicle pose estimate. This advantage is commonly counteracted by the well known disadvantage of reduced computational efficiency of redundant sys-



**Figure 9.1** Feasible localization system design decisions are based on the localization requirements and constraints such as available bandwidth for data transfer. Statistical models are used in this section to support the design process by quantification of expert knowledge about the interrelations between design parameters and constraints.

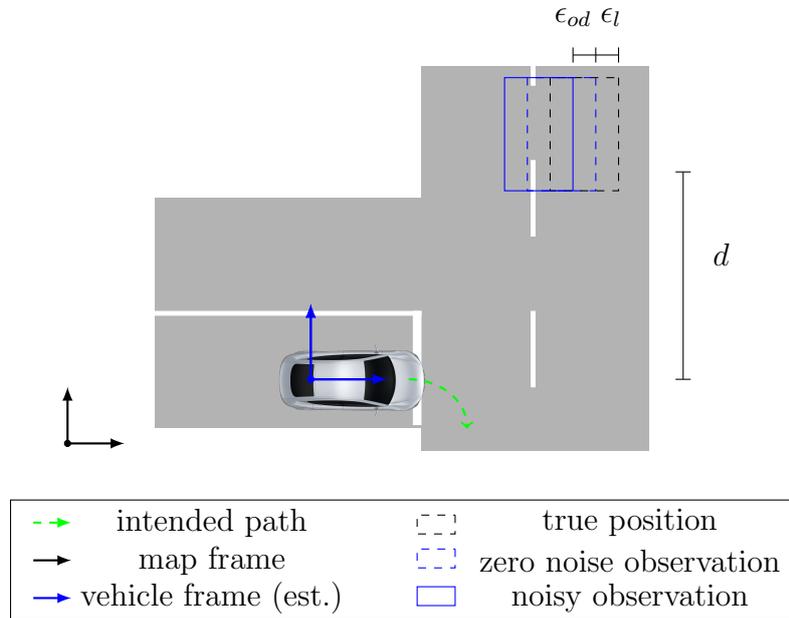
tem architectures. Consequently, during the development of the vehicle localization system, a trade-off between the gain in robustness by the introduction of redundancy and the computational costs involved has to be found. Thus, it is argued here that model-based approaches can support to find such a beneficial trade-off. In this context, the interrelation of matching rates, matching accuracy and the localization performance are discussed and respective requirements are derived in Section 9.2. As another important aspect of efficiency and robustness, the derivation of landmark density requirements and the selection of suitable detection algorithms is considered in Section 9.3. Additionally, the influence of vehicle-relative landmark arrangements is discussed. In addition to the selection of feature detectors, the choice of intrinsic sensor parameters is of crucial importance for the development of perception systems. The derivation of accuracy requirements for LiDAR based detection of specific feature types is discussed in Section 9.4.1. Intrinsic sensor parameters are derived for stereo video camera based localization. The impact of the sensor field of view on the localization accuracy is discussed in Section 9.3.2. A discussion of the presented results and their contribution to the design of the localization systems introduced in Part IV is given in Section 9.5.

## 9.1 Localization requirements

The definition of localization requirements are strongly coupled to the design of the automated vehicle system. The intended level of automation (Section 2.3) determine the requirements concerning normal system operation and availability of safety strategies within defined boundaries. Requirements commonly enclose accuracy of the pose estimate and the availability of correct estimates under well defined circumstances.

For normal vehicle operation, localization accuracy requirements for UAD typically arise from associating the ego vehicle and other traffic participants to driving lanes. This is of special significance for joining priority roads as well as for the detection of oncoming traffic for overtaking and at left turns. Lane marking detection could enable the association problem on the sole basis of the automated vehicle system. However, the required high detection ranges would likely result in potentially infeasible sensor requirements and might not be possible for specific sensor setups. Thus, map-relative localization in combination with dynamic object detection is commonly utilized. This approach leads to a summation of the detection and localization errors for highly accurate maps. Thereby, the localization requirements are increased compared to the assumption of zero noise in dynamic object detection. In the work at hand, both uncertainty sources are taken into account for localization requirement derivation. In the following, a right turn scenario is considered for the derivation of localization requirements and results are derived for the setup shown in Figure 9.2. The localization requirements are highly sensitive to the detection distance  $d$  which depends on the use case. In the right turn scenario, the ego vehicle is standing at the turning point and the system has to decide whether to accelerate or wait for another vehicle to pass. For a velocity of 50 km/h, an ego vehicle acceleration of  $2 \text{ m/s}^2$  and a non-cooperative behavior of the approaching vehicle, a result of  $d = 87 \text{ m}$  is obtained. This result also includes dead time and a gap between the ego and the approaching vehicle of 1.5 s. The requirement for the lateral detection accuracy of 1.25 m for 2.5 m broad lanes is now symmetrically divided into error contributions of  $\epsilon_{l,max} = \epsilon_{od,max} = 0.625 \text{ m}$  from localization and dynamic object detection. The error contribution from localization  $\epsilon_l$  is the sum of the orientation  $\epsilon_o = d \tan(\epsilon_\Theta)$  and position  $\epsilon_x$  estimation error contributions. The approximation  $\epsilon_{l,max} \approx \epsilon_x + d\epsilon_{\Theta,rad}$  suggests an asymmetric assignment of error contributions. To this point,  $\epsilon_o$  is required to be less than 0.35 m and  $\epsilon_x$  less than 0.25 m. The requirement for  $\epsilon_o$  corresponds to  $\epsilon_\Theta \leq 0.23^\circ$ .

Requirements concerning the vehicle pose estimation accuracy are complemented by availability requirements. The latter occur in several different scenarios. The matching rate is an important parameter during regular operation of the localization system. It has to be sufficiently high in order to meet the localization accuracy requirements. Potentially higher requirements arise from scenarios where an initialization of the localization system has to be conducted. This is typically the case at the beginning of a automated drive where an accurate pose estimate has to be determined or verified. Furthermore, intended safety strategies might require

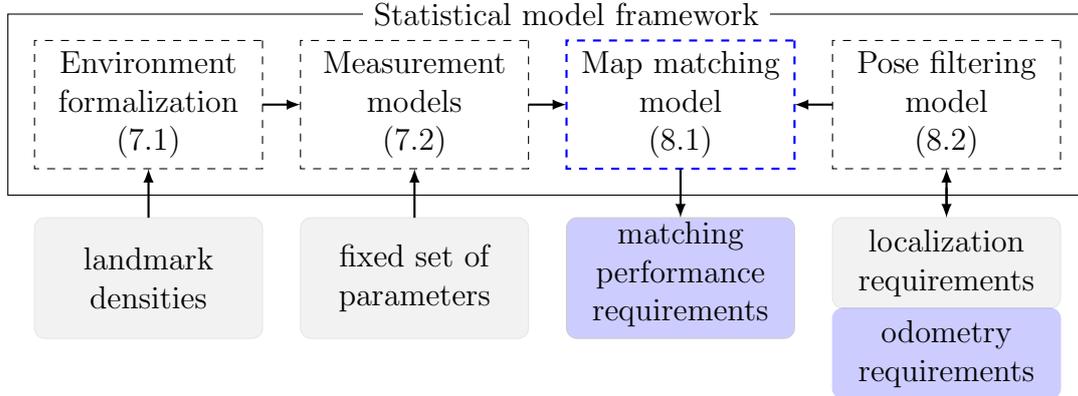


**Figure 9.2** A common root cause for demanding localization accuracy requirements is the association of dynamic objects to lanes, as in the depicted right turn scenario. Object detection and localization uncertainties contribute to the lane-relative object perception uncertainty  $\epsilon_{od} + \epsilon_l$ .

a initialization capability or increased accuracy requirements for critical areas in the operation environment, like intersections. The same argumentation holds for scenarios where a system crash requires fast reinitialization. These scenarios are discussed in later sections. It is argued, that the utilization of the developed statistical models contribute to the identification of localization system design parameters which lead to a fulfillment of the localization requirements. Furthermore, the models can contribute to find the parameters which foster an efficient way of achieving the localization requirements as compared to other parameter choices.

## 9.2 Matching rate and accuracy requirements

In literature there is an ongoing discussion on the often map matching algorithm inherent trade-off between accuracy and robustness manifested as increased matching rates compared to the non-robust counterparts. The current trend of data transmission based on wireless communication technology forms another source of motivation for this section. In the context of automated vehicle operation, many approaches include a wireless transmission of localization map data to the vehicle. Therefore, the loss of localization map fragments has to be added to the set of possible causes of map matching failures which has significant implications on the localization system design. The aforementioned topics are now addressed by evaluation of the interrelation between the matching probability  $P_m$  and the map matching



**Figure 9.3** Overview over the statistical model evaluation strategy for the derivation of map matching and odometry requirements.

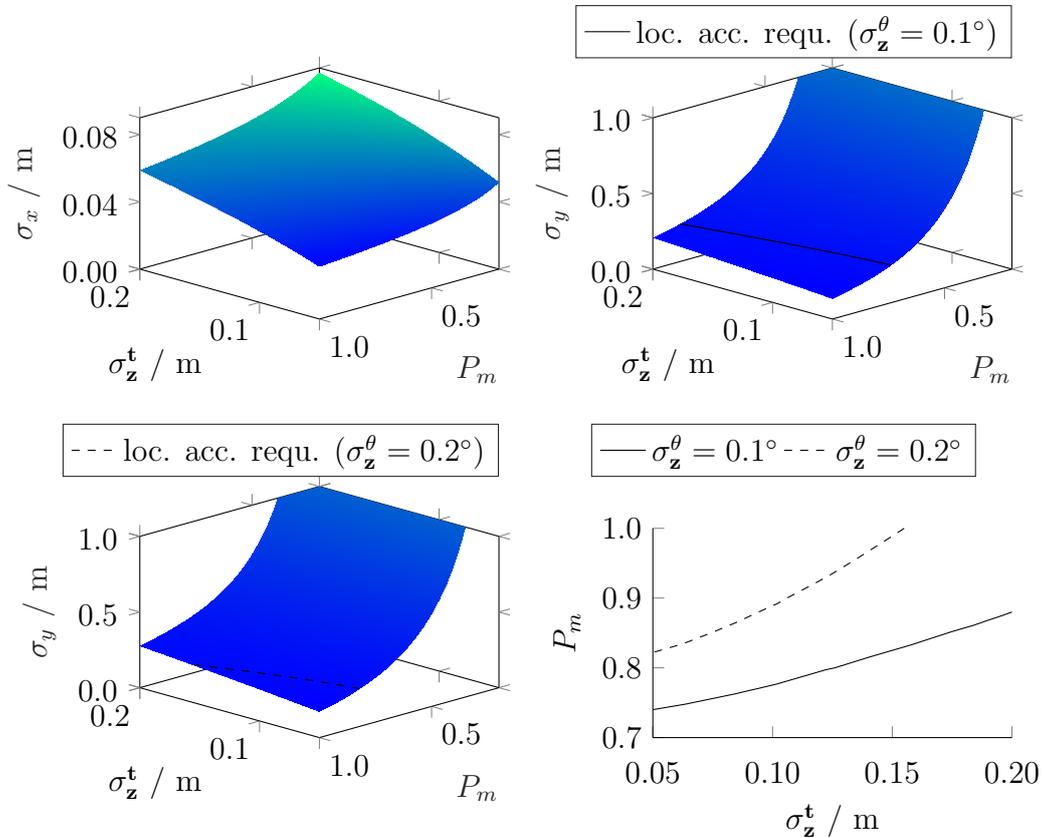
accuracy. An overview over this section is given in Figure 9.3.

A given localization accuracy requirement as derived in Section 9.1, necessitates a minimum rate of information gain about the absolute vehicle pose. In the latter context, the matching rate and accuracy as well as the accuracy of vehicle motion estimation are the main parameters. An increased matching rate counteracts a decrease in matching accuracy and vice versa. The impact and the interrelation between these two major influencing factors are discussed in Section 9.2.1. Furthermore, the role of relative motion estimation is discussed in Section 9.2.2.

### 9.2.1 Interrelation between matching rate and accuracy

The results presented in this section quantify the interrelation between matching rate  $P_m$  and map matching accuracy  $\Sigma_z$ . For the sake of readability, the evaluation of the localization system model from Section 8 is conducted on the basis of the parameters proposed in Section 8.2.2. This choice corresponds to a drive with a steering angle of  $0^\circ$ . The most significant simulation outcomes for the translational localization accuracy components  $\sigma_x$  and  $\sigma_y$  are summarized in Figure 9.4. The data in the upper row of Figure 9.4 illustrate the interrelation between the translational ( $\sigma_z^t$ ) and rotational matching accuracy ( $\sigma_z^\Theta$ ), the matching rate  $P_m$  and the resulting localization accuracy  $\sigma_x$ ,  $\sigma_y$  of the absolute vehicle pose estimate. The  $\sigma_x$  indicates a high accuracy while map matching requirements arise from  $\sigma_y$ . The surface is divided by a curve, indicating the localization accuracy requirement of 0.25 m and shown in the lower right figure. All combinations of matching rate and accuracy on the left side of the curve are sufficient in order to meet the localization accuracy requirements. Accordingly, a matching rate of approximately 78% in combination with an average map matching error of  $\sigma_z^t = 0.1$  m is sufficient. Cutting the rotation estimation accuracy in map matching to half from  $0.1^\circ$  to  $0.2^\circ$ , the requirements for  $P_m$  rises to 90% for the same  $\sigma_z^t$ . An increase in the map matching frequency from 10 Hz to 20 Hz leads to significantly decreased map matching requirements

(Figure 9.5).

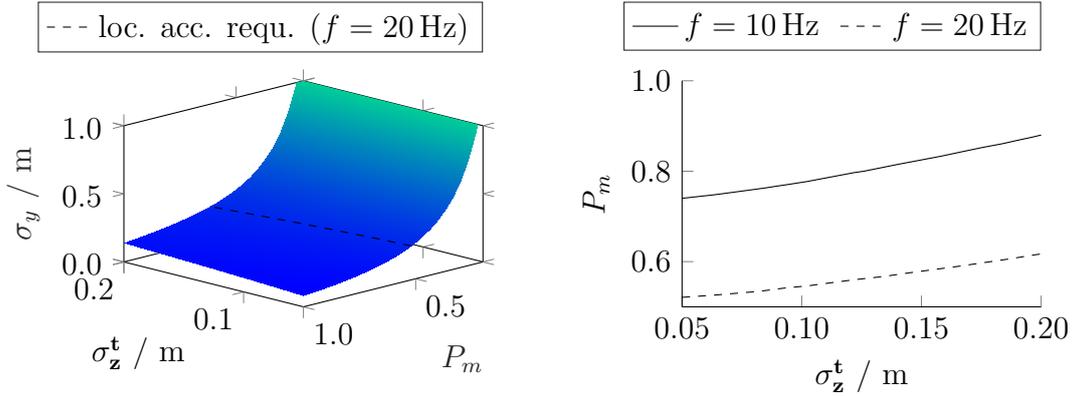


**Figure 9.4** The trade-off between accuracy and matching rate is fundamental for map matching algorithm selection. The rotational matching accuracy has greater influence on the localization accuracy than the translational. These results indicate that the area covered by a scan shall be sufficiently high which is an important observation for the selection of a grid dimension and resolution pair for spectral registration based localization.

This observation holds for regular operation of localization systems and equally distributed and random map matching failure occurrences. In summary, for normal operation of the localization system, the matching accuracy can be low compared to the localization accuracy requirement for high matching rates and high accuracy can compensate low matching rates - as long as the matching failures are random. The role of relative motion estimation is further discussed in the following section.

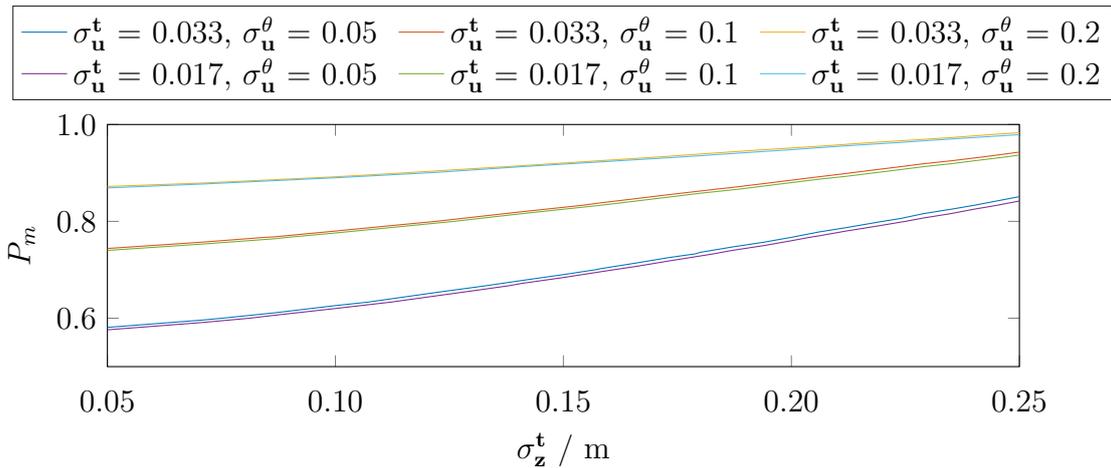
### 9.2.2 Odometry accuracy

The results in this section describe the interrelation between map matching performance requirements and the odometry accuracy. According to the results in Figure 9.6, the odometry significantly contributes to the localization accuracy. Due



**Figure 9.5** Increased map matching frequencies significantly contribute to relaxed map matching performance requirements.

to the additional information from relative motion estimation which is fused with the map matching results, the matching accuracy requirements are decreased. To this point, the rotational accuracy  $\sigma_{\mathbf{u}}^\theta$  has a stronger impact than the translation estimation performance  $\sigma_{\mathbf{u}}^t$ . This can be observed from significantly decreased map matching requirements for higher precisions in rotation estimation, whereas an increase in translation estimation accuracy results in minor requirement changes. Minimum odometry accuracy can be derived for a given map matching accuracy and rate. For the design of the scan-based odometry in Section 10.1, the focus is put on the reduction of rotation errors which is a particular challenge due to the utilization of discretized scan data for spectral registration.



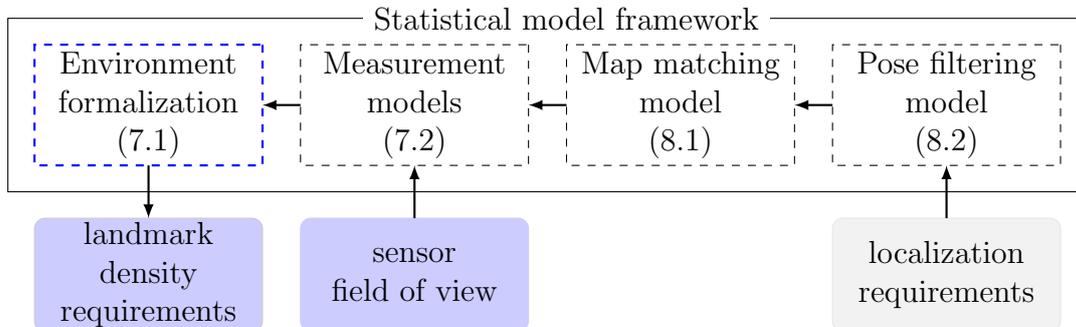
**Figure 9.6** The rotation estimation accuracy has significantly higher impact on the localization accuracy than the translational accuracy. This observation will be considered during the development of a scan-based odometry in Section 10.1.

In addition to the beneficial influence on localization accuracy, odometry mea-

surements can compensate prolonged periods of map matching failures. In order to compensate for prolonged map matching failures, the map matching requirements can be increased, such that the pose estimate at the beginning of the matching failure sequence is highly accurate. Consequently, the allowed gain in uncertainty from motion would be extended. Otherwise, an accurate odometry can reduce the amount of uncertainty induced by motion and thereby extend the operation range of the vehicle system. The decision about the compensation strategy has to be considered during design and can lead to asymmetric accuracy requirements, e.g. if a highly accurate odometry was required in combination with an inaccurate map matching procedure. Since, prolonged failures of map matching shall be an exception, the strategy with an accurate odometry is chosen in this work.

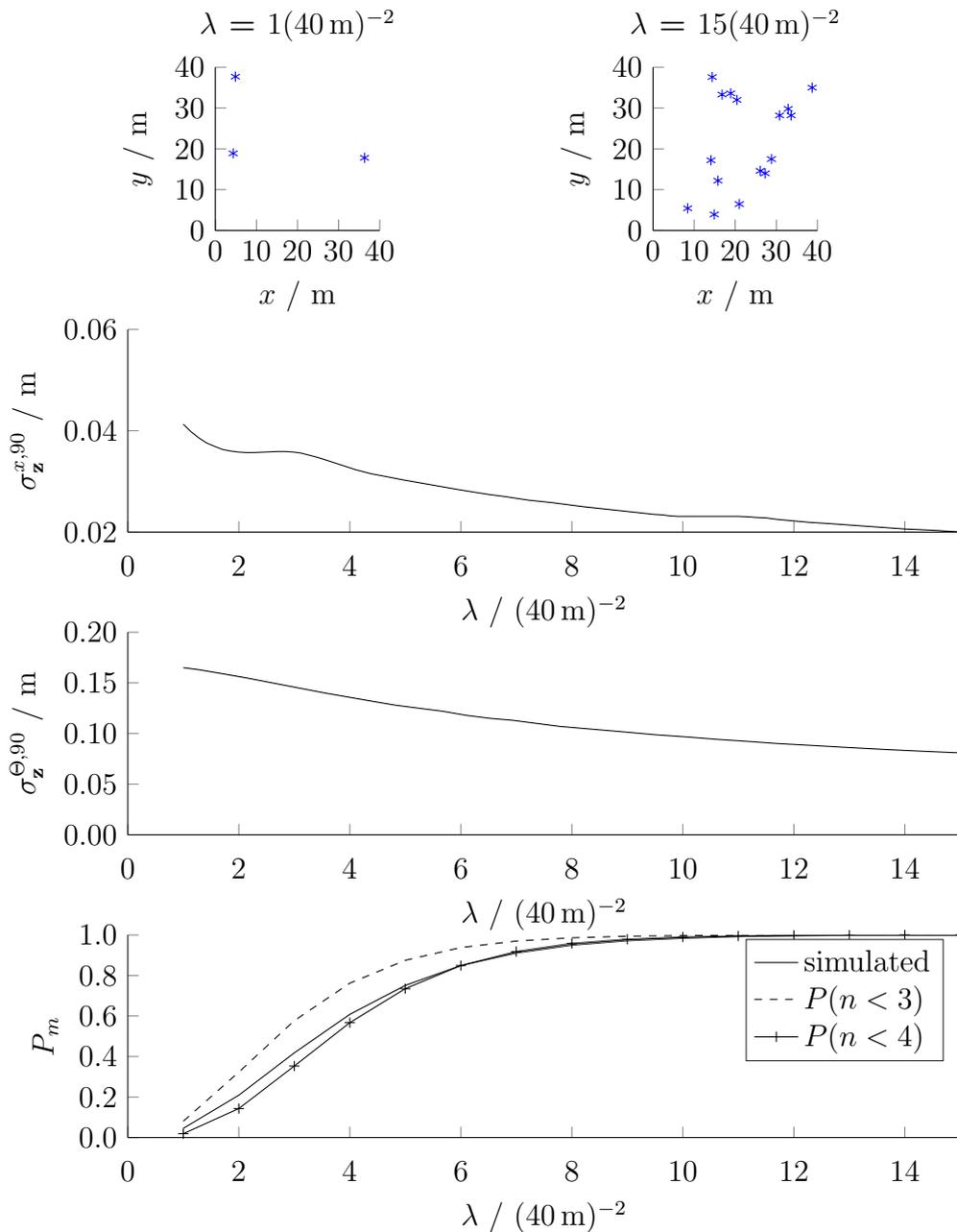
### 9.3 Landmark distribution requirements

The accuracy of vehicle pose estimates provided by a localization system strongly depends on the available density of features and their vehicle-relative arrangement. The feature distribution has to be sufficiently dense, while excessively high feature densities lead to unnecessarily high storage space demands, communication bandwidth requirements and computational costs (SLAM:  $\sim O(N^2)$ ). Therefore, a localization map is not a complete representation of the urban environment, but is a model which entails a set of descriptive features. In this section, the interrelation between landmark densities, arrangements and localization system parameters are studied based on the introduced statistical models as shown in Figure 9.7.



**Figure 9.7** Overview over the statistical model evaluation strategy for the selection of feature detectors and a discussion on the impact of changes to the sensor fields of view.

The influence of landmark density and matching performance is shown in Figure 9.8 and were obtained by Monte Carlo simulation with  $10^6$  point process realizations. For this evaluation, matching performance is expressed as the 90% quantiles of the map matching accuracy  $\sigma_{\mathbf{z}}^t$ ,  $\sigma_{\mathbf{z}}^\theta$  and matching rate  $P_m$ . The landmark density is modeled by the intensity  $\lambda$  of a homogeneous PPP. Two PPP realizations for different  $\lambda$  are shown in the top row of Figure 9.8.



**Figure 9.8** Influence of the landmark density on the map matching accuracy and matching probability.

For homogeneous Poisson point processes, the matching probability  $P_m$  can be approximated in closed-form under the assumption that at least  $n_{\min} = 3$  landmarks have to be visible for obtaining a map matching result. The intensity  $\lambda_{\mathcal{S}}$  denotes the point process intensity normalized to the area  $\mathcal{S}$  covered by the sensor FoV.

Then  $P_m$  has the upper bound

$$P_m = P(n \geq n_{\min}) = 1 - \sum_{l=0}^{n_{\min}-1} P(l) \quad (9.1)$$

$$= 1 - \sum_{l=0}^{n_{\min}-1} \frac{(\int_{\mathcal{S}} \lambda(\mathbf{s}) d\mathbf{s})^l}{l!} \exp\left(-\int_{\mathcal{S}} \lambda(\mathbf{s}) d\mathbf{s}\right) \quad (9.2)$$

$$= 1 - \left(\frac{1}{0!} + \frac{\lambda_{\mathcal{S}}}{1!} + \frac{(\lambda_{\mathcal{S}})^2}{2!}\right) \exp(-\lambda_{\mathcal{S}}) .$$

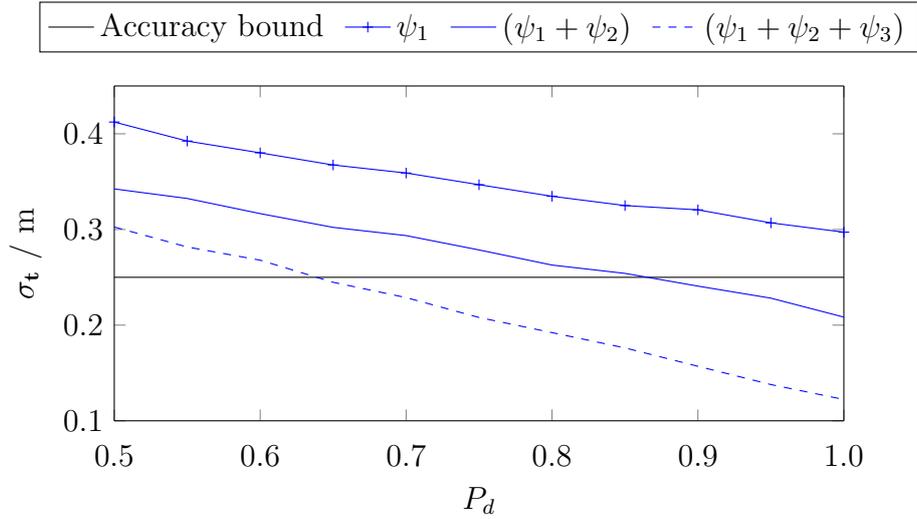
For  $n_{\min} = 4$  the lower bound is obtained. These bounds can be used for an efficient calculation of design parameters without the necessity of Monte Carlo simulations.

The results in Figure 9.8 are intuitive, since higher landmark densities lead to increased map matching performance and decrease the feature detection requirements. Nonetheless, efficient system designs shall consider the landmark densities as a design parameter. The latter can be adapted to the capability of other submodules like the sensor setup to avoid asymmetric requirement assignments. This is accomplished by the model-based selection of feature detection algorithms that enable the detection of landmarks, characterized by a sufficiently high density in the operation environment. The selection of detection algorithms is studied in Section 9.3.1. The influence of vehicle-relative landmark arrangements on the map matching accuracy is studied in Section 9.3.2. The field of view of the perception system typically has significant influence the number of landmark detections. This aspect is studied in Section 9.3.2 as well.

### 9.3.1 Detector selection

The density of landmarks for localization is limited by the operation environment as long as no artificial landmarks are available. Typically, the number of detectable landmarks is significantly higher than required by the localization system. Hence, the derivation of minimum landmark densities is here interpreted as the task of selecting a set of landmark detectors. This approach allows to combine different sparse or heterogeneously distributed landmark types to an average cumulative density which meets the requirements. On the perception side, the density of detected landmarks is also influenced by detection performance of the detectors.

The interrelations between the localization performance and landmark densities are studied under the assumption of landmark distribution according to a homogeneous Poisson point process. The results are shown in Figure 9.9. In this example, utilization of one landmark type with intensity  $\psi_1 = 30$  per unit area of  $1000 \text{ m}^{-2}$  does not suffice in order to meet the localization accuracy requirement. The combination of two detectors for landmark types 1 and 2 enables a sufficiently high localization accuracy  $\sigma_t$ . The choice of this landmark detector set requires a minimum detection rate of 86%, including missed detection and occlusions as caused by other traffic participants. Depending on the mounting position, traffic density and



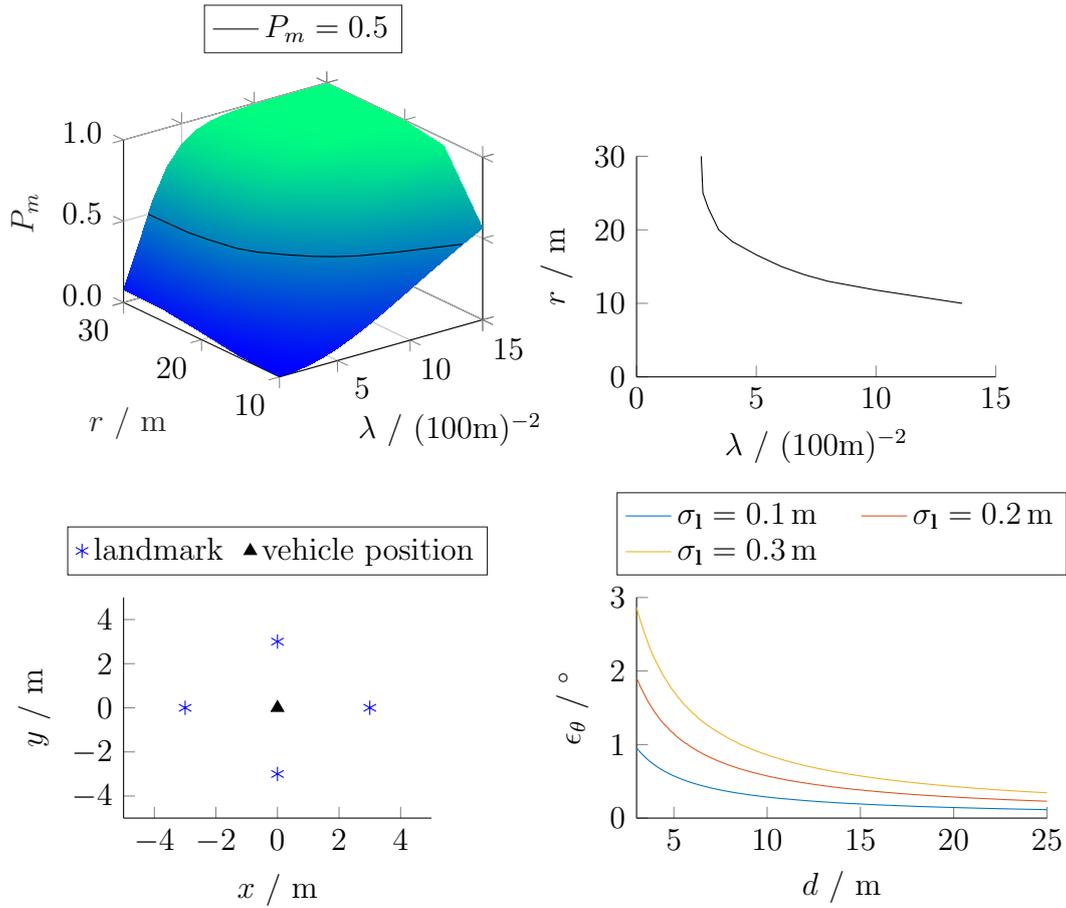
**Figure 9.9** The robustness of a localization system against occlusions is coupled to the available landmark density which can be influenced by the selection of detector algorithms. The densities per unit area of  $1000 \text{ m}^2$  are selected as  $\psi_1 = 30$ ,  $\psi_2 = 15$  and  $\psi_3 = 15$ . Utilization of all landmark detectors leads to a robustness against a missed detection rate of 35%.

weather conditions, this design can likely lead to localization failures. Consequently, the localization system would be sensitive to small occlusion and missed detection rates. From a robustness perspective, the utilization of a combination of the three hypothetical landmark detectors is advisable. By the additional increase in landmark density, the allowed false negative rate is increased by more than 20%. This underlines the inherent trade-off between efficient system design and robustness. High landmark densities and detector performances can be traded for increased localization robustness. Statistical models can be utilized in the proposed manner to identify feasible sets of detectors for the intended accuracy and robustness level.

For the pole-based map matching approach presented in Section 11, a microscopic view on the problem is taken by consideration of the actual use case. The matching modality is utilized in order to localize within intersection areas where the pole landmark density is commonly sufficiently high. The important design parameter in this case is the sensor FoV whose role is discussed in Section 9.3.2.

### 9.3.2 Influence of the sensor field of view

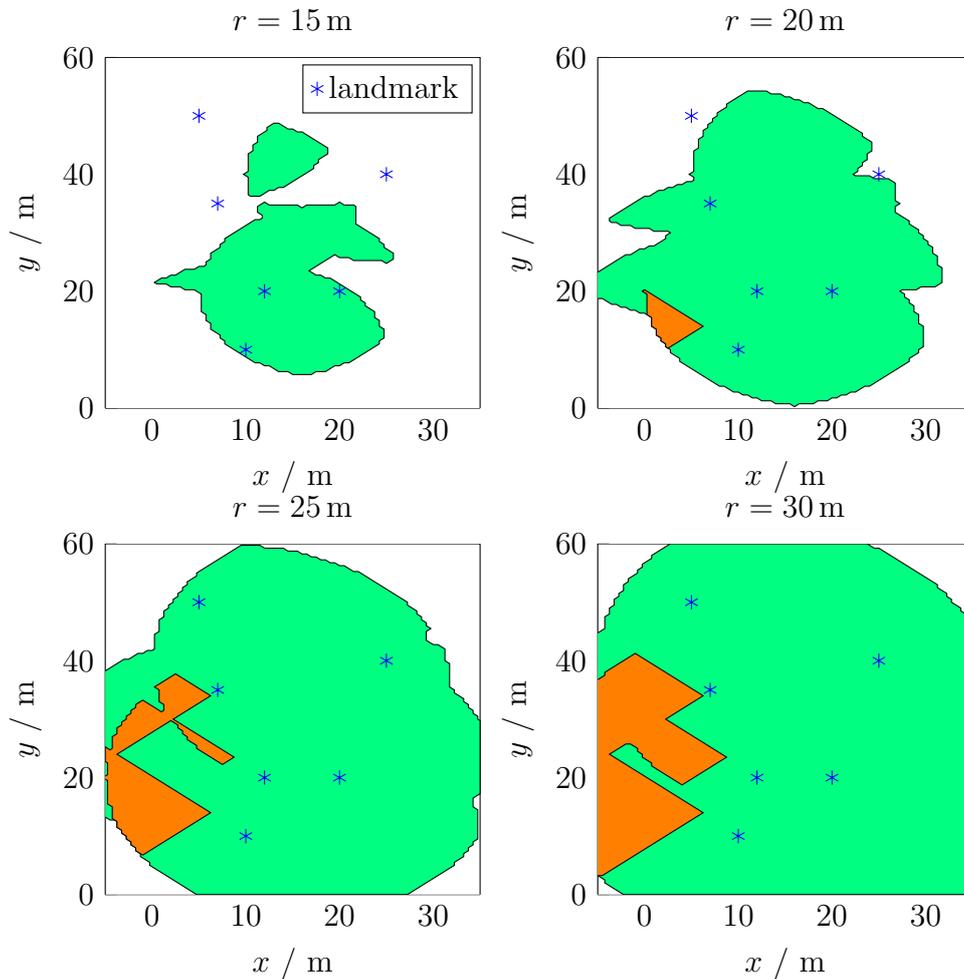
For pole landmark based localization for intersection areas, the landmark detection range of the Velodyne is influenced by several aspects. The first limitation is imposed by the sensor design, namely the limited angular resolution which leads to a maximum detection range for pole objects. This limitation subsumes the maximum LiDAR detection range which is far higher. Beside the physical limitations, the false positive and missed detection rates increase with the detection range. This likely



**Figure 9.10** An increased detection range  $r$  can lead to an improved matching rate  $P_m$  and enables the detection of landmarks in further distances. The latter has a positive influence on the angular matching error  $\epsilon_\theta$  and can reduce the impact of feature measurement uncertainty  $\sigma_1$ . As a consequence, the maximal range for the pole detection is expanded despite the increased amount of false positive detections to be processed.

leads to higher numbers of landmark measurements in general and an increased false detection rate in particular. Both aspects account for increased computational costs and false data association results might become more likely. In order to decide whether the benefits of a high detection range justify the arising disadvantages, a qualitative and quantitative understanding of the benefits is elaborated in this section.

First important results are summarized in Figure 9.10. Increased  $r$  likely lead to more landmark detections and matching rates  $P_m$  (top left figure in Figure 9.10). Differently put, a matching rate requirement can be achieved by increased detection ranges (top right figure), given a respective landmark distribution. Additionally, the vehicle-relative landmark arrangement has an influence on the matching accuracy.



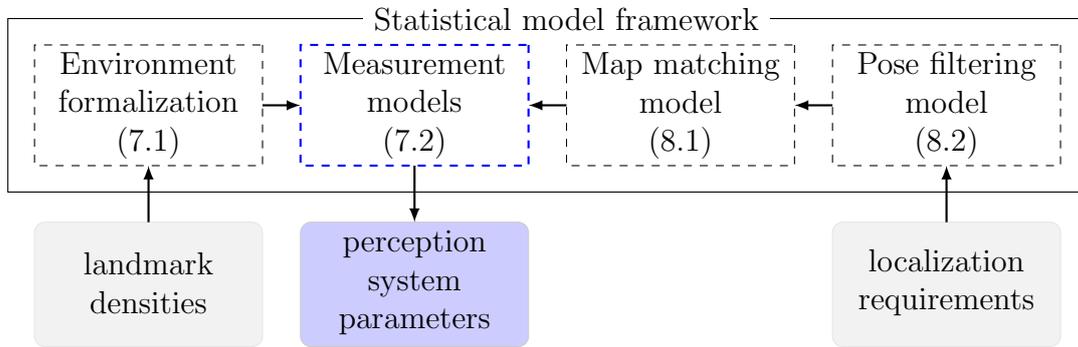
**Figure 9.11** The sensor field of view has significant influence on the achievable localization system performance (green area:  $360^\circ$  opening angle; orange area:  $90^\circ$  opening angle at  $\theta = 0^\circ$ ).

As shown in the lower row of Figure 9.10, the angular matching error  $\epsilon_\theta$  becomes smaller with increasing distance  $d$  between the sensor and landmarks. This influence is weakened by sensor technologies with detection errors which increase with  $d$ , such as stereo-video systems (Section 7.3). Localization requirements might justify the expansion of the detection range in order to make use of this effect for respective landmarks types. For  $r > 20$  m, the slope of the curves decrease, yielding minor further accuracy gains. Another significant influence of the sensor FoV can be seen from exemplary landmark arrangements shown in Figure 9.11. For safety purposes, localizability in the whole intersection area and for all vehicle orientations is considered favorable. Accordingly, a FoV of  $360^\circ$  in combination with a detection range of at least 20 m is required for a good coverage in the considered example. In summary, additional advantages of increased FoVs can outweigh disadvantages. To this point,

the detection range for pole-based map matching will be chosen as  $r \geq 20$  m.

## 9.4 Derivation of sensor parameters

One aspects of perception system design is the selection of sensor parameters for the support of sensor selection and the design of new sensors. An overview is provided in Figure 9.12. The derivation of pole detection accuracy requirements is conducted in Section 9.4.1. Parameters for a stereo video camera are derived in Section 9.4.2. The results were obtained under the following adaptations to the map matching model (Section 8.1).



**Figure 9.12** Overview over the statistical model evaluation strategy for the derivation of perception system parameters.

To enable an efficient evaluation of the measurement models from Section 7.2 and Section 7.3, these are rewritten as follows:

$$\Sigma_1 = \xi \begin{bmatrix} \sigma_{x,\xi}^2 & \sigma_{xy,\xi}^2 \\ \sigma_{yx,\xi}^2 & \sigma_{y,\xi}^2 \end{bmatrix}. \quad (9.3)$$

The scalar factor  $\xi$  can be interpreted in different ways. As part of the stereo video (SVC) depth measurement model, it entailed sensor parameters (compare Section 7.3)

$$\xi = \frac{\sigma_d^2}{(c \cdot b_w)^2}. \quad (9.4)$$

In the case of LiDAR based pole detection it takes the role of a covariance scaling factor  $\xi = \sigma_l^2$  which can also be applied to the SVC model. Consecutively, the map matching accuracy model from Equation 8.14 is simplified to allow for a fast derivation of  $\xi$ . Lengthy, but straight forward calculation leads to the following formulation

$$\begin{aligned} 0 &= a_{t11}\xi^2 + b_{t11}\xi - \sigma_{t11}^2, \\ 0 &= a_{t22}\xi^2 + b_{t22}\xi - \sigma_{t22}^2, \\ 0 &= a_{t12}\xi^2 + b_{t12}\xi - \sigma_{t12}^2, \\ 0 &= a_{t21}\xi^2 + b_{t21}\xi - \sigma_{t21}^2. \end{aligned} \quad (9.5)$$

With

$$\Sigma_{\mathbf{p}} = \begin{bmatrix} \mathbf{E}_{11} & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{E}_{22} \end{bmatrix} \quad (9.6)$$

from Equation 8.16, the coefficients in Equation 9.5 are given as

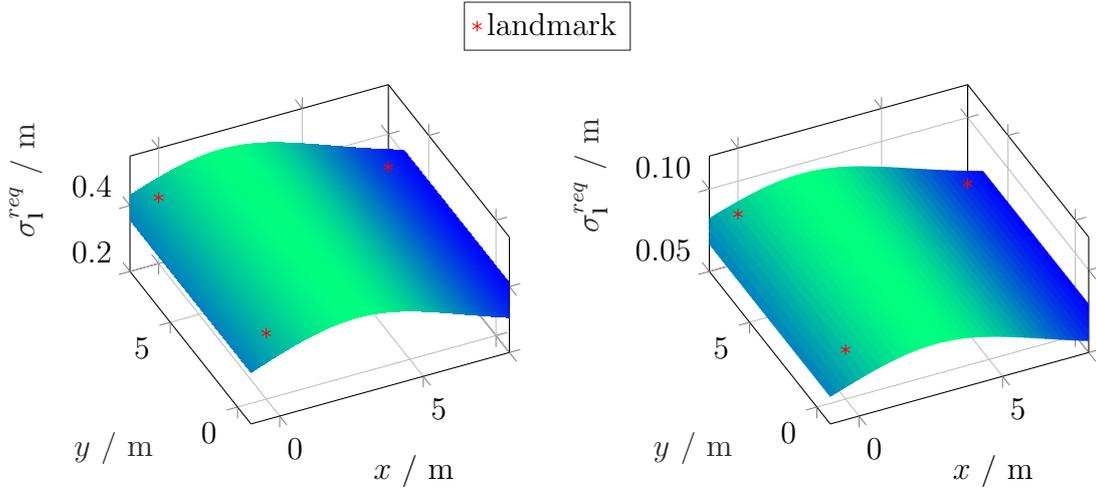
$$\begin{aligned} a_{t11} &= \text{trace}(\mathbf{D}_1 \mathbf{E}_{22} \mathbf{D}_1 \mathbf{E}_{11}) + \text{trace}(\mathbf{D}_1 \mathbf{E}_{12} \mathbf{D}_1 \mathbf{E}_{12}) \\ b_{t11} &= ([c_k \ s_k] \mathbf{D}_1 \mathbf{E}_{12} + \tilde{\mathbf{x}}_k^\top \mathbf{D}_1 \mathbf{E}_{22}) \mathbf{D}_1 \tilde{\mathbf{x}}_k + ([c_k \ s_k] \mathbf{D}_1 \mathbf{E}_{11} + \tilde{\mathbf{x}}_k^\top \mathbf{D}_1 \mathbf{E}_{21}) \mathbf{D}_1 [c_k \ s_k]^\top \\ a_{t12} &= a_{t21} = \text{trace}(\mathbf{D}_1 \mathbf{E}_{11} \mathbf{D}_2 \mathbf{E}_{22}) \\ b_{t12} &= b_{t21} = ([c_k \ s_k] \mathbf{D}_1 \mathbf{E}_{12} + \tilde{\mathbf{x}}_k^\top \mathbf{D}_1 \mathbf{E}_{22}) \mathbf{D}_2 \tilde{\mathbf{x}}_k + ([c_k \ s_k] \mathbf{D}_1 \mathbf{E}_{11} + \tilde{\mathbf{x}}_k^\top \mathbf{D}_1 \mathbf{E}_{21}) \mathbf{D}_2 [c_k \ s_k]^\top \\ a_{t22} &= \text{trace}(\mathbf{D}_2 \mathbf{E}_{22} \mathbf{D}_2 \mathbf{E}_{11}) + \text{trace}(\mathbf{D}_2 \mathbf{E}_{12} \mathbf{D}_2 \mathbf{E}_{12}) \\ b_{t22} &= ([c_k \ s_k] \mathbf{D}_2 \mathbf{E}_{12} + \tilde{\mathbf{x}}_k^\top \mathbf{D}_2 \mathbf{E}_{22}) \mathbf{D}_2 \tilde{\mathbf{x}}_k + ([c_k \ s_k] \mathbf{D}_2 \mathbf{E}_{11} + \tilde{\mathbf{x}}_k^\top \mathbf{D}_2 \mathbf{E}_{21}) \mathbf{D}_2 [c_k \ s_k]^\top . \end{aligned}$$

Accuracy requirements for the translation  $\mathbf{t}$  are expressed by inserting the bounding variances  $\sigma_{t11,b}^2$ ,  $\sigma_{t22,b}^2$  and  $\sigma_{t12,b}^2$ . The parameter  $\xi$  is then obtained for the given requirement by the solution of the quadratic equation. This formulation has the advantage over the previous model formulation that the sensor parameter  $\xi$  can be directly calculated for a given localization requirement by solving the set of quadratic equations.

#### 9.4.1 Landmark detection accuracy

Once a decision about the landmark type for localization and the sensor setup has been taken, it can be of interest to determine the sensor parameters which lead to the goal localization accuracy with a sufficiently high probability. For the derivation of pole feature detection accuracy requirements, a worst case landmark configuration is used as reference. Feasible configurations depend on the utilized feature type and other design aspects like a potentially required recovery capability in intersection areas or safe operation modes in case of system faults. For this experiment, a configuration of three proximate pole landmarks is considered as it is often encountered in narrow intersections (Figure 9.13). The low distances between sensor and landmarks are particularly challenging for rotation estimation and are therefore considered part of a possible worst case scenario. The vehicle shall be able to localize from a single map matching step at every possible pose in the whole intersection area with an accuracy of 0.25 m and  $0.5^\circ$ . For brevity, it is assumed that all landmarks from the reference configuration are detected by the perception system. The experimental results are given in Figure 9.13.

A maximum landmark detection error of 0.05 m is derived from the angular matching accuracy requirement (Figure 9.13, left) which is considerably higher than for the translational component (Figure 9.13, right). The static error for pole detection during vehicle standstill was determined as 0.025 m in Section 7.2.1 and consequently the requirement is met for a standing vehicle. The difference of 0.025 m corresponds



**Figure 9.13** The minimum pole detection accuracy of  $\sigma_1^{req,min} = 0.04$  m for the considered landmark arrangement is driven by the vehicle orientation estimation accuracy requirement of  $0.23^\circ$  (right). Sole consideration of translational localization accuracy requirements of  $0.25$  m result in relaxed requirements (left).

to the maximally allowed error in motion compensation of the LiDAR scan. According to Equation 7.12, the velocity estimation error has to be below  $0.45$  km/h for an angular motion estimation error of  $0.12^\circ$ . For a velocity of  $50$  km/h, this requirement does approximately corresponds to a maximal odometry drift of  $0.9\%$ . The calculated parameters depend on the considered worst case landmark arrangement and requirements for vehicle localization. To this point, the results might be overly pessimistic, especially if scan matching and odometry measurements are available as additional sources of vehicle pose information. Nonetheless, the high impact of errors in LiDAR scan motion compensation on the localization accuracy remains significant. Consequently, the landmark detection accuracy requirements support the results from Section 9.2.2. These results will be considered by the design of an accurate scan-based odometry procedure in Section 10.1.

#### 9.4.2 Intrinsic sensor parameters for robust localization

A more microscopic view on perception requirements than provided in the last section is elaborated in the following. Intrinsic sensor parameters are calculated for the case of partially occluded FoV and missed landmark detections. As a basis for the subsequent contents, a strategy for consideration of perception errors and degradation was discussed in Section 8.3. The results from this section support the design of perception systems which are robust against adverse environment conditions and perception errors. Sensor parameters are derived for a stereo video camera that is modeled as described in Section 7.3. Furthermore, it is assumed that the vehicle moves in a straight line between dashed lane boundary markings similar to the lines depicted in Figure 7.2. For map-relative localization, the starting and ending points

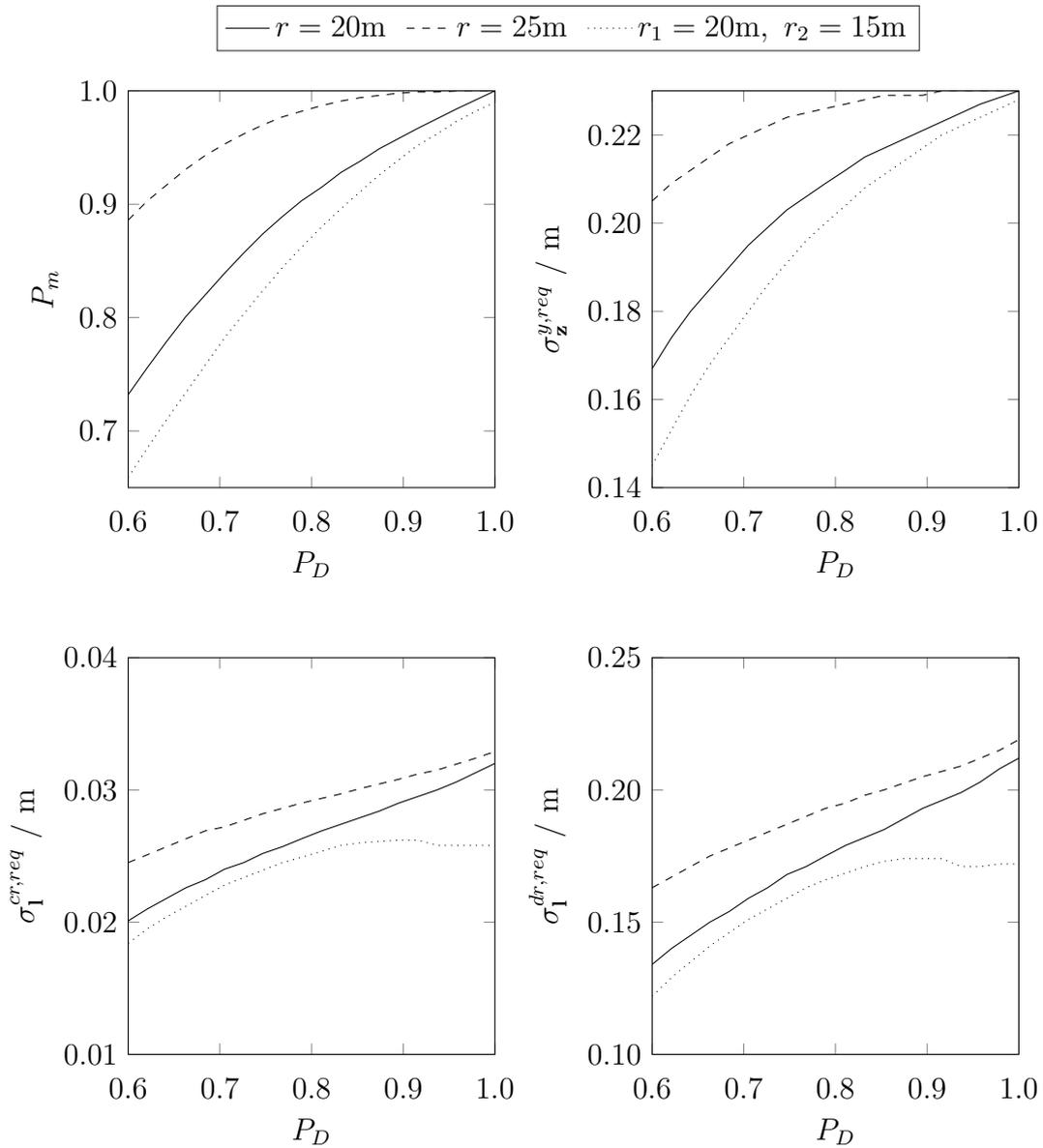
of the line segments are matched with a suitable localization map. Thereby, a simplified version of the popular lane marking based localization system from [SKF13] is simulated. Missed detection lead to a detection probability  $P_d < 1$  and map matching failures as caused by missing feature detections to  $P_m < 1$ . Occlusions of parts of the sensor FoV are represented by an adaptation of the FoV during evaluation.

For the simulation study, an upper bound for motion estimation covariance  $\mathbf{Q}_u \geq \mathbf{Q}_k$ ,  $\forall k$  is defined as a diagonal matrix with the entries  $\sigma_{\mathbf{u},x} = 0.361\text{m}$ ,  $\sigma_{\mathbf{u},y} = 0.01\text{m}$  and  $\sigma_{\mathbf{u},\theta} = 0.1^\circ$ . The values of  $\sigma_{\mathbf{u},x}$  and  $\sigma_{\mathbf{u},y}$  were calculated based on the assumption of a matching rate of 10Hz, a velocity of 130km/h and an error of 1% of the driven distance. The localization requirement is set to 0.2 m orthogonal to the driving direction, hereafter denoted as  $x$ -direction. The uncertainty due to camera calibration is set to  $\sigma_d = 0.25$  and  $\sigma_u = 0.50$ . The models in Section 7 and Section 8 are evaluated by execution of the following steps:

1. Feature distributions are modeled on the basis of a deterministic model for the size of and distance between the road markings.
2. Different sensor FoVs are evaluated in combination with the assumed sensor-relative feature distribution. Occlusions are modeled by splitting the FoV in half and assuming an individual range for each half ( $r_1, r_2$  in Figure 9.14). The result is a FoV dependent map matching probability  $P_m$  (Figure 9.14, upper left).
3. The required map matching performance  $\Sigma_{\mathbf{z},max}$  is derived for a given  $P_m$ , motion estimation covariance  $\mathbf{Q}_u$  and  $\mathbf{A}$  ( $y$ -component shown in Figure 9.14, upper right).
4. The 95% percentile of the sensor measurement covariance scaling factor (Section 7.3)  $\xi_{95}$  is calculated for  $\Sigma_{\mathbf{z}}^{req}$ . From  $\xi_{95}$ , suitable values for the base width  $b_w$  and the camera constant  $c$  can be derived. For brevity and as a descriptive example,  $\xi_{95}$  is inserted in Equation 7.15 and landmark detection accuracies are calculated in the subsequent step.
5. The maximum measurement uncertainty  $\sigma_1^{cr,req}$  (cross range: Figure 9.14, lower left) and  $\sigma_1^{dr,req}$  (down range: Figure 9.14, lower right) are obtained from  $\xi_{95}$  for a reference point at (10 m, 1.5 m) in sensor coordinates.

Varying detection rates lead to distinct map matching accuracy requirements  $\sigma_{\mathbf{z}}^{y,req}$ . Again, the FoV has a significant influence on the resulting accuracy requirements. This can be seen from the lower maximum standard deviation in map matching. The effect of  $P_D$  on the feature detection accuracy requirements  $\sigma_1^{dr,req}$  in direction towards the sensed object and  $\sigma_1^{cr,req}$  orthogonal to this direction depends on the sensor FoV. From the results it can be seen that the detection range has a significant influence on the arising perception requirements. For decreasing  $r$ , the probability of detecting a sufficient number of landmarks for map matching is decreased as well. Consequently, the assurance against partial occlusions of the sensor

field of view leads to significantly increased parameter requirements. According to the simulation results, the detection rate decrease of 10% has small impact on the accuracy requirements.



**Figure 9.14** The sensor FoV has significant influence on the sensitivity of the map matching rate  $\lambda$  against the feature detection probability  $P_D$ . In this simulation run, the opening angle was set to  $40^\circ$ . Occlusions cause a decreased field of view and consequently lead to increased feature detection probability requirements.

## 9.5 Conclusion

This part was concerned with the derivation of statistical models of localization systems and operation environments. The contents from this section demonstrated how these models can enable the quantification of expert knowledge and a system design starting from localization requirements. Hence, the research question 2 (*rq2*, Section 1) could be answered.

For the example of a right turn scenario at an intersection, a localization accuracy requirement of 0.25 m was derived. In order to achieve this target accuracy, a manifold of localization system parameters have to be designed. The following significantly important design parameters for localization systems were identified:

- map matching accuracy,
  - data discretization (e.g. grid resolution for spectral registration)
  - feature detection accuracy
  - feature detection rate
  - matching algorithm selection (e.g. selection of cost functions)
- map matching rate,
- odometry accuracy and
- sensor parameters.
  - field of view

The discussion of angular odometry accuracy underlined its high impact on the localization accuracy. Consequently, a decrease of angular errors in scan-based odometry is fostered in Section 10.1. Additionally, research question 6 (*rq6*, Section 1) of how to add an efficient scan-based odometry to the localization system will be emphasized.

It was shown, that the impact of the average matching rate  $P_m$  depends on the matching failure distribution. Clusters of matching failures have to be compensated by odometry measurement and can drastically increase the respective accuracy requirements. On the contrary, equally distributed matching failures have a less significant impact. This observation underlined the advantages of hybrid localization approaches for vehicle pose estimation in urban environments (Section 11.2). In heterogeneous operation environments, diverse localization modalities lead to an increased probability of achieving the required matching rate and matching failures are more likely to be approximately uniformly distributed. Nonetheless, computational costs increase with the number of localization modalities. In order to decrease the number of required modalities, map matching algorithms with high tolerance against stressful effects commonly encountered in urban environments might be used. Due to the increase in robustness, the rate of matching failures can be reduced, yielding higher matching rates for each localization modality. Obviously, this is a trade-off

between robust matching algorithms with potentially less efficient runtime behavior than the non-robust counterparts and the overall number of required localization modalities. To this point, the combination of spectral and pole-based matching will be used in Section 11.

The required matching rate  $P_m$  is correlated with the accuracy of relative pose estimation. Consequently, there is another trade-off between accuracy requirements for map matching and odometry measurements. From a system design perspective, it is proposed to choose the odometry accuracy based on a threshold for maximally tolerable distances without availability of global pose updates. Alternatively, the requirement from scan motion compensation can be utilized. The proposed model was utilized to derive map matching and odometry parameters. One example is the selection of a suitable spacial discretization of the grid map used for spectral map matching. Given a suitable pre-processing of the sensor observations, the accuracy of the LiDAR-based map matching is limited by the grid resolution. By adaptation of the grid resolution to this requirement, the computational costs and storage requirements can be improved.

In conclusion and as an answer to research question 2 (*rq2*, Section 1), statistical models can provide valuable insights into the design of localization systems. Complex interrelations between design parameters can be captured, quantified and utilized during localization system design.

**Part IV**  
**Vehicle localization with enhanced  
robustness for UAD**



## 10 LiDAR-based odometry and SLAM

An accurate and robust estimation of the vehicle motion is a fundamental component of localization and mapping systems. This has been underlined by the model-based studies conducted in Section 9 as well. Beside the often utilized wheel odometry, light detection and ranging (LiDAR) scan based odometry can contribute to a gain in accuracy, redundancy and for omitting performance degradation through wheel slippage. To this point, a scan-based odometry is developed in Section 10.1 which makes use of the spectral registration procedure from Section 6.3. In addition to the compensation of missing absolute pose estimates, the localization system has to be able to operate in unmapped areas. Therefore, simultaneous localization and mapping (SLAM) is used in order to extend the operation environment of the vehicle system. The latter is relevant in areas where potentially no maps exist, e.g. private ground including parking lots, and the realization of safety strategies, e.g. in case of map errors or communication failures. The SLAM framework makes use of the spectral registration algorithm for the introduction of relative motion and loop closure constraints and is described in Section 10.2. The utilization of the spectral registration algorithm for scan-based odometry and SLAM leads to an overall system implementation consisting of a low number of submodules as compared to most algorithms from the state of the art where specialized algorithms are required for different tasks. In combination with an efficient hardware implementation of the spectral registration procedure, a overall localization and mapping framework with high efficiency and comparatively low architectural complexity is realizable.

### 10.1 Scan-based odometry

The role of accurate relative motion estimation in vehicle localization was underlined in Section 9. Sharp turns, wheel slippage and low velocities are frequently encountered during operation of an automated vehicle system in urban environments and pose a challenge to wheel odometry. In such situations, relative motion estimation based on LiDAR scans or alternative sensor measurements has the advantage of being insensitive to these influences. In order to account for potentially imprecise wheel odometry, a scan-based odometry is developed in this section. The scan-based odometry makes use of the spectral matching procedure introduced in Section 6 and uses LiDAR scans as only inputs.

The most basic method for relative motion estimation is the matching of consecutive scans for every time step  $k$ . This approach suffers from inaccurate results in comparison to more sophisticated algorithms and high sensitivity to punctual scan matching failures. The introduction of local loop closure constraints between scans of a scan history

$$\mathbf{S}_k = \{\mathbf{s}_{k-j}\}_{j=0}^S \quad (10.1)$$

leads to increased estimation accuracy and enhanced robust against partial occlusions and weakly structured environments. By accumulation of LiDAR scan se-

quences to a more detailed local map, negative effects caused by inaccurate motion compensation of the scans and errors in scan registration between two time steps can be reduced. To this point, every new scan is matched against  $\mathbf{S}_k$  instead of a single scan. In the following, a scan-odometry is developed which makes use of the distinct characteristics of spectral registration. An important advantage of correlation-based procedures like spectral registration is the independence of the matching error from the transformation prior. Furthermore, the transformation errors are typically in the order of magnitude of the grid discretization of the input scans. This is of special interest for the rotational transformation component which leads to a fast accumulation of pose estimation errors as the trajectory length grows. Based on these observations, it is argued that the registration of a current scan  $\mathbf{s}_k$  with a scan  $\mathbf{s}_{k-1-h}$  in distance  $-(h+1)$  leads to more accurate motion estimates as  $h$  is increased. For a gain in efficiency, the matching of scan pairs  $\{\mathbf{s}_k, \mathbf{s}_{k-1-h}\}$  or  $\{\mathbf{s}_k, \mathbf{S}_{k-1-h}\}$  with high offset can be conducted. Thereby, the highest information gain to computational cost ratio is achieved. Different parameterizations of the odometry framework and for the spectral registration algorithm are experimentally evaluated in the following section.

### 10.1.1 Experimental evaluation

The scan-based odometry is evaluated on the basis of the urban data sets presented in Section 4.3. Due to the special challenges (Table 10.1) for scan-based odometry, data sets 1018-11 and 0816-31 are selected for a detailed evaluation and comparison of different odometry parameterizations. 1018-11 is the longest urban data set and is used in order to test the sensitivity to broad roads and the mixture of high and low curvatures. 0816-31 is an interesting data set for the evaluation due to the high number of sharp turns.

Evaluation is conducted relative to the reference vehicle poses which were determined using the method presented in Section 5.3. The Euclidean distance between endpoints of the reference and odometry trajectory is evaluated and averaged over the whole trajectory. This commonly used evaluation method gives an approximate impression of the odometry accuracy and the differences between the respective implementations. This evaluation is only approximate since the matching errors along the driven trajectory can cancel each other out which is not captured by this evaluation strategy. Nonetheless, it enables a comprehensive analysis of the strengths and weaknesses of the different odometry parameterizations and utilized sensors sets.

Figure 10.1 contains an overview over odometry results for different choices of the number  $S$  of accumulated scans, the maximum detection range  $r$ , grid resolution  $\Delta_g$  and the distance  $h+1$  to the scan from which the matching result with the highest weighting was obtained. Quantitative results for the two considered data sets are stated in Figure 10.2. From the results of the first part of data set 1018-11 in Figure 10.1, it can be observed that a lower position drift  $\delta$  is achieved by matching against accumulated scan data. For 1018-11, the position drift decreased from 0.832% to 0.374% for otherwise unchanged parameters. At some sections of

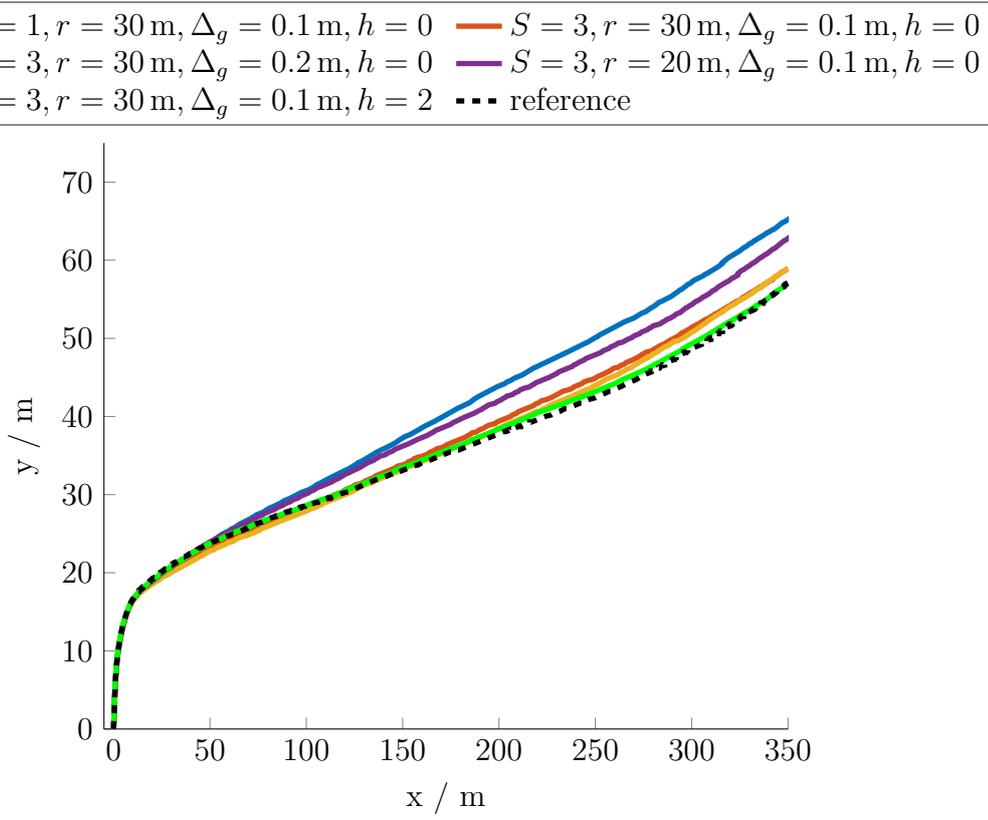
**Table 10.1** Overview over the data sets (Section 4.3) used for the experimental evaluation and their characteristics.

Data set	Characteristics
1018-11	<ul style="list-style-type: none"> <li>• roundabout</li> <li>• broad main roads</li> <li>• mixture of high and low curvatures</li> </ul>
0316-31	<ul style="list-style-type: none"> <li>• most turns</li> </ul>

the trajectory, scan accumulation can compensate a reduced scanner range and grid resolution.

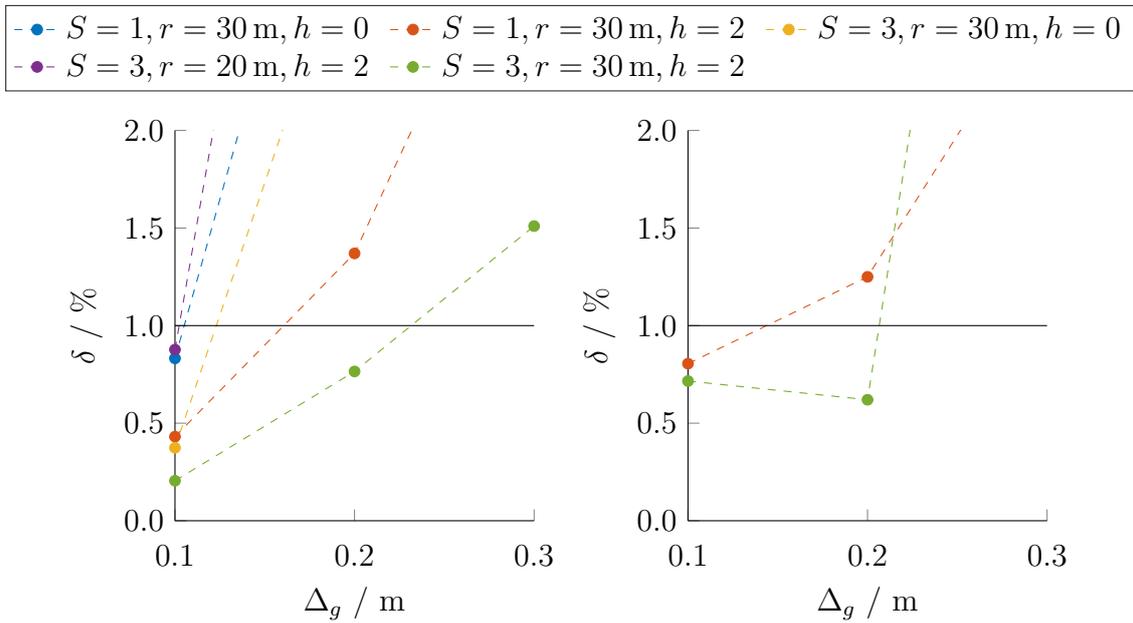
By utilization of accumulated scans for matching, the drift for a decreased scanner range of 20m was reduced from well above 2% to 0.877% which is similar to the drift for  $r = 30$  m and no accumulation. However, such parameterizations are prone to drastically increased local errors in intersection and other expansive areas. In the case of data set 1018-11, a decreased grid resolution and scanner range has no negative impact, whereas the intersections in the remaining part of the trajectory are not well handled and a significant increase in orientation drift occurs. For data set 0316-31, the drift for parameter sets with  $\Delta_g > 0.1$  m was always well above 2% which is most likely due to the high number of turns. Discretization of the input scans leads to increased position drifts for road curvatures, as seen in the right part of the trajectory in Figure 10.1. This effect is compensated by an increase in  $h$ . While the impact is low for sharp turns like at the beginning of the trajectory, mentionable differences occur for small curvatures as encountered in the second half of the shown trajectory. Without the utilization of local loop closures, small curvatures would lead to a fast growth in pose estimation uncertainty, because the estimated rotation is biased against zero. Consequently, the best setting which gives the most accurate and reliable motion estimates is given for high detection ranges and grid resolution, matching against accumulated scan data and the introduction of local loop closures. This parameterization led to a drift of 0.21% for data set 1018-11 and 0.72% for 0316-31. A comparison between wheel and the scan-based odometry is shown in Figure 10.3. The results indicate, that scan-based odometry can lead to increased motion estimation accuracy in the presence of frequently occurring sharp turns.

A Velodyne scan from the considered data set is shown in Figure 10.4 together with an accumulated scan from the Ibeo sensor setup with six sensors. Based on both data sources, the odometry data are calculated for data set 1018-11. The respective best results are shown in Figure 10.5. The performance of the Ibeo is comparable to the Velodyne-based odometry with a detection range of 20 m or with a decreased grid resolution between 0.2 m and 0.3 m. One reason for this observation is the higher point density of the Velodyne sensor. Furthermore, the Velodyne is

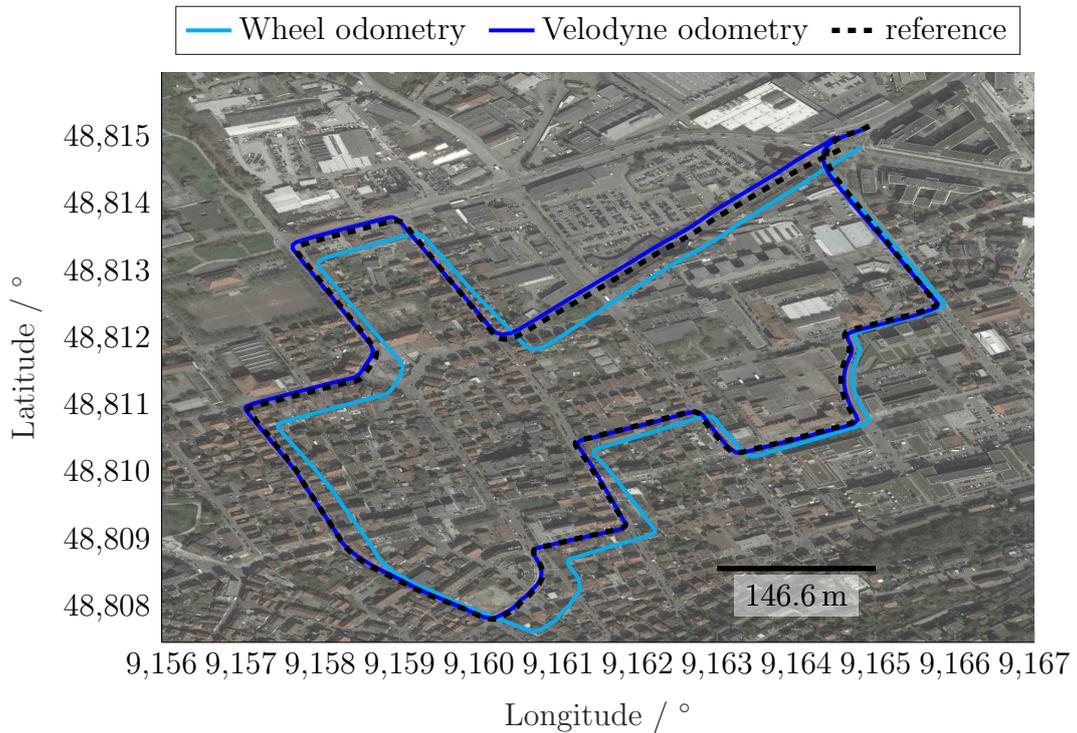


**Figure 10.1** Selected odometry results for data set 1018-11. The accumulation of scans leads to a reduced error in angle estimation. Negative effects of discretization for small changes in angle can be significantly reduced by the introduction of constraints between non adjacent scans ( $h > 0$ , local loop closures).

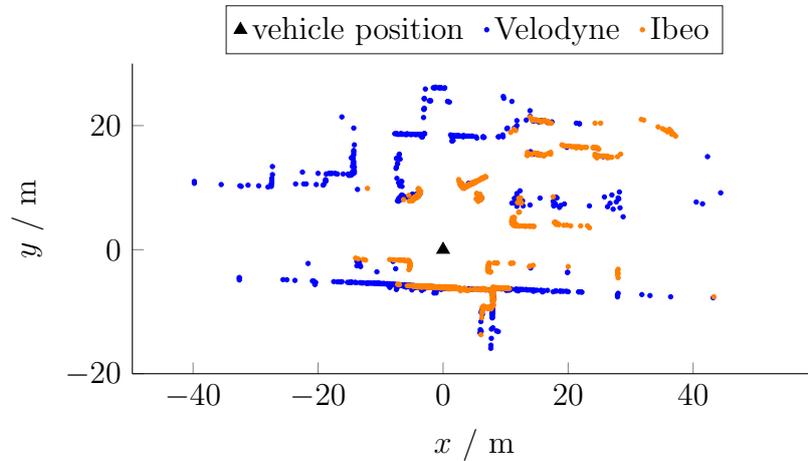
mounted at the top of the vehicle and, therefore, occlusions have smaller impact on the maximum detection range. The latter observation is significant, hence the studies in Section 9.3 on the influence of landmark configurations on the matching accuracy become relevant in this case. The increased insensitivity of the maximum detection range due to the high mounting position of the Velodyne sensor to occlusions by other traffic participants contributes to the matching accuracy. With a suitable parameterization, good odometry performance can be achieved for urban environments with frequently occurring sharp turns (Figure 10.3). In summary, the higher number of point measurements and increased insensitivity of the maximum detection range due to the high mounting position of the Velodyne sensor to occlusions by other traffic participants contributes to the matching accuracy.



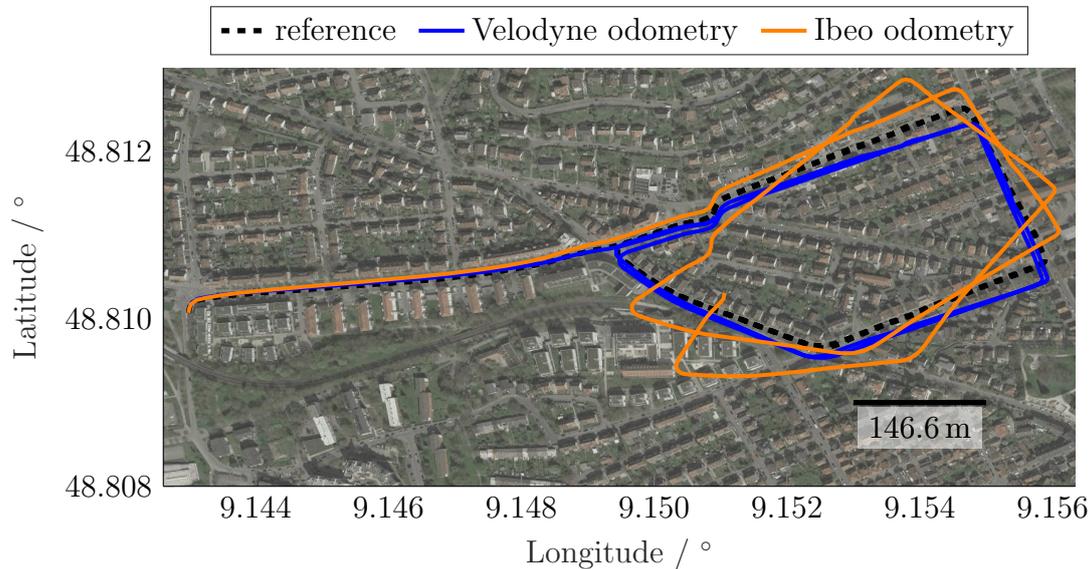
**Figure 10.2** Local loop closures and utilization of accumulated scans for motion estimation result in the highest estimation accuracy (left: 1018-11; right: 0316-31).



**Figure 10.3** The results for data set 0316-31 underline the good performance of scan-based odometry in the presence of high numbers of sharp turns in comparison to wheel odometry. (source of satellite image: Google, Imagery ©2017, DigitalGlobe, GeoBasis-DE/BKG, GeoContent, Landeshauptstadt Stuttgart)



**Figure 10.4** Due to the low mounting position, the Ibeo sensor setup suffers from occlusions caused by near-distant objects. Consequently, the likelihood of high-range measurements which significantly contribute to the rotation estimation accuracy (compare Figure 9.10) is decreased.



**Figure 10.5** The best results obtained from the Velodyne- and Ibeo-based odometry for the 3.01 km long trajectory in data set 1018-11. (source of satellite image: Google, Imagery ©2017, DigitalGlobe, GeoBasis-DE/BKG, GeoContent, Landeshauptstadt Stuttgart)

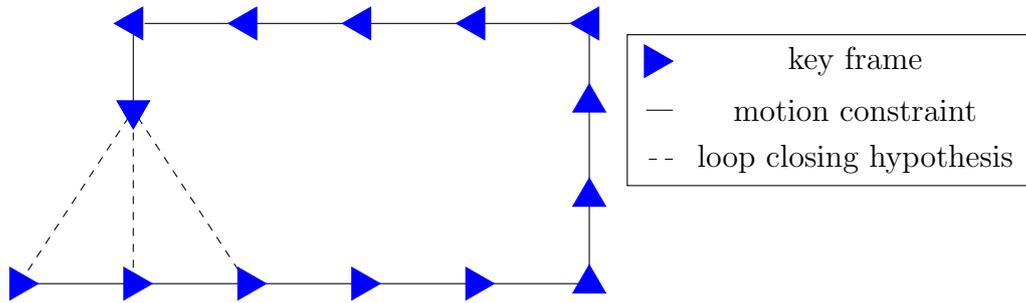
## 10.2 Spectral SLAM

Simultaneous localization and mapping (SLAM) enables the operation of mobile robots in previously unknown environments. In the context of urban automated driving (UAD) such cases might likely occur on private ground such as parking garages. Therefore, a SLAM framework on the basis of the spectral registration algorithm from Section 6 is introduced in this section. The proposed SLAM framework uses an ordered sequence of 2D LiDAR scans as input and does not require odometry or other motion sensors. Optimization is conducted based on the optimization framework provided in [KGS<sup>+</sup>11]. A robust kernel is used for graph optimization.

Place recognition is of crucial importance in mapping applications as it enables the recognition of previously visited places. Thereby, it allows to establish transformation constraints between current and previous observations. In the context of simultaneous localization and mapping, this task is denoted loop closing. Without loop closing, the accuracy of the absolute vehicle pose estimate would decrease without a bound and the map would become globally inconsistent. With increasing number of accurate and correct loop closures, the map and pose estimate becomes more accurate and less prone to drift.

Especially for large loops, loop closing in real-time remains a challenging task. This is mainly due to the accumulation of errors in relative motion estimation. In this section, a real-time loop closing strategy based on spectral matching is introduced. Due to the broad basin of convergence of the matching algorithm, loop closing hypotheses can be generated from inaccurate priors. To this point, the SLAM graph is in every time step checked for key frames within a predefined search radius which might have an overlap with the current observation. Two poses  $\mathbf{x}_k$  and  $\mathbf{x}_a$  with  $a \leq k$  are considered neighbors if their Euclidean distance  $d_{k,a}$  is smaller than the threshold  $d_{max}$ . For each key frame in this neighborhood, a transformation is calculated based on the spectral matching procedure as shown in Figure 10.6. The broad basin of convergence of spectral registration leads to reliable loop closing performance. Furthermore, the robust spectral registration procedure, developed in Section 6.3 is utilized. Due to the potentially catastrophic impact of false loop closures on the mapping process, a consistency check is performed for this set of hypotheses. Therefore, the consistency requirement is expressed by a threshold on the Euclidean distance between the translation hypotheses. The set of consistent constraints is given as the subset with the highest number of members which all fulfill the consistency requirement. By requiring at least 3 consistent loop closing constraints, the likelihood of false loop closure constraints is decreased. The consistent constraints are then averaged and a constraint is added to the SLAM graph. For the reduction of uncertainty accumulation in the vehicle pose estimate, local loop closure constraints between consecutive key frames are added as well.

Throughout the last decade the update of localization maps has found increasing interest throughout the research community. Although not being part of this thesis, it shall be mentioned here, that the broad basin of convergence of spectral matching algorithm can also be incorporated into map update procedures. Interna-



**Figure 10.6** Spectral registration is used for loop closure hypotheses generation. Candidate key frames are matched and the results checked for consistency. Due to the broad basin of convergence of the spectral registration procedure, the transformation prior can be inaccurate.

tional standardization documents as ISO 17572-3:2015-01 define a map-independent geo-referencing approach for sensor measurements, e.g. based on consumer grade GPS measurements, which can then be used for map updating. One advantage of this approach is the avoidance of closed loops in the data processing chain. More specifically, the update of a potentially corrupted localization map shall not depend on the map itself and the utilized localization system.

### 10.2.1 Experimental evaluation

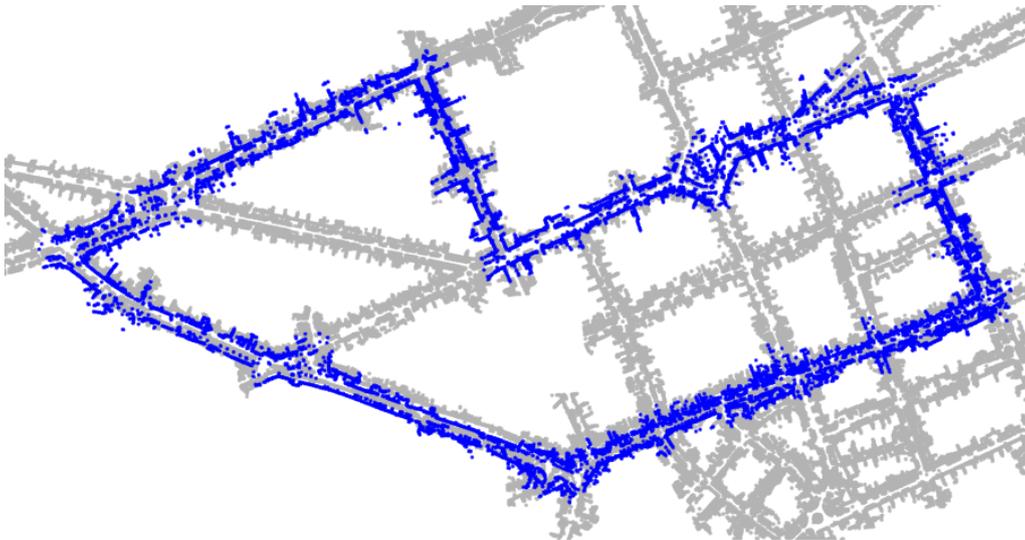
The evaluation of the spectral SLAM approach is based on data sets 0823-71 and 0316-61. Data set 0823-71 is characterized by a long loop and a challenging loop closure in an area with sparse structures and was chosen as a test of the drift per driven distance and the loop closing performance for large loops. 0316-61 contains the most loop closures and was chosen to test the loop closing reliability. Relative motion constraints are solely based on scan matching results obtained from spectral registration. Consequently, wheel odometry is exclusively used for LiDAR scan deskewing. Spectral registration is also successfully used for loop closing. The result of the mapping run is compared to the localization map described in Section 5.2. Since, this map was recorded using a highly accurate IMU sensor and LiDAR sensors and contains several loop closures, it is considered accurate and suitable as a reference for the evaluation of the developed SLAM framework.

The results (Figure 10.7) show that the map which was created from the single test run are similar to the prerecorded localization map. It can be seen, that significant changes between the two maps occurred. The high amount of dynamic objects in the measurement data did not lead to mapping failures. Loop closing was successfully established in a challenging intersection area from an initial offset of 4.57 m after driving the entire loop as shown in Figure 10.8. This corresponds to a drift of 0.18% of the driven distance. The mapping is executed at a real time factor of 1.43 running on robot operating system (ROS) under Ubuntu. A computer with an Intel i5-4690 ( $4 \times 3.5$  GHz) central processing unit (CPU) and 32 GB RAM was used for

the calculations.

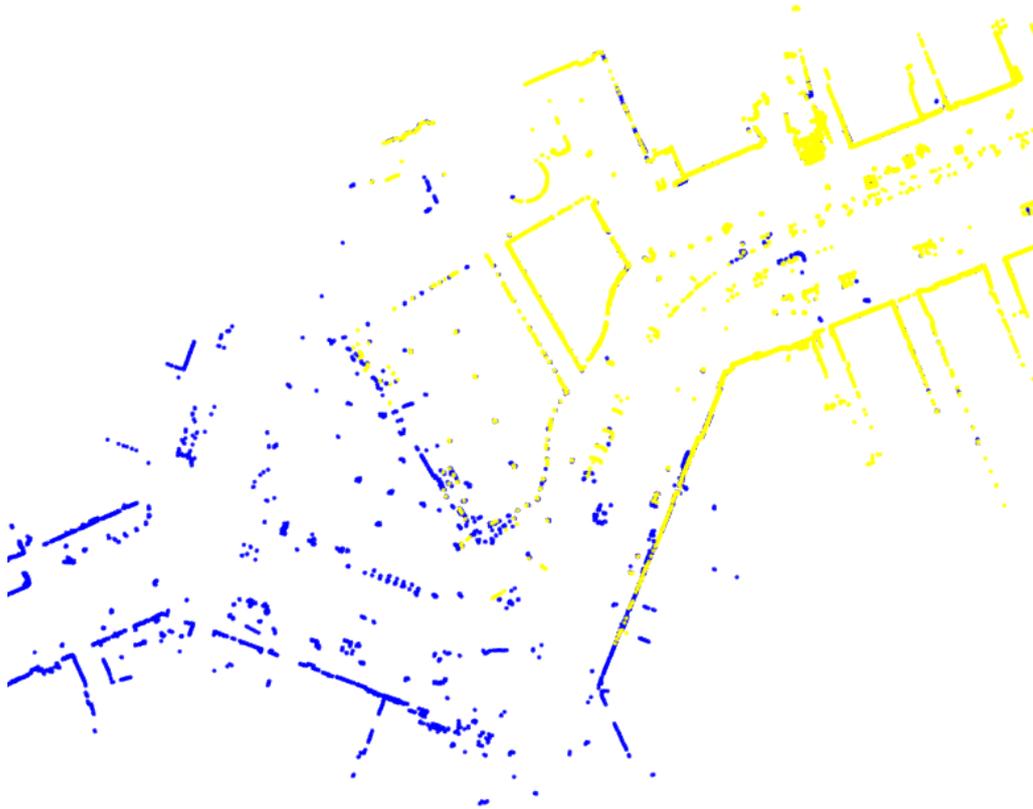
**Table 10.3** Overview over the data sets (Section 4.3) used for experimental evaluation and their characteristics.

Data set	Characteristics
0823-71	<ul style="list-style-type: none"> <li>• loop closure in expansive intersection area</li> <li>• roundabout</li> <li>• long loop</li> </ul>
0316-61	<ul style="list-style-type: none"> <li>• most urban loop closures</li> </ul>
campus-4	<ul style="list-style-type: none"> <li>• most loop closures</li> <li>• longest trajectory</li> <li>• high heterogeneity (vegetation, free space, building canyons)</li> </ul>



**Figure 10.7** The map built from a single test drive (data set 0823-71) and with FMT-SLAM (blue) plotted over a three year old version of a highly accurate localization map (light gray). All constraints in the SLAM graph were created by spectral registration.

The accuracy of the resulting map is measured in two different way. Firstly, the estimated trajectory is aligned with the reference trajectory for the considered data set (Section 5.3). Since, a reference pose for each key frame in the SLAM graph exists and the correspondences are known, the transformation between both trajectories can be directly calculated by solving the orthogonal Procrustes problem (OPP).

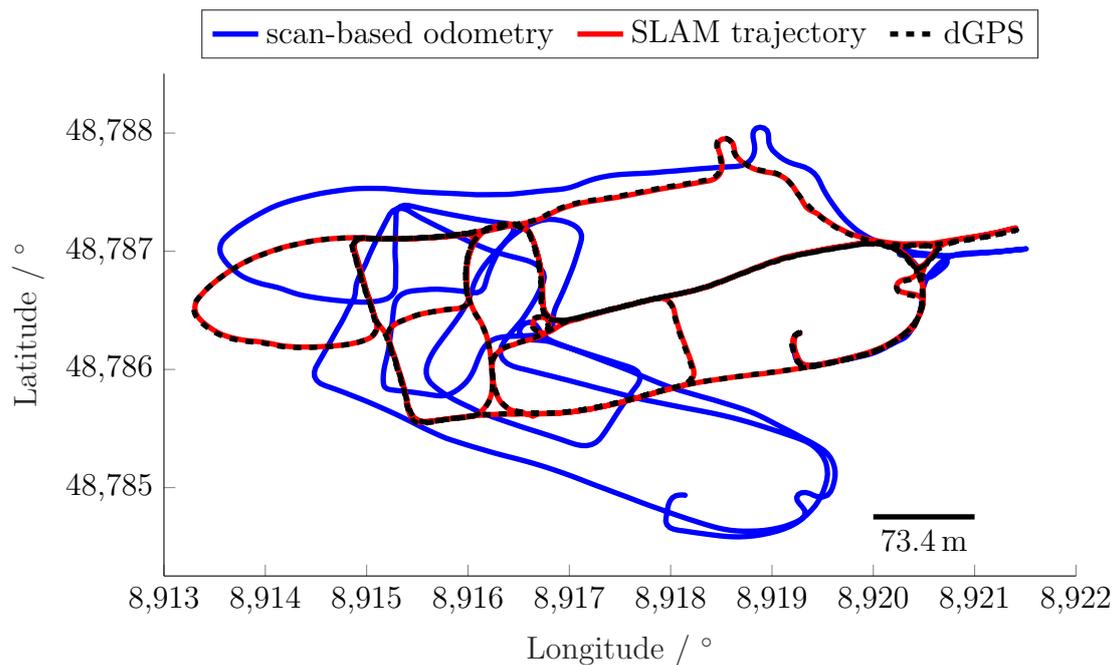


**Figure 10.8** A detailed view of the challenging loop closing area contained in data set 0823-71 (blue: start of counterclockwise test drive). Loop closing was achieved on the basis of spectral registration from an initial offset of 4.57 m that corresponds to an average position drift of 0.18% over the trajectory.

The localization map is considered more precise than the newly recorded one due to highly accurate inertial measurement unit (IMU) measurements which were used for recording and the high amount of global loop closures. Consequently, the accuracy of the mapping run can be accurately estimated as the standard deviation over all Euclidean distances between corresponding position pairs after alignment. For data set 0823-71, the standard deviation is 0.87 m and 0.62 m for 0316-61. Secondly, the estimated trajectory is manually aligned with the collected differential global positioning system (dGPS) measurements. Then, a robust iterative closest point (ICP) variant is used for refinement of the alignment and the resulting mean squared error is used as measure for mapping accuracy. This approach is independent of the localization map and enables a comparison to the results from the campus data set for which no prerecorded map exists. The error for data set 0823-71 is 0.98 m and 0.81 m for 0316-61. In both cases, the results show that loop closures led to an increased mapping accuracy, showing the conformity of the evaluation results.



**Figure 10.9** The calculated map of the campus is in agreement with the satellite image. The alignment with the WGS48 coordinate frame was conducted in a post-processing step. (source of satellite image: Google, Imagery ©2017, DigitalGlobe, GeoBasis-DE/BKG, GeoContent)



**Figure 10.10** The scan-odometry prior is calculated with a scanner range of 30 m. Areas with sparse structures and curvatures lead to an increased drift in the estimated vehicle pose in comparison to the urban data sets. Nonetheless, all loop closures were detected correctly. Due to the high number of loop closures, the optimized trajectory is highly accurate.

The Campus environment contains significantly more loop closings than the urban data set and more heterogeneities. Special challenges are expansive free spaces in the middle of the campus and high amounts of vegetation. The length of the driven trajectory is 3.7 km (Campus-4, Section 4.3). For this mapping run, the scanner range for the scan-odometry measurements is set to 40 m in order to deal with expansive free spaces. Despite the large field of view, the scan-odometry has a higher drift as for the previously considered urban data set which is due to the combination of sparse structures and high curvatures. Loop closures and scan matching constraints are again calculated on the basis of the spectral matching procedure. After several loop closures, the estimated trajectory is consistent with the dGPS signal as shown in Figure 10.10. The resulting map of the campus is shown in Figure 10.9 and plotted over a satellite image of the area. The map contains the main structures from the campus environment. The accuracy is again measured relative to the dGPS trajectory. The standard deviation is 0.34 m and hence more accurate than for both urban data sets which is likely due to the high amounts of loop closures. In all cases, the mapping accuracy is high enough for the operation in areas like parking garages, where no high localization accuracy requirements are given.

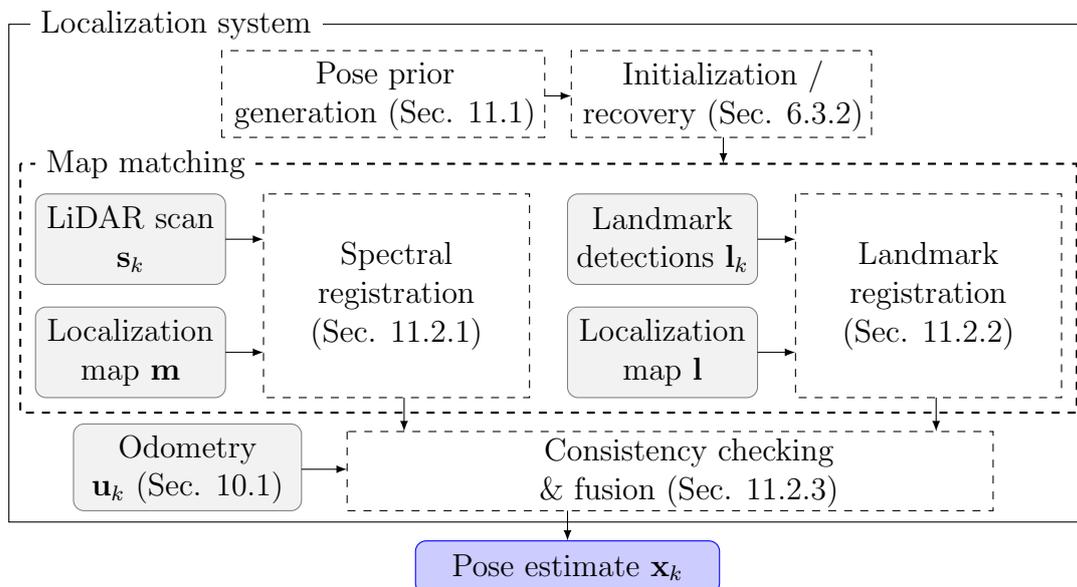
### 10.3 Conclusion

In this section, a scan-based odometry and SLAM framework on the basis of spectral registration was developed. Different parameterizations of the matching algorithm were used to adapt it to the operation environment and to the specific application. Experimental evaluation was conducted with a selection of challenging data sets from urban environments. Problems arising from the discretization of the input scans were addressed in the odometry system by accumulation of scan data and introduction of local loop closures. For the Velodyne LiDAR sensor, an odometry position drift of 0.21% could be reached which is sufficiently accurate in the context of vehicle localization (compare Section 9). The Ibeo-based solution suffered from low measurement ranges due to occlusions which led to increased angular errors. As outlined in Section 9, angular uncertainty has significant influence on the localization performance and shall be kept at a low value. For a data set with a high number of sharp turns, the scan-based odometry outperformed a wheel odometry from wheel speed and steering angle sensor measurements. In the SLAM framework, relative motion constraints were exclusively based on scan matching and hence did not require proprioceptive sensors except for scan deskewing. Furthermore, robust spectral registration was utilized for the calculation of loop closing hypotheses. Thereby, real-time loop closing with deterministic outcomes, i.e. no randomization and heuristics are introduced to the system, was enabled. The experimental evaluation with data sets from urban and campus environments showed the reliable loop closing performance even in challenging intersection areas and high initial transformation offsets after long loops. The accuracy of the mapping result was found to be sufficiently high for operation in parking garages or comparable operation environments.

## 11 LiDAR-based vehicle localization

The fundamental importance of map-relative vehicle localization has been outlined in this thesis, starting from the introduction and underlined by the respective definitions as well as the context provided in Section 2.3. The utilization of suitable subsystems, like the spectral registration approach (Section 6) and accurate odometry as well as SLAM (Section 10) contributes to the achievement of this goal. Model-based design on the basis of newly developed statistical models (Section 9) was introduced as an additional component for increasing the localization system robustness by a quantification of expert knowledge.

In this section, previous results are combined for the development of a framework for map matching based vehicle localization with enhanced robustness (Figure 11.1). To this point, spectral registration (Section 6) is utilized for the alignment of LiDAR scans and outdated localization maps (Section 5) as a first source of absolute vehicle pose estimates. In order to increase the matching failure detection performance and decrease the probability of prolonged periods without valid matching results, analytical redundancy is introduced. This is approached by adding a pole-based map matching modality for the alignment of detected pole-features (Section 5.1.2) and the corresponding localization map. The sparsity of pole features as compared to other feature types like Harris corners and the relinquishment of elaborated descriptors, yield a resource efficient vehicle pose estimation modality. Furthermore, pole-based map matching is shown to perform well in intersections and parking lots, where spectral registration might suffer from sparse structures (Section 6.4).



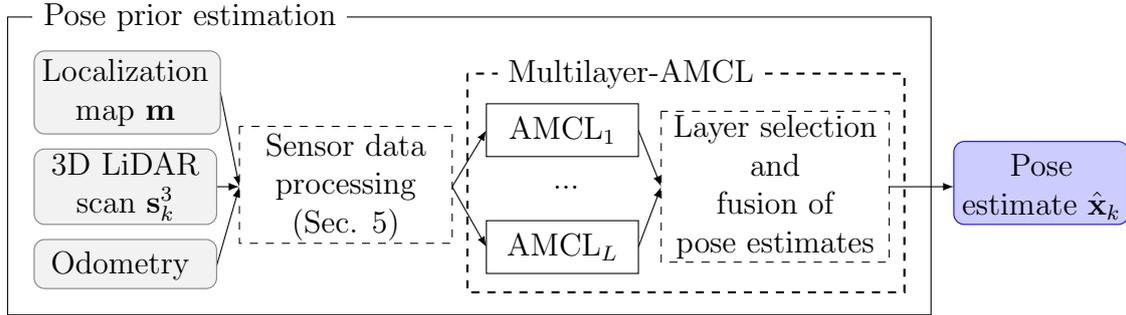
**Figure 11.1** The system architecture of the proposed localization system on the basis of scan- and pole-based map matching. Map matching failures are detected in a consistency checking step.

Due to the orthogonality, the combination of both approaches is well suited for an usage in a redundant localization system. The results from both map matching modalities are checked for consistency against runs of odometry measurements. Consistent matching results are selected and fused to obtain an accurate vehicle pose estimate. This localization framework is developed in Section 11.2.

Another contribution is concerned with the calculation of vehicle pose priors which have a outstanding role in vehicle localization systems. The design of the proposed localization system is conducted with a strong focus on the avoidance of localization failures. Nonetheless, recovery strategies for the case of localization failures are of crucial importance for the safe operation of automated vehicle systems (compare Section 2.3). Therefore, a recovery and initialization procedure on the basis of robust spectral registration algorithm (Section 6.3) is proposed due to the encouraging results presented in Section 11.3.1. For its applicability, a rough initial pose estimate is required which is then refined during the recovery step. Solutions from the state of the art include GPS/INS-based systems which incorporate receiver autonomous integrity monitoring (RAIM), place recognition and infrastructure-based systems that provide pose information to vehicle systems. Many of the aforementioned approaches are characterized by low localization accuracy, high computational costs or lack availability. To this point, a localization map based framework for global prior acquisition is developed in Section 11.1. The focus is put on the efficiency and robustness of the prior and accuracy is considered as less important. The pose refinement is left to an efficient initialization procedure (Section 6.3.2) and the LiDAR based localization framework developed in Section 11.2. Experimental evaluation results are presented in Section 11.3.

## 11.1 Vehicle pose prior generation

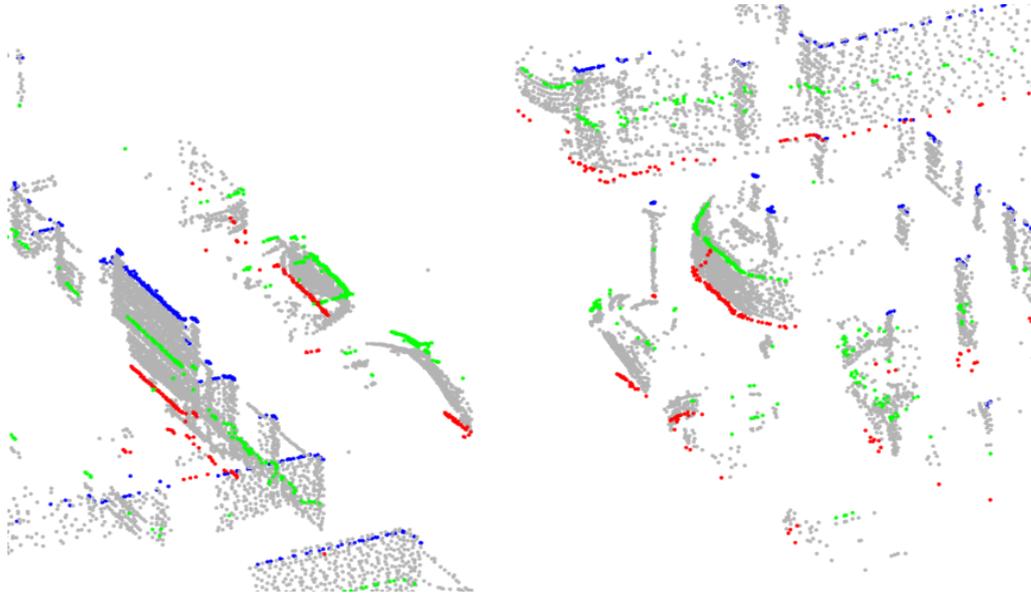
The main challenge in the generation of vehicle pose priors is the need to perform global localization during initialization. Global localization denotes the task of vehicle pose estimation under the absence of prior information. A global positioning system (GPS) provides a convenient source of absolute vehicle pose estimates but is often subject to degraded availability and precision in urban environments. In cases with GPS availability, initialization can be conducted on the basis of robust spectral registration. Otherwise, particle filter based localization approaches enable global localization, given the computational capacity to evaluate high amounts of particles. Once, initialization is conducted, the number of particles can be significantly reduced during tracking of the vehicle pose. To this point, a GPS-independent system for the generation of vehicle pose priors in urban environments on the basis of a particle filter based localization framework is developed in this section. The well known Monte-Carlo localization algorithm (Section 2.1.3) is utilized and its applicability to cluttered sensor measurements and outdated localization maps is extended. Accuracy of the result is traded in for computational efficiency by utilization of a small particle set.



**Figure 11.2** In a data preprocessing step, 3D LiDAR scans are divided into a set of layers. Each layer is an input for a separate instantiation of the AMCL procedure. Consistency checking between the resulting set of pose estimates and runs of odometry measurements is conducted. Consistent estimates are fused by covariance intersection to obtain a vehicle pose estimate.

In the following, a framework for global prior generation, called multilayer adaptive Monte Carlo localization (ML-AMCL) is developed. An overview over the ML-AMCL framework is provided in Figure 11.2. The basic idea of ML-AMCL is to perform a consistency check for separate localization system instances which use different measurement subsets as inputs. The outcomes are subsequently compared to a local reference. Under the assumption of a linearly increasing error in relative motion estimation, e.g. based on wheel or LiDAR odometry, short runs of odometry measurements can be used as a local references. On the contrary, errors in absolute localization often do not occur as linear drifts but as a high-frequent jumps. Consequently, these errors can often be detected by evaluation of residuals relative to the local reference. The developed algorithm was published in [RJMZ16] and is described in the following.

In a first step, 3D LiDAR measurements are divided into a set of measurement subsets, i.e. a set of layers  $\{\mathbf{z}_k^i\}_{i=1}^L$  which are then projected to the  $xy$ -plane (see Figure 11.3). Although applied to LiDAR data and AMCL, this approach can be used in combination with other sensor technologies and localization frameworks as well. A 3D LiDAR scan is shown in Figure 11.3 which entails characteristics that are symptomatic for urban data sets. Firstly, the scan contains measurement points originating from static objects like house walls and poles that are well suited for localization in urban environments. Secondly, the amount of clutter, i.e. measurements due to dynamic and semi-static objects, e.g. parked vehicles, are entailed. These measurement characteristics complicate the localization task as it causes inconsistencies between sensor observation and localization map. The utilization of layers instead of other scan partitioning strategies is chosen based on the observation, that occlusions in urban environments often occur up to a specific height, commonly due to parked or driving cars (see Figure 11.3). However, measurement points from lower layers like curb stones and fences often contain valuable structural

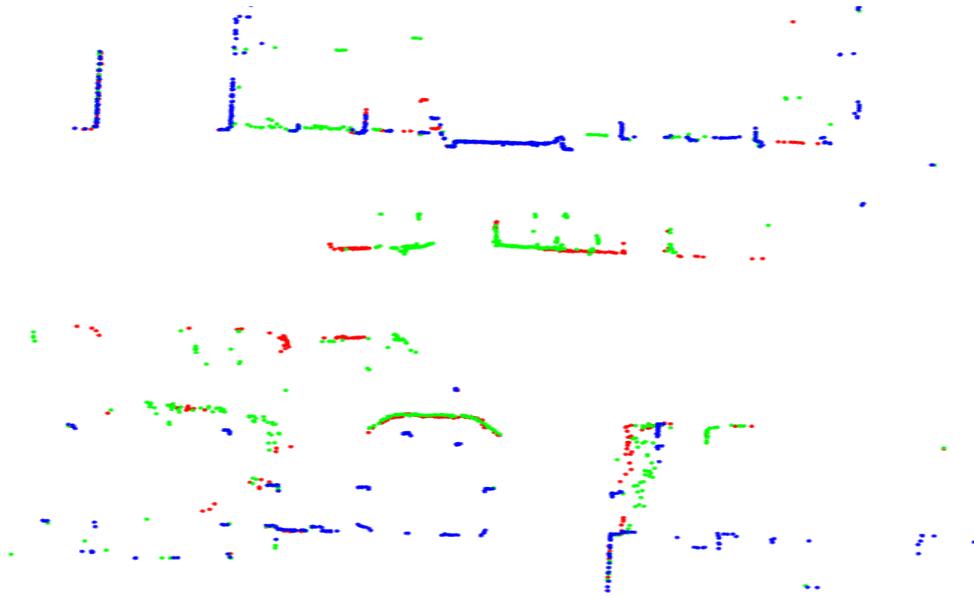


**Figure 11.3** Multilayer adaptive Monte Carlo localization (ML-AMCL) [RJMZ16] follows the intuition, that inconsistencies between sensor observation (gray) and the localization map are limited to measurement subsets. In urban environments, parking cars are a significant source of inconsistencies and hence, the scan is divided in horizontal layers (red, green, blue) which are used in separate localization algorithm instances. Consistency checking is utilized to identify valid localization results that are fused by covariance intersection.

information (see Figure 11.4) and, therefore, shall not be neglected. This is where the key contribution of ML-AMCL is made by providing a strategy for the dynamical selection of measurement subsets which enable localization with the underlying, potentially non-robust localization system (i.e. AMCL). This aspect is discussed and analyzed in the context of the state of the art in the following paragraph.

Current localization systems are often highly specialized for a specific operation environment. In the process of specialization, assumptions about the expected environment characteristics are incorporated into the localization system design. This is typically done by implicit or explicit assumptions on the measurement signals and the respective choice of signal processing algorithms. Many state of the art localization algorithms instrument the static world assumption and therefore might not sufficiently account for inconsistencies between LiDAR scans and the localization map. Signal processing algorithms are typically utilized in order to remove dynamic and semi-static objects from the sensor measurements. The remaining objects are then used for localization. Once the clutter removal fails, inconsistencies between the observed and mapped environment representation occur. Depending on the localization algorithm characteristics, this can lead to localization failures.

This event can be partially prevented by selection of specific subsets of the available sensor measurements in which clutter is unlikely to occur. In this context, the



**Figure 11.4** The frequency of occurrence of static localization features like curb stones and house walls often varies between the scan layers (red, green, blue). Therefore, ML-AMCL decides on the removal of layers based on their momentary usability for localization and not by static rules.

point measurements with an origin on the ground plane can be a suitable measurement subset. The authors in [LT10] use this observation to remove all dynamic and semi-static objects by extraction of the ground plane and rejection of all remaining measurements. While this approach seems feasible for highways and main roads with sufficient road markings, in urban environment it is likely to be overly pessimistic (compare Figure 11.4). This approach trades in the adherence to the static world assumption for the need of distinctive features on the ground plane. In urban environments, the latter requirement is frequently not met and whole areas without road marking might not be covered. Additionally, information about static structures like house walls are lost. Another approach for semi-static object would be the selection of measurements above the height of parked cars. Thereby, some valuable structures like low walls are at the height of parked vehicles and would thereby be erased from the measurement as well. Further works use machine learning approaches in order to learn static selection criteria for measurement subsets or feature types, e.g. in [SSB09]. In the context of the previously discussed solutions, it is argued that a static selection of measurement subsets can lead to insufficient adaptation to and coverage of heterogeneous operation environments. To this point, an approach is developed which enables the consideration of the whole measurement. This approach can be beneficial in order to prevent excessive information loss and contains comparatively mild assumptions about the clutter characteristics.

The set of scan layers is incorporated into a separate instance of the target localization system. Each AMCL instance takes a 2D localization map  $\mathbf{m}$ , 2D LiDAR

scans  $\mathbf{s}_k^i$  and odometry measurements  $\mathbf{u}_k$  as inputs. The map is considered to be 2D with no separate layers. This assumption is realistic, since in most cases only one map is provided and the utilization of several layers would result in increased computational costs and bandwidth requirements for handling and transmitting the additional amount of data. AMCL instances provide sets of pose estimates  $\{\hat{\mathbf{x}}_{k_i}^i\}_{i=1}^L$  and covariance estimates  $\Sigma_{\mathbf{x},k_i}$ . Every estimate is evaluated against the odometry measurements between the latest two absolute pose update at  $k_i - 1$  and  $k_i$ . The odometry measurements arrive at a significantly higher rate  $1/T_u$  than the pose estimates. Hence, the delay caused by waiting for the next odometry measurement is sufficiently low and the odometry measurement is calculated by linear interpolation.

$$\mathbf{u}_{k_i} = \mathbf{u}_{k_{u-1}} + (\mathbf{u}_{k_u} - \mathbf{u}_{k_{u-1}}) \frac{t_i - t_{u-1}}{T_u}, t_{u-1} < t_i < t_u. \quad (11.1)$$

After calculation of  $\mathbf{u}_{k_i}$ , the motion between the last time steps is determined as

$$\Delta_{k_i}^{\mathbf{u}} = \mathbf{u}_{k_i} - \mathbf{u}_{k_{i-1}} \quad (11.2)$$

and serves as reference for the difference of absolute pose estimates

$$\Delta_{k_i}^{\mathbf{x}} = \mathbf{x}_{k_i} - \mathbf{x}_{k_{i-1}}. \quad (11.3)$$

The subsequent step is the determination of a set with  $F_k \leq L$  (here:  $L = 3$ ) pose estimates which are consistent with the odometry measurements. This step can be compared to the detection of outlying pseudo-range measurements in GPS/INS based navigation. Detected outliers are removed from the overall fusion step. For a fast detection of outliers, a simple threshold is applied to the Mahalanobis distance  $\epsilon$  between  $\Delta_{k_i}^{\mathbf{x}}$  and  $\Delta_{k_i}^{\mathbf{u}}$

$$\mathbf{S}_{k_i} = \mathbf{Q}(T_i) + \Sigma_{\mathbf{x},k_{i-1}} + \Sigma_{\mathbf{x},k_i} \quad (11.4)$$

$$\Delta_{k_i} = \Delta_{k_i}^{\mathbf{x}} - \Delta_{k_i}^{\mathbf{u}} \quad (11.5)$$

$$\epsilon_{k,i} = \Delta_{k_i}^{\top} \mathbf{S}_{k_i} \Delta_{k_i}. \quad (11.6)$$

If  $\epsilon_{k,i}$  exceeds the threshold  $\epsilon_{th}$ , the corresponding pose estimate  $\mathbf{x}_{k,i}$  is rejected. An alternative implementation would be a majority vote between the interpolated outcomes of the AMCL instances. However, this approach would increase the sensitivity against influences of inconsistent observations existent in several measurement subsets and causing a correlation between the pose estimates. One example is the scene from Figure 11.3, where semi-static objects, i.e. parked cars are contained in the lower two layers. For these layers, the objects might lead to consistent, yet false localization results. In this case, a majority vote would have outvoted the correct results corresponding to the highest layer.

After the selection step, the set of consistent pose estimates is fused in order to obtain the overall pose estimate  $\hat{\mathbf{x}}_k$ . The aforementioned correlations between the

pose estimates are considered by the selection of a suitable fusion procedure. To this point, covariance intersection (Equation 11.7) is utilized for fusion

$$\hat{\mathbf{x}}_k = \mathbf{P}_k \left( \sum_{i=1}^{F_k} w_i \hat{\Sigma}_{\mathbf{x},k_i}^{-1} \hat{\mathbf{x}}_i \right)^{-1} \quad (11.7)$$

$$\mathbf{P}_k = \left( \sum_{i=1}^{F_k} w_i \hat{\Sigma}_{\mathbf{x},k_i}^{-1} \right). \quad (11.8)$$

The weights  $w_i$  are subject to the constraint

$$\sum_{i=1}^{F_k} w_i = 1 \quad (11.9)$$

and are found by optimization.

Prolonged periods of faulty pose estimates from an AMCL instance are an indicator for filter divergence and as a consequence a reinitialization can be advisable. This aspect of the ML-AMCL approach differs from previous works on GPS/INS navigation and necessitates a reinitialization decision for each AMCL instance has to be made. In order to detect small drifts and limit the rate of unnecessary reinitialization steps, a variant of the sequential probability ratio testing (SPRT), the cumulative sum (CUSUM) test [Gus00] (compare Section 2.1.4) is employed. Firstly, the test statistic  $g_{k_i}$  is calculated as the sum of  $g_{k_i-1}$  and  $\epsilon_{k,i}$  which is compensated by the drift constant  $d$

$$g_{k_i} = g_{k_i-1} + \epsilon_{k,i} - d \quad (11.10)$$

$$g_{k_i} = 0, \text{ and } \hat{k}_i = k_i \text{ if } g_{k_i} < 0 \quad (11.11)$$

$$g_{k_i} = 0, \text{ and } k_{a,i} = \hat{k}_i \text{ and reinitialization if } g_{k_i} > \epsilon_{th}^s > 0. \quad (11.12)$$

If the test statistic is smaller than zero, it is set to zero and the time is stored as a time  $\hat{k}_i$  of a potential failure occurrence. If the threshold  $\epsilon_{th}^s$  is exceeded, the reinitialization of the respective AMCL instance is triggered. The time of the localization failure occurrence is denoted  $k_{a,i}$ . The information from the layer selection step can be utilized in other procedures, such as a pose refinement step on the basis of non-robust map matching procedures (see [RJMZ16]).

## 11.2 Precise vehicle localization

A vehicle localization system with enhanced robustness (compare Figure 11.1) is developed in this section. Therefore, robustness is mainly considered as decreased dependence on updated localization maps, explicit consideration of the heterogeneity of urban environments and insensitivity to map matching failures. The latter is fostered by utilization of an accurate scan-based odometry introduced in Section 10.1 and consistency checking. Insensitivity against inconsistent sensor observations is achieved by spectral registration (Section 6) and a suitable state of the art point matching algorithm applied to pole measurements (Section 5.1.2).

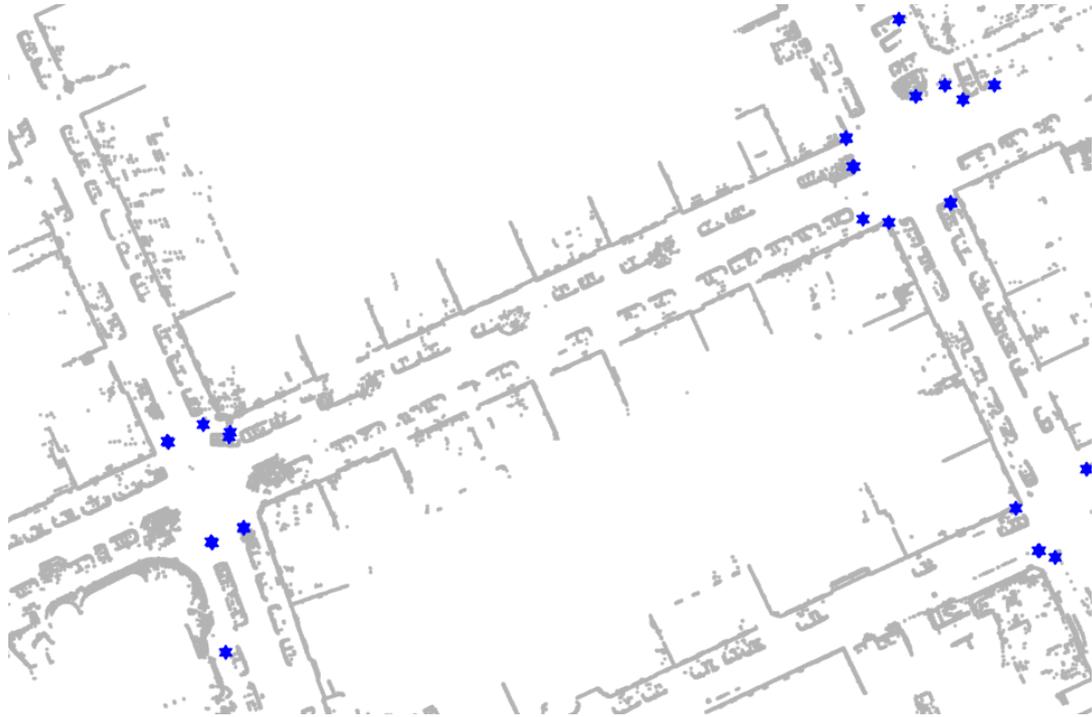
This combination of matching algorithms is motivated by the observation that urban areas road sides can often be categorized by their respective density of structure (Figure 11.5). Sparse structures as encountered in intersection areas are commonly accompanied by pole-shaped objects like traffic sign and traffic light posts and pole-based map matching is therefore well suited for vehicle pose estimation. In densely structured areas, spectral registration provides good matching results while making mild assumptions on the quality of the measurement data and the localization map. The combination of scan- (Section 11.2.1) and pole-based vehicle pose estimation (Section 11.2.2) is used in a hybrid vehicle localization system (Section 11.2.3) and accounts for the heterogeneity of the urban environment.

### 11.2.1 Scan-based map matching

Narrow inner-city roads with densely parked cars as shown in Figure 11.5 and missing road markings are a great challenge in vehicle localization. The spectral matching algorithm (Section 6) is considered for the scan-based map matching in such densely structured environments. Two operation modes are defined:

1. Standard mode (accurate prior is used): standard spectral registration with utilization of prior for pose estimation
2. Initialization mode: utilization of robust spectral registration under consideration of inaccurate vehicle pose priors

The standard mode shall be as efficient as possible while providing the required matching accuracy. The latter has been addressed in Section 9.2. Accordingly, for a localization accuracy requirement of 0.25 m, a grid resolution of 0.2 m suffices to compensate map matching failure rates of over 30 %. The initialization mode utilizes the proposed robust spectral registration algorithm in combination with a 0.1 m grid resolution. Thereby, it can be assured that the initialization of the vehicle pose is accurate enough for compensation of consecutive map matching failures. In areas with increased dependability requirements such as intersections, the initialization mode or any other parameterization between both proposed modes could be used. The introduction of automated parameterization is an interesting starting point for future work.



**Figure 11.5** The localization maps for scan-based (gray) and pole-based (blue) localization. Pole-features are often encountered in expansive intersection areas where scan-based map matching is prone to degradation. On the contrary, poles might not be available between the intersection areas.

### 11.2.2 Landmark-based map matching

Expansive areas in urban environments are often characterized by the existence of pole shaped objects like posts of traffic signs and traffic lights (compare Figure 11.1). Nonetheless, long road segments without poles, as seen in Figure 11.5, might render purely pole-based localization infeasible. In this work, pole-based map matching is introduced as an efficient method to increase the robustness of the scan-based localization system. This section is intended to give a short overview over the utilized registration approach on the basis of a state of the art correspondence estimation procedure.

The detection of poles is based on the algorithm developed in Section 5.1.2. The number of detected landmarks in one time step is typically smaller than ten. Due to the sparseness of pole landmarks, a matching procedure with exhaustive correspondence search is preferential, since it decreases the probability of false data association and is computationally feasible for the expected number of pole measurements. For a given set of correspondences, the registration parameters are calculated by the solution of the scalar-weighted OPP (Section 8.1.1). This decision is feasible, since the measurement covariance for the LiDAR-based pole detection is weakly distance dependent and approximately isotropic as shown in Section 7.2.1. The weighting

matrix in the cost function  $J$

$$J = \sum_{i=1}^n \|\mathbf{W}_i (\mathbf{R}\mathbf{z}_{l,i} + \mathbf{t} - \mathbf{m}_{k,i})\|^2 \quad (11.13)$$

is chosen as  $\mathbf{W}_i = w_i \mathbf{I} = \left( \sqrt{\max \text{eig}(\boldsymbol{\Sigma}_{l_i})} \right)^{-1} \mathbf{I}$ . The incorporation of additional information of measurement noise leads to an increase in matching accuracy. Now, it is left to determine an efficient procedure.

As opposed to point measurements in scans, accurate correspondences between the extracted environment features and the map can be established. To this point, a comparison between random sample consensus (RANSAC) [FB81] and single cluster graph partitioning (SCGP) [OWTL05] is conducted.

For RANSAC, the number of drawn correspondence pair samples is often set to three [HMHS14]. From the three samples, the transformation parameters are estimated and the fitness score for the respective parameter set is calculated. This procedure leads to increased runtime requirements. This effect is decreased by adapting the number of iterations which, however, increases the probability of not drawing the correct set of correspondences. SCGP in general or single cluster graph partitioning (SCGP) [OWTL05] in specific can be used as well and is proposed as a beneficial alternative for the given use case of correspondence estimation for sparse landmarks. For SCGP, only two samples are drawn and the mean squared error is minimized. The resulting mean squared errors are added to a consistency matrix from which, by means of eigenvalue decomposition, the set of consistent correspondence pairs is determined (compare Section 2.1.4). This procedure makes the matching algorithm more computationally efficient than standard RANSAC. This leads to implementations which favor an exhaustive search over the correspondence space and thereby to higher probability of finding the correct set of correspondences.

### 11.2.3 Hybrid vehicle localization

In this section, the scan- and pole-based matching procedures are combined in a localization framework with consistency checking capability. Consistent matching results are fused in an extended Kalman filter (EKF), yielding efficient pose estimation and precise vehicle pose estimates. The proposed map matching procedure combination has advantages over a solely scan-based approach which justify the utilization of an additional pole-based matching procedure given the high matching rate of spectral registration:

- increased detectability of matching failures due to the gain in analytical redundancy (compare results from Section 6.3.3)
  - increased failure detection rate
  - decreased false rejection rate
- increased matching rates due to

- partial orthogonality of the matching procedures
- decreased false rejection rate
- decreased impact of short-term localization failures due to the utilization of matching algorithms with broad basin of convergence
- avoidance of errors in synchronization of map matching results due to the utilization of a single sensor modality

The positive impact of aforementioned aspects on the localization robustness has been studied during model-based localization system design throughout Section 9. Accordingly, the matching rate has a considerable influence on the localization performance and the perception accuracy requirements. Hence, the matching rate should be increased in order to enhance the localization robustness and potentially decrease the perception system requirements. The performance of local confidence measures like the signal-to-noise ratio (SNR) for detection of outlying map matching results is limited. The enforcement of a SNR threshold for map matching result rejection would result in a decreased matching rate due to false positives. Another significant influencing factor was identified by evaluation of the statistical localization system models:

- convergence to a spatially uniform distribution of matching results leading to
  - increased average localization accuracy
  - decreased length of matching failure sequences (relaxed odometry accuracy requirements)

Furthermore, the calculation of vehicle pose estimates on the basis of pole measurements can be efficiently integrated in the signal processing chain for spectral registration based localization and, consequently, results in only slightly increased computational costs.

Generally, matching failures can be distinguished by their frequency of occurrence and require different diagnostic approaches. Jumps can be detected by classical hypotheses tests. Low-frequent drifts require the evaluation of time series of residuals as conducted in Section 11.1. Under the assumption of a approximately bias-free odometry, the utilization of map matching algorithms with broad basin of convergence decreases the likelihood of slow drifts in the vehicle pose estimate. Therefore, the residuals from one time step are considered for consistency checking which corresponds to a jump detection. In the consistency checking step, the odometry is used as a local reference. Due to the drift of the odometry measurements, the global odometry pose does prohibit the calculation of residuals between absolute poses. Consequently, differences of poses are considered for consistency checking that can be obtained from runs of odometry measurements and map matching results (FMT

index spectral registration)

$$\Delta_k^{FMT} = \|\mathbf{x}_k^{FMT} - \mathbf{x}_{k-1}^{FMT}\|_2 \quad (11.14)$$

$$\Delta_k^{SCGP} = \|\mathbf{x}_k^{SCGP} - \mathbf{x}_{k-1}^{SCGP}\|_2 \quad (11.15)$$

$$\Delta_k^{\mathbf{u}} = \|\mathbf{u}_k - \mathbf{u}_{k-1}\|_2 . \quad (11.16)$$

The distance between the motion hypotheses is then expressed by the following differences  $d\mathbf{t}_i$  in the translation estimates

$$d\mathbf{t}_1 = |\Delta_k^{FMT,\mathbf{t}} - \Delta_k^{SCGP,\mathbf{t}}| \quad (11.17)$$

$$d\mathbf{t}_2 = |\Delta_k^{FMT,\mathbf{t}} - \Delta_k^{\mathbf{u},\mathbf{t}}| \quad (11.18)$$

$$d\mathbf{t}_3 = |\Delta_k^{SCGP,\mathbf{t}} - \Delta_k^{\mathbf{u},\mathbf{t}}| . \quad (11.19)$$

Additionally, the respective differences  $d\Theta_i$  for the orientation estimates could be considered. A spectral clustering algorithm is then used in order to determine the set of consistent hypotheses (see Section 2.1.4). The consistency checking step requires an accurate temporal calibration between the measurement sources. This is achieved by utilization of the same LiDAR scan for map matching and relative motion estimation.

### 11.3 Experimental evaluation

This section contains the evaluation results for the prior generation on basis of ML-AMCL in Section 11.3.2 and the LiDAR localization in Section 11.3.3. In both cases data sets 0823-71 and 0316-11 are used (Table 11.1). Data set 0823-71 is challenging due to the high outdatedness of the corresponding localization map, including several building sites (the example in Figure 1.2 was taken from this data set). Furthermore, it contains the highest number of intersection which makes it well suited for demonstrating the orthogonality of the pole- and scan-based map matching in the proposed localization framework. Data set 0316-11 is slightly more diverse with respect to road widths (the narrow road example in Figure 4.3d was taken from this data set) and vegetation than data set 0823-71 at a decreased level of outdatedness. These data sets are selected to provide the highest challenges for the localization algorithms. In this aspect, the other data sets are subsumed.

The localization map (compare Section 5.2) was recorded more than 2 years before the evaluation data sets (Section 4.3). During the mapping process, no procedures for the detection and removal of semi-static objects were incorporated. The remaining clutter in the map complicates the localization procedure. Since, changes to the localization map can occur at high frequencies, in practical application it cannot be guaranteed that measurements originating from semi-static, dynamic objects and other artifacts remain in the map. To this point, the ability of a localization system for dealing with such clutter is an important aspect of its robustness. Hence, it is argued that the cluttered map is a suitable proving ground for the robustness of

**Table 11.1** Overview over the data sets (Section 4.3) used for experimental evaluation of the localization frameworks and their characteristics.

Data set	Characteristics
0823-71	<ul style="list-style-type: none"> <li>• expansive intersection area</li> <li>• high outdatedness of the localization map (building sites)</li> </ul>
0316-11	<ul style="list-style-type: none"> <li>• narrow urban canyons</li> <li>• expansive intersections</li> </ul>

a localization system and is used as such in the following sections. In the case of LiDAR based localization, evaluation of robustness is conducted by comparing the localization accuracy for an outdated and updated version of the same localization map. Localization is further complicated by utilization of a different LiDAR sensor setup for localization and mapping. All algorithms are implemented in C++ and are running in real-time on the Robot operating system (ROS) under Ubuntu. As the basis for all results, reference vehicle poses were calculated as described in Section 5.3. By utilization of the same data for reference generation and localization, errors in time stamp synchronization can be omitted. The vehicle pose initialization on the basis of robust spectral registration is evaluated in the following section.

### 11.3.1 Vehicle pose initialization and recovery

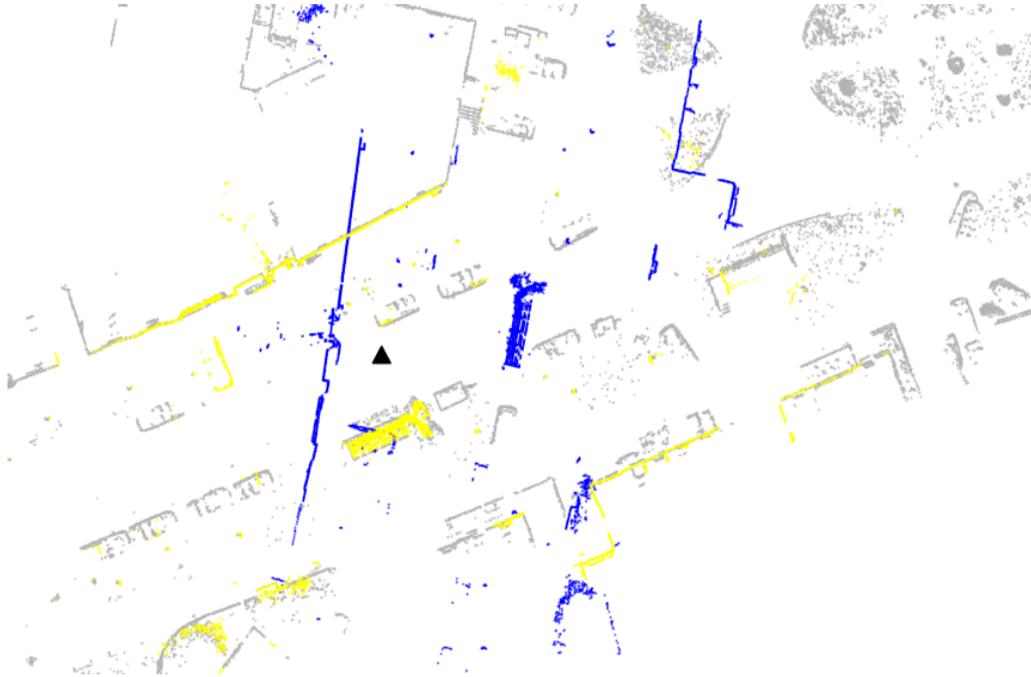
During the initialization of an automated vehicle system and in cases of localization failures, the calculation of a precise vehicle pose estimate from an inaccurate prior is of crucial importance. In these scenarios, a high availability of a precise vehicle pose estimate outweighs the real time capability of the matching system. Selected particle filter (PF) based localization procedures and SLAM frameworks are able to perform initialization, but typically require to move the vehicle in order to gather additional environment information (compare respective results in [ORD13]). Especially for the recovery of vehicle systems which are in a safe state, e.g. after stopping upon detection of a localization failure, moving the vehicle might not be possible. Furthermore, a direct system initialization without the need for driver intervention might provide additional safety and comfort to the passenger. In this context, rFMT can be beneficially utilized for initialization and recovery of localization systems. Due to the aforementioned decreased runtime requirements, higher grid dimensions and resolutions can be used. Transformation priors are generated from adding Gaussian noise with a standard deviation of 10 m to the reference transformation in order to simulate the initialization from an inaccurate GPS prior. The two urban data sets were selected to have the fewest overlap in order to capture a manifold of different areas. The initialization capability of the robust spectral regis-

tration algorithm is evaluated in an experiment with data sets 0823-42 and 0316-31 (Table 11.3, Section 4.3).

**Table 11.3** Overview over the data sets (Section 4.3) used for experimental evaluation of the initialization capability of rFMT.

Data set	Characteristics
0823-42	<ul style="list-style-type: none"> <li>• high outdatedness</li> <li>• roundabout</li> <li>• narrow urban canyon</li> </ul>
0316-31	<ul style="list-style-type: none"> <li>• longest loop without self overlap</li> <li>• narrow urban canyon</li> </ul>

In the following, results for an example implementation of the developed prior-free hypothesis selection procedure is evaluated and discussed. Firstly, a set of  $U_k = 50$  point measurements is selected from the scan  $\mathbf{s}_k$  and different numbers of hypotheses are chosen. The evaluation results are given in Table 11.5.



**Figure 11.6** Initialization and recovery of accurate vehicle pose estimates (vehicle position: black triangle; transformed scan: yellow) from inaccurate transformation priors (blue), as obtained from GPS or ML-AMCL, is enabled by robust spectral registration.

**Table 11.5** The results from the pose recovery rate ( $P_r$ ) evaluation for data sets 0823-42 and 0316-31 (Section 4.3). The proposed hypothesis selection procedure was evaluated for  $P_k = 50$  sample points.

Hypotheses set	$P_r / \%$	
	0823-42	0316-31
single hypothesis (SNR)	84.3	82.7
$H_{k,rot} = 2, H_{k,trans} = 4, H_k = 16$	96.0	95.3
$H_{k,rot} = 2, H_{k,trans} = 8, H_k = 32$	96.3	95.6
$H_{k,rot} = 4, H_{k,trans} = 2, H_k = 16$	97.1	97.6
$H_{k,rot} = 6, H_{k,trans} = 4, H_k = 48$	97.8	97.7
$H_{k,rot} = 8, H_{k,trans} = 2, H_k = 32$	99.2	99.3
accurate prior	99.9	99.9

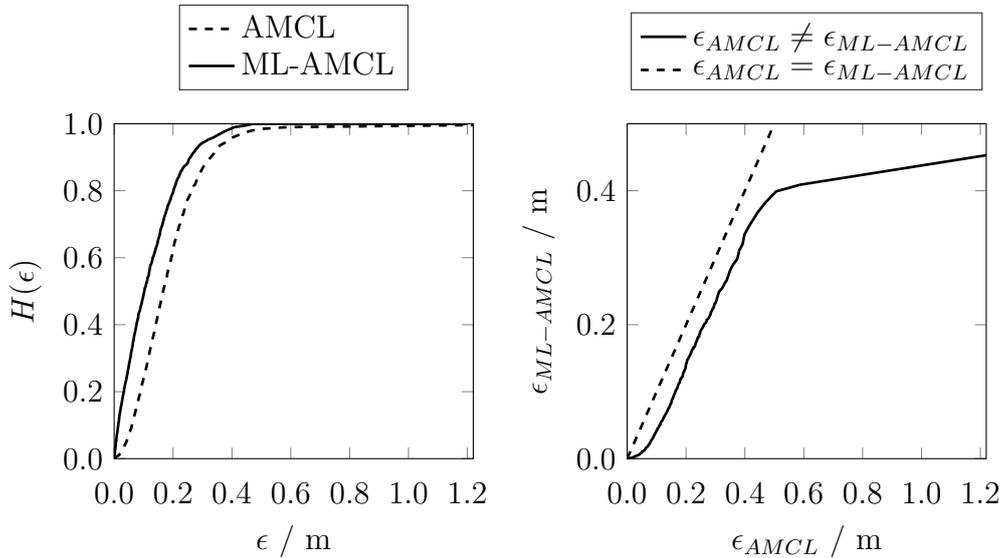
The evaluation for  $P_k = 50$  sample points for all sets of hypotheses takes less than 0.5 ms. According to the results, the hypothesis for translation hypotheses are more distinctive compared to the rotation hypotheses (compare Section 6.3.1). This observation can be explained by the frequent occurrence of walls or other structures beside the road which potentially lead to ambiguities in the rotation parameter domain. Especially in intersection areas as shown in Figure 6.5, the rotational transformations are typically characterized by a  $\pi/2$  periodicity due to the structural ambiguities and additional side peaks as shown in Figure 6.8. This effect can be seen from the recovery rate for  $H_{k,rot} = 4$  and  $H_{k,trans} = 2$  of 97.1% which is 1.1% higher than for  $H_{k,rot} = 2$  and  $H_{k,trans} = 4$  at the same absolute number of hypotheses. The best result among all considered parameterizations of 99.3% are obtained for  $H_{k,rot} = 8$  and  $H_{k,trans} = 2$ . In all cases, the matching rate is significantly increased in comparison to the single-hypothesis algorithm. The incorporation of prior information about the transformation parameters leads to a matching rate of 99.9%. Consequently, the incorporation of an accurate prior results in the best matching performance. The remaining number of outliers can be due to cases with extremely noisy or corrupted LiDAR scans. The robust version of spectral registration leads to comparable results, hence enabling a highly available system for recovery and initialization from inaccurate transformation priors. Thereby, the registration algorithm contributes to the goal of robust vehicle localization in challenging operation environments and under the assumption of imperfect measurement and localization map data processing.

### 11.3.2 Multilayer adaptive Monte Carlo localization

For the experimental evaluation of the ML-AMCL framework, the publicly available ROS AMCL package [Ger] is used which implements an extended version of the

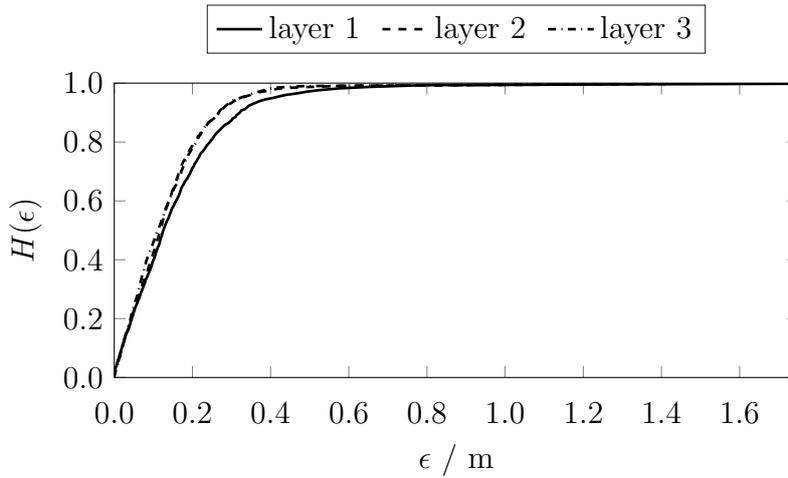
Monte Carlo localization approach introduced in [DFBT99]. The ML-AMCL is evaluated on an urban test track with high buildings, narrow roads and an outdated localization map. GPS measurements are utilized for initialization of the AMCL instances. In order to evaluate the gain in robustness for ML-AMCL relative to AMCL, the localization accuracies are compared. For AMCL, the whole 3D scan  $\mathbf{s}_k^3$  is projected to the  $xy$ -plane as described in Section 5.1.1. For ML-AMCL,  $\mathbf{s}_k^3$  is divided into three layers which are then separately projected to the  $xy$ -plane. For the AMCL, a minimum and maximum particle number of 300 and 3000 was selected which were divided by three for ML-AMCL.

The comparison of accuracies is conducted on the basis of the empirical cumulative density functions (ECDF) of the vehicle pose estimation errors as shown in Figure 11.7 for data set 0823-71. It can be observed, that for ML-AMCL, the ECDF has a steeper incline than for AMCL. Consequently, the experiments indicate an increased precision of the ML-AMCL in direct comparison to AMCL. The evaluation



**Figure 11.7** In the presence of outdated localization maps and cluttered LiDAR scans, ML-AMCL leads to an increased accuracy in pose estimation as compared to AMCL. This can be seen from the empirical cumulative density functions. ML-AMCL results in a sharper decline of the curve than AMCL. This observation is equivalent to a lower rate of pose estimates with high uncertainty for ML-AMCL.

of the layer-specific results in Figure 11.8 shows that the top layer is selected in 84.5% of the time steps. The lower and middle layer are selected in 54.1% and 59.8% of the selection steps. These results were expectable since inconsistencies frequently occur due to building sites or other traffic participants which infrequently occur above the mounting height of the sensor. For the data set 0361-11, the application of ML-AMCL did not result in a considerable accuracy gain which might be due to the decreased outdatedness of the corresponding localization map.



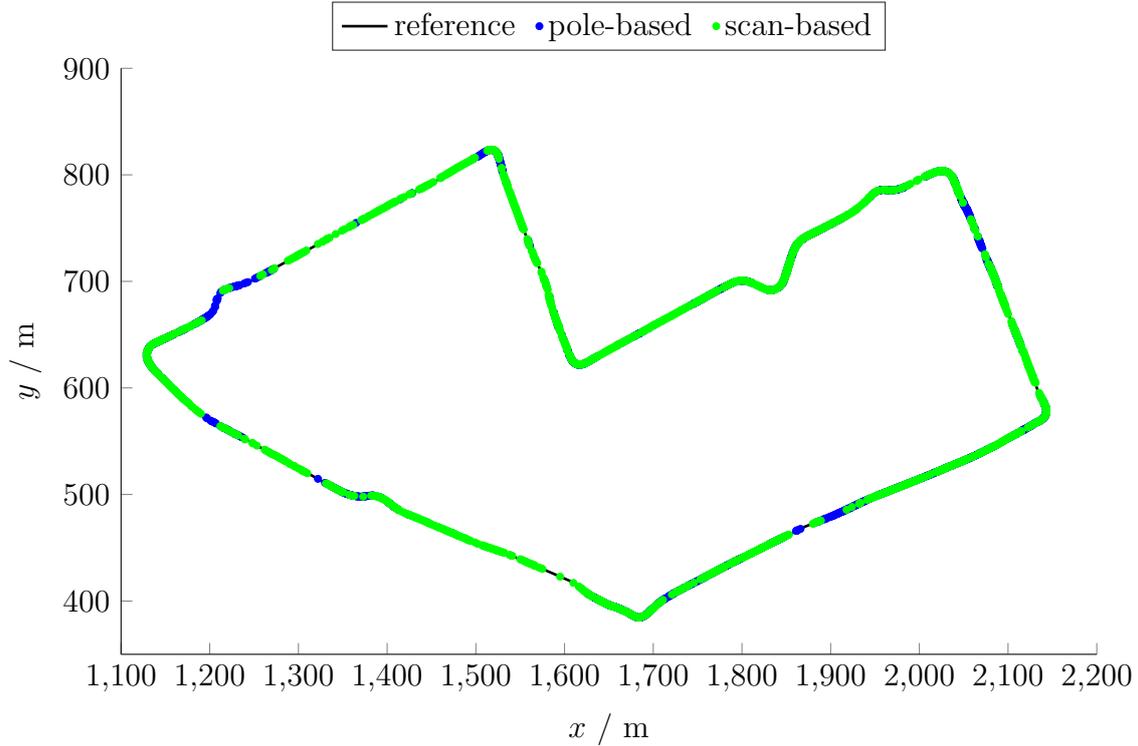
**Figure 11.8** The results for the upper two layers are similarly accurate, whereas the utilization of the lower layer leads to a lower localization performance.

The results presented and discussed in [RJMZ16] indicate that the information from layer selection can lead to increased accuracy in consecutive pose refinement steps on the basis of non-robust matching procedures. In the experimental evaluation, this information is used to construct a 3D scan used for map matching with G-ICP. Due to the rejection of layers with high amounts of inconsistent sensor observations, the matching results have an increased accuracy as compared to the utilization of the whole scan. The applicability of ML-AMCL for measurement selection benefits from similar sensitivity characteristics of the utilized localization framework and matching algorithm.

### 11.3.3 Hybrid vehicle localization

The proposed localization framework makes use of pole- and scan-based map matching procedures. The results in Figure 11.9 show the good matching rate for pole-feature registration in intersection areas. In this aspect, pole-based localization is orthogonal to scan-based matching which has the most matching failures in intersection areas (compare Section 6.2.3). In some areas, both matching algorithms did not give valid matching results. Therefore, the accurate LiDAR scan odometry or wheel odometry is required. If the distances between valid matching results are too long, the robust spectral registration algorithm can be utilized for recovery. Thereby, the gain in efficiency is decreased, but the functioning of the localization system can be ensured. This strategy is especially interesting, if energy shall be saved, e.g. in electric vehicles, and therefore a low matching performance is used. If required, the matching performance is increased at the cost of additional energy consumption. At this point, the adaptivity of the spectral matching algorithm is of significant advantage.

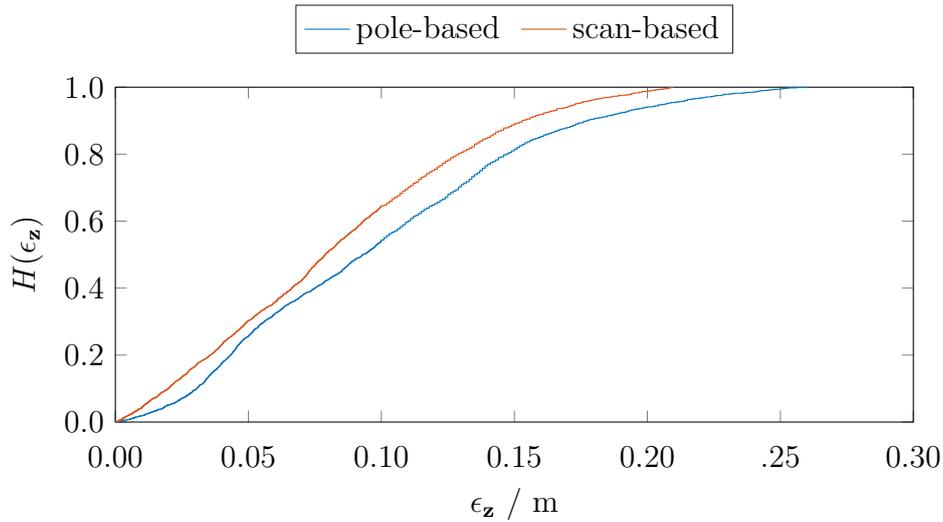
The scan processing parameters are set to a scanner range of 20 m and a grid



**Figure 11.9** Results from the matching of pole-features ( $P_m = 51.3\%$ ) and scans ( $P_m = 63.6\%$ ) for data set 0823-71. An accumulated matching rate of 87.9% is achieved.

resolution of 0.1 m which enables a fast execution of the map matching procedure on the target system. Due to the selected setting, the matching rate is decreased to  $P_m = 63.6\%$  for data set 0823-71 and 76.2% for 0316-11. The introduction of pole-based map matching with  $P_m = 51.3\%$  (0823-71) and  $P_m = 48.1\%$  (0316-11) leads to an increased overall matching rate 87.9% (0823-71) and 91.3% (0316-11). The empirical cumulative density function (ECDF) in Figure 11.10 contains the results from both data sets and shows the higher matching accuracy of the spectral matching approach in comparison to the pole-based matching. The mean matching error for spectral registration is 0.098 m with a standard deviation of 0.058 m for 0823-71 and 0.095 m as well as 0.056 m for 0316-11. Additionally, more than 90% of the considered matching results had an accuracy of higher than 0.20 m which, in combination with the high matching rate, yields a sufficiently high localization accuracy.

The robustness of the localization procedure against inconsistencies between sensor observations and the localization map is evaluated. To this point, the vehicle pose estimation errors for the outdated localization map and an updated version are compared. A submap from both localization maps is shown in Figure 11.11. For localization algorithm with a perfect robustness against the inconsistencies between the data set and the localization map, it is expected that the localization accuracy



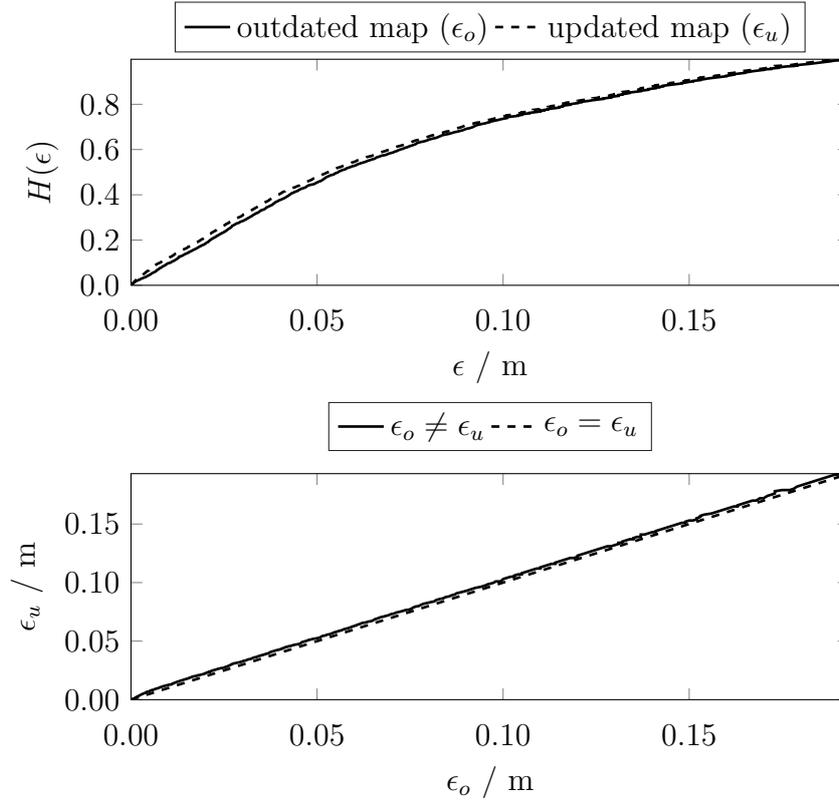
**Figure 11.10** Results from scan-matching are characterized by a higher accuracy than pole-based matching (combined results for data sets 0823-71 and 0316-11).

is identical in both cases.



**Figure 11.11** Two different versions of the localization map are used to evaluate the localization robustness against outdated maps. To this point, an outdated localization map (left image) and an updated version thereof are used.

Figure 11.12 contains the results from two different test runs. Firstly, the localization is conducted on the basis of the outdated localization map which contains building sites and parked cars. This test run resulted in a mean error of 0.092 m with a standard deviation of 0.06 m for data set 0823-71 and 0.071 m with a standard deviation of 0.05 m for data set 0316-11. Additionally, the localization was conducted on the basis of the updated version of the localization map, where the semi-static objects are removed. The accuracy was slightly increased, resulting in a mean error of 0.087 m with a standard deviation of 0.06 m for data set 0823-71 and



**Figure 11.12** Results for the localization accuracy from the proposed LiDAR based localization procedure with enhanced robustness (data sets 0823-71 and 0316-11). The accuracy for localization with an outdated and updated localization map show no significant differences. Therefore, the proposed system can be considered robust against outdated localization maps.

0.069 m with a standard deviation of 0.05 m for data set 0316-11. The similarity of the results from both test runs is evaluated on the basis of the ECDF shown in Figure 11.12. show only small deviations between the ECDF.

Consequently, the robustness of the proposed localization algorithm against inconsistencies is found to be given. In summary, the highest localization accuracy and efficiency is reached for updated maps accurate priors for map matching which leads to the highest matching rate (Section 11.3.1) and lowest runtime (Section 6.2.3). Robustness is gained at the cost of runtime performance. In the case of the proposed hybrid localization framework, the additional costs for robustness are kept at a low level and the level of robustness can be adapted to the intended use case or current situational context.

## 11.4 Conclusion

Vehicle localization with enhanced robustness for UAD was addressed in this part. The developed systems were based on the components previously described in part II and the design principles and parameters derived in part III were considered. Different parameterizations of the spectral registration algorithm enabled a variety of different applications, including map matching, scan matching, initialization and loop closing. The introduced LiDAR scan based odometry showed accurate motion estimates with 0.21% average position drift per driven distance. From the results it could be observed, that the mounting height and angular resolution of the utilized LiDAR sensors has significant influence on the achievable accuracy. The higher detection range of the Velodyne sensor enabled a more accurate orientation estimation than the Ibeo sensors which have to deal with frequent occlusions due to their mounting at bumper-height. Higher detection ranges also rendered possible the matching of scans from several time steps which lead to a further increased orientation estimation accuracy. The latter is fundamental as uncertainty in orientation estimation accumulate to high errors, especially for long trajectories. In a SLAM framework, the good scan matching performance was further instrumented to establish high amounts of local and global loop closing hypotheses. Without incorporation of additional sensor readings, the developed extensions of the spectral registration method enabled accurate mapping of urban and campus environments. Due to the determinism of the transformations involved in spectral registration, the mapping results for a given data set are identical between two runs. Since, the same matching algorithm with different parameterizations could be used, the additional efforts for adding scan-based odometry and SLAM capability to the overall localization was kept negligible. Hence, an answer to research question 6 (*rq6*, Section 1) could be given.

The localization framework on the basis of pole- and scan-based map matching, resulted in pose estimates with the required localization accuracy of 0.1 m and a standard deviation of 0.06 m. The initialization and recovery of accurate vehicle pose estimates was based on the developed robust spectral registration algorithm (Section 6.3). By application of this algorithm, the recovery rate of precise vehicle pose estimates from inaccurate GPS measurements was increased from 82.7% for the state of the art implementation to 99.3% for the proposed algorithm. For localization, the resolution of the grid for map matching was selected according to the accuracy requirement. Thereby, the size of the localization map could be reduced to less than 1 MB for an area of several square-kilometers. The decreased grid resolution led to a decreased matching rate in intersection areas. Hence, the combination with pole-based map matching procedure served two main purposes. Firstly, it increased the matching rate in intersection areas, where an accurate localization result is mandatory and the spectral matching algorithm shows some disadvantages. Secondly, it was used for the introduction of analytical redundancy to the localization system, enabling increased consistency checking capability. Consequently, the answer to research question 4 (*rq4*, Section 1) of how to detect map

matching failures given in Section 6.4 was further elaborated. The overhead for the pole-based matching step is small due to the sparsity of the landmarks. In combination with the orthogonality to scan-based matching, this makes the utilization of scan- and pole-based map matching a well suited combination for localization in urban environments. By consideration of consistent information, an efficient EKF based implementation of the fusion of relative and absolute pose measurements did suffice for a high level of localization robustness. The *ML-AMCL* approach developed in Section 11.1 served as an answer to research question 5 (*rq5*, Section 1). For *ML-AMCL*, the LiDAR scans are separated into layers and running several *AMCL* instances with the different layers as an input. Consistency checking against odometry measurements was then used to identify the scan layers which provided consistent localization results. These results were then fused by means of covariance intersection. The experimental evaluation was conducted with data sets from areas with highly outdated localization maps which included building sites and a high number of parked vehicles. The changes to the *AMCL* system architecture lead to increased insensitivity against outdated localization maps which was manifested as an increased localization accuracy. The *ML-AMCL* reliably provided vehicle pose priors, accurate enough for a pose refinement with spectral matching.

A final conclusion and combined discussion of the results of this thesis are given in the following section. Based on this discussion, the main research question from Section 1 is answered and future research perspectives are provided.

---

## Conclusion and future work

The main goal of this thesis was the development of a system for map-relative vehicle localization in urban environments with enhanced robustness. The research was guided by the main research question, formulated in the introduction: How can GPS-less vehicle localization for urban automated driving be achieved on the basis of an in-vehicle sensor setup and a localization map? Along this guideline, the key contributions were made. This led to the development of a model-based design framework and a localization system which is intended to meet the requirements imposed by the automated operation of vehicle systems. This explicitly includes the requirement of providing an enhanced level of robustness against adverse environment conditions frequently encountered in urban areas. It was argued that the utilization of model knowledge supports the design of localization systems that are intended to meet specified performance requirements while avoiding overly pessimistic design decisions.

The key contributions, a discussion of the overall results and concluding remarks are summarized in the following section. This section allows to give an answer to the main research question. A perspective for future research directions is provided in the consecutive section.

### Conclusion

Contributions to increasing the robustness of localization system have been made throughout this thesis. The developed system architecture entails robust map matching, analytical redundancy as well as recovery and consistency checking capability as core components.

Data association remains a fundamental problem in automated driving and vehicle localization in specific. In the context of map-relative vehicle localization, data association occurs when the correspondences between sensor observations and the localization map or between several sets of sensor observations have to be determined. This task becomes more challenging as the inconsistencies between the data sources increases. This is often given for outdated localization maps and respective sensor observations. Furthermore, urban environments are heterogeneous that can be encountered when driving from an area with dense vegetation to a road with house wall on both sides of the road. A comparison of different matching algorithms was conducted in this thesis in order to determine a well suited class of algorithms for this challenging task. On this basis, it was proposed to utilize spectral registration based on the Fourier-Mellin transformation (FMT). This approach was based on the fast Fourier transformation (FFT) for which efficient software and hardware implementations for real-time operation are available which distinguishes it from other correlation-based registration procedures. The selection was further motivated by the observation that the algorithm has the potential to deal with inconsistent input data and varying structural properties of the vehicle surrounding. This approach

was adapted to the requirements of automated driving (AD) (*rq1*, *kc1*, Section 1). For the task of vehicle pose initialization from inaccurate pose information, a robust spectral registration algorithm was developed (*rq1*, *kc1*). The developed algorithm led to a significant increase in the matching rate from approximately 80 % to above 99 %. Different parameterizations of the spectral registration algorithm lead to significant changes in the algorithm characteristics. This enables a broad utilization of the developed registration algorithm. Thus, the robust registration algorithm was used for initialization of the localization system (*rq3*, *kc1*) and for map matching during normal operation beside other applications.

The required localization accuracy has great impact on the localization system design. Consequently, the development of the localization system started with an analysis of the localization requirements. The resulting accuracy of approximately 0.2 m translational accuracy was used as basis for the derivation of localization system design parameters. Due to complex interrelations between design parameters, the urban operation environment and the resulting localization accuracy, statistical model were developed to support design decisions (*rq2*, *kc2*) and gain a deeper understanding of localization systems. The model was obtained by error propagation from uncertain sensor observations to the end of the whole signal processing chain of the localization system. Important model parameters are the map matching accuracy and the matching rate. Beside other applications, this approach was utilized to determine minimum feature detection rates and accuracies. Additionally, the matching rate requirements were utilized to support threshold selection for the detection of invalid map matching results based on the signal-to-noise ratio (SNR). By enabling the quantification of expert knowledge, utilization of the statistical model framework contributed to an enhancement of the localization robustness. Simultaneously, the gain in robustness can be traded against computational efficiency, e.g. by increasing or decreasing the accuracy of the map matching algorithm.

The model-based design step underlined the importance of accurate motion estimation for cases of prolonged map matching failure sequences. Therefore, a light detection and ranging (LiDAR) scan based odometry was developed on the basis of spectral registration (*rq6*, *kc1*). Experimental evaluations on the basis of challenging urban data sets resulted in an average position drift of 0.21%. For data sets with high numbers of sharp turns, the scan-based odometry outperformed classical wheel odometry. It was shown that the mounting height of the LiDAR sensor has a decisive influence on the motion estimation accuracy. For low mounting heights, occlusions led to a decreased detection range which had negative influence on the turning angle estimation accuracy. LiDAR scans are the only required inputs and thereby, the approach is well suited as a source of motion estimates independent of the wheel odometry that might suffer from wheel slippage. For vehicle system operation in previously unmapped areas like parking garages, a SLAM framework was developed. With different parameterizations of the spectral registration algorithm, consecutive LiDAR scans could be matched and loop closure hypotheses be generated. The resulting mapping accuracy was shown to provide sufficiently accu-

rate results in challenging urban environments. By utilization of the same matching algorithm throughout several processing steps in the localization system, a low number of different algorithms could be achieved. This might have positive influence on the system complexity and could thereby form a further contribution to robustness improvement.

The developed multilayer adaptive Monte Carlo localization (ML-AMCL) framework (*rq5*, *kc1*) enabled the generation of global vehicle pose estimates. Separate instance of a localization algorithm were executed in parallel and subsets of the current LiDAR measurement were used as inputs. Consistency checking against sequences of odometry measurements then enabled the identification of consistent localization results which were fused to the final localization output. Thereby, the sensitivity against inconsistent sensor observations could be reduced relative to the standard adaptive Monte Carlo localization (AMCL) implementation. A pose refinement step on the basis of spectral registration and pole-based map matching was then executed to obtain an accurate estimate of the vehicle pose. Undetected invalid map matching results can cause localization system failures and shall be avoided. In areas with sparse structures like expansive intersections, the spectral registration procedure tends to show performance degradation. Therefore, analytical redundancy and consistency checking capability were introduced to the localization system to increase the availability of valid matching results and detect invalid results (*rq4*, *kc1*). The selection of matching algorithms for densely and sparsely structured areas has proven to be a promising approach, due to the orthogonality of the respective approaches. This led to the incorporation of a pole-based map matching approach which was well suited for localization in intersection areas. Experimental evaluations showed, that the resulting localization algorithm gives almost identical results for significantly outdated and updated localization maps. This result indicates the robustness of the developed approach against outdated localization maps.

In conclusion, the main research question can be positively answered on the basis of the previously described contributions. A global estimate of the vehicle pose was provided by ML-AMCL. Localization system initialization and recovery can be reliably performed by utilization of the proposed matching procedure. The newly developed localization system was then used for increasing the vehicle pose estimation accuracy. Challenges of urban environments were addressed by robust data association and consistency checking. The goal was not to achieve the centimeter-precision of a differential global positioning system (dGPS) system, but to meet the localization accuracy requirements. This was enabled by the model-based design of localization system parameters.

## Future work

The development of automated vehicle systems begins to speed up as more research institutes and companies start new related activities. Connectivity and increased environment perception capabilities, fostered by machine learning, are two influential

directions of development. Both have an impact on future research perspectives for vehicle localization.

Connectivity might open up new interesting research perspectives for automated driving in general and vehicle localization in particular. Currently, communication technology starts to enter modern vehicle systems. Therefore, it is argued that connected vehicle systems can and will greatly contribute to further increase the performance of today's localization systems. By enabling the communication between different vehicles and with infrastructure elements, an additional source of vehicle pose information can be made accessible. First research results strengthen this hypothesis. The introduction of communication technology to the automated vehicle system in general and the localization system in specific, leads to new challenging questions regarding system robustness and functional safety. These new challenges include data security which is mandatory for avoiding external attacks of any kind.

Environment perception is likely to further improve throughout the next years by the introduction of innovative sensor technologies and their combination. This development is accompanied by new results in the area of signal processing. The latter is more frequently left to machine learning based approaches, allowing for a better situational awareness as well as an increased amount of extractable information from uncertain sensor observations. Consequently, the dependence on additional information from digital maps that motivated the importance of vehicle localization in Section 1 might be reduced. Thus, this trend could lead to reduced requirements on the localization accuracy. Thereby, the focus is shifted from highly accurate vehicle pose estimation to the fusion of heterogeneous information sources, their temporal synchronization and calibration.

## Literature

### Publications

- [RMZ17] Jan Rohde, Holger Mielenz, and J. Marius Zöllner. *Innovative Produkte und Dienstleistungen in der Mobilität*, chapter Perzeption für robuste Fahrzeuglokalisierung, pages 1–11. Springer Fachmedien Wiesbaden GmbH, 2017.
- [RSMZ16] Jan Rohde, Jan Erik Stellet, Holger Mielenz, and J. Marius Zöllner. Localization accuracy estimation with application to perception design. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4777–4783, 2016.
- [RVMZ16] Jan Rohde, Benjamin Völz, Holger Mielenz, and J. Marius Zöllner. Precise vehicle localization in dense urban environments. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 853–858, 2016.
- [RMH<sup>+</sup>16] Jan Rohde, Holger Mielenz, Carsten Hasberg, Frank Niewels, and J. Marius Zöllner. Towards robust localization in urban environments. In *International VDI Conference on Automated Driving*, 2016.
- [RJMZ16] Jan Rohde, Inga Jatzkowski, Holger Mielenz, and J. Marius Zöllner. Vehicle pose estimation in cluttered urban environments using multilayer adaptive monte carlo localization. In *2016 19th International Conference on Information Fusion (FUSION)*, pages 1774–1779, 2016.
- [RSMZ15] Jan Rohde, Jan Erik Stellet, Holger Mielenz, and J. Marius Zöllner. Model-based derivation of perception accuracy requirements for vehicle localization in urban environments. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC)*, pages 712–718, 2015.
- [MHR<sup>+</sup>13] Holger Mielenz, Christian Heigele, Jan Rohde, Philipp Lehner, Frank Beruscha, and Jörg Heckel. Homezone - Automatische Querführungsassistentz für alltäglich wiederkehrende Fahrhandlungen des ruhenden Verkehrs. In *14. Braunschweiger Symposium (AAET), Automatisierungs-, Assistenzsysteme und eingebettete Systeme für Transportmittel*, 2013.

**Supervised theses**

- [Jat16] Inga Jatzkowski. LiDAR reference localization for automated driving. Master's thesis, Leibniz University Hannover, 2016.
- [Lat16] Liana Lattin. Radar-based mapping in urban environments. Master's thesis, Hochschule Esslingen, 2016.
- [Höf16] Marc Oliver Höfling. Entwicklung von Methoden zur Erkennung und Beseitigung von semi-statischen Objekten in Lokalisierungskarten für urbanes automatisiertes Fahren. Bachelor's thesis, Duale Hochschule Baden-Württemberg Stuttgart, 2016.
- [Jes16] Markus Jesenski. Entwicklung von Methoden zur Fahrzeuglokalisierung auf Basis landmarkenbasierter Karten für urbanes automatisiertes Fahren. Bachelor's thesis, Duale Hochschule Baden-Württemberg Stuttgart, 2016.
- [Meh16] Benedikt Mehler. Implementierung eines zweidimensionalen Map-Matching Algorithmus zur Fahrzeuglokalisierung für urbanes hochautomatisiertes Fahren. Bachelor's thesis, Duale Hochschule Baden-Württemberg Ravensburg, 2016.
- [Vol15] Laura Volkenborn. Entwicklung und Simulation verschiedener Lösungsansätze zur exakten Selbstlokalisierung von hochautomatisierten Fahrzeugen in urbanem Umfeld. Bachelor's thesis, Duale Hochschule Baden-Württemberg Stuttgart, 2015.

## References

- [AAHAJ10] T. S. Abuhashim, M. F. Abdel-Hafez, and M. A. Al-Jarrah. Building a robust integrity monitoring algorithm for a low cost gps-aided-ins system. *International Journal of Control, Automation and Systems*, 8(5):1108–1122, 2010.
- [AAS15] A. Alba and E. R. Arce-Santana. Robust rigid registration by scanning multiple phase correlation peaks. *Optik-International Journal for Light and Electron Optics*, 126(24):5115–5118, 2015.
- [AB13] S. Anderson and T. D. Barfoot. Ransac for motion-distorted 3d visual sensors. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2093–2099, Nov 2013.
- [AGDT<sup>+</sup>14] P. Agarwal, G. Grisetti, G. Diego Tipaldi, L. Spinello, W. Burgard, and C. Stachniss. Experimental analysis of dynamic covariance scaling for robust map optimization under bad initial estimates. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 3626–3631, May 2014.
- [AHB87] K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(5):698–700, Sept 1987.
- [ATS<sup>+</sup>13] P. Agarwal, G. Tipaldi, L. Spinello, C. Stachniss, and W. Burgard. Robust map optimization using dynamic covariance scaling. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pages 62–69, May 2013.
- [Bad09] H. Badino. *Binocular ego-motion estimation for automotive applications*. PhD thesis, Goethe Universität in Frankfurt am Main, 2009.
- [BB13] H. Bülow and A. Birk. Spectral 6dof registration of noisy 3d range data with partial overlap. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(4):954–969, April 2013.
- [BBH14] A. Bansal, H. Badino, and D. Huber. Understanding how camera configuration and environmental conditions affect appearance-based localization. In *2014 IEEE Intelligent Vehicles Symposium Proceedings*, pages 800–807, June 2014.
- [BD09] Y. Boers and H. Driessen. A note on bounds for target tracking with pdj1. *IEEE Transactions on Aerospace and Electronic Systems*, 45(2):640–646, April 2009.

- [BDW06] T. Bailey and H. Durrant-Whyte. Simultaneous localization and mapping (slam): part ii. *Robotics Automation Magazine, IEEE*, 13(3):108–117, Sept 2006.
- [Bei14] M. Beinhofer. *Landmark Placement for Mobile Robot Navigation*. PhD thesis, University of Freiburg, Department of Computer Science, September 2014.
- [BF00] P. Batchelor and J. Fitzpatrick. A study of the anisotropically weighted procrustes problem [optical image-guided surgery application]. In *Mathematical Methods in Biomedical Image Analysis, 2000. Proceedings. IEEE Workshop on*, pages 212–218, 2000.
- [BG97] M. Betke and L. Gurvits. Mobile robot localization using landmarks. *Robotics and Automation, IEEE Transactions on*, 13(2):251–263, Apr 1997.
- [BGU13] M. A. Brubaker, A. Geiger, and R. Urtasun. Lost! leveraging the crowd for probabilistic visual self-localization. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3057–3064, June 2013.
- [BM92] P. Besl and N. D. McKay. A method for registration of 3-d shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(2):239–256, Feb 1992.
- [BMB11] M. Beinhofer, J. Muller, and W. Burgard. Near-optimal landmark selection for mobile robot navigation. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 4744–4749, May 2011.
- [BMKB13] M. Beinhofer, J. Muller, A. Krause, and W. Burgard. Robust landmark selection for mobile robot navigation. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 3637–2643, Nov 2013.
- [BS03] P. Biber and W. Strasser. The normal distributions transform: a new approach to laser scan matching. In *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, volume 3, pages 2743–2748 vol.3, Oct 2003.
- [BS06] P. Biber and W. Strasser. nscan-matching: simultaneous matching of multiple scans and application to slam. In *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pages 2270–2276, May 2006.

- [BSLK04] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.
- [BSMW14] B. Bellekens, V. Spruyt, and R. B. Maarten Weyn. A survey of rigid 3d pointcloud registration algorithms. In *Fourth International Conference on Ambient Computing, Applications, Services and Technologies, Proceedings*, pages 8–13. IARA, 2014.
- [BW16] M. Buczko and V. Willert. How to distinguish inliers from outliers in visual odometry for high-speed automotive applications. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, pages 478–483, June 2016.
- [CB09] L. Carlone and B. Bona. A comparative study on robust localization: Fault tolerance and robustness test on probabilistic filters for range-based positioning. In *Advanced Robotics, 2009. ICAR 2009. International Conference on*, pages 1–8, June 2009.
- [CC09] A. Censi and S. Carpin. Hsm3d: Feature-less global 6dof scan-matching in the hough/radon domain. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 3899–3906, May 2009.
- [CCD14] L. Carlone, A. Censi, and F. Dellaert. Selecting good measurements via l1 relaxation: A convex approach for robust estimation over graphs. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pages 2667–2674, Sept 2014.
- [CDD94] Q.-S. Chen, M. Defrise, and F. Deconinck. Symmetric phase-only matched filtering of fourier-mellin transforms for image registration and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(12):1156–1168, Dec 1994.
- [Cen07a] A. Censi. An accurate closed-form estimate of icp’s covariance. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 3167–3172, April 2007.
- [Cen07b] A. Censi. On achievable accuracy for range-finder localization. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 4170–4175, April 2007.
- [Cho14] J. Choi. Hybrid map-based slam using a velodyne laser scanner. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 3082–3087, Oct 2014.

- [CIG05] A. Censi, L. Iocchi, and G. Grisetti. Scan matching in the hough domain. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 2739–2744, April 2005.
- [CM87] E. D. Castro and C. Morandi. Registration of translated and rotated images using finite fourier transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(5):700–703, Sept 1987.
- [CM91] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. In *Proceedings. 1991 IEEE International Conference on Robotics and Automation*, pages 2724–2729, Apr 1991.
- [CT98] M. T. Chu and N. T. Trendafilov. On a differential equation approach to the weighted orthogonal procrustes problem. *Statistics and Computing*, 8(2):125–133, 1998.
- [DB13] V. Drevelle and P. Bonnifait. Localization confidence domains via set inversion on short-term trajectory. *IEEE Transactions on Robotics*, 29(5):1244–1256, Oct 2013.
- [DFBT99] F. Dellaert, D. Fox, W. Burgard, and S. Thrun. Monte carlo localization for mobile robots. In *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No.99CH36288C)*, volume 2, pages 1322–1328 vol.2, 1999.
- [DRN<sup>+</sup>14] H. Deusch, S. Reuter, D. Nuss, M. Fritzsche, K. Dietmayer, et al. Multi-sensor self-localization based on maximally stable extremal regions. In *2014 IEEE Intelligent Vehicles Symposium Proceedings*, pages 555–560. IEEE, 2014.
- [DWB06] H. Durrant-Whyte and T. Bailey. Simultaneous localization and mapping: part i. *Robotics Automation Magazine, IEEE*, 13(2):99–110, June 2006.
- [DWS<sup>+</sup>04] R. Dearden, T. Willeke, R. Simmons, V. Verma, F. Hutter, and S. Thrun. Real-time fault detection and situational awareness for rovers: report on the mars technology program task. In *Aerospace Conference, 2004. Proceedings. 2004 IEEE*, volume 2, pages 826–840 Vol.2, March 2004.
- [EHH13] H. ElSawy, E. Hossain, and M. Haenggi. Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey. *IEEE Communications Surveys Tutorials*, 15(3):996–1019, Third 2013.

- [FB81] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, jun 1981.
- [Fit03] A. W. Fitzgibbon. Robust registration of 2d and 3d point sets. *Image and Vision Computing*, 21(13):1145–1153, 2003.
- [Fox01] D. Fox. Kld-sampling: Adaptive particle filters. In *NIPS*, volume 14, pages 713–720, 2001.
- [FS12] F. Fraundorfer and D. Scaramuzza. Visual odometry : Part ii: Matching, robustness, optimization, and applications. *IEEE Robotics Automation Magazine*, 19(2):78–90, June 2012.
- [FSK06] F. Farshidi, S. Sirouspour, and T. Kirubarajan. Optimal positioning of multiple cameras for object recognition using cramer-rao lower bound. In *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, pages 934–939, May 2006.
- [FTY<sup>+</sup>15] A. Fujii, M. Tanaka, H. Yabushita, T. Mori, and T. Odashima. Detection of localization failure using logistic regression. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 4313–4318, Sept 2015.
- [FWM98] J. Fitzpatrick, J. West, and J. Maurer, C.R. Predicting error in rigid-body point-based registration. *Medical Imaging, IEEE Transactions on*, 17(5):694–702, Oct 1998.
- [FZB02a] H. Foroosh, J. B. Zerubia, and M. Berthod. Extension of phase correlation to subpixel registration. *IEEE Transactions on Image Processing*, 11(3):188–200, Mar 2002.
- [FZB02b] H. Foroosh, J. B. Zerubia, and M. Berthod. Extension of phase correlation to subpixel registration. *IEEE transactions on image processing*, 11(3):188–200, 2002.
- [GB05] A. Gning and P. Bonnifait. Dynamic vehicle localization using constraints propagation techniques on intervals a comparison with kalman filtering. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 4144–4149, April 2005.
- [GBFK98] J. S. Gutmann, W. Burgard, D. Fox, and K. Konolige. An experimental comparison of localization methods. In *Proceedings. 1998 IEEE/RSJ International Conference on Intelligent Robots*

- and Systems. Innovations in Theory, Practice and Applications (Cat. No.98CH36190)*, volume 2, pages 736–743 vol.2, Oct 1998.
- [Gen] Genesys. *Genesys Adma*.
- [Ger] B. P. Gerkey. amcl: Package summary. <http://wiki.ros.org/amcl>. last accessed on April 26 2017.
- [GF02] J. S. Gutmann and D. Fox. An experimental comparison of localization methods continued. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 454–459 vol.1, 2002.
- [GI96] R. Greiner and R. Isukapalli. Learning to select useful landmarks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(3):437–449, Jun 1996.
- [GKSK11] G. Grisetti, R. Kümmerle, H. Strasdat, and K. Konolige. g2o: A general framework for (hyper) graph optimization. Technical report, Technical report, 2011.
- [GL10] C. Glennie and D. D. Lichti. Static calibration and analysis of the velodyne hdl-64e s2 for high accuracy mobile scanning. *Remote Sensing*, 2(6):1610–1624, 2010.
- [GLU12] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3354–3361, June 2012.
- [GMBN14] C. Gläser, T. P. Michalke, L. Bürkle, and F. Niewels. Environment perception for inner-city driver assistance and highly-automated driving. In *2014 IEEE Intelligent Vehicles Symposium Proceedings*, pages 1270–1275, June 2014.
- [Gos12] A. A. Goshtasby. *Image registration: Principles, tools and methods*. Springer Science and Business Media, 2012.
- [Gus00] F. Gustafsson. *Adaptive filtering and change detection*, volume 1. Wiley New York, 2000.
- [GVI13] W. S. Grant, R. C. Voorhies, and L. Itti. Finding planes in lidar point clouds for real-time registration. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4347–4354, Nov 2013.

- [GY03] M. Greenspan and M. Yurick. Approximate k-d tree search for efficient icp. In *Fourth International Conference on 3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings.*, pages 442–448, Oct 2003.
- [HD07] M. Huang and S. Dey. Stability of kalman filtering with markovian packet losses. *Automatica*, 43(4):598 – 607, 2007.
- [HJL<sup>+</sup>89] R. Haralick, H. Joo, D. Lee, S. Zhuang, V. Vaidya, and M. Kim. Pose estimation from corresponding point data. *Systems, Man and Cybernetics, IEEE Transactions on*, 19(6):1426–1446, Nov 1989.
- [HMHS14] C. Heigle, H. Mielenz, J. Heckel, and D. Schramm. Accurate and fast localization in unstructured environment based on shape context keypoints. In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–7, July 2014.
- [Hor87] B. K. Horn. Closed-form solution of absolute orientation using unit quaternions. *JOSA A*, 4(4):629–642, 1987.
- [HRFT04] M. Hernandez, B. Ristic, A. Farina, and L. Timmoneri. A comparison of two crame acute;r-rao bounds for nonlinear filtering with pdj1. *Signal Processing, IEEE Transactions on*, 52(9):2361–2370, Sept 2004.
- [HS09] S. Hochdorfer and C. Schlegel. Landmark rating and selection according to localization coverage: Addressing the challenge of lifelong operation of slam in service robots. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 382–387, Oct 2009.
- [Hub11] P. J. Huber. *Robust statistics*. Springer, 2011.
- [Ibe] Ibeo. Ibeo. <https://www.ibeo-as.com/ibeoreference/reference-sensor-system/>. Accessed: 2017-05-30.
- [Int14] S. International. Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems, 01 2014.
- [ISO10] I. ISO. Ieee, systems and software engineering–vocabulary. *ISO/IEC/IEEE 24765: 2010 (E) Piscataway, NJ: IEEE computer society, Tech. Rep.*, 2010.
- [ISO12] D. ISO. 8855: Straßenfahrzeuge-fahrzeugdynamik und fahrverhalten. *ISO/IEC/IEEE 8855: Straßenfahrzeuge-Fahrzeugdynamik und Fahrverhalten*, 2012.

- [KASS05] W. S. Kim, A. I. Ansar, R. D. Steele, and R. C. Steinke. Performance analysis and validation of a stereo vision system. In *2005 IEEE International Conference on Systems, Man and Cybernetics*, volume 2, pages 1409–1416, Oct 2005.
- [KBN04] S. Kumar, M. Biswas, and T. Q. Nguyen. Efficient phase correlation motion estimation using approximate normalization. In *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004.*, volume 2, pages 1727–1730 Vol.2, Nov 2004.
- [KG17] R. Karlsson and F. Gustafsson. The future of automotive localization algorithms: Available, reliable, and scalable localization: Anywhere and anytime. *IEEE Signal Processing Magazine*, 34(2):60–69, March 2017.
- [KGL10] B. Kitt, A. Geiger, and H. Lategahn. Visual odometry based on stereo image sequences with ransac-based outlier rejection scheme. In *2010 IEEE Intelligent Vehicles Symposium*, pages 486–492, June 2010.
- [KGS<sup>+</sup>11] R. Kummerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard. G2o: A general framework for graph optimization. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 3607–3613, May 2011.
- [KH90] B. V. Kumar and L. Hassebrook. Performance measures for correlation filters. *Applied optics*, 29(20):2997–3006, 1990.
- [KKM03] S. Kaneko, T. Kondo, and A. Miyamoto. Robust matching of 3d contours using iterative closest point algorithm improved by m-estimation. *Pattern Recognition*, 36(9):2041–2047, 2003.
- [KS91] M. A. Koschat and D. F. Swayne. A weighted procrustes criterion. *Psychometrika*, 56(2):229–239, 1991.
- [LA99] H. Lester and S. R. Arridge. A survey of hierarchical non-linear medical image registration. *Pattern recognition*, 32(1):129–149, 1999.
- [LCI<sup>+</sup>04] B. Lussier, R. Chatila, F. Ingrand, M. Killijian, and D. Powell. On fault tolerance and robustness in autonomous systems, 2004.
- [LCJS04] Y. Lu, E. G. Collins Jr, and M. F. Selekwa. Parity relation based fault detection, isolation and reconfiguration for autonomous ground vehicle localization sensors. Technical report, DTIC Document, 2004.

- [LCN14] Y. Latif, C. Cadena, and J. Neira. Robust graph slam back-ends: A comparative analysis. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2683–2690, Sept 2014.
- [LDW91] J. Leonard and H. Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *Robotics and Automation, IEEE Transactions on*, 7(3):376–382, Jun 1991.
- [LL02] G. M. C. L. Lucchese, G. Doretto. A frequency domain technique for range data registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(11):1468–1484, Nov. 2002.
- [LM97] F. Lu and E. Milios. Robot pose estimation in unknown environments by matching 2d range scans. *Journal of Intelligent and Robotic Systems*, 18(3):249–275, 1997.
- [LMT07] J. Levinson, M. Montemerlo, and S. Thrun. Map-based precision vehicle localization in urban environments. In *Proceedings of Robotics: Science and Systems*, Atlanta, GA, USA, June 2007.
- [LRS07] R. Lerner, E. Rivlin, and I. Shimshoni. Landmark selection for task-oriented navigation. *Robotics, IEEE Transactions on*, 23(3):494–505, June 2007.
- [LS12] H. Lategahn and C. Stiller. City gps using stereo vision. In *Vehicular Electronics and Safety (ICVES), 2012 IEEE International Conference on*, pages 1–6, July 2012.
- [LSL76] R. W. Lissitz, P. H. Schönemann, and J. C. Lingo. A solution to the weighted procrustes problem in which the transformation is in agreement with the loss function. *Psychometrika*, 41(4):547–550, 1976.
- [LSZS13] H. Lategahn, M. Schreiber, J. Ziegler, and C. Stiller. Urban localization with camera and inertial measurement unit. In *Intelligent Vehicles Symposium (IV), 2013 IEEE*, pages 719–724, June 2013.
- [LT10] J. Levinson and S. Thrun. Robust vehicle localization in urban environments using probabilistic maps. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 4372–4378, May 2010.
- [MBIGB09] O. L. Marchand, P. Bonnifait, J. Ibañez-Guzmán, and D. Bétaille. Vehicle localization integrity based on trajectory monitoring. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3453–3458, Oct 2009.

- [MDHGB10] D. Meyer-Delius, J. Hess, G. Grisetti, and W. Burgard. Temporary maps for robust localization in semi-static environments. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 5750–5755, Oct 2010.
- [MLD07] M. Magnusson, A. Lilienthal, and T. Duckett. Scan registration for autonomous mining vehicles using 3d-ndt. *Journal of Field Robotics*, 24(10):803–827, 2007.
- [MMR07] F. Mirzaei, A. Mourikis, and S. Roumeliotis. On the performance of multi-robot target tracking. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 3482–3489, April 2007.
- [MNL<sup>+</sup>09] M. Magnusson, A. Nuchter, C. Lorken, A. J. Lilienthal, and J. Hertzberg. Evaluation of 3d registration reliability and speed—a comparison of icp and ndt. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 3907–3912. IEEE, 2009.
- [MR06] A. I. Mourikis and S. I. Roumeliotis. Performance analysis of multi-robot cooperative localization. *IEEE Transactions on Robotics*, 22(4):666–681, Aug 2006.
- [MR11] F. Mirzaei and S. Roumeliotis. Optimal estimation of vanishing points in a manhattan world. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2454–2461, Nov 2011.
- [MS12] Y. Mo and B. Sinopoli. Kalman filtering with intermittent observations: Tail distribution and critical value. *IEEE Transactions on Automatic Control*, 57(3):677–689, March 2012.
- [MVS<sup>+</sup>15] M. Magnusson, N. Vaskevicius, T. Stoyanov, K. Pathak, and A. Birk. Beyond points: Evaluating recent 3d scan-matching algorithms. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3631–3637, May 2015.
- [NDH16] J. Nygård, V. Deleskog, and G. Hendeby. Safe fusion compared to established distributed fusion methods. In *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 265–271, Sept 2016.
- [NWBS01] R. Niu, P. Willett, and Y. Bar-Shalom. Matrix crlb scaling due to measurements of uncertain origin. *IEEE Transactions on Signal Processing*, 49(7):1325–1335, Jul 2001.
- [oEA14] U. N. D. of Economic and S. Affairs. *World Urbanization Prospects, the 2014 Revision: Highlights*. 2014.

- [Ols08] E. B. Olson. *Robust and efficient robotic mapping*. PhD thesis, Massachusetts Institute of Technology, 2008.
- [Ols09] E. B. Olson. Real-time correlative scan matching. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 4387–4393, May 2009.
- [OLT06] E. Olson, J. Leonard, and S. Teller. Fast iterative alignment of pose graphs with poor initial estimates. In *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, pages 2262–2269, May 2006.
- [ORD13] J. Oberlander, A. Roennau, and R. Dillmann. Hierarchical slam using spectral submap matching with opportunities for long-term operation. In *Advanced Robotics (ICAR), 2013 16th International Conference on*, pages 1–7, Nov 2013.
- [OWTL05] E. Olson, M. R. Walter, S. J. Teller, and J. J. Leonard. Single-cluster spectral graph partitioning for robotics applications. In *Robotics: Science and Systems*, pages 265–272, 2005.
- [Pad12] D. Padfield. Masked object registration in the fourier domain. *Image Processing, IEEE Transactions on*, 21(5):2706–2718, May 2012.
- [Pag54] E. Page. Continuous inspection schemes. *Biometrika*, 41(1/2):100–115, 1954.
- [PBB12] M. Pfingsthorn, A. Birk, and H. Bülow. Uncertainty estimation for a 6-dof spectral registration method as basis for sonar-based underwater 3d slam. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 3049–3054, May 2012.
- [PBS<sup>+</sup>10] M. Pfingsthorn, A. Birk, S. Schwertfeger, H. B $\tilde{A}$  $\frac{1}{4}$ low, and K. Pathak. Maximum likelihood mapping with spectral image registration. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 4282–4287, May 2010.
- [PBVP10] K. Pathak, A. Birk, N. Vaskevicius, and J. Poppinga. Fast registration based on noisy planes with unknown correspondences for 3-d mapping. *IEEE Transactions on Robotics*, 26(3):424–441, 2010.
- [PCS15] F. Pomerleau, F. Colas, and R. Siegwart. A review of point cloud registration algorithms for mobile robotics. *Foundations and Trends in Robotics*, 4(1):1–104, 2015.

- [Pin10] O. Pink. *Bildbasierte Selbstlokalisierung von Strassenfahrzeugen*. PhD thesis, Institute of Measurement and Control Systems, Karlsruhe Institute of Technology, Karlsruhe, Germany, #November# 2010.
- [PKB14] V. Peretroukhin, J. Kelly, and T. D. Barfoot. Optimizing camera perspective for stereo visual odometry. In *2014 Canadian Conference on Computer and Robot Vision*, pages 1–7, May 2014.
- [PMB09] O. Pink, F. Moosmann, and A. Bachmann. Visual features for vehicle localization and ego-motion estimation. In *Intelligent Vehicles Symposium, 2009 IEEE*, pages 254–260, June 2009.
- [PP12] K. B. Petersen and M. S. Pedersen. *The matrix cookbook*, 2012.
- [PSB06] C. Plagemann, C. Stachniss, and W. Burgard. Efficient failure detection for mobile robots using mixed-abstraction particle filters. In *European Robotics Symposium 2006*, pages 93–107. Springer, 2006.
- [RBH<sup>+</sup>15] M. Rapp, M. Barjenbruch, M. Hahn, J. Dickmann, and K. Dietmayer. Clustering improved grid map registration using the normal distribution transform. In *Intelligent Vehicles Symposium (IV), 2015 IEEE*, pages 249–254, June 2015.
- [RC96] B. Reddy and B. Chatterji. An fft-based technique for translation, rotation, and scale-invariant image registration. *Image Processing, IEEE Transactions on*, 5(8):1266–1271, Aug 1996.
- [RC11] R. B. Rusu and S. Cousins. 3d is here: Point cloud library (pcl). In *2011 IEEE International Conference on Robotics and Automation*, pages 1–4, May 2011.
- [RL00] T. Rupp and P. Levi. Optimized landmark arrangement for absolute localization—a practical approach. In *Intelligent Robots and Systems, 2000. (IROS 2000). Proceedings. 2000 IEEE/RSJ International Conference on*, volume 1, pages 448–453 vol.1, 2000.
- [RRKB11] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. Orb: An efficient alternative to sift or surf. In *2011 International Conference on Computer Vision*, pages 2564–2571, Nov 2011.
- [SASL13] J. Saarinen, H. Andreasson, T. Stoyanov, and A. Lilienthal. Normal distributions transform monte-carlo localization (ndt-mcl). In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 382–389, Nov 2013.

- [SB14] A. Schlichting and C. Brenner. Localization using automotive laser scanners and local pattern matching. In *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*, pages 414–419, June 2014.
- [SHK<sup>+</sup>14] J. Stellet, C. Heigle, F. Kuhnt, J. Zollner, and D. Schramm. Performance evaluation and statistical analysis of algorithms for ego-motion estimation. In *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*, pages 2125–2131, Oct 2014.
- [SHT09] A. Segal, D. Haehnel, and S. Thrun. Generalized-icp. In *Proceedings of Robotics: Science and Systems*, Seattle, USA, June 2009.
- [SJ06] P. Sundvall and P. Jensfelt. Fault detection for mobile robots using redundant positioning systems. In *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, pages 3781–3786, May 2006.
- [SKF13] M. Schreiber, C. Knoppel, and U. Franke. Laneloc: Lane marking based localization using highly accurate maps. In *Intelligent Vehicles Symposium (IV), 2013 IEEE*, pages 449–454, June 2013.
- [SMAL11] T. Stoyanov, M. Magnusson, H. Almqvist, and A. J. Lilienthal. On the accuracy of the 3d normal distributions transform as a tool for spatial representation. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 4080–4085, May 2011.
- [SP12] N. Sünderhauf and P. Protzel. Switchable constraints for robust pose graph slam. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1879–1884, Oct 2012.
- [SRA15] X. Shen, D. Rus, and M. H. Ang. Bounds for kalman filtering with intermittent observations. In *2015 European Control Conference (ECC)*, pages 2842–2846, July 2015.
- [SS01] L. Shapiro and G. Stockman. Textbook: computer vision, 2001.
- [SSB09] H. Strasdat, C. Stachniss, and W. Burgard. Which landmark is useful? learning selection policies for navigation in unknown environments. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 1410–1415, May 2009.
- [SSF<sup>+</sup>04] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry. Kalman filtering with intermittent observations. *Automatic Control, IEEE Transactions on*, 49(9):1453–1464, Sept 2004.

- [SSSD06] P. Sala, R. Sim, A. Shokoufandeh, and S. Dickinson. Landmark selection for vision-based navigation. *Robotics, IEEE Transactions on*, 22(2):334–349, April 2006.
- [TAT95] K. A. Tarabanis, P. K. Allen, and R. Y. Tsai. A survey of sensor planning in computer vision. *IEEE Transactions on Robotics and Automation*, 11(1):86–104, Feb 1995.
- [TB13] C. H. Tong and T. D. Barfoot. Gaussian process gauss-newton for 3d laser-based visual odometry. In *2013 IEEE International Conference on Robotics and Automation*, pages 5204–5211, May 2013.
- [TBA14] G. Tipaldi, M. Braun, and K. Arras. Flirt: Interest regions for 2d range data with applications to robot navigation. In O. Khatib, V. Kumar, and G. Sukhatme, editors, *Experimental Robotics*, volume 79 of *Springer Tracts in Advanced Robotics*, pages 695–710. Springer Berlin Heidelberg, 2014.
- [TBCC07] C. Tessier, M. Berducat, R. Chapuis, and F. Chausse. A new landmark and sensor selection method for vehicle localization and guidance. In *2007 IEEE Intelligent Vehicles Symposium*, pages 123–129, June 2007.
- [TBF05] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005.
- [TENAH<sup>+</sup>14] N. Tmazirte, M. El Najjar, J. Al Hage, C. Smaili, and D. Pomorski. Fast multi fault detection exclusion approach for gnss integrity monitoring. In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–6, July 2014.
- [Thr98] S. Thrun. Finding landmarks for mobile robot navigation. In *Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on*, volume 2, pages 958–963 vol.2, May 1998.
- [TKZS16] Ö. S. Tas, F. Kuhnt, J. M. Zöllner, and C. Stiller. Functional system architectures towards fully automated driving. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, pages 304–309, June 2016.
- [TMN98] P. Tichavsky, C. H. Muravchik, and A. Nehorai. Posterior cramer-rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on Signal Processing*, 46(5):1386–1396, May 1998.
- [TMZIUMGS07] R. Toledo-Moreo, M. A. Zamora-Izquierdo, B. Ubeda-Minarro, and A. F. Gomez-Skarmeta. High-integrity imm-ekf-based road

- vehicle navigation with low-cost gps/sbas/ins. *IEEE Transactions on Intelligent Transportation Systems*, 8(3):491–511, Sept 2007.
- [TSG08] D. Tornqvist, T. B. Schon, and F. Gustafsson. Detecting spurious features using parity space. In *Control, Automation, Robotics and Vision, 2008. ICARCV 2008. 10th International Conference on*, pages 353–358, Dec 2008.
- [vdBAG11] J. van den Berg, P. Abbeel, and K. Goldberg. LQG-MP: Optimized path planning for robots with motion uncertainty and imperfect state information. *International Journal of Robotic Research*, 30:895–913, 2011.
- [Vela] Velodyne. Velodyne. <http://velodynelidar.com/hdl-64e.html>. Accessed: 2017-05-30.
- [Velb] Velodyne. *Velodyne HDL-64E*.
- [VL07] U. Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007.
- [WBK<sup>+</sup>15] K. Werber, M. Barjenbruch, J. Klappstein, J. Dickmann, and C. Waldschmidt. Roughcough - a new image registration method for radar based vehicle self-localization. In *Information Fusion (Fusion), 2015 18th International Conference on*, pages 1533–1541, July 2015.
- [WBSS04] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.
- [WCR92] J. Weng, P. Cohen, and N. Rebibo. Motion and structure estimation from stereo image sequences. *Robotics and Automation, IEEE Transactions on*, 8(3):362–382, Jun 1992. test.
- [WHLS15] H. Winner, S. Hakuli, F. Lotz, and C. Singer. Handbuch fahrerassistenzsysteme: Grundlagen, komponenten und systeme für aktive sicherheit und komfort (3. überarbeitete und ergänzte auflage aufl.), 2015.
- [WRK<sup>+</sup>15] K. Werber, M. Rapp, J. Klappstein, M. Hahn, J. Dickmann, K. Dietmayer, and C. Waldschmidt. Automotive radar gridmap representations. In *Microwaves for Intelligent Mobility (ICMIM), 2015 IEEE MTT-S International Conference on*, pages 1–4, April 2015.

- [WSG10] K. M. Wurm, C. Stachniss, and G. Grisetti. Bridging the gap between feature- and grid-based slam. *Robotics and Autonomous Systems*, 58(2):140–148, 2010.
- [WSV91] M. W. Walker, L. Shao, and R. A. Volz. Estimating 3-d location parameters using dual number quaternions. *CVGIP: image understanding*, 54(3):358–367, 1991.
- [WWBN15] S. Worrall, J. Ward, A. Bender, and E. M. Nebot. Gps/gnss consistency in a multi-path environment and during signal outages. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 2505–2511, Sept 2015.
- [YX11] K. You and L. Xie. Minimum data rate for mean square stabilizability of linear systems with markovian packet losses. *IEEE Transactions on Automatic Control*, 56(4):772–785, April 2011.
- [YYM<sup>+</sup>15] K. Yoneda, C. Yang, S. Mita, T. Okuya, and K. Muto. Urban road localization by using multiple layer map matching and line segment matching. In *2015 IEEE Intelligent Vehicles Symposium (IV)*, pages 525–530, June 2015.
- [ZBIG16] C. Zinoune, P. Bonnifait, and J. Ibanez-Guzman. Sequential fdia for autonomous integrity monitoring of navigation maps on board vehicles. *Intelligent Transportation Systems, IEEE Transactions on*, 17(1):143–155, Jan 2016.
- [ZCO<sup>+</sup>11] Z. Zhu, H. P. Chiu, T. Oskiper, S. Ali, R. Hadsell, S. Samarasekera, and R. Kumar. High-precision localization using visual landmarks fused with range data. In *CVPR 2011*, pages 81–88, June 2011.
- [ZF03] B. Zitová and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977 – 1000, 2003.
- [Zha94] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International journal of computer vision*, 13(2):119–152, 1994.
- [ZLS<sup>+</sup>14] J. Ziegler, H. Lategahn, M. Schreiber, C. Keller, C. Knoppel, J. Hipp, M. Haueis, and C. Stiller. Video based localization for bertha. In *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*, pages 1231–1238, June 2014.
- [ZS15] J. Zhang and S. Singh. Visual-lidar odometry and mapping: low-drift, robust, and fast. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2174–2181, May 2015.

- 
- [ZS16] J. Zhang and S. Singh. Low-drift and real-time lidar odometry and mapping. *Autonomous Robots*, pages 1–16, 2016.