

**Automatische Erzeugung langzeitverfügbarer
Punktmerkmalskarten zur robusten
Lokalisierung mit Multi-Kamera-Systemen für
automatisierte Fahrzeuge**

Zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

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

angenommene

Dissertation

von

M. Sc. Marc Sons

Tag der mündlichen Prüfung: 29. Oktober 2019

Referent: Prof. Dr.-Ing. Christoph Stiller

Korreferent: Prof. Dr.-Ing. Michael Heizmann

Danksagung

Die vorliegende Dissertation resultiert aus meiner Tätigkeit am Institut für Mess- und Regelungstechnik (MRT) des Karlsruher Instituts für Technologie (KIT). An erster Stelle bedanke ich mich bei Herrn Prof. Dr.-Ing. Christoph Stiller für die vielseitige Unterstützung und die Freiheiten hinsichtlich der Orientierung meiner Forschung. Ohne diese vorteilhaften Rahmenbedingungen wäre diese Arbeit so nicht entstanden. Des Weiteren danke ich Herrn Prof. Dr.-Ing. Michael Heizmann für sein Interesse an dieser Arbeit und der Übernahme des Korreferats.

Ich danke allen Kolleginnen und Kollegen für die äußerst angenehme und entspannte Arbeitsatmosphäre, die inspirierenden gemeinsamen Sommerseminare und die erhaltene Unterstützung. Außerordentlicher Dank geht an Henning Lategahn für die richtungswisende Betreuung und Tobias Strauß für die tiefgründigen fachlichen Diskussionen zu Beginn meiner Promotionszeit. Im gleichen Maße bedanke ich mich bei Martin Lauer für die herausragende wissenschaftliche Betreuung und die Unterstützung beim Schreiben dieser Arbeit sowie verschiedener Veröffentlichungen. Für das Korrekturlesen dieser Arbeit bedanke ich mich bei Johannes Beck, Florian Wirth, Sascha Wirges, Julius Kümmerle, Christoph Burger und Christian Kinzig. Auch danke ich den Teilnehmern meiner Besprechungsgruppe, die mich in den wöchentlichen Gruppentreffen durch das kritische Hinterfragen meiner Methoden und die vielen hilfreichen Hinweise außerordentlich unterstützt haben. Philipp Bender und Ömer Sahin Tas danke ich für die administrative Rückendeckung in den anstrengenden Tagen zur Vorbereitung der zahlreichen Live-Demonstrationen. Günter Barth möchte ich danken für die professionelle Instandhaltung meines Fahrrads, das mich Tag für Tag zuverlässig ins Büro getragen hat. Ebenso gilt mein Dank dem gesamten Werkstatt- und Sekretariatspersonal.

Schlussendlich bedanke ich mich ganz besonders bei meiner Freundin und meinen Eltern, die mich stets bedingungslos unterstützt haben und mir den notwendigen Zuspruch gegeben haben, um auch in den anstrengsten Phasen meiner Zeit als Doktorand meine Motivation aufrechtzuerhalten.

Kurzfassung

Gemäß des aktuellen Stands der Technik benötigen automatisierte Fahrzeuge eine präzise kartenrelative Eigenlokalisierung, um sich zuverlässig, sicher und effizient im dynamischen Straßenverkehr zu bewegen. Karten können frühzeitig Informationen aus der Umgebung bereitstellen, die mit der Bordsensorik aufgrund von Verdeckungen oder begrenzter Wahrnehmungsdistanz nicht erfasst werden können.

Diese Aufgabe kann im vor- und innerstädtischen Umfeld aufgrund von Signalabschattung nicht zuverlässig durch globale Satellitennavigationssysteme gelöst werden. Bildbasierte Lokalisierungsmethoden stellen eine robuste und kostengünstige Alternative dar. Jedoch hat eine Karte nur eine begrenzte Gültigkeitsspanne, da sich die kartierte, als statisch angenommene Umwelt aufgrund einer Vielzahl äußerer Einflüsse stetig verändert. Ein weiteres Problem aktueller Lokalisierungsansätze ergibt sich durch den begrenzten Wahrnehmungsbereich eines monoskopischen Kamerasystems. Ist die Sicht auf die kartierte Umwelt gestört, führt dies häufig zu einem Ausfall der Lokalisierung. Die vorliegende Arbeit setzt sich mit diesen beiden Problemfällen auseinander. Dem Problem der Kartenalterung wird durch eine kontinuierliche Aktualisierung der bestehenden Karte durch Integration weiterer Merkmale aus neuen Befahrungen des kartierten Gebiets entgegengewirkt. Der Fokus liegt dabei auf einer robusten, effizienten und skalierbaren Methode, die eine fortlaufende Kartenaktualisierung unabhängig von der Länge der Aufnahmefahrten und unabhängig von der Anzahl der Kartierungsiterationen ermöglicht. Den zweiten Schwerpunkt der Arbeit stellt die Verwendung eines Multi-Kamera-Systems zur Kartengenerierung und zur Lokalisierung dar, wodurch die Genauigkeit, Verfügbarkeit und Flexibilität im Vergleich zu monoskopischen Systemen gesteigert werden kann.

Grundbaustein der Methode sind Merkmalspunkte, die in den aufgezeichneten Kamerabildern detektiert und zu in der Karte gespeicherten Merkmalspunkten assoziiert werden. Dafür wird eine universelle Assoziationsmethode vorgestellt, die In-

formation über die geometrische Anordnung der Kameras nutzt um den Suchraum effizient einzuschränken. So ist auch bei starken perspektivischen Veränderungen eine robuste Merkmalsassoziation möglich.

Ein fundamentales Problem der Kartenaktualisierung ist die stetig wachsende Anzahl der Kartenparameter und -merkmale. Aus theoretischer Sicht können die Kartenparameter am besten gemeinsam durch Lösen eines einzelnen, umfassenden Bündelblockausgleichsproblems bestimmt werden. Aufgrund begrenzter Hardwareressourcen ist dies praktisch jedoch nicht möglich. Daher wird das kartierte Gebiet in kleine Bereiche unterteilt und für jeden Bereich unabhängig ein separates Bündelblockausgleichsproblem gelöst. Die Teillösungen werden schlussendlich durch Lösen eines Posendifferenzenausgleichsproblems zu einer konsistenten Kartenrepräsentation zusammengeführt. Die so erstellte Kartenbasis stellt die Grundlage für eine Merkmalsselektion dar, wodurch für die Lokalisierung geeignete Merkmale auf Basis einer Wiedererkennungsbewertung identifiziert werden. Auf diese Weise bleibt die zu speichernde Informationsmenge der Lokalisierungskarte gebietsweise begrenzt und es wird eine zuverlässige Lokalisierung über lange Zeiträume ermöglicht.

Die entwickelte Multi-Kamera-Lokalisierung umfasst zwei Teilsysteme. Das erste Teilsystem basiert auf dem Prinzip der merkmalsbasierten Ortswiedererkennung und bestimmt fortlaufend den Aufenthaltsort des Egofahrzeugs in der Karte. Das zweite Teilsystem schätzt eine hochgenaue kartenrelative Fahrzeugpose in sechs Freiheitsgraden. Dafür wird eine Kartenstruktur vorgestellt, welche die kartierten Merkmale effizient bereitstellt. Die Verwendung eines Multi-Kamera-Systems steigert bei beiden Teilsystemen die Genauigkeit und Robustheit der Lokalisierung gegenüber der Verwendung einer einzelnen Kamera. Außerdem ermöglicht die entwickelte Methode eine präzise Lokalisierung bei Befahrung des kartierten Gebiets in beliebiger Orientierung.

Die vorgestellte Kartengenerierungs- und Lokalisierungsmethode wurde in mehreren Projekten zum automatisierten Fahren erfolgreich eingesetzt und ist seit mehreren Jahren das Fundament eines Projekts zur Erprobung des kartenbasierten vollautomatischen Fahrens in innerstädtischen Gebieten in Karlsruhe.

Abstract

Automated vehicles heavily rely on accurate self-localization to operate safely and efficiently in dynamic road traffic scenarios. Using maps enables to provide information of the surrounding structure prematurely which cannot be measured with on-board sensors due to occlusions or limited sensor ranges.

Global navigation satellite systems are not reliable in urban environments due to multipath and shadowing effects. Vision-based localization upon a previously created map offers a promising alternative. However, environmental changes lead to obsolescence of a once created map. A further problem of current vision-based localization approaches arises from the bounded field of view of a monoscopic camera system which is subject to failure in case of poor visibility conditions. The present work proposes solutions to overcome these two drawbacks. The problem of the obsolescence of a static map is tackled by updating the map iteratively after passing the mapped area again. In this work, an iterative and robust approach to create highly accurate maps comprising any number of drives of arbitrary length with constant computation complexity is proposed. Furthermore, a high-precision self-localization method using multiple cameras covering the surrounding environment is presented.

The basic element of the proposed method are feature points which are detected in the recorded images and matched to features stored in the map. A universal matching approach which efficiently shrinks the search space based on knowledge of the geometric configuration of the cameras is proposed in this work. By this, robust feature matching can be achieved under large perspective variations.

Bundling more and more passes of partly overlapping areas into a single map leads to an unbounded growing number of poses, landmarks and features. Theoretically, solving a single joint bundle adjustment problem achieves the best possible estimate of these map parameters. However, this is unfeasible in practice due to limited hardware resources. Therefore, the map is divided into disjoint clusters. The poses

and landmarks of each cluster are estimated independently by solving bundle adjustment problems of constant complexity for each cluster. To achieve a consistent map representation, all clusters are interconnected afterwards by solving a pose-graph adjustment problem. This enables to select features in order to create a localization map which has locally a bounded size and preserves reliable localization estimates in different environmental conditions over long periods of time.

The proposed surround view localization system consists of two subsystems. The first one is a feature-based multi view place recognition approach which continuously estimates a rough position of the ego-vehicle in the map. The other subsystem estimates a high-precision six degrees of freedom pose estimate. For that, a novel map structure which efficiently provides the mapped features with respect to a surround view camera system is proposed. For both systems, it is shown that using multiple cameras leads to superior results in terms of accuracy, availability and robustness compared to a monoscopic system. Furthermore, it is demonstrated that the proposed method enables precise localization while passing the mapped area at any vehicle orientation.

The methods for localization and mapping proposed in this work were successfully applied in several automated driving projects and are the basis of a project to test algorithms for map-based automated driving in inner-city areas in Karlsruhe for several years.

Inhaltsverzeichnis

Akronyme	
1 Einleitung	1
1.1 Ziele der Arbeit	3
1.1.1 Kartengenerierung	4
1.1.2 Multi-Kamera-Lokalisierung	7
1.2 Struktur der Arbeit	8
2 Stand der Technik	11
2.1 Kartengenerierung	11
2.2 Visuelle Lokalisierung	13
3 Grundlagen	17
3.1 Sensorkonfiguration	17
3.2 Isometrische Transformationen	19
3.3 Kalibrierung	22
4 Merkmalsassoziation	23
4.1 Suchraumeinschränkung	24
4.1.1 Bildsequenzen	25
4.1.2 Merkmal-Landmarke Assoziation	26
4.1.3 Geometrische Assoziation ohne Strukturwissen	27
4.2 Deskriptor-Vergleich	33
4.3 Globale Assoziation	34
5 Robuste Parameterschätzung	37
5.1 Nichtlineare Ausgleichsrechnung	38

5.1.1	Bündelblockausgleich	39
5.1.2	Clusterausgleich	40
5.1.3	Posendifferenzenausgleich	42
5.1.4	Ausreißerbehandlung	43
5.2	Initialisierungsmethoden	45
5.2.1	Odometrie	45
5.2.2	Lokalisierung	46
5.2.3	Aufspannende Bäume	47
5.2.4	Landmarkentriangulation	50
5.3	Trajektorienvalidierung	51
5.3.1	Momentanpolbewegung	52
5.3.2	Polynomielle Approximation	53
5.3.3	Trajektorienvergleich	54
6	Kartengenerierung	57
6.1	Ablauf	58
6.1.1	Sequenzaufbereitung	60
6.1.2	Sequenz-Karte-Assoziation	61
6.1.3	Finale Parameterschätzung	62
6.2	Analyse der Optimierungsstrategie	63
6.2.1	Gleitender Bündelblockausgleich	64
6.2.2	Assoziationstopologie	69
6.2.3	Lokaler Clusterausgleich	70
6.2.4	Robuste Kartenposenschätzung	73
6.3	Experimente	76
7	Multi-Kamera-Lokalisierung	85
7.1	Überblick	86
7.2	Ortswiedererkennung	90
7.2.1	Ablauf	90
7.2.2	Experimente	93
7.3	Metrische Lokalisierung	103
7.3.1	Kartengestützte Merkmalsassoziation	105
7.3.2	Experimente	108

8	Merkmalsselektion	119
8.1	Selektionsmethode	120
8.2	Landmarkenselektion	121
8.3	Experimente	123
9	Anwendungsfälle	131
9.1	A2B	131
9.2	Future Bus	134
9.3	Adenauer	136
10	Zusammenfassung	139

Akronyme

2D/3D	2-dimensional/3-dimensional
BBF	Bertha-Benz-Fahrt
BLUE	Best Linear Unbiased Estimator
BOW	Bag Of Words
BRIEF	Binary Robust Independent Elementary Features
BRISK	Binary Robust Invariant Scalable Keypoints
CNN	Convolutional Neural Network
DIRD	Dird is an Illumination Robust Descriptor
DOF	Degrees-Of-Freedom
EKF	Extended Kalman Filter
GNSS	Global Navigation Satellite System
IMU	Inertial Measurement Unit
LIDAR	LIght Detection And Ranging
MKS	Multi-Kamera-System
NLS	Non-Linear Least Squares
ORB	Oriented FAST and rotated BRIEF
RANSAC	Random Sampling Consensus
SIFT	Scale Invariant Feature Transform
SLAM	Simultaneous Localization And Mapping
SSD	Solid-State-Drive
SURF	Speeded up Robust Features
SVO	Stereo Visuelle Odometrie
UKF	Unscented Kalman Filter
VO	Visuelle Odometrie

1 Einleitung

Das Automobil hat sich im Laufe des 20. Jahrhunderts als das Mittel der Wahl zur individuellen Mobilität etabliert. Eine weltweit engmaschig ausgebaute Straßeninfrastruktur ermöglicht eine flexible Beförderung von Personen und Gütern über große Distanzen. Aufgrund der hohen Komplexität und Diversität des modernen Straßenverkehrs werden heutzutage alle Kraftfahrzeuge auf öffentlichen Straßen von Menschen gesteuert. Das Verständnis der wahrgenommenen Umgebung und die Ableitung und Ausführung daraus resultierender angemessener Handlungen sind kognitiv höchst anspruchsvoll. Aufgrund dieser hohen Anforderungen können automatisierte Fahrzeuge diese Aufgabe noch nicht vollumfänglich bewältigen. Jedoch zeigt beispielsweise die Unfallstatistik [1] von 2016 auf deutschen Straßen mit über 3.000 Unfällen mit Todesfolge, dass der Faktor Mensch im Straßenverkehr ein hohes Gefahrenpotenzial darstellt. Durch zunehmende Automatisierung des Automobils kann der Faktor Mensch zunehmend ausgeschlossen werden und so langfristig ein sicherer und effizienterer Straßenverkehr erreicht werden. Daher ist die Entwicklung von höheren Automatisierungsstufen bis hin zu einer vollständigen Automatisierung des Straßenverkehrs im öffentlichen Interesse [2].

Gemäß des aktuellen Stands der Technik sind präzise verortete Kartendaten unverzichtbar für das hochautomatisierte Fahren [3]. Karten ermöglichen eine schnelle und effiziente Bereitstellung von geometrisch verorteten Informationen der quasi-statischen Umwelt. Die in einer Karte abgelegten Informationen können ohne zusätzlichen Berechnungsaufwand zur Laufzeit bereitgestellt werden. Auch können Karten frühzeitig Informationen bereitstellen, die mit der verfügbaren Sensorik beispielsweise aufgrund von Verdeckungen oder begrenzter Wahrnehmungsdistanz nicht erfasst werden können. Heutzutage benötigt bereits eine Vielzahl von Anwendungen, beispielsweise die Erkennung und Prädiktion anderer Verkehrsteilnehmer, die Umgebungsinterpretation, die Verhaltensentscheidung oder die Manöverpla-

nung, aus Karten bereitgestelltes Wissen, um robust und zuverlässig die jeweiligen Aufgaben bewältigen zu können [4].

Um die in einer Karte abgelegten Informationen nutzen zu können, ist eine Schätzung der eigenen Position und Orientierung innerhalb der Karte notwendig. Je genauer und detaillierter die in der Karte abgelegten Informationen dabei sind, desto höher sind die Anforderungen an die Genauigkeit und Robustheit der Lokalisierung. Üblicherweise werden zu diesem Zweck inertielle Messeinheiten (IMU) und globale Satellitennavigationssysteme (GNSS) kombiniert. Diese Systeme sind jedoch in urbanen Gebieten, aufgrund von Abschattung und Mehrwegeausbreitung, nicht zuverlässig. Da Kameras günstig, wartungsarm und im Vergleich zu Laserscannern platzsparend sind, werden bereits heute viele Fahrzeuge mit Kameras ausgestattet, sodass es günstig ist, diese auch zur Lokalisierung zu verwenden.

Aktuelle Ansätze zur visuellen Lokalisierung im Bereich des automatischen Fahrens [5], [6] zeigen, dass in Umgebungen, in denen selbst präzise GNSS-/IMU-Ansätze versagen, eine zuverlässige Lokalisierung gewährleistet werden kann. Dabei hat es sich etabliert, eine vorab erstellte Lokalisierungskarte zu verwenden. Im Gegensatz zu den Simultaneous-Localization-and-Mapping (SLAM) [7]–[11] Ansätzen trennt dieses Vorgehen die Lokalisierung und die Kartengenerierung voneinander, wodurch eine akasale Datenverarbeitung möglich ist und keine Echtzeitverarbeitung zur Kartengenerierung erforderlich ist. Dadurch kann die Karte robuster und genauer als bei klassischen SLAM-Methoden geschätzt werden. Des Weiteren werden zur Lokalisierungszeit Ressourcen gespart, da keine Karte parallel berechnet oder aktualisiert werden muss.

Zu Beginn der Arbeit haben Lategahn et al. [6] im Rahmen der Bertha-Benz-Fahrt (BBF) [4] gezeigt, dass das kartenbasierte automatische Fahren im öffentlichen Straßenverkehr mit Punktmerkmalskarten möglich ist. In dicht bebauten Gebieten mit viel markanter Struktur werden genug Punktmerkmale mit einer einzelnen Kamera detektiert und korrekt zu Kartenlandmarken assoziiert, sodass eine präzise Lokalisierungspose geschätzt werden kann. Andererseits musste für das Fahren in ländlichen Gebieten und auf größeren Zubringerstraßen eine zusätzliche Lokalisierungsmethode unter Verwendung von Fahrbahnmarkierungen und GNSS verwendet werden, da in diesen Bereichen nur wenig markante Umgebungsstruktur vorhanden ist. Der

Wechsel zwischen den verschiedenen Lokalisierungssystemen ist allerdings problematisch und muss sorgfältig vorbereitet werden.

Ein weiteres Problem ergibt sich dadurch, dass Karten i. A. eine statische Umwelt repräsentieren. Die Annahme einer statischen Umwelt hat auf öffentlichen Straßen jedoch nur eine begrenzte Gültigkeitsspanne, insbesondere bei der Verwendung von Punktmerkmalen [12]. Für eine Anwendung wie der Bertha-Benz-Fahrt muss die Lokalisierungskarte stets deckungsgleich mit weiteren Kartenschichten, wie beispielsweise der Planungskartenschicht, sein. Daher ist die Erzeugung einer gänzlich neuen Karte problematisch, da alle anderen Kartenschichten ebenfalls neu erstellt bzw. angepasst werden müssen.

Die vorliegende Arbeit basiert auf der Arbeit von Lategahn et al., greift die zuvor genannten Schwächen auf und stellt Methoden bereit, um sowohl das Problem der Kartenalterung zu lösen als auch die Robustheit und Verfügbarkeit der Lokalisierung durch Verwendung mehrerer Kameras zu steigern.

1.1 Ziele der Arbeit

Die vorliegende Arbeit hat zum einen die Entwicklung einer vollumfänglichen Methode zur Erstellung und Aktualisierung von hochgenauen 3D-Lokalisierungskarten und zum anderen die Umsetzung einer Methode zur präzisen und robusten Lokalisierung zum Ziel. Die Methode erfordert die Verwendung von mindestens einer Kamera, die am zu lokalisierenden Fahrzeug angebracht ist. Der Fokus liegt jedoch auf der Verwendung eines Multi-Kamera-Systems (MKS) [5], welches große Teile des Fahrzeugumfelds wahrnimmt. Durch die Verwendung eines Multi-Kamera-Systems sollen die Genauigkeit, Verfügbarkeit und Flexibilität der Lokalisierung im Vergleich zu monoskopischen Systemen verbessert werden.

Sowohl die Kartengenerierung als auch die Lokalisierung umfassen mehrere aufeinanderfolgende Prozessschritte. Ausgehend von der Bildaufnahme bis hin zur Bestimmung und Validierung von Fahrzeugposen und Landmarken ergibt sich eine Vielzahl unterschiedlicher Themenfelder und Potenziale zur Verbesserung bestehender Methoden. Diese Arbeit fokussiert dabei:

- die vollautomatische, robuste und skalierbare Erzeugung und Aktualisierung einer Kartendatenbank durch iterative Integration weiterer Aufnahmesequenzen von Befahrungen des gleichen Gebiets (engl. *large-scale multi-session mapping*).
- die Entwicklung einer Methode zur robusten kamerabasierten Ortswiedererkennung und initialen Orientierungsschätzung des Egofahrzeugs (engl. *multi-view place recognition*).
- die Umsetzung einer genauen und vielseitig verfügbaren Lokalisierung mittels eines MKS, welches zusätzlich die Möglichkeit bietet, die Karte in unterschiedliche Richtungen zu befahren (engl. *multi-view 6-DOF localization*).
- die statistikbasierte Merkmalsselektion zur Bestimmung einer langzeitverfügbaren Lokalisierungskarte (engl. *life-long localization*).

Die vorgestellten Methoden basieren im Wesentlichen auf geometrischen Betrachtungen und sind daher unabhängig von den verwendeten Merkmalen und Merkmalsdeskriptoren. Bei den durchgeführten Experimenten wurden stets Ecken- und Blob-Merkmaldetektoren [13] sowie der DIRD-Deskriptor [14] verwendet. Es ist aber auch die Verwendung einer beliebigen anderen Kombination aus Detektoren und Deskriptoren möglich, da alle Varianten den gleichen projektiven und geometrischen Einschränkungen unterliegen.

In Abschnitt 1.1.1 werden zunächst die Aspekte der Kartierung motiviert und abgegrenzt. Darauffolgend werden in Abschnitt 1.1.2 die in der Arbeit entwickelten Lokalisierungskomponenten aufgezeigt.

1.1.1 Kartengenerierung

In der Arbeit von Lategahn et al. [6] wurde die Lokalisierungskarte ausschließlich aus einer einzelnen Befahrung erzeugt. Eine einmalig erzeugte Karte hat jedoch nur eine begrenzte Gültigkeitsspanne, da sich die Umgebung beispielsweise durch Witterungsänderungen, Baustellen oder aufgrund parkender Fahrzeuge verändert. Je nach Jahreszeit, Umgebung und Art der verwendeten Merkmale führt dies mit zunehmendem Alter der Karte zu einer Verschlechterung bzw. zum Ausfall der Lokalisierung. Durch eine erneute Befahrung des kartierten Gebiets kann die bestehende Karte erweitert und aktualisiert werden und so langfristig eine aktuelle Karte

zur Verfügung gestellt werden [15]–[17]. Die Aktualisierung soll bereits kartierte, langzeitverfügbare Merkmale beibehalten und aktualisieren, nicht wiedererkannte Merkmale entfernen und neu detektierte Merkmale der Karte hinzufügen. Wird zusätzlich bisher unkartiertes Gebiet befahren, wird die bestehende Karte um den neu befahrenen Bereich erweitert.

Das im Rahmen der Arbeit entwickelte Kartierungsframework ermöglicht die vollautomatische iterative Generierung von hochgenauen 3D-Lokalisierungskarten unter Verwendung aufgezeichneter Kamerabilder aus verschiedenen Befahrungen von teilweise überlappenden Gebieten. Werden zusätzliche Informationen der Fahrzeugbewegung beispielsweise aus Fahrzeugodometern oder GNSS-Messungen bereitgestellt, können diese zur Bewegungsschätzung bzw. zur geodäsischen Verortung der Karte verwendet werden. Der Fokus der Kartierungsmethode liegt dabei auf den folgenden Gesichtspunkten:

Skalierbarkeit

Durch die Hinzunahme immer weiterer Befahrungen in die bestehende Karte wächst der Speicher- und Rechenaufwand zur Schätzung der Kartenlandmarken und Kartenposen stetig. In der Arbeit von Muehlfellner et al. [18] wurde durch Landmarkenselektion ein begrenzter Speicherbedarf erzielt. Die Bestimmung der Kartenposen und -landmarken wurde jedoch auf einem Rechencluster durchgeführt, da der Rechenaufwand zur numerischen Lösung des zugrunde liegenden Bündelblockausgleichsproblems die Kapazitäten von durchschnittlichen Rechnern weit übersteigt. Diese Arbeit stellt demhingegen ein iteratives und flexibles Konzept zur hochgenauen Referenzierung beliebig vieler Befahrungen mit stets begrenztem Speicher- und Rechenaufwand bereit. Zur Bestimmung der Kartenparameter werden mehrere lokale Bündelblockausgleichsprobleme gelöst, die jeweils nur einen begrenzten Bereich des kartierten Gebiets umfassen. Die Ergebnisse dieser lokalen Optimierungen werden in Form von Posendifferenzen gespeichert und mit diesen darauffolgend die finalen Kartenposen geschätzt. Die abgeleiteten Posendifferenzen überführen dabei die erzielte lokale Genauigkeit der lokalen Optimierung in die finale Trajektorien-schätzung. Durch dieses Vorgehen ist der Speicher- und Rechenaufwand aller Kartierungsschritte deterministisch begrenzt. Daher können beliebig viele und beliebig lange Befahrungen in die Karte integriert werden. Die durchgeführten Experi-

mente zeigen, dass durch dieses Vorgehen lokal hochgenaue Kartenposen geschätzt werden. Das vorgestellte Kartierungsframework ermöglicht somit die Schätzung von konsistenten dreidimensionalen Mehrfachbefahrungskarten für beliebig große Gebiete auf Basis beliebig langer Aufnahmesequenzen aus beliebig vielen Befahrungen.

Effizienz

Die Optimierung der lokalen Bündelblockausgleichsfenster ist der Kartierungsschritt mit dem mit Abstand höchsten Rechenaufwand. Da aus dieser Optimierung nur relative Informationen in Form abgeleiteter Posendifferenzen und Merkmalskorrespondenzen gespeichert und weiterverwendet werden, kann jedes lokale Bündelblockausgleichsproblem parallel und ohne wechselseitige Abhängigkeit zu anderen Optimierungsfenstern berechnet werden. Somit kann die verfügbare Rechenkapazität bestmöglich ausgenutzt und die gemeinsame Trajektorien-schätzung effizient beschleunigt werden. Das entwickelte Kartierungsframework unterstützt so die effiziente Berechnung von Karten auf handelsüblichen Desktop-Computern, auf verteilten Systemen als auch auf massiv parallelisierenden Rechenclustern sowie Cloud-Computing.

Robustheit

Sowohl in [6] als auch in [18] werden hauptsächlich abgegrenzte innerstädtische Bereiche mit ausreichend markanter Umgebungsstruktur kartiert. Werden größere Gebiete mit mehr Umgebungsvariation vollautomatisch kartiert, erhöht sich die Wahrscheinlichkeit, dass, aufgrund von zu wenig verfügbaren oder falsch assoziierten Merkmalskorrespondenzen, die Kartenschätzung fehlerhaft wird. Das in dieser Arbeit entwickelte Kartierungsframework analysiert daher zum einen die berechneten Zwischenergebnisse auf Fehler und Ungenauigkeiten. Zum anderen ermöglicht die vorgestellte Optimierungsstrategie, dass die bestehende Karte innerhalb von wenigen Minuten auf Basis der abgeleiteten Posendifferenzen vollständig neu geschätzt werden kann. So können detektierte Kartenfehler mit vernachlässigbarem Aufwand beispielsweise durch Anpassung von Optimierungsgewichten oder Hyperparametern korrigiert werden.



Abbildung 1.1: Verschiedene Problemszenarien. Links: Großflächige Verdeckung durch dynamische Objekte. Rechts: Schlechte Sichtbedingungen durch starke Sonneneinstrahlung.

1.1.2 Multi-Kamera-Lokalisierung

Die in dieser Arbeit entwickelte Lokalisierungsmethode erzielt eine echtzeitfähige und zentimetergenaue Lokalisierung für beliebige Fahrzeugorientierungen, sofern Beobachtungen aus verschiedenen Blickrichtungen vorliegen. Dies ermöglicht auch dann eine zuverlässige Lokalisierung, wenn die kartierte Strecke in entgegengesetzter Richtung befahren wird. Außerdem wird die Lokalisierung durch die Verwendung mehrerer Kameras robuster. Bei einem Ausfall einzelner Kameras beispielsweise durch einen technischen Defekt, intensive Sonneneinstrahlung oder großflächige Verdeckung kann die Lokalisierung mit den verbleibenden, funktionierenden Kameras aufrechterhalten werden. Abbildung 1.1 zeigt exemplarisch solche Problemfälle. Durch den erweiterten Erfassungsbereich eines MKS wird zusätzlich die Genauigkeit der Lokalisierung verbessert, da in der Regel mehr Kartenlandmarken aus einem größeren Winkelspektrum als bei einem monoskopischen Kamerasystem erfasst werden.

Die Lokalisierungsmethode durchläuft in jedem Schritt zwei Phasen. In der ersten Phase wird der ungefähre Aufenthaltsort des Fahrzeugs in der Karte bestimmt. Diese Schätzung wird darauffolgend als Initialisierung für eine hochgenaue metrische Lokalisierung verwendet. Für beide Phasen wurde im Rahmen der Arbeit eine Methode entwickelt, die das erweiterte Sichtfeld eines MKS effizient ausnutzt.

Ortswiedererkennung

Zur erstmaligen Initialisierung der Lokalisierung wird der wahrscheinlichste Aufenthaltsort des Fahrzeugs durch eine bildbasierte Ortswiedererkennung bestimmt. Dazu speichert die Karte holistische Merkmale zu allen kartierten Orten. Diese Merkmale werden aus den aufgezeichneten Kamerabildern gewonnen. Durch einen Abgleich der Merkmale aus den aktuellen Kamerabildern mit allen in der Karte gespeicherten Merkmalen wird der wahrscheinlichste Aufenthaltsort bestimmt. Die hier vorgestellte Methode evaluiert alle zu einem Zeitpunkt aufgezeichneten Bilder des MKS und erhöht somit die Verfügbarkeit und Robustheit der Ortswiedererkennung gegenüber monoskopischen Ansätzen. Außerdem wird durch eine eindeutige Zuordnung der Merkmale zu den verschiedenen Kameras des MKS zusätzlich eine initiale Schätzung der Fahrzeugorientierung ermöglicht.

Metrische Lokalisierung

Auf Basis einer initialen Lokalisierungsposenschätzung werden in der zweiten Phase robust Kartenlandmarken zu aktuell wahrgenommenen Merkmalen assoziiert. Darauf basierend wird die kartenrelative Fahrzeugpose durch robustes Lösen eines Ausgleichsproblems hochgenau geschätzt. Die metrische Lokalisierungskarte speichert dazu punktförmige Landmarken sowie dazu korrespondierende Merkmale in einer neu entwickelten Rasterkartenstruktur. Diese Struktur ermöglicht eine effiziente und dynamische Bereitstellung der Landmarken und Merkmale in Echtzeit. Die bereitgestellten Merkmale werden dafür räumlich gruppiert und nach Relevanz sortiert bereitgestellt. Dadurch wird zum einen nur eine geeignete Teilmenge aller bereitgestellten Landmarken zu den verschiedenen Kameras zugeordnet und zum anderen alle Raumrichtungen, die das MKS abdeckt, unabhängig von der Umgebungsstruktur gleichverteilt abgetastet. Die entwickelte Kartenstruktur verringert so zum einen den Aufwand der Landmarkenassoziation und erhöht zum anderen die Robustheit und Genauigkeit der metrischen Lokalisierung.

1.2 Struktur der Arbeit

Diese Arbeit setzt sich mit verschiedenen Aspekten der iterativen Generierung einer Lokalisierungskarte aus mehreren Befahrungen sowie der dazu korrespondieren-

den Lokalisierungsmethode auseinander. Obwohl die Kartengenerierung und die Lokalisierung eng miteinander verzahnt sind, werden diese in separaten, in sich abgeschlossenen Kapiteln vorgestellt. Alle durchgeführten Experimente und Evaluationen sind daher und der Übersichtlichkeit halber in die entsprechenden Kapitel eingebettet.

Nach einer Übersicht und Abgrenzung zu verwandten Arbeiten in Kapitel 2 werden in Kapitel 3 Rahmenbedingungen definiert, verwendete Versuchsträger und Sensor-konfigurationen vorgestellt und Transformationskonventionen definiert.

Die Kartengenerierung und die Lokalisierung basieren auf ähnlichen zugrunde liegenden Prinzipien und Verarbeitungsschritten. Eine fundamentale gemeinsame Grundlage ist die Detektion und Beschreibung von Punktmerkmalen und deren Assoziation zueinander. Die verwendeten Methoden und geleisteten Beiträge zur Merkmalsassoziation werden in Kapitel 4 detailliert beleuchtet. Ein Schwerpunkt dieses Kapitels ist die robuste Assoziation von Punktmerkmalen aus verschiedenen Bildern unter Ausnutzung von Information über die Position und Blickrichtung der Kameras.

Ein wesentlicher Verarbeitungsschritt der Kartengenerierung und der Lokalisierung ist die robuste und effiziente Schätzung von Fahrzeugposen und Landmarken durch numerisches Lösen nichtlinearer Gleichungssysteme. Mit der Lösung dieser Gleichungssysteme befasst sich das Kapitel 5. Eine robuste Initialisierung der zu schätzenden Parameter ist von fundamentaler Bedeutung für die Lösung der aufgestellten Ausgleichsprobleme. Die verwendeten Initialisierungsmethoden sowie Modelle zur Validierung der geschätzten Parameter stellen den zweiten Teil dieses Kapitels dar.

Kapitel 6 stellt das erste Hauptthema, die effiziente und skalierbare Kartengenerierung aus mehreren Befahrungen des gleichen Gebiets, dar. Zunächst wird die gesamte Prozesskette beschrieben und danach die wesentlichen Elemente dieser Kette analysiert und evaluiert. Zur Demonstration der Leistungsfähigkeit der vorgestellten Kartierungsmethode wird eine aus vierzehn Befahrungen erstellte Karte empirisch beleuchtet und die erzielte Genauigkeit evaluiert.

Kapitel 7 zeigt die geleisteten Beiträge zur Lokalisierung mittels eines MKS auf. Es wird zunächst der Lokalisierungsablauf und das Zusammenspiel der Ortswiederer-

kennung und der metrischen Lokalisierung erläutert. Darauffolgend werden beide Komponenten separat im Detail beschrieben, analysiert und evaluiert.

Die Selektion von Kartenmerkmalen, die in die Lokalisierungskarte überführt werden, wird in in Kapitel 8 vorgestellt. Auch wenn dies nicht im Fokus dieser Arbeit stehen soll, werden geleistete Arbeiten und die grundlegenden Konzepte präsentiert. Zusätzlich wird exemplarisch eine empirische Analyse der Selektion von Kartenlandmarken gezeigt. Da die Merkmalsselektion die Schnittstelle zwischen der Kartengenerierung und der Lokalisierung darstellt, wird diese außerdem zur Darstellung eines umfassenden Bildes des Gesamtsystems benötigt.

Die in dieser Arbeit vorgestellte Kartengenerierung und Lokalisierung wurde in verschiedenen Projekten unter schwierigen Bedingungen erfolgreich angewendet und stellt aktuell die Grundlage zum automatischen Fahren in innerstädtischen Bereichen in Karlsruhe dar. Diese Anwendungsfälle werden in Kapitel 9 aufgezeigt. Abschließend wird die Arbeit in Kapitel 10 zusammengefasst.

2 Stand der Technik

Dieses Kapitel gibt eine Übersicht zu aktuellen Arbeiten, die thematisch zu dieser Arbeit in Beziehung stehen. Abschnitt 2.1 gibt eine Übersicht über aktuelle Arbeiten aus dem Forschungsbereich SLAM. Im zweiten Abschnitt 2.2 werden verwandte Arbeiten zur visuellen Lokalisierung aufgezeigt.

2.1 Kartengenerierung

SLAM befasst sich mit dem Problem der gleichzeitigen Lokalisierung und Kartengenerierung. Die Lokalisierungsschätzung basiert auf einer bestehenden Karte, die während der Lokalisierung parallel erzeugt und aktualisiert wird. Erste SLAM-Ansätze basieren auf rekursiven Bayes-Filtern, beispielsweise dem Extended-Kalman-Filter (EKF) [7], [19]. Jede neue Landmarke erweitert dabei den zu schätzenden Zustandsvektor. In der Praxis ist der Ansatz jedoch nur begrenzt verwendbar, da der Zustandsvektor fortlaufend wächst und somit auch der Rechen- und Speicheraufwand. Neuere Ansätze [7] formulieren das SLAM-Problem als Parametergraph [20]. Bei dieser Formulierung werden alle Landmarken und Posen mittels eines Nonlinear-Least-Squares (NLS)-Verfahren gleichzeitig geschätzt. Zur Lösung solcher Probleme wird häufig der Levenberg-Marquardt-Algorithmus verwendet. Die dünnbesetzte und nur lokal korrelierte Struktur der Messmatrix des SLAM-Problems ermöglicht unter Ausnutzung des Schur-Komplement-Tricks [21] und der Verwendung von Sparse-Matrix-Solvern eine effiziente Schätzung aller Posen und Landmarken. Diese Methoden haben dazu geführt, dass sich diese SLAM-Variante durchgesetzt hat. Eine detailliertere Übersicht zum Forschungsfeld SLAM findet sich in [22]. Die in dieser Arbeit vorgestellten Methoden zur Parameterschätzung basieren auf den Ideen dieser Ansätze.

Wie bereits in Kapitel 1 beschrieben, hat es sich im Bereich des automatischen Fahrens etabliert, eine vorab erstellte Karte zur Lokalisierung zu verwenden [5], [6],

[18], [23]. Eine vorab erzeugte Karte unterliegt jedoch den bereits erwähnten Alterungseffekten und wird ohne regelmäßige Aktualisierung obsolet und unzuverlässig. Burki et al. [15] haben gezeigt, dass sich die Gültigkeitsspanne einer Punktmerkmalskarte nur über wenige Wochen erstreckt. Wird eine bereichsweise Verschlechterung der Lokalisierung festgestellt, wird dieser Teil der Karte durch Hinzufügen aktueller Merkmale innerhalb dieses Gebietes die bestehende Karte aktualisiert. Die Landmarken- und Merkmalsselektion zur Kartenaktualisierung und zur Verbesserung der Lokalisierung ist der Schwerpunkt weiterer Arbeiten [16], [24]–[28]. Die Merkmalsselektion kann zusätzlich durch das Entfernen dynamischer Objekte aus den aufgezeichneten Bildern unterstützt werden [29]. Das wesentliche Ziel dieser Methoden ist es, die Informationsdichte und -güte zu maximieren und gleichzeitig die Menge der zu verarbeitenden Daten zu minimieren.

Abgrenzend zu den genannten Arbeiten fokussiert die vorliegende Arbeit die effiziente Schätzung von Kartenposen und -landmarken bei stets begrenztem Rechen- und Speicheraufwand innerhalb eines iterativen Kartierungszyklus. Diese Arbeit stellt somit eine robuste und skalierbare Grundlage für die zuvor genannten Arbeiten dar und erweitert somit deren Anwendungsfeld auf den Bereich ganzer Großstädte. Außerdem wird durch Referenzierung der Karte zu GNSS-Messungen oder Luftbildreferenzmerkmalen [30] eine global konsistente Kartenschätzung erzielt.

Die theoretisch beste Methode zur gemeinsamen Schätzung von Posen und Landmarkenparametern ist der Bündelblockausgleich [20], [21], [31], [32]. Jedoch übersteigt der Verarbeitungsaufwand zur gemeinsamen Schätzung aller Parameter, insbesondere bei der Kartierung von stadtdichten Gebieten, in der Regel die verfügbaren Hardwareressourcen [33], [34]. Daher wird im Rahmen der Arbeit eine Optimierungsstrategie vorgestellt, die das Gesamtproblem explizit in kleinere Teilprobleme aufteilt, diese separat optimiert und schließlich iterativ zusammensetzt. Ähnliche Ansätze (bekannt als Smoothing-And-Mapping (SAM) bzw. Sub-Mapping) wurden bereits in anderen Arbeiten vorgestellt [35]–[37]. Die Aufteilung in die Teilprobleme wird dabei implizit durch die Aufteilung der dünnbesetzten Messmatrix bzw. Jacobi-Matrix realisiert. Ni et al. [33] haben eine Methode vorgestellt, welche die Linearisierung der Teilprobleme zwischenspeichert und bei der Zusammenführung für die globale Schätzung wiederverwendet. Ein ähnliches Problem ergibt sich

bei der Zusammenführung von Teilkartenschätzungen verschiedener Agenten zu einer globalen Schätzung im Rahmen von Multi-Agenten-SLAM Methoden [11], [38]. Cieslewski et al. [39] haben ein dezentralisiertes Karten-Versionsmanagement vorgestellt, wodurch verschiedene Agenten parallel die gleiche Karte manipulieren können. Diese wird dann konsistent zusammengeführt. Ein wesentlicher Aspekt dieser Arbeit ist die Reduktion des Informationstransfers zwischen den verschiedenen Agenten. Die vorgestellte Kartierungsmethode unterstützt durch die akausale Datenverarbeitung ebenso eine effiziente Zusammenführung von Teilkarten verschiedener Agenten.

2.2 Visuelle Lokalisierung

Dieser Abschnitt gibt eine Übersicht über bestehende aktuelle Ansätze zur visuellen Lokalisierung. Visuelle Lokalisierung für das automatisierte Fahren umfasst im Wesentlichen zwei unterschiedliche Klassen. Die erste Klasse bilden die sogenannten Ansätze zur Ortswiedererkennung (englisch: *place recognition*) [40]. Dabei wird aus der Menge der in der Karte gespeicherten Kandidatenposen jener Ort geschätzt, der den aktuellen Bildern am ähnlichsten ist [41]. Die zweite Klasse stellen die metrischen Lokalisierungsmethoden dar, die darauf abzielen, eine möglichst genaue kartenrelative metrische Position und Orientierung des Egofahrzeugs zu schätzen.

Ortswiedererkennung

Die meisten Arbeiten zur Ortswiedererkennung fokussieren sich auf die Verbesserung der verwendeten Bilddeskriptoren. Naseer et al. [42] assoziieren Sequenzen von aufeinanderfolgenden Bildern durch einen minimalen Assoziationskostenfluss. Dies ermöglicht eine effiziente Zurückweisung von falschen Assoziationen durch Teilverdeckungen. Lynen et al. [43] formulieren die Wiedererkennung von kartierten Orten als Wahrscheinlichkeitsdichtenschätzungsproblem unter Verwendung der ähnlichsten Kartenorte. Eine wichtige Klasse von Ansätzen basiert auf dem sogenannten Bag-of-Images-Ansatz [44]. Bei diesem Ansatz werden Bilder anhand einer Selektion von diskriminativen visuellen Wörtern identifiziert. Dazu werden Vektoren der visuellen Wörter aus den Bildern extrahiert und zueinander assoziiert. Raul et al. [9] verwenden einen BOW-Ansatz zur robusten Ringschluss-Detektion innerhalb von ORB-SLAM2.

Die junge Erfolgsgeschichte von tiefen Faltungsnetzen (englisch: Convolutional Neural Network (CNN)) hat dazu geführt, dass in den letzten Jahren mehrere Ansätze von CNNs gelernten Deskriptoren veröffentlicht wurden [40], [45]. Das Training der Deskriptoren zielt in diesen Arbeiten auf eine möglichst hohe Robustheit gegen Umweltveränderungen ab. Sünderhauf et al. [46] detektieren Objektvorschläge mittels eines R-CNN-Faltungsnetzes [47] und beschreiben diese durch Extraktion von Merkmalen aus vortrainierten Objektklassifizierungs-Netzen. Alle Objektvorschläge werden wechselseitig auf Basis der Kosinus-Distanz assoziiert. Auf Basis von Formähnlichkeit der assoziierten Objekte wird dann ein Ähnlichkeitswert zwischen den kartierten Orten und dem aktuellen Bild bestimmt.

Lategahn et al. [48] haben maschinell einen Deskriptor trainiert, der insbesondere robust gegen Beleuchtungsvariationen ist, da unterschiedliche Lichtverhältnisse einen wesentlichen Störfaktor für die visuelle Lokalisierung darstellen. Durch dynamische Programmierung werden konsistente Bildhypotheseensequenzen geschätzt und so falsche Schätzungen zurückgewiesen. In der vorliegenden Arbeit wird eine Methode vorgestellt, die den gleichen Deskriptor verwendet. Im Unterschied dazu wird eine Verfolgung von konsekutiven Aufenthaltsposen vorgestellt. Das hat zum Vorteil, dass die Methode robuster wird, Multi-Kamera-Systeme unterstützt werden und eine Orientierung geschätzt werden kann. Die hier vorgestellte Methode ermöglicht die gemeinsame Nutzung mehrerer Kamerabilder zu einem Zeitpunkt. Es wird gezeigt, dass sich die Wiedererkennungsrobustheit und -genauigkeit auf diese Weise erhöht. Da die Methode unabhängig vom verwendeten Deskriptor ist, stellt sie eine Ergänzung der zuvor genannten Verfahren dar und kann mit beliebigen Merkmals-typen verwendet werden.

Metrische Lokalisierung

Ein wesentlicher Aspekt der visuellen metrischen Lokalisierung sind die verwendeten Merkmale. In den Ansätzen von Schreiber et al. [49] und Poggenhans et al. [50] werden Straßenmarkierungen im Kamerabild detektiert und mit kartierten Merkmalen assoziiert. Die Straßenmarkierungen werden aus virtuellen Luftbildansichten, die aus einem Stereo-Kamerabild erzeugt werden, extrahiert. Spurmarkierungen haben jedoch aufgrund der hohen Mehrdeutigkeit eine hohe Wahrscheinlichkeit, falsch assoziiert zu werden. Außerdem sind Spurmarkierungen in innerstädtischen Berei-

chen oft nicht gut erkennbar oder nicht vorhanden, sodass in diesen Bereichen keine Lokalisierung möglich ist. Diesem Problem kann durch Detektion zusätzlicher semantischer Merkmale aus anderen Sensoren wie beispielsweise Laserscannern begegnet werden [51].

Punktmerkmalsbasierte Ansätze hingegen erfordern keine definierten makroskopischen Strukturtypen, sondern lediglich die Existenz von markanter Umgebungsstruktur, wodurch sich ein großer Anwendungsbereich ergibt. Mühlfellner et al. [12] haben gezeigt, dass mit seriennahen MKSen prinzipiell eine Lokalisierung möglich ist. Grundlegende Konzepte sowie eine Evaluierung der Lokalisierung auf Basis punktförmiger Landmarken mit Weitwinkel Multi-Kamera-Systemen zum automatischen Einparken finden sich in [52].

Auch im Bereich der metrischen Lokalisierung sind insbesondere in den letzten Jahren CNN-basierte Ansätze veröffentlicht worden. Im Vergleich zu klassischen Ansätzen sind diese Methoden aus aktueller Sicht jedoch zu ungenau und nicht robust genug für eine Anwendung im Bereich des automatischen Fahrens. Stenborg et al. [53] haben eine metrische Lokalisierung präsentiert, die mittels pixelweiser semantischer Information Landmarkenkorrespondenzen zu einer bestehenden Karte bestimmt. Die semantische Information wird in der Arbeit durch ein CNN bestimmt, wobei eine mittlere Ungenauigkeit von $< 1\text{m}$ erreicht wird. Kendall et al. [54]–[56] haben mit PoseNet ein Ende-zu-Ende Lokalisierungs-CNN vorgestellt. Hierzu wurde ein Regressions-CNN gelernt, das als Eingabe ein Bild erhält und als Ausgabe eine kartenrelative Pose schätzt. Das neuronale Netz stellt somit gleichzeitig die Karte als auch den metrischen Schätzer dar. Die Lokalisierungsungenauigkeit dieses Ansatzes ist jedoch im Mittel mehrere Meter.

Orthogonal zu den genannten Arbeiten existieren nur wenige Ansätze zur Lokalisierung mit Multi-Kamera-Systemen. Diese Arbeit stellt dafür grundlegende Konzepte bereit, die sich auch auf andere Merkmalstypen und Lokalisierungsvarianten übertragen lassen.

3 Grundlagen

3.1 Sensorkonfiguration

Grundlage der in der Arbeit vorgestellten Konzepte ist ein Versuchsfahrzeug, an dem n Kameras befestigt sind. Die n Kameras zeichnen die Bilder I_1, \dots, I_n zeitlich synchron zu einem Zeitpunkt auf, wobei jedes Bild I_k zu einer der Kameras korrespondiert. Die Aufnahme­frequenz der Bilder ist beliebig. Für die Merkmalsassoziation zwischen Bildern von aufeinanderfolgenden Aufnahmezeitpunkten ist es vorteilhaft, die Aufnahme­frequenz an die Bewegungsgeschwindigkeit des Fahrzeugs anzupassen, um so einen konstanten Positionsunterschied zwischen zwei aufeinanderfolgenden Bildaufnahmen zu gewährleisten. Jedoch sind auch variierende Positionsdifferenzen zwischen aufeinanderfolgenden Bildern für die vorgestellte Methodik unproblematisch, sofern die zurückgelegte Distanz des Fahrzeugs zwischen zwei Aufnahmezeitpunkten nicht zu groß ist. Für eine zuverlässige konsekutive Merkmalsassoziation ist die maximale Distanz u. a. abhängig von der Szene, dem Szenenfluss im Kamerabild, den Eigenschaften der Kameraprojektion, dem verwendeten Deskriptor und der Merkmalsassoziationsmethodik. Bei den Experimenten zu dieser Arbeit wurden Bilder mit konstanter Aufnahme­frequenz von 10 Hz bei Geschwindigkeiten bis maximal 70 km/h aufgezeichnet. Daher treten Distanzen von bis zu ~ 2 m zwischen aufeinanderfolgenden Bildaufzeichnungen auf. Durch sporadische Bildausfälle wird dieser Wert in einigen Fällen überschritten.

Die in der Arbeit entwickelten Methoden zur Kartengenerierung und Lokalisierung wurden mit unterschiedlichen Versuchsträgern und Kamera­konfigurationen evaluiert und angewendet. Abbildung 3.1 zeigt zwei verschiedene Kamera­aufbauten, mit denen die Methodik erfolgreich angewendet wurde. Die folgende Auflistung beschreibt die beiden dargestellten Systeme:

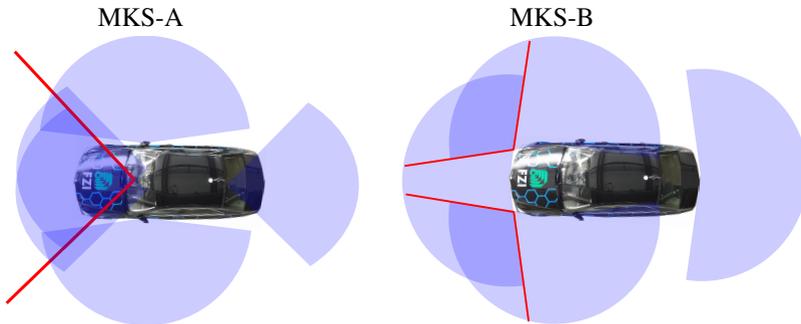


Abbildung 3.1: Verschiedene Kameraaufbauten, die für die Experimente in dieser Arbeit verwendet wurden. In Blau dargestellt sind die Sichtkegel der einzelnen Kameras. Die roten Linien zeigen relevante Überlappungsbereiche der einzelnen Sichtkegel.

- **MKS-A**

Abbildung 3.1 zeigt den Aufbau von MKS-A, der fünf teilweise unterschiedliche Kameras umfasst. Zwei Kameras mit einem Öffnungswinkel von $\sim 110^\circ$ sind hinter der Windschutzscheibe angebracht. Eine weitere Kamera mit einem Öffnungswinkel von $\sim 110^\circ$ ist hinter der Heckscheibe montiert. Außerdem sind zwei Kameras mit einem Öffnungswinkel von $\sim 170^\circ$ seitlich über den vorderen Radkästen angebracht.

- **MKS-B**

In Abbildung 3.1 ist der Kameraaufbau von MKS-B dargestellt. Jede der vier Kameras hat einen Öffnungswinkel von $\sim 190^\circ$. Die vordere Kamera ist ungefähr auf der Höhe des Kühlergrills und die hintere Kamera auf Höhe der Anhängerkupplung montiert. Zwei weitere Kameras sind wie bei MKS-A seitlich über den vorderen Radkästen angebracht.

Die verwendeten Kamerasysteme sind seriennah und ähnliche Konfigurationen werden in anderen aktuellen Arbeiten [12], [57] bezüglich der Kartengenerierung und Lokalisierung untersucht. Die rot skizzierten Bereiche in Abbildung 3.1 zeigen relevante Überlappungsbereiche der Sichtkegel verschiedener Kameras. Diese Überlappungsbereiche sind für eine skalentreue Schätzung der gefahrenen Trajektorien

wichtig, da Merkmale zwischen Bildern aus unterschiedlichen Kameras zum gleichen Zeitpunkt assoziiert werden können. Aufgrund der bekannten Kalibrierung wird dadurch die Skale der Fahrzeugbewegung beobachtbar. Ist dies nicht gegeben, müssen zusätzliche Sensoren oder Annahmen über die Einbaupositionen der Kameras für eine skalentreue Trajektorien-schätzung vorausgesetzt werden. Bis auf die Skalenschätzung ist die in dieser Arbeit entwickelte Methode unabhängig von der Anbringungsposition und der Anzahl der Kameras. Prinzipiell ist auch die Verwendung von nur einer einzelnen Kamera möglich.

Die in den folgenden Kapiteln dargestellten Experimente sind mit den beiden beschriebenen Konfigurationen MKS-A und MKS-B durchgeführt worden. In der Beschreibung der Experimente wird daher auf dieses Kapitel und das verwendete MKS verwiesen.

3.2 Isometrische Transformationen

Zur Beschreibung der Methoden und Konzepte in dieser Arbeit sind isometrische Transformationen im dreidimensionalen Raum von fundamentaler Bedeutung. In diesem Kapitel wird die Notation und Interpretation der verwendeten Transformationsbegriffe definiert und beschrieben.

Eine Transformation

$$\mathbf{T}_{D,S} \in SE(3) \quad (3.1)$$

beschreibt im Rahmen der Arbeit stets eine isometrische Abbildung von Koordinatensystem S nach Koordinatensystem D , wobei $SE(3)$ die euklidische Gruppe bezeichnet, welche die Bewegung von Starrkörpern im dreidimensionalen Raum beschreibt [58]. Eine isometrische Transformation setzt sich daher ausschließlich aus einer Rotationskomponente $\mathbf{R}_{\mathbf{T}_{D,S}} \in SO(3)$ und einer Translationskomponente $\mathbf{t}_{\mathbf{T}_{D,S}} \in \mathbb{R}^3$ zusammen. Die spezielle euklidische Gruppe $SO(3)$ stellt hierbei die Gruppe aller Rotationen im dreidimensionalen Raum dar. Die inverse Transformation $\mathbf{T}_{D,S}^{-1} = \mathbf{T}_{S,D} \in SE(3)$ beschreibt die Umkehrabbildung von D nach S und es gilt

$$\mathbf{T}_{D,S}^{-1} \cdot \mathbf{T}_{D,S} = \mathbf{T}_{D,S} \cdot \mathbf{T}_{D,S}^{-1} = \mathbf{I}, \quad (3.2)$$

wobei $\mathbf{I} \in SE(3)$ die Identitätstransformation beschreibt. Sei des Weiteren $\mathbf{T}_{S,E} \in SE(3)$ eine Transformation von E nach S , so ist

$$\mathbf{T}_{D,E} = \mathbf{T}_{D,S} \cdot \mathbf{T}_{S,E} \quad (3.3)$$

auch eine isometrische Transformation, die sich aus der Komposition von $\mathbf{T}_{D,S}$ und $\mathbf{T}_{S,E}$ ergibt und die Transformation von Koordinatensystem E nach D beschreibt. Sei außerdem $\mathbf{p}_S \in \mathbb{R}^3$ ein Raumpunkt bezüglich Koordinatensystem S , so ergibt sich durch

$$\mathbf{p}_D = \mathbf{T}_{D,S} \cdot \mathbf{p}_S \in \mathbb{R}^3 \quad (3.4)$$

die Darstellung dieses Raumpunktes bezüglich Koordinatensystem D . Im Folgenden wird ein Raumpunkt als Landmarke bezeichnet.

Nomenklatur

In den folgenden Kapiteln werden unterschiedliche Begriffe verwendet, die eine isometrische Transformation zwischen unterschiedlichen Koordinatensystemen darstellen. Abbildung 3.2 zeigt die Transformationen und Koordinatensysteme, die für die Betrachtungen in dieser Arbeit relevant sind. Die folgende Auflistung beschreibt diese im Detail:

- **Fahrzeugpose \mathbf{P}**

Eine Fahrzeugpose $\mathbf{P} = \mathbf{T}_{G,R}$ beschreibt die Transformation vom Referenzkoordinatensystem R des Versuchsträgers in ein globales Koordinatensystem G , wobei sich G aus dem Kontext der Betrachtungen ergibt. Eine Fahrzeugpose korrespondiert stets zu einer Menge $\Sigma = \{I_1, \dots, I_m\}$ von m synchron zu einem Zeitpunkt aufgezeichneten Kamerabildern.

- **Kamerapose \mathbf{C}**

Eine Kamerapose $\mathbf{C} = \mathbf{T}_{G,K}$ stellt stets die Transformation von einem Kamera-Koordinatensystem K in ein globales Koordinatensystem G dar. Da die Kameras rigid am Versuchsträger angebracht sind, kann eine Kamerapose \mathbf{C} stets als Komposition $\mathbf{P} \cdot \mathbf{K} = \mathbf{T}_{G,R} \cdot \mathbf{T}_{R,K}$ einer Fahrzeugpose \mathbf{P} und der extrinsischen Transformation \mathbf{K} dargestellt werden. Die Transformation \mathbf{K} beschreibt hierbei die isometrische Abbildung vom Kamerakoordinatensystem

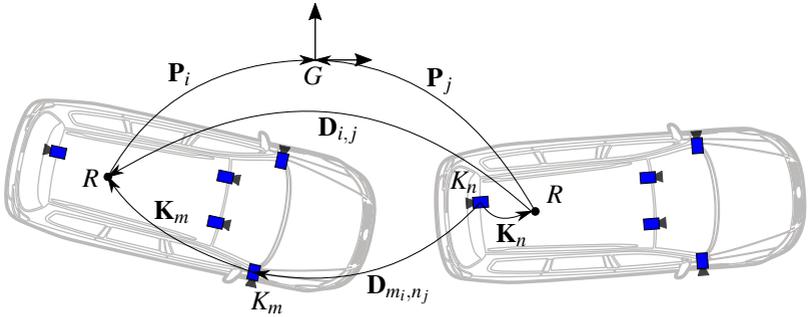


Abbildung 3.2: Darstellung der verwendeten Koordinatensysteme, der Fahrzeugposen und der extrinsischen Kameratransformationen.

K in das Referenzkoordinatensystem R des Versuchsträgers. Eine Kamera-
pose korrespondiert in der Arbeit immer zu einem aufgezeichneten Bild I .

- **Posendifferenz $D_{s,d}$**

Eine Posendifferenz $D_{i,j} = \mathbf{P}_i^{-1} \cdot \mathbf{P}_j$ beschreibt die isometrische Transfor-
mation von Fahrzeugpose \mathbf{P}_j nach Fahrzeugpose \mathbf{P}_i . In einigen Kapiteln wer-
den Posendifferenzen D_{m_i,n_j} zwischen Kameraposen betrachtet, für die die
gleiche Nomenklatur verwendet wird. Dies ist jedoch aus dem Kontext der
Betrachtungen ersichtlich und nachvollziehbar.

- **Posenmenge P**

Eine Posenmenge $P = \{\mathbf{P}_1, \dots, \mathbf{P}_n\}$ beschreibt eine Menge von Fahrzeu-
gposen. In den folgenden Kapiteln ergibt sich der Umfang einer betrachteten
Posenmenge aus dem Kontext der Betrachtungen.

- **Trajektorie T**

Eine Trajektorie $T = \{\mathbf{P}_1, \dots, \mathbf{P}_n\}$ bezeichnet im Kontext der Arbeit eine zeit-
diskrete Abtastung von Posen $\mathbf{P}_1, \dots, \mathbf{P}_n$ der gefahrenen Raumkurve des Fahr-
zeugs. Eine Trajektorie T ist also eine geordnete Posenmenge von konsekuti-
ven Fahrzeugposen. Die Abtastzeitpunkte korrespondierenden dabei zu den
Aufnahmezeitpunkten der synchron aufgezeichneten Bildmengen $\Sigma_1, \dots, \Sigma_n$
aus einer einzelnen Aufnahmesequenz S . Da eine hinreichend hohe Aufnah-
mefrequenz und eine maximale Bewegungsgeschwindigkeit des Fahrzeugs

angenommen wird, kann zwischen aufeinanderfolgenden Posen einer Trajektorie T eine glatte Bewegung angenommen werden. Die Posen einer Trajektorie lassen sich daher durch Bewegungsmodelle erklären und validieren.

3.3 Kalibrierung

Im Rahmen der Arbeit wird eine gemeinsame geometrische und projektive Kalibrierung aller Kameras vorausgesetzt [59]. Durch die Kalibrierung ist für jede montierte Kamera K_i eine Projektionsfunktion

$$\pi_i(\mathbf{l}) = \mathbf{z} \quad (3.5)$$

gegeben, die eine Landmarke $\mathbf{l} \in \mathbb{R}^3$ im Kamerakoordinatensystem K_i auf einen Bildpunkt $\mathbf{z} \in \mathbb{R}^2$ in der Sensorebene abbildet. Bei der Projektion einer Landmarke ins Kamerabild geht die Information des Abstandes der Landmarke zum Kameraursprung verloren. Daher ist Gleichung (3.5) nicht eindeutig invertierbar. Stattdessen stellt die Kalibrierung eine Abbildungsfunktion

$$\kappa_i(\mathbf{z}) = \mathbf{r} \quad (3.6)$$

bereit, die einen Bildpunkt \mathbf{z} auf einen Richtungsvektor $\mathbf{r} \in \mathbb{R}^3, |\mathbf{r}| = 1$ im Kamerakoordinatensystem abbildet. Der Richtungsvektor \mathbf{r} wird im Folgenden als Sichtstrahl von \mathbf{z} bezeichnet. Des Weiteren stellt die Kalibrierung eine Transformation $\mathbf{K}_i = \mathbf{T}_{R,K_i} \in SE(3)$ vom Koordinatensystem der i -ten Kamera in ein fahrzeugfestes Referenzkoordinatensystem R bereit. Wie in Abbildung 3.2 dargestellt, liegt der Ursprung des Referenzkoordinatensystem R üblicherweise im Mittelpunkt der Fahrzeughinterachse.

4 Merkmalsassoziation

Die Detektion, Beschreibung und Assoziation (engl. *feature matching*) von Bildmerkmalen ist die Grundlage der Kartengenerierung und der Lokalisierung. Merkmale werden an markanten Bildpunkten $\mathbf{z} \in \mathbb{R}^2$ detektiert und mittels eines Deskriptors $\mathbf{v} \in \mathbb{R}^n$ beschrieben. Über einen Vergleich der Deskriptoren werden Merkmale aus verschiedenen Bildern zueinander assoziiert, wobei die Merkmale im Falle einer korrekten Zuordnung dieselbe Landmarke in der Szene abbilden. Im Rahmen dieser Arbeit ist ein Merkmal

$$m = \{\mathbf{z}, \mathbf{v}\} \quad (4.1)$$

durch ein Tupel aus einem Bildpunkt \mathbf{z} und einem Deskriptor \mathbf{v} definiert. Zur Merkmalsdetektion wird jedes Bild mit einer Blob- und einer Eckendetektionsmaske diskret gefaltet [13]. Die extremalen Antworten dieser Faltung ergeben die markanten Bildpunkte. Als Deskriptor wird der DIRD-Deskriptor [14] verwendet, der sich im Gegensatz zu prominenten Alternativen wie dem ORB- [60], SIFT- [61], BRIEF- [62] oder BRISK- [63] Deskriptor durch seine Robustheit gegenüber Beleuchtungsveränderungen auszeichnet. Der DIRD-Deskriptor ermöglicht eine robuste Assoziation von Merkmalen aus Bildern, welche die gleiche Szene bei unterschiedlichen Lichtverhältnissen aus einer ähnlichen Perspektive darstellen. Er ist daher besonders zur Merkmalsassoziation zwischen Bildern aus verschiedenen Aufnahmesequenzen zu verschiedenen Tages- und Jahreszeiten geeignet.

Das Ziel der Merkmalsassoziation ist die korrekte Zuordnung eines Kandidatenmerkmals m_c aus einem Kandidatenbild I_c zu einem Merkmal m_v aus einem Vergleichsbild I_v , das die gleiche Landmarke abbildet wie m_c . Der erste Schritt der Merkmalsassoziation ist die Bestimmung von Bildpaaren, zwischen denen Merkmale überhaupt assoziiert werden können. Ein weiterer fundamentaler Schritt ist die Suchraumeinschränkung im Vergleichsbild, um eine möglichst hohe Assoziationsrobustheit zu erzielen. Insbesondere zu diesen beiden Aspekten wurden im Rahmen

der Arbeit Beiträge geleistet und Untersuchungen durchgeführt, die in Abschnitt 4.1 beschrieben sind. Die letztendliche Auswahl der richtigen Korrespondenz erfolgt über einen Vergleich des Deskriptors v_c des Kandidatenmerkmals m_c mit den Deskriptoren der Vergleichsmerkmale innerhalb des zuvor bestimmten Suchraums im Vergleichsbild I_v . Dieser Schritt ist in Abschnitt 4.2 im Detail beschrieben. Die paarweisen Merkmalskorrespondenzen werden zusätzlich auf Basis einer eindeutigen Identifikation einzelner Merkmale weiter verknüpft, wodurch ein höherer Verknüpfungsgrad erzielt wird. Dieser Prozess wird im Folgenden als globale Assoziation bezeichnet und ist in Abschnitt 4.3 beschrieben.

4.1 Suchraumeinschränkung

In diesem Abschnitt werden die Strategien zur Suchraumeinschränkung für eine möglichst robuste und effiziente Merkmalsassoziation beschrieben. Die Suchraumeinschränkung umfasst zwei Schritte:

1. Bestimmung von geeigneten Bildpaaren

Der erste Schritt ist die Zuordnung zweier Bilder, welche die gleiche Szene aus einer ähnlichen Perspektive abbilden, sodass überhaupt Merkmale detektiert werden können, die die gleiche Landmarke abbilden. Der maximale Unterschied der Aufnahmeposen hängt dabei vom verwendeten Detektor und Deskriptor ab. Dieser Prozessschritt schränkt somit die Menge der Vergleichsbilder für das Kandidatenbild I_c ein.

2. Einschränkung des Suchbereichs in der Bildebene

Der zweite Schritt ist die Einschränkung des Suchbereichs im Vergleichsbild I_v . Die Suchraumeinschränkung zielt hierbei stets darauf ab, dem Kandidatenmerkmal m_c eine möglichst kleine Menge $\Theta_{\text{cmp}} = \{m_1, \dots, m_N\}$ von Vergleichsmerkmalen aus dem Vergleichsbild I_v zuzuordnen. Je mehr der Suchraum eingeschränkt werden kann, desto geringer sind die Anforderungen an die Robustheit und Invarianzen des verwendeten Deskriptors.

Der DIRD-Deskriptor ist zwar robust gegen Beleuchtungsveränderungen, jedoch empfindlich gegenüber unterschiedlichen Projektionsverhalten der Kameras und Veränderungen in der Perspektive. Insbesondere wenn das MKS unterschiedliche Kameratypen umfasst oder wenn der Abstand zwischen den

Aufnahmeposen größer wird, beispielsweise wenn die gleiche Strecke mehrmals auf unterschiedlichen Fahrspuren befahren wird, sinkt die Anzahl richtiger Merkmalskorrespondenzen signifikant. Um auch in solchen Fällen eine möglichst hohe Assoziationsgüte zu erzielen, wurden im Rahmen der Arbeit Methoden entwickelt, die das zur Verfügung stehende Wissen über die räumliche Anordnung und das Projektionsverhalten der Kameras nutzen, um den Suchraum bestmöglich einzuschränken.

Die drei folgenden Unterabschnitte 4.1.2, 4.1.1 und 4.1.3 beschreiben die angewendeten Suchraumeinschränkungen für die drei in der Kartengenerierung und Lokalisierung auftretenden Anwendungsfälle. Für die Merkmalsassoziation von Merkmalen aus verschiedenen Aufnahmesequenzen im Rahmen der Kartengenerierung ist insbesondere die geometrische Suchraumeinschränkung, die in Abschnitt 4.1.3 beschrieben ist, von hoher Bedeutung.

4.1.1 Bildsequenzen

Die Assoziation von Merkmalen zwischen konsekutiven Bildern aus Bildsequenzen ist für den ersten Schritt der Kartengenerierung, der initialen Trajektorien-schätzung, relevant. In diesem Anwendungsfall ist kein Wissen über die umgebende Struktur bekannt. Das Wissen über die Posendifferenz der Kameraposen zu konsekutiven Kamerabildern basiert auf der Annahme, dass die Kamera zwischen den Aufnahmezeitpunkten, wie in Abschnitt 3.1 beschrieben, maximal wenige Meter weiterbewegt wurde. Die Auswahl der Vergleichsbilder ergibt sich in diesem Fall direkt aus der Aufnahme- und Abfolge der Bilder.

Für ein Kandidatenmerkmal ergibt sich in diesem Fall die Suchraumeinschränkung durch eine kreisförmige Suchregion, deren Mittelpunkt sich am Endpunkt eines prädierten Flussvektors befindet [13]. Die prädierten Flussvektoren ergeben sich dabei aus dem geschätzten optischen Fluss der vorhergegangenen Assoziationsschritte. Abbildung 4.1 zeigt den optischen Fluss einer vorwärts aus dem Fahrzeug gerichteten Kamera. Die in Grün dargestellten Merkmalsspuren (engl. *feature tracks*) visualisieren den optischen Szenenfluss.



Abbildung 4.1: Darstellung des optischen Flusses einer in Fahrtrichtung ausgerichteten Kamera in einem typischen Stadtszenario.

4.1.2 Merkmal-Landmarke Assoziation

Der zweite Fall ist die Assoziation von Merkmalen zur einer existierenden Landmarke \mathbf{l} . Dieser Fall tritt typischerweise bei der Merkmalsassoziation im Rahmen der metrischen Lokalisierung auf, bei der aus einer bestehenden Lokalisierungskarte Landmarken bereit gestellt werden. Zusätzlich ist eine ungefähre Schätzung der kartenrelativen Pose des Fahrzeugs bekannt. Die Landmarke \mathbf{l} korrespondiert dabei zu einer Menge $\Theta_{\mathbf{l}}$ von Merkmalen, aus denen die Landmarke vorab trianguliert wurde. Die Vergleichsbilderauswahl ergibt sich hierbei durch die folgenden Bedingungen:

- Die Aufnahmepose \mathbf{C}_c eines Kandidatenbildes I_c ist in der Nähe von einer der zu $\Theta_{\mathbf{l}}$ korrespondierenden Aufnahmeposes.
- Die Projektion $\mathbf{z}_l = \pi_c(\mathbf{C}_c^{-1}\mathbf{l})$ der Landmarke \mathbf{l} in das Kamerabild I_c erfüllt die Chiralitätsbedingung [64] – die Landmarke liegt also vor der Bildebene. Außerdem muss die Projektion \mathbf{z}_l innerhalb der Bildfläche von I_c liegen.

Sind diese beiden Kriterien erfüllt, existiert aus geometrischer Sicht eine hohe Wahrscheinlichkeit, dass ein Merkmal aus I_c zur Landmarke \mathbf{l} assoziiert werden kann, wobei eine Verdeckung der Landmarke mit diesen Kriterien nicht ausgeschlossen werden kann.

Die Suchraumeinschränkung im Kandidatenbild I_c ergibt sich durch eine kreisförmige Suchregion um die Projektion \mathbf{z}_l herum. Je unsicherer das Wissen über \mathbf{l} und \mathbf{C}_c hierbei ist, desto größer wird der Suchradius gewählt. Als Vergleichsmerkmal $m_v \in \Theta_{\mathbf{l}}$ wird das ausgewählt, dessen korrespondierende Kamerapose den kleinsten euklidischen Abstand zu \mathbf{C}_c hat.

4.1.3 Geometrische Assoziation ohne Strukturwissen

Dieser Fall beschreibt die Merkmalsassoziation zwischen zwei Bildern zu beliebigen Zeitpunkten aus unterschiedlichen Kameras der gleichen bzw. verschiedener Aufnahmesequenzen. Hierbei ist angenommen, dass Kameraposenschätzungen zu den Kamerabildern gegeben sind, jedoch kein Wissen über die umgebende Struktur vorhanden ist. Dieser Fall deckt eine Vielzahl von möglichen Anwendungen ab und ist in dieser Arbeit insbesondere wichtig für die Assoziation von Merkmalen zwischen Bildern aus verschiedenen Aufnahmesequenzen im Rahmen der Kartengenerierung.

Zuordnung von Vergleichsbildern

Für die Zuordnung von Vergleichsbildern zu einem Kandidatenbild I_c werden im ersten Schritt Bilder bestimmt, die in der Nähe des Kandidatenbildes aufgezeichnet wurden. Dazu wird zunächst eine Radiussuche [65] in einem k -d-Baum, der alle Kamerapositionen umfasst, durchgeführt. Ein k -d-Baum ist ein k -dimensionaler unbalancierter Suchbaum zur Speicherung von Punkten aus dem \mathbb{R}^k . Er ermöglicht effiziente Gebietsanfragen wie beispielsweise eine Radius- oder Nächste-Nachbarn-Suche. Wie in Abbildung 4.2 dargestellt, ermöglicht die Radiussuche im k -d-Baum ($k = 2$) eine effiziente Bestimmung der Bilder $\mathcal{Y}_{\text{near}} = \{I_1, \dots, I_n\}$, die innerhalb eines Umkreises mit Radius r um die zu I_c korrespondierende Kamerapose \mathbf{C}_c aufgezeichnet wurden. Die Suche erstreckt sich dabei über die zweidimensionale Befahrungsebene, die näherungsweise durch die Fahrzeugposen aufgespannt wird.

Darauffolgend wird für jedes Bild $I_k \in \mathcal{Y}_{\text{near}}$ der Blickwinkelunterschied

$$\alpha_{c,k} = \angle(\mathbf{C}_c, \mathbf{C}_k) \quad (4.2)$$

bestimmt, wobei \mathbf{C}_k die zu I_k korrespondierende Kamerapose ist. Ist $\alpha_{c,k}$ kleiner als ein Grenzwert, wird das Bild zur Menge \mathcal{Y}_{cmp} der Vergleichsbilder hinzugefügt, zwischen denen nachfolgend Merkmale assoziiert werden. Durch diesen Schwellwertvergleich wird sichergestellt, dass sich die zu den Bildern korrespondierenden Sichtkegel hinreichend überlappen. Dies ist in Abbildung 4.2 illustriert.

Suchraumeinschränkung durch Epipolargeometrie

Nach der Bestimmung der Vergleichsbilder \mathcal{Y}_{cmp} werden Merkmalsassoziationen

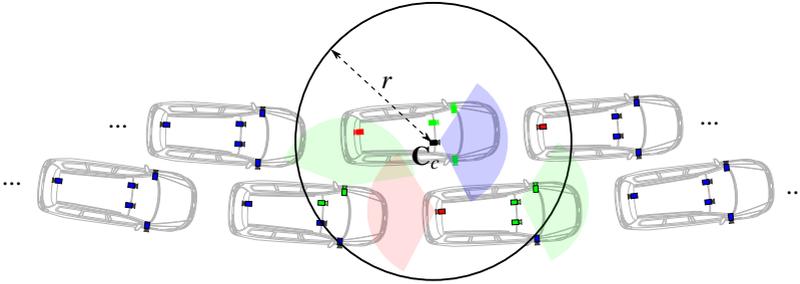


Abbildung 4.2: Bestimmung der Vergleichsbilder \mathcal{I}_{cmp} bei der geometrischen Assoziation ohne Strukturwissen. Die Bilder $\mathcal{I}_{\text{near}}$ zu den in Grün und Rot dargestellten Kameras werden durch die Radiussuche bestimmt. Die Bilder der in Rot dargestellten Kameras werden zurückgewiesen, da diese keine ausreichende Sichtkegelüberlappung mit dem Sichtkegel (Blau) des Kandidatenbildes haben.

zwischen dem Kandidatenbild I_c und den Bildern \mathcal{I}_{cmp} bestimmt. Da zu jedem Bild Wissen über die korrespondierende Kamerapose gegeben ist, kann die Kameraposendifferenz $\mathbf{D}_{c,v} = \mathbf{C}_c^{-1} \mathbf{C}_v$ zwischen der zum Kandidatenbild I_c korrespondierenden Kamerapose \mathbf{C}_c und der zum Vergleichsbild $I_v \in \mathcal{I}_{\text{cmp}}$ korrespondierenden Kamerapose \mathbf{C}_v bestimmt werden. Auf Basis der Kameraposendifferenz $\mathbf{D}_{c,v}$ kann der Suchraum im Vergleichsbild I_v unter Ausnutzung der Epipolargeometrie [64] beschränkt werden.

Die Suchraumeinschränkung durch Epipolargeometrie bestimmt einen Suchschlauch um die Projektion des Sichtstrahls des Kandidatenmerkmals m_c ins Vergleichsbild I_v herum. Im Folgenden wird die Projektion des Sichtstrahls als Epipolarkurve bezeichnet. Da eine allgemeine Projektionsfunktion (3.5) und eine beliebige Kameraposendifferenz $\mathbf{D}_{c,v}$ angenommen wird, ist die Epipolarkurve i. A. gekrümmt. Abbildung 4.3 demonstriert dies beispielhaft an Epipolarkurven von Merkmalen zwischen zwei orthogonal zueinander ausgerichteten Weitwinkelkameras. Als Projektionsfunktion (3.5) wurde in diesem Fall eine sphärische Abbildungsfunktion [59] verwendet.

Die Umsetzung dieser Suchraumeinschränkung ergibt sich durch geometrische Betrachtungen im dreidimensionalen Raum. Zunächst wird einmalig ein Referenznormalenvektor \mathbf{n}_0 einer Referenzepipolarebene bestimmt. Der Normalenvektor

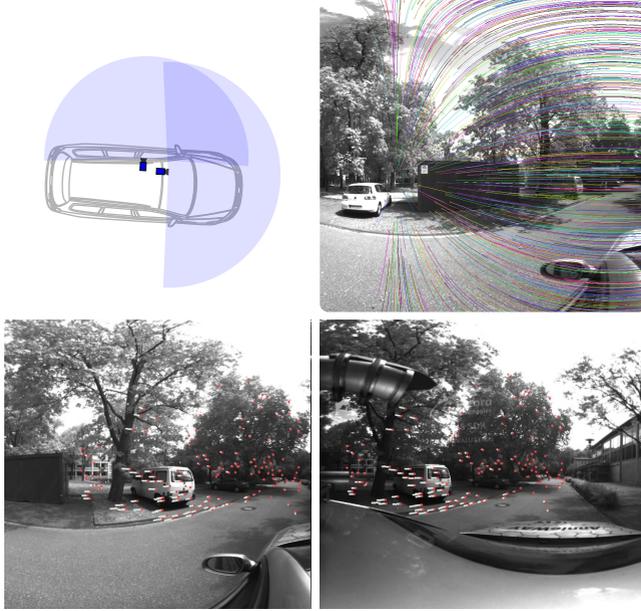


Abbildung 4.3: Oben: Epipolarlinien zwischen zwei orthogonal zueinander ausgerichteten Weitwinkelkameras. Unten: Flussvektoren von Korrespondenzen, für die eine Merkmalsassoziation zwischen den beiden Kameras erfolgreich bestimmt werden konnte.

$\mathbf{n}_0 = \mathbf{r}_0 \times \mathbf{r}_{epi}$ ergibt sich aus dem Kreuzprodukt eines beliebigen Sichtstrahls \mathbf{r}_0 und dem Verbindungsvektor \mathbf{r}_{epi} der Epipole. Darauf basierend wird jedem detektierten Merkmal m aus I_c und I_v der Winkel

$$\alpha_m = \angle(\mathbf{n}_0, \mathbf{n}_m) = \arccos\left(\frac{\mathbf{n}_0^T \mathbf{n}_m}{\|\mathbf{n}_0\| \|\mathbf{n}_m\|}\right) \quad (4.3)$$

zwischen dem Referenznormalenvektor \mathbf{n}_0 und dem Normalenvektor \mathbf{n}_m zugeordnet. Der Normalenvektor \mathbf{n}_m ergibt sich dabei aus dem Kreuzprodukt des Verbindungsvektors \mathbf{r}_{epi} und des Sichtstrahls \mathbf{r}_m des Merkmals m . Durch dieses Vorgehen ergibt sich für das Kandidatenmerkmal m_c der Winkel α_{m_c} und für ein Vergleichsmerkmal m_v der Winkel α_{m_v} . Basierend auf diesen Winkeln kann durch einen skalaren Schwellwertvergleich $|\alpha_{m_c} - \alpha_{m_v}| < \alpha_{\max}$ der Differenz der Winkel geprüft wer-

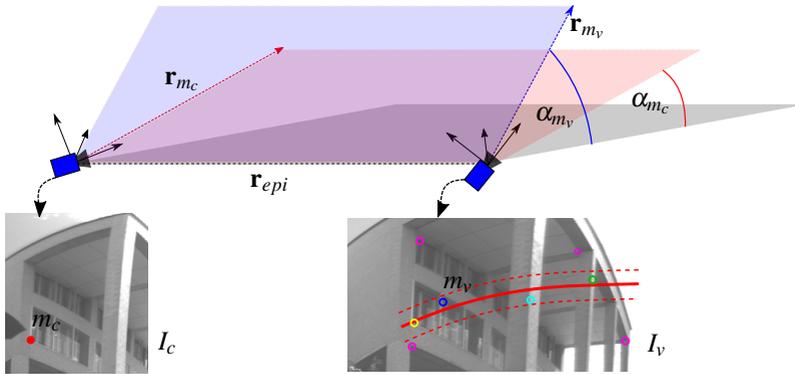


Abbildung 4.4: Geometrische Verhältnisse bei der Suchraumeinschränkung durch Epipolargeometrie.

den, ob beide Epipolarebenen eine ähnliche Orientierung im Raum haben. Das ist gleichbedeutend dazu, dass m_v innerhalb des Suchschlauchs um die Epipolarcurve des Merkmals α_{m_c} im Vergleichsbild I_v liegt. Abbildung 4.4 stellt dieses Vorgehen schematisch dar. Die graue Ebene stellt die Referenzebene dar, die zum Referenznormalenvektor \mathbf{n}_0 korrespondiert. Die in Rot dargestellte Ebene korrespondiert zum Kandidatenmerkmal m_c und die in Blau dargestellte Ebene zu einem Vergleichsmerkmal m_v aus I_v . Die Vergleichsmerkmale sind als farblich hervorgehobene Kreise im Vergleichsbild I_v dargestellt. Alle Merkmale innerhalb des in Rot gestrichelt dargestellten Suchschlauchs um die in Rot dargestellte Epipolarcurve in I_v werden durch die epipolargeometrische Suchraumeinschränkung ausgewählt. Das in Gelb dargestellte Merkmal entspricht der korrekten Korrespondenz zu m_c . Die Breite des Suchschlauchs ist proportional zum Schwellwert α_{\max} . Je ungenauer die Schätzung von $\mathbf{D}_{c,k}$ ist, desto größer wird α_{\max} gewählt.

Der Vorteil der Verwendung eines Referenznormalenvektors \mathbf{n}_0 im Vergleich zu einem direkten Vergleich der Normalenvektoren \mathbf{n}_{m_c} und \mathbf{n}_{m_v} der beiden zu vergleichenden Merkmale m_c und m_v ist, dass die Winkelberechnung in Gleichung (4.3) nur ein einziges Mal pro Merkmal durchgeführt werden muss. Die paarweisen Vergleiche können darauffolgend durch eine einfache Differenzbildung berechnet werden,



Abbildung 4.5: Exemplarische Auswahl von Bildpaarkombinationen, zwischen denen erfolgreich Merkmale mit der epipolargeometrischen Suchraumeinschränkung bestimmt wurden. Die Bilder entspringen dabei jeweils aus unterschiedlichen Aufnahmesequenzen.

wodurch Rechenzeit eingespart wird.

Evaluation

Die beschriebene Suchraumeinschränkung wurde im Rahmen der Arbeit in einem Monte-Carlo Experiment untersucht. Dazu wurden zufällig Kamerabildpaare gewählt, deren Kameraposen ähnlich zueinander sind. Die abgebildeten Szenen in diesen Experimenten umfassen innerstädtische Gebiete, Überlandstraßen und Wald-durchfahrten aus insgesamt 10 verschiedenen Aufnahmesequenzen. Das Experiment umfasst Aufnahmesequenzen, die sowohl mit MKS-A als auch mit MKS-B aufgezeichnet wurden (siehe Abschnitt 3.1). Zu allen Kamerabildern existieren Kameraposen, die mit der in Abschnitt 6 beschriebenen Kartierungsmethode bestimmt wurden. Diese werden als Grundwahrheit bei der Evaluation verwendet. Für die Zuordnung von Vergleichsbildern wurde eine maximale euklidische Distanz von 5 m und ein maximaler Blickwinkelunterschied von 90° verwendet. Abbildung 4.5 zeigt ver-

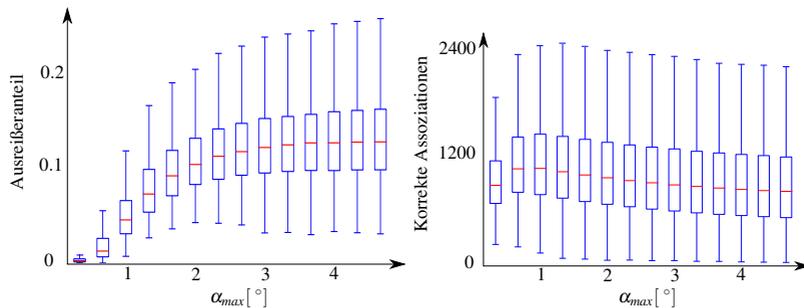


Abbildung 4.6: Boxplots des Ausreißeranteils (Links) und der Anzahl erfolgreicher Korrespondenzen (Rechts) für steigende Schwellwertwinkel α_{\max} .

schiedene Bildpaarungen, zwischen denen Korrespondenzen mit der epipolargeometrischen Suchraumeinschränkung bestimmt wurden. Alle Bilder wurden mit einem sphärischen Projektionsmodell entzerrt.

In dem Experiment wurden für jedes Kamerabildpaar Korrespondenzen bestimmt, wobei die Assoziation pro Bildpaar mit verschiedenen Winkelschwellwerten α_{\max} durchgeführt wurde. Die finale Korrespondenz wurde durch den Deskriptorvergleich gemäß Abschnitt 4.2 bestimmt. Zur Entscheidung, ob eine so bestimmte paarweise Korrespondenz korrekt ist, wurde eine Landmarke auf Basis der Grundwahrheitskameraposen trianguliert und in die Ursprungsbilder der beiden Merkmale projiziert. Eine Korrespondenz wird als korrekt klassifiziert, wenn der Rückprojektionsfehler in beiden Bildern < 2 Pixel ist. Der Rückprojektionsfehler ergibt sich hierbei durch den Betrag der Differenz der Pixelpositionen des detektierten Merkmals und der Position der Landmarkenprojektion im Kamerabild. Eine falsch klassifizierte Korrespondenz wird im Folgenden als Ausreißer (engl. *outlier*) bezeichnet. Abbildung 4.6 zeigt die Ergebnisse dieser Untersuchung. In Abbildung 4.6 ist links der Boxplot des Ausreißeranteils und rechts der Boxplot der Anzahl korrekter Korrespondenzen für verschiedene Winkelschwellwerte α_{\max} dargestellt. Die Boxplots umfassen dabei die Ergebnisse aus allen paarweisen Merkmalsassoziationen zwischen zwei verschiedenen Bildern. Für sehr kleine α_{\max} -Schwellwerte sinkt die Anzahl korrekter Korrespondenzen als auch der Anteil falscher Korrespondenzen. Für steigen-

de α_{\max} -Schwellwerte bleibt die Anzahl gefundener Korrespondenzen nahezu konstant, jedoch steigt der Ausreißeranteil an. Für $\alpha_{\max} \approx 1^\circ$ ist die Anzahl gefundener Korrespondenzen maximal, wobei der Ausreißeranteil kleiner ist als bei größeren α_{\max} -Werten. Zusätzlich steigt die Rechenzeit deutlich an, da aufgrund des breiteren Suchschlauchs mehr Deskriptoren von Vergleichsmerkmalen mit Deskriptoren der Kandidatenmerkmale verglichen werden. Da eine möglichst hohe Anzahl korrekter Korrespondenzen, ein möglichst geringer Ausreißeranteil und eine möglichst geringe Laufzeit von Vorteil ist, stellt $\alpha_{\max} \approx 1^\circ$ die optimale Wahl in diesem Experiment dar.

Das Experiment zeigt, dass die epipolare geometrische Merkmalsassoziation unter Ausnutzung des geometrischen Vorwissens eine robuste und effiziente Merkmalsassoziation ermöglicht, da bei geeigneter Wahl des Parameters α_{\max} die Anzahl gefundener Korrespondenzen maximal ist und gleichzeitig ein geringerer Ausreißeranteil erzielt wird als bei größeren α_{\max} -Werten. Ein Vorteil dieser Methode ist, dass sie unabhängig vom verwendeten Deskriptor und Detektor ist – also in einem allgemeinen Sinne supplementär ist.

4.2 Deskriptor-Vergleich

Durch die Suchraumeinschränkung für ein Kandidatenmerkmal m_c im Kandidatenbild I_c , verbleibt eine Menge $\Theta_{\text{cmp}} = \{m_1, \dots, m_n\}$ von n Vergleichsmerkmalen im Vergleichsbild I_v . Das finale Assoziationspaar wird mittels Deskriptorvergleich des Deskriptors \mathbf{v}_c mit den Deskriptoren der Merkmale aus Θ_{cmp} bestimmt. Dazu wird eine symmetrische Distanzfunktion

$$d_{m_i, m_j} = \delta_m(\mathbf{v}_i, \mathbf{v}_j) \quad (4.4)$$

verwendet, die ein Unterschiedlichkeitsmaß $d_{m_i, m_j} \in \mathbb{R}^+$ auf Basis der beiden korrespondierenden Deskriptor-Vektoren $\mathbf{v}_i, \mathbf{v}_j$ bestimmt. Auf Basis der Distanzfunktion (4.4) kann dem Kandidatenmerkmal m_c die Korrespondenz

$$m_a = \arg \min_{m_i \in \Theta_{\text{cmp}}} d_{m_c, m_i} \quad (4.5)$$

zugeordnet werden, die sich am wenigsten vom Deskriptor \mathbf{v}_c des Kandidatenmerkmals m_c unterscheidet. Da im Rahmen der Arbeit die byte-quantisierte Variante des DIRD-Deskriptors [14] verwendet wird, wird die L_1 -Norm der Deskriptordifferenz $\mathbf{v}_i - \mathbf{v}_j$ als Distanzfunktion verwendet.

Um die Robustheit der Merkmalsassoziation zu steigern, werden weitere Vergleiche durchgeführt, die in der folgenden Auflistung beschrieben sind:

- Die Deskriptordistanz d_{m_c, m_a} des besten Assoziationspaars wird mit einem Schwellwert $T_a \in \mathbb{R}$ verglichen. Wenn $d_{m_c, m_a} > T_a$ gilt, wird die Korrespondenz verworfen.
- Die Eindeutigkeit der besten Korrespondenz $\{m_c, m_a\}$ wird mittels

$$u_{m_c, m_a} = 1 - \frac{d_{m_c, m_a}}{d_{m_c, m_b}} \quad (4.6)$$

bewertet, wobei d_{m_c, m_b} die Distanz von m_c zum zweitähnlichsten Merkmal $m_b \in \Theta_{\text{cmp}}$ ist. Wenn der Deskriptordistanzunterschied u_{m_c, m_a} kleiner als ein Schwellwert $T_u \in [0, 1]$ ist, wird die Korrespondenz verworfen.

- Es wird ein sogenannter Links-Rechts-Vergleich durchgeführt. Das bedeutet, dass für jede zugeordnete Assoziationen im Vergleichsbild eine Korrespondenzsuche im Kandidatenbild durchgeführt wird. Wird dabei ein anderes Merkmal als Kandidatenmerkmal m_c assoziiert, wird die Korrespondenz verworfen.

Nur wenn alle zuvor beschriebenen Überprüfungen erfolgreich sind, wird die Merkmalskorrespondenz akzeptiert.

4.3 Globale Assoziation

Zur Steigerung des Verknüpfungsgrads der Korrespondenzen werden die paarweisen Merkmalskorrespondenzen, die durch die zuvor beschriebene bildbasierte Merkmalsassoziation bestimmt wurden, weiter verknüpft. Dazu ist jedes Merkmal eindeutig durch seine diskrete Position im Kamerabild und einen eindeutigen Bildindex identifizierbar. Sind beispielsweise $\{m_i, m_j\}$ und $\{m_i, m_k\}$ zwei verschiedene Merkmalskorrespondenzen, die beide das Merkmal m_i umfassen, so werden diese beiden

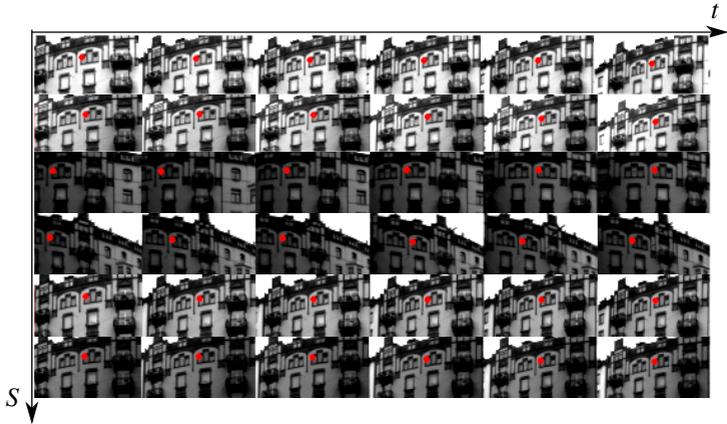


Abbildung 4.7: Exemplarische Darstellung einer einzelnen globalen Merkmalskorrespondenz. Die Korrespondenz erstreckt sich dabei über mehrere Zeitschritte (Horizontal) und verschiedene Kameras aus unterschiedlichen Aufnahmesequenzen (Vertikal).

Korrespondenzen zu $\{m_i, m_j, m_k\}$ durch die globale Assoziation zu einer gemeinsamen Korrespondenz zusammengefasst.

Die globale Assoziation lässt sich als Graphenproblem formulieren. Jedes Merkmal stellt dabei einen Knoten und jede Korrespondenz zwischen zwei Merkmalen eine ungerichtete Kante dar. Jede verknüpfte Komponente (engl. *connected component*) dieses Graphen entspricht somit einer globalen Korrespondenz. Abbildung 4.7 zeigt exemplarisch Ausschnitte aus unterschiedlichen Bildern. Die in Rot hervorgehoben Bildpunkte zeigen die Merkmalspunkte einer einzelnen Korrespondenz, die durch die globale Assoziation von paarweisen Merkmalskorrespondenzen zusammengesetzt wurde.

5 Robuste Parameterschätzung

In diesem Kapitel werden die eingesetzten Methoden zur robusten Bestimmung der Kartenparameter und zur robusten Lokalisierungsposenschätzung beschrieben. Die Kartenparameter umfassen Fahrzeugposen und Landmarken. Die Parameterschätzung entspricht in beiden Fällen dem Lösen eines überbestimmten nichtlinearen Gleichungssystems. Zur Lösung solcher Ausgleichsprobleme werden in dieser Arbeit gradientenabstiegsbasierte Optimierungsverfahren eingesetzt. Die zu schätzenden Modellparameter werden dabei durch Minimierung einer Fehlerfunktion bestimmt, wodurch die Messungen bestmöglich beschrieben werden. Abschnitt 5.1 befasst sich mit den aufgestellten Ausgleichsproblemen und Lösungsmethoden im Rahmen dieser Arbeit.

Wesentlich für eine robuste Parameterschätzung durch numerische Optimierung sind die folgenden zwei Aspekte:

- **Ausreißerbehandlung**

In der Praxis treten häufig fehlerbehaftete Messungen und falsche Korrespondenzen auf. Bereits wenige Ausreißer haben einen signifikanten Effekt auf die Optimierung und führen i. A. zu einer falschen Parameterschätzung. Daher müssen Ausreißer vor der Optimierung eines Ausgleichsproblems identifiziert und entfernt oder ihr Einfluss auf die Fehlerfunktion minimiert werden. Dafür werden in dieser Arbeit unterschiedliche Methoden eingesetzt, die in Abschnitt 5.1.4 beschrieben werden.

- **Initialisierung**

Die betrachteten Ausgleichsprobleme sind nichtlinear, wodurch eine i. A. nicht konvexe Kostenfunktion minimiert wird. Da numerische Gradientenabstiegsverfahren zur Lösung verwendet werden, besteht die Gefahr, in einem lokalen Minimum zu konvergieren. Eine geeignete Initialisierung der zu optimierenden Parameter in der Nähe des korrekten Minimums ist daher von fun-

damentaler Bedeutung. Die verwendeten Initialisierungsmethoden werden in Abschnitt 5.2 beschrieben.

Ein weiteres Problem sind Gebiete mit nur wenigen markanten Strukturen, Szenarien, in denen überwiegend dynamische Objekte beobachtet werden, oder schlechte Sichtbedingungen. In solchen Fällen können oft nur wenige richtige Merkmalskorrespondenzen bestimmt werden. Das kann dazu führen, dass Parameter nicht vollständig beobachtbar sind oder dass der Ausreißeranteil sehr groß ist, sodass die Kartenparameter in diesen Gebieten falsch geschätzt werden. Daher werden die geschätzten Parameter an verschiedenen Stellen im Kartierungsprozess durch Modellannahmen validiert und gegebenenfalls korrigiert. Diese Validierung steigert die Robustheit der Kartengenerierung, da falsch geschätzte Parameter frühzeitig identifiziert werden. Die Modellannahmen können zusätzlich als Kostenterme in die Ausgleichsprobleme integriert werden, wodurch eine Abweichung der Parameter von den Modellen bestraft wird. Die verwendeten Modelle und Validierungsmethoden werden in Abschnitt 5.3 ausführlich beschrieben.

5.1 Nichtlineare Ausgleichsrechnung

Die Schätzung der Kartenposen und -landmarken als auch die Schätzung der Lokalisierungspose wird durch die Minimierung einer Summe von quadratischen Fehlertermen erzielt. Dieses Vorgehen ist in der Literatur bekannt als die Methode der kleinsten Fehlerquadrate (engl. *least-squares*; LS). Sei $\mathbf{x} \in \mathbb{R}^n$ der zu optimierende n -dimensionale Parametervektor und $\mathbf{r}(\mathbf{x}) = (r_1(\mathbf{x}), \dots, r_m(\mathbf{x}))^T$ eine m -dimensionale Residuenfunktion, dann ist

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{r}(\mathbf{x})\|_{\Omega}^2 \quad (5.1)$$

die Parameterschätzung, die das durch $\mathbf{r}(\cdot)$ beschriebene Beobachtungsmodell bestmöglich approximiert. Durch die Gewichtungsmatrix Ω werden die einzelnen Residuen gewichtet. Die Quadrierung der Residuenfunktion $\mathbf{r}(\cdot)$ ist durch den Satz von Gauß-Markov motiviert. Für den Fall einer linearen Residuenfunktion besagt der Satz, dass ein Schätzer, der den quadratischen Fehler minimiert, der beste lineare erwartungstreue Schätzer (engl. *best linear unbiased estimator*; BLUE) im Sinne der kleinsten Varianz ist [31].

Da die Fehlerterme im Rahmen der Arbeit nichtlinear sind, ergibt sich ein nichtlineares Ausgleichsproblem (engl. *nonlinear least-squares*; NLS). In diesem Fall kann Problem (5.1) i. A. nicht geschlossen gelöst werden. Stattdessen werden dafür iterative numerische Optimierungsverfahren eingesetzt. In dieser Arbeit wird der weit verbreitete Levenberg-Marquardt Algorithmus (LM) [21] verwendet. Des Weiteren gilt für alle im Rahmen der Arbeit verwendeten Ausgleichsprobleme, dass die Jacobi-Matrix der Residuenfunktion $\mathbf{r}(\cdot)$ dünn besetzt (engl. *sparse*) ist. Diese Eigenschaft nutzen die verwendeten *Sparse-Matrix-Solver* [32] aus, wodurch auch Probleme mit sehr vielen Parametern und Residuen effizient gelöst werden können. Zur Lösung der Ausgleichsprobleme in dem zu dieser Arbeit entwickelten Softwareframework wird die C++-Bibliothek *Ceres* [66] verwendet.

5.1.1 Bündelblockausgleich

Das am häufigsten angewendete Verfahren zur gemeinsamen Schätzung von Posen, Kameramodellparametern und Landmarken auf Basis von assoziierten Merkmalen ist der sogenannte Bündelblockausgleich (engl. *bundle adjustment*).

Eine Landmarke \mathbf{l}_i korrespondiert stets zu einer Menge $Z_i = \{\mathbf{z}_1, \dots, \mathbf{z}_N\}$ von Bildpunkten, die sich aus der zugehörigen Merkmalskorrespondenz $\Theta_i = \{m_1, \dots, m_N\}$ ergibt. Der dem Merkmal m_i zugeordnete Bildpunkt \mathbf{z}_i entspringt aus dem Bild I_{z_i} , das Kamera K_{z_i} aufgezeichnet hat. Außerdem sei \mathbf{P}_{z_i} die Pose des Fahrzeugs zum Zeitpunkt der Aufnahme des Bildes I_{z_i} . Damit kann für jedes Merkmal der zweidimensionale Rückprojektionsfehler

$$r_{z_i} = \mathbf{z}_i - \pi_{z_i}(\mathbf{K}_{z_i}^{-1} \cdot \mathbf{P}_{z_i}^{-1} \cdot \mathbf{l}_i) \quad (5.2)$$

im Bildraum bestimmt werden, wobei $\pi_{z_i}(\cdot)$ die Projektionsfunktion (3.5) und \mathbf{K}_{z_i} die extrinsische Transformation der Kamera K_{z_i} darstellt. Durch Einsetzen von Gleichung (5.2) in Problemformulierung (5.1) ergibt sich das Bündelblockausgleichsproblem. Sei P die Menge der zu optimierenden Fahrzeugposen und L die Menge der zu optimierenden Landmarken, so ergibt sich durch Minimierung

$$\hat{P}, \hat{L} = \arg \min_{P, L} \sum_{j=1}^{|L|} \sum_{\mathbf{z}_i \in Z_j} \|r_{z_i}\|_{\Omega_{z_i}}^2 \quad (5.3)$$

die Menge an Posen \hat{P} und Landmarken \hat{L} , welche die involvierten Bildpunkt-Korrespondenzen $Z_1, \dots, Z_{|L|}$ am besten beschreibt. Jeder Messung ist hierbei eine Gewichtungsmatrix $\Omega_{z_i} \in \mathbb{R}^{2 \times 2}$ zugeordnet.

5.1.2 Clusterausgleich

In Anlehnung an das Bündelblockausgleichsproblem (5.3) wird in dieser Arbeit ein weiteres Problem formuliert, das im Folgenden als Clusterausgleichsproblem bezeichnet wird. Die Idee dabei ist, lokale Posensegmente, die im Folgenden als Cluster bezeichnet werden, als “starre Körper” aufzufassen und nur die Cluster aneinander anzupassen.

Ein Cluster $C \subseteq P$ bezeichnet i. A. eine Teilmenge der Menge aller betrachteten Posen P . Im Rahmen der Kartengenerierung ist ein Cluster jedoch stets ein lokales Trajektoriensegment von konsekutiven Fahrzeugposen aus einer einzelnen Aufnahme-sequenz. Seien C_1, \dots, C_n disjunkte Cluster, wobei o. B. d. A. $|C_i| = |\{\mathbf{P}_1, \dots, \mathbf{P}_m\}_i|$ die Anzahl der Posen eines Clusters ist. Abbildung 5.1 zeigt beispielsweise $n = 2$ Cluster mit jeweils $|C_1| = 3$ bzw. $|C_2| = 4$ Posen. Wie in Abbildung 5.1 dargestellt, wird für jedes Cluster C_i die erste Pose als Basispose \mathbf{B} definiert. Alle verbleibenden Posen $\mathbf{P}_2 = \mathbf{B} \cdot \mathbf{D}_2, \dots, \mathbf{P}_m = \mathbf{B} \cdot \mathbf{D}_m$ in Cluster C_i können dann relativ zur Basispose \mathbf{B} formuliert werden. Somit kann C_i durch die Menge $\{\mathbf{B}, \mathbf{D}_2, \dots, \mathbf{D}_m\}_i$ repräsentiert werden. Sei des Weiteren $\Theta = \{\Theta_1, \dots, \Theta_{|L|}\}$ die Menge der Merkmalskorrespondenzen, die mindestens zwei verschiedene Cluster verknüpfen. Eine Merkmalskorrespondenz Θ_i verknüpft dabei k Cluster, wenn sie $k \leq |\Theta_i|$ Merkmale umfasst, deren Ursprungskamerabilder zu k verschiedenen Fahrzeugposen korrespondierenden, die alle einem unterschiedlichen Cluster zugeordnet sind. Außerdem sei L die Menge der zu Θ korrespondierenden Landmarken. In Anlehnung an Gleichung (5.2) kann so der Rückprojektionsfehler

$$r_{z_i} = \mathbf{z}_i - \pi_{z_i}(\mathbf{K}_{z_i}^{-1} \cdot \mathbf{D}_{z_i}^{-1} \cdot \mathbf{B}_{z_i}^{-1} \mathbf{l}_i) \quad (5.4)$$

bezüglich der Basispose \mathbf{B}_{z_i} des zu \mathbf{z}_i korrespondierenden Clusters C_{z_i} formuliert werden. Des Weiteren bezeichnet \mathbf{D}_{z_i} die Relativpose innerhalb des Clusters C_{z_i} und \mathbf{l}_i die Landmarke, die zum Bildpunkt \mathbf{z}_i korrespondiert. Wie bei der Formulierung des Bündelblockausgleichsproblems bezeichnet $\pi_{z_i}(\cdot)$ die Projektionsfunktion und

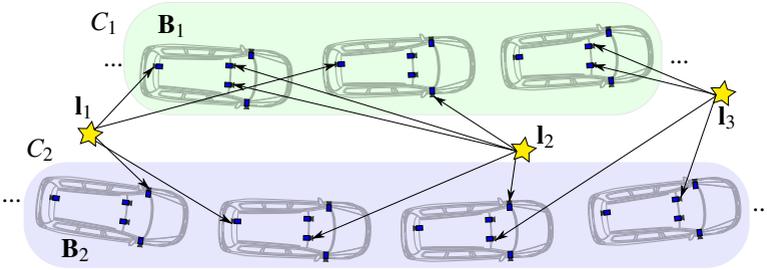


Abbildung 5.1: Schematische Darstellung des Clusterausgleichsproblems (5.5) für zwei Cluster und drei Landmarken, welche die beiden Cluster auf Basis von Merkmalskorrespondenzen (Kanten) verknüpfen.

\mathbf{K}_{z_i} die extrinsische Transformation der Kamera K_{z_i} , die das Kamerabild, aus dem der Bildpunkt \mathbf{z}_i entspringt, aufgezeichnet hat. In Analogie zum Bündelblockausgleichsproblem (5.3) ergibt sich durch Minimierung

$$\hat{B}, \hat{L} = \arg \min_{B, L} \sum_{j=1}^{|L|} \sum_{z_i \in Z_j} \|r_{z_i}\|_{\Omega_{z_i}}^2, \quad (5.5)$$

die optimale Schätzung \hat{B} der Basisposen B aller involvierten Cluster und die optimale Schätzung \hat{L} der Landmarken L , die mindestens in zwei der $|B|$ verschiedenen Cluster beobachtet wurden. Im Falle von $|C_i| = 1, \forall i \in \{1, \dots, n\}$ geht Problem (5.5) in das Bündelblockausgleichsproblem (5.3) über.

Durch den Clusterausgleich werden im Vergleich zum Bündelblockausgleich sowohl Parameter als auch Residuen eingespart. Dadurch reduziert sich die Dimension der Jacobi-Matrix der Residuenfunktion, da sich diese aus dem Produkt aus der Anzahl der Messungen und der Anzahl der Parameter im Ausgleichsproblem ergibt. Der Berechnungsaufwand zur Lösung von Ausgleichsproblemen durch numerische Optimierungsverfahren ist wiederum maßgeblich von der Dimension der Jacobi-Matrix der Residuenfunktion abhängig, da diese in jedem Optimierungsschritt im aktuellen Zustand ausgewertet und invertiert werden muss [21]. Daher wird auch der Berechnungsaufwand zur Lösung eines Clusterausgleichsproblems im Vergleich zum Bündelblockausgleich signifikant reduziert. Die wesentliche Reduktion des Berech-

nungsaufwands ergibt sich dabei durch die reduzierte Anzahl der Landmarken, da nur jene Landmarken geschätzt werden, welche die verschiedenen Cluster auf Basis ihrer zugehörigen Merkmalskorrespondenzen miteinander verbinden.

5.1.3 Posendifferenzenausgleich

Ein integraler Bestandteil der Kartengenerierung ist die gemeinsame Schätzung von Fahrzeugposen P auf Basis gegebener Posendifferenzen D . Im Rahmen der Kartengenerierung wird die Trajektorie einer neuen Aufnahmesequenz, welche der Karte hinzugefügt werden soll, durch Lösen eines Posendifferenzenausgleichsproblems an die bestehende Karte angepasst.

Eine Posendifferenz $\mathbf{D}_{i \rightarrow j} \in D$ assoziiert jeweils zwei Fahrzeugposen $\mathbf{P}_i, \mathbf{P}_j \in P$ aus der gleichen oder unterschiedlichen Aufnahmesequenzen. Das Posendifferenzresiduum

$$r_{\mathbf{D}_{i,j}} = \xi \left(\mathbf{D}_{i \rightarrow j} \cdot \mathbf{P}_j^{-1} \cdot \mathbf{P}_i \right) \quad (5.6)$$

bewertet den Unterschied zwischen der Posendifferenzmessung $\mathbf{D}_{i,j}$ und der durch die Fahrzeugposen \mathbf{P}_i und \mathbf{P}_j induzierte Posendifferenz $\mathbf{P}_i^{-1} \cdot \mathbf{P}_j$ [67]. Die Funktion $\xi : SE(3) \rightarrow \mathbb{R}^6$ bildet eine isometrische Transformation auf einen sechsdimensionalen Vektor ab. Die ersten drei Komponenten dieses Vektors repräsentieren die Rotation in einer Winkelachsendarstellung [68] und die verbleibenden drei Komponenten die Translation. Im Falle von $\mathbf{D}_{i,j} = \mathbf{P}_i^{-1} \cdot \mathbf{P}_j$ gilt daher $r_{\mathbf{D}_{i,j}} = \mathbf{0}$.

Der Posendifferenzenausgleich

$$\hat{P} = \arg \min_P \sum_{\mathbf{D}_{i,j} \in D} \|r_{\mathbf{D}_{i,j}}\|_{\Omega_{\mathbf{D}_{i,j}}}^2, \quad (5.7)$$

ergibt die optimale Posenschätzung \hat{P} der Posen P bezüglich der Posendifferenzmessungen D . Zusätzlich wird angenommen, dass für jede Posendifferenz eine positiv semi-definite symmetrische Gewichtungsmatrix $\Omega_{\mathbf{D}_{i,j}} \in \mathbb{R}^{6 \times 6}$ gegeben ist. Stellt die Struktur des Parametergraphen von Posendifferenzenausgleichsproblem (5.7) eine Baumstruktur dar, ist das Problem eindeutig bestimmt und hat daher ein verbleibendes Gesamtresiduum von 0. Die optimale Lösung kann in diesem Fall durch Integration (mittels Traversierung durch Breiten- oder Tiefensuche; siehe Abschnitt 5.2.3) der Posendifferenzen bestimmt werden.

Im Rahmen der Kartengenerierung wird eine Teilmenge $P_{\text{fix}} \subset P$ beim Lösen von Ausgleichsproblem (5.7) fixiert. Dadurch passen sich die verbleibenden Posen $P \setminus P_{\text{fix}}$ an die konstant gehaltenen Posen P_{fix} an. Des Weiteren werden zur Kartengenerierung GNSS-Positionsmessungen

$$r_{\mathbf{p}_i, \text{GNSS}} = \mathbf{p}_i - \mathbf{p}_{i, \text{GNSS}} \quad (5.8)$$

zur globalen Korrektheit der Kartenposenschätzung und zur Georeferenzierung in den Posendifferenzenausgleich (5.7) integriert, sofern diese verfügbar sind. Dabei ist $\mathbf{p}_i \in \mathbb{R}^3$ die Translationskomponente von Pose \mathbf{P}_i und $\mathbf{p}_{i, \text{GNSS}}$ eine gemessene GNSS-Position, deren Aufnahmezeitpunkt zum Aufnahmezeitpunkt von \mathbf{P}_i korrespondiert. Die Positionsmessung $\mathbf{p}_{i, \text{GNSS}}$ ist dabei in kartesischen Koordinaten (beispielsweise UTM-Koordinaten [69]) repräsentiert. Für jedes GNSS-Messungsresiduum existiert zusätzlich eine Gewichtungsmatrix $\Omega_{\mathbf{p}_i, \text{GNSS}} \in \mathbb{R}^{3 \times 3}$, die den Einfluss der Messung $r_{\mathbf{p}_i, \text{GNSS}}$ auf die Schätzung der Kartenposen bestimmt.

5.1.4 Ausreißerbehandlung

Neben einer guten Initialisierung ist die Identifikation und Entfernung von Ausreißern von fundamentaler Wichtigkeit bei der Lösung von nichtlinearen Ausgleichsproblemen. Im Folgenden werden die verwendeten Methoden beschrieben:

- **RANSAC**

RANSAC [70] ist in seiner ursprünglichen Formulierung ein iterativer Algorithmus, der in jeder Iteration zufällig eine repräsentative Teilmenge aus allen verfügbaren Messungen auswählt und mit diesen das zu lösende Ausgleichsproblem löst. Danach werden alle Residuen mit den in der Iteration geschätzten Parametern evaluiert und schwellwertbasiert die Menge von Messungen bestimmt, welche die geschätzten Parameter unterstützen (engl. *consensus set*). Dieser Vorgang wird bis zum Erreichen eines Abbruch-Kriteriums wiederholt und es werden die Parameter mit den meisten Unterstützern weiterverwendet.

- **Vorauswahl**

Bei der Vorauswahl wird eine gute Initialisierung der zu schätzenden Parame-

ter vorausgesetzt. Dabei werden vorab alle Residuen evaluiert und nur jene mit nicht zu großen Initialkosten für die Parameterschätzung verwendet.

- **Verlustfunktionen**

Im Gegensatz zum RANSAC und der Vorauswahl werden bei der Anwendung von Verlustfunktionen (engl. *loss functions*) Ausreißer nicht entfernt, sondern ihr Einfluss auf den Gradientenabstieg abgeschwächt [66]. Dazu wird jeder Summand $\|r_i(\mathbf{x})\|^2$ in (5.1) durch $\rho(\|r_i(\mathbf{x})\|^2)$ ersetzt, wobei $\rho: \mathbb{R} \rightarrow \mathbb{R}$ die Verlustfunktion darstellt. Eine prominente Verlustfunktion ist die Cauchy-Verlustfunktion $\rho(r_i(\mathbf{x})) = \log(1 + \|r_i(\mathbf{x})\|^2)$, bei welcher der Einfluss des Gradienten des Residuums $r_i(\mathbf{x})$ auf den Gradientenabstieg für steigende Kosten von $r_i(\mathbf{x})$ nahezu gänzlich verschwindet.

- **Deaktivierbare Residuen**

Deaktivierbare Residuen (engl. *switchable constraints*) schwächen, ähnlich wie Verlustfunktionen, den Einfluss von Ausreißern auf den Gradienten ab [71]. Im Unterschied zu Fehlerfunktionen geht jedoch die Abschwächung einzelner Residuen mit zusätzlichen Kosten einher, sodass Residuen nur dann abgeschwächt werden, wenn die zusätzlichen Kosten zu geringeren Gesamtkosten führen. Dazu wird das Residuum $r_i(\mathbf{x})$ mit einer Aktivierungsfunktion $\Psi: \mathbb{R} \rightarrow [0, 1]$ multipliziert, wodurch sich das deaktivierbare Residuum $\Psi(s_i) \cdot r_i(\mathbf{x})$ ergibt. Die Aktivierungsfunktion hängt dabei von einer Aktivierungsvariable $s_i \in \mathbb{R}$ ab, die dem Ausgleichsproblem als zusätzlicher Parameter hinzugefügt wird. Des Weiteren wird dem Ausgleichsproblem für jede hinzugefügte Aktivierungsvariable ein weiteres Residuum $s_i - \gamma_i$ hinzugefügt, das eine Änderung der Aktivierungsvariable s_i bestraft. Dies führt bei geeigneter Initialisierung dazu, dass Ausreißerkanten im Parametergraph dynamisch herabgewichtet werden. Gemäß [71] wird $s_i = \gamma_i = 1$ initialisiert und

$$\Psi(s_i) = \begin{cases} 0, & s_i < 0 \\ s_i, & 0 \leq s_i \leq 1 \\ 1, & s_i > 1 \end{cases} \quad (5.9)$$

als beschränkt lineare Funktion gewählt.

5.2 Initialisierungsmethoden

Nichtlineare Ausgleichsprobleme sind i. A. keine konvexen Optimierungsprobleme und haben oft mehrere lokale Minima. Dadurch besteht die Gefahr, dass iterative, numerische Optimierungsverfahren in ein falsches Minimum konvergieren und so die zu optimierenden Parameter schlecht oder falsch geschätzt werden. Maßgeblich für ein korrektes Konvergenzverhalten ist die Initialisierung der zu schätzenden Parameter. Entscheidend dabei ist, dass die initialen Parameter in der Nähe des korrekten Minimums liegen, sodass der Gradientenabstieg dieses erreichen kann. Die Methoden zur Initialschätzung der Kartenparameter werden in den folgenden Abschnitten detailliert beschrieben.

5.2.1 Odometrie

Der erste Schritt zur Integration einer neuen Aufnahmesequenz S in die bestehende Karte ist die Schätzung T der gefahrenen Trajektorie durch einen gefensterten Bündelblockausgleich. Die Initialisierungsschätzung T_{odom} für den Bündelblockausgleich wird hierbei durch Odometrie berechnet.

Das entwickelte Kartierungsframework ermöglicht die modulare Einbettung unterschiedlicher Odometrieverfahren. Es werden die folgenden Odometrieverfahren unterstützt:

- **Visuelle Odometrie**

Stehen ausschließlich Kamerabilder zur Verfügung, wird T_{odom} mittels visueller Odometrie (VO) bestimmt. Stellt die verwendete Kamerakonfiguration mindestens zwei Kameras mit hinreichend überlappenden Sichtfeldern bereit, kann stereoskopische VO verwendet werden [72]. Falls kein Kamerapaar mit hinreichendem Sichtfeldüberlapp vorhanden ist, wird die Odometrie mittels monokularer VO bestimmt [72], [73]. Die Skalenschätzung erfolgt dabei durch Messungen aus einem Laserscanner oder durch Wissen der Höhe der Kamera über der Fahrbahn.

- **Integration von Bewegungsinformation**

Sind zusätzlich zu den Kamerabildern Messungen des Lenkwinkels und der linearen Beschleunigung aus der Fahrzeugsensoren verfügbar, können diese zur Odometrieschätzung verwendet werden. Mit diesen Messungen wird

unter Verwendung eines dynamisches Einspurmodells [74] die Odometrie-trajektorie T_{odom} durch Vorwärtsintegration bestimmt. Im Gegensatz zu VO schätzt das verwendete Fahrzeugmodell nicht die Roll-, Nick- und Vertikalbewegung des Fahrzeugs. Daher werden diese Parameter mit 0° initialisiert. Da ein Fahrzeug jedoch nur sehr kleine Rollwinkeländerungen, Nickwinkeländerungen und vertikale Positionsänderungen zwischen konsekutiven Fahrzeugposen während einer durchschnittlichen Fahrt erfährt, ist die Initialisierung des Bündelblockausgleichs durch das Fahrzeugmodell ausreichend genau.

5.2.2 Lokalisierung

Zur Erstellung von Lokalisierungskarten aus mehreren Aufnahmesequenzen müssen diese geometrisch zueinander referenziert werden. Dazu wird eine neue Aufnahmesequenz S initial durch Lokalisierung in die bestehende Karte registriert. Dafür werden die aufgezeichneten Bilder in der gleichen Art und Weise wie bei der Echtzeitanwendung im Fahrzeug lokalisiert. Aus der Lokalisierung ergibt sich eine Sequenz von Lokalisierungsposen T_{loc} , die zu einer Teilmenge der Aufnahmepunkte in S korrespondiert. Dies ist in Abbildung 5.2 beispielhaft dargestellt. Die rote Trajektorie stellt die Kartentrajektorie dar und die blaue die Trajektorie der zu kartierenden Aufnahmesequenz S . Die durch Kanten verbunden Posen, die als blaue Punkte in Abbildung 5.2 dargestellt sind, zeigen Posen aus der Aufnahmesequenz S , die erfolgreich lokalisiert werden konnten.

Ringschluss-Lokalisierung

Ein Spezialfall ist die Referenzierung von Ringschlüssen innerhalb einer neuen Aufnahmesequenz S . Diese werden durch eine Ringschlussdetektion mittels GNSS oder der entwickelten Ortswiedererkennung detektiert. Das Resultat sind Paare von Aufnahmezeitpunkten $\{\gamma_i, \gamma_j\} \in S$, die den gleichen Ort zu unterschiedlichen Zeitpunkten aus einer ähnlichen Beobachtungspose abbilden. Da sporadisch falsche Ringschlussdetektionen auftreten können, wird eine Lokalisierung an allen gefundenen Ringschlusspaaren initialisiert und konsekutiv fortgeführt. Dadurch werden zum einen Ringschlusskandidaten validiert und zum anderen weitere Ringschlusspaare durch eine fortlaufende metrische Lokalisierung bestimmt. Für die neue Aufnahmesequenz ist jedoch noch keine Lokalisierungskarte verfügbar. Daher wird eine tem-

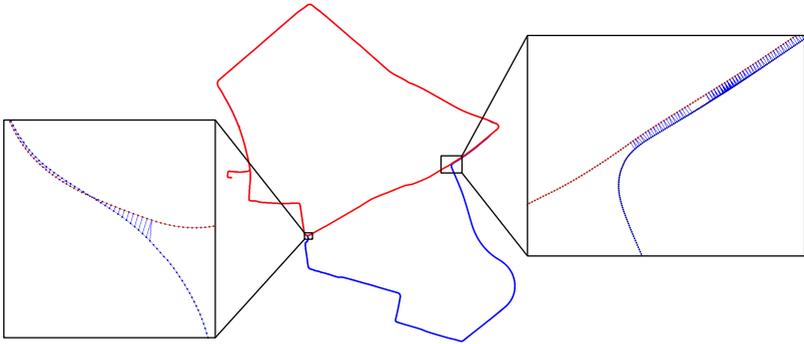


Abbildung 5.2: Trajektorie (Blau) einer neuen Aufnahmesequenz S und einer bereits kartierten Trajektorie (Rot), die sich teilweise überlappen. Die dargestellten Kanten in den hervorgehobenen Bereichen zeigen beispielhaft die Zuordnung von Lokalisierungsposen T_{loc} (Blaue Punkte) zu ihren nächsten Kartenposen (Rote Punkte). Für Posen der Aufnahmesequenz S , die mit keiner Kante verbunden sind, konnte keine Lokalisierung bestimmt werden.

poräre lokale Lokalisierungskarte in der Umgebung von \mathcal{Y}_i berechnet und an Stelle \mathcal{Y}_i die Lokalisierung initialisiert. Zur Erstellung der temporären Lokalisierungskarten wird eine initiale Trajektorien schätzung T_{odom} vorausgesetzt. Abbildung 5.3 zeigt dieses Vorgehen schematisch.

5.2.3 Aufspannende Bäume

Der Ausgangspunkt für das Aufstellen und Lösen eines Posendifferenzenausgleichsproblems (5.7) im Rahmen der Kartengenerierung ist eine gegebene Menge von Posendifferenzen D . Der korrespondierende Parametergraph verknüpft jeweils zwei Posen \mathbf{P}_i und \mathbf{P}_j aus der Menge der zu schätzenden Posen P durch eine Posendifferenz $\mathbf{D}_{i \rightarrow j} \in D$. Die Menge D umfasst i. A. konsequente Posendifferenzen, Ringschlussdifferenzen als auch Differenzen zwischen Posen verschiedener Aufnahmesequenzen, wodurch sich i. A. eine nicht zyklentreue Graphenstruktur ergibt. Zur Initialisierung von P wird der Graph daher ausgehend von einer Startpose \mathbf{P}_0 mittels Breitensuche [75] traversiert, wie in Abbildung 5.4 dargestellt ist. Alle weiteren Posen werden dann ausgehend von \mathbf{P}_0 gemäß der Traversierungsreihenfolge durch

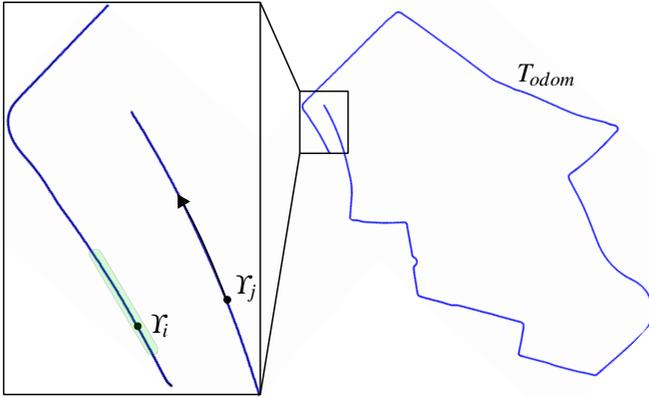


Abbildung 5.3: Darstellung der initialen Trajektorienschätzung T_{odom} mit Ringschluss, der aufgrund von Odometriedrift nicht korrekt geschlossen wurde. Im hervorgehobenen Bereich ist beispielhaft ein detektiertes Ringschlusspaar $\{Y_i, Y_j\}$ dargestellt. Der grüne Bereich zeigt den Trajektorienabschnitt, für den eine temporäre Lokalisierungskarte berechnet wird. Die Lokalisierung startet initial mit den Bildern Σ_j von Aufnahmeort Y_j und bewegt sich entlang der gefahrenen Trajektorie weiter. Wird das Ende der lokalen Karte erreicht, wird der Kartenbereich der temporären Lokalisierungskarte entlang der Trajektorie T_{odom} verschoben.

Integration der entsprechenden Posendifferenzen initialisiert.

Die Initialisierung des Posendifferenzenausgleichs durch Breitensuche ist besser geeignet als eine Initialisierung durch Tiefensuche, da sich bei der Breitensuche einzelne Integrationspfade nur bis zur nächsten Verzweigung ausbreiten. Der Odometriedrift einzelner Pfade prägt sich daher auch nur bis zur nächsten Verzweigung aus. Abbildung 5.5 zeigt vergleichsweise initialisierte Posen der dreifachen Umrundung eines Rundkurses durch Breiten- und Tiefensuche. Die Initialisierung durch Tiefensuche zeigt deutlich den Odometriedrift entlang einzelner Integrationspfade. Das initiale Gesamtresiduum gemäß (5.7) ist im Falle der Initialisierung durch Breitensuche in diesem Beispiel wesentlich geringer.

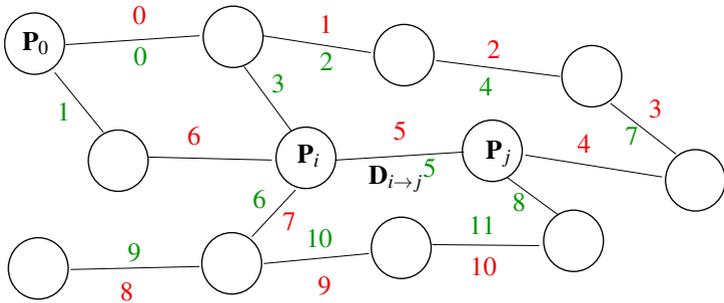


Abbildung 5.4: Parametergraph eines exemplarischen Posenausgleichsproblems. In Rot dargestellt ist eine mögliche Traversierungsreihenfolge durch Tiefensuche und in Grün eine mögliche Variante der Traversierungsreihenfolge der Breitensuche. Es ist ersichtlich, dass einzelne Traversierungspfade bei der Breitensuche deutlich kürzer sind.

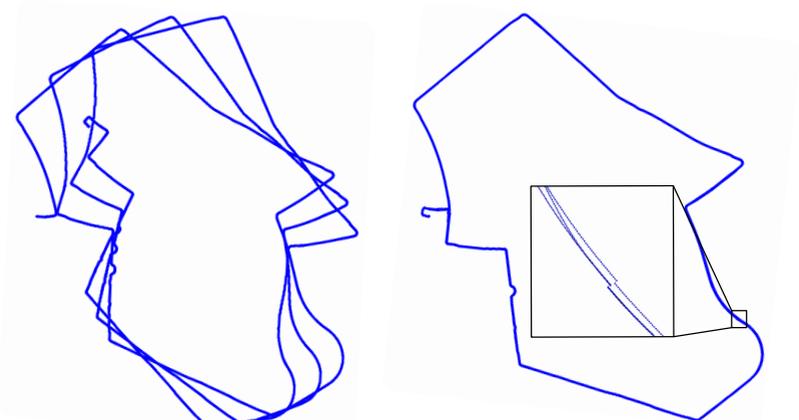


Abbildung 5.5: Initialisierung durch Tiefensuche (Links) und durch Breitensuche (Rechts). Während sich bei der Traversierung durch Tiefensuche der Odometriedrift integrieren kann, bleiben einzelne Pfade bei der Integration entlang der Traversierungsreihenfolge durch Breitensuche näher beieinander. In der Detailsicht sieht man den verbleibenden Sprung, an dem die Enden der Traversierung wieder zusammenkommen. Das initiale Residuum von Problem (5.7) ist bei der Initialisierung durch Breitensuche bedeutend kleiner.

Existiert eine nichtleere Menge $P_{\text{fix}} \in P$ von Posen, die während der Optimierung des Posendifferenzenausgleichs konstant gehalten werden, gilt $\mathbf{P}_0 \in P_{\text{fix}}$. Auf diese Weise startet die Integration aus der Menge der fixierten Posen heraus, sodass sich die nicht fixierten Posen an P_{fix} anpassen. Erreicht die Traversierung eine weitere Pose $\mathbf{P}_1 \in P_{\text{fix}}$, bleibt diese unverändert und die Integration setzt sich von \mathbf{P}_1 aus fort. Das Gleiche gilt für die Verwendung von GNSS-Messungen. Hier startet die Integration von einer Pose, der einer GNSS-Messung zugeordnet wurde, sodass bereits durch die Initialisierung durch Breitensuche eine Georeferenzierung erzielt wird.

5.2.4 Landmarkentriangulation

Die Initialisierung der Landmarken für den Bündelblockausgleich bzw. Clusterausgleich wird durch Triangulation berechnet. Eine Landmarke $\mathbf{l} \in \mathbb{R}^3$ korrespondiert stets zur einer Merkmalskorrespondenz $\Theta = \{m_1, \dots, m_N\}$, $N > 1$ und es wird vorausgesetzt, dass bereits eine Schätzung der Kameraposen $C_\Theta = \{\mathbf{C}_1, \dots, \mathbf{C}_N\}$ bekannt ist. Dabei ist $\mathbf{C}_i \in C_\Theta$ die Kamerapose, die zu jenem Bild korrespondiert, in dem $m_i \in \Theta$ detektiert wurde. Mittels des Rotationsanteils $\mathbf{R}_i \in SO(3)$ von \mathbf{C}_i und der Projektionsfunktion (3.6) kann der zu Merkmal m_i korrespondierende Sichtstrahl

$$\mathbf{r}_i = R_i \cdot \kappa_i(\mathbf{z}_i) \quad (5.10)$$

im globalen Koordinatensystem dargestellt werden, wobei \mathbf{z}_i die zu m_i korrespondierende Bildkoordinate ist. Des Weiteren stellt der Positionsanteil $\mathbf{p}_i \in \mathbb{R}^3$ von \mathbf{C}_i den Aufpunkt des Sichtstrahls im globalen Koordinatensystem dar.

Zur Triangulation wird angenommen, dass die Beobachtungspositionen $\{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ unterschiedlich und die korrespondierenden Sichtstrahlen $\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ nicht kollinear sind. Im Idealfall schneiden sich alle Sichtstrahlen exakt in \mathbf{l} . Aufgrund von Schätzungenauigkeiten in der Kalibrierung und in den Kameraposen C_Θ und aufgrund von Detektorrauschen ist dies jedoch in der Praxis i. A. nicht erfüllt. Daher wird zur Triangulation einer Landmarke der senkrechte Abstand der Landmarke \mathbf{l} zu allen Sichtstrahlen minimiert, wie in Abbildung 5.6 dargestellt ist. Der senkrechte Abstandsvektor des i -ten Sichtstrahls zur Landmarke \mathbf{l} ergibt sich durch

$$\mathbf{d}_i = (\mathbf{l} - \mathbf{p}_i) - \mathbf{r}_i^T (\mathbf{l} - \mathbf{p}_i) \mathbf{r}_i. \quad (5.11)$$

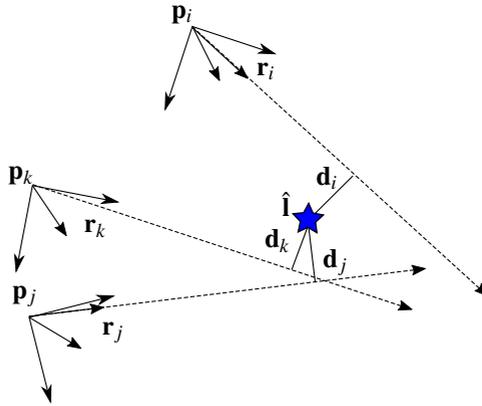


Abbildung 5.6: Darstellung des Triangulationsfehlers (5.11). Die Sichtstrahlen sind i. A. windschiefe dreidimensionale Geraden.

Somit lässt sich \mathbf{I} durch

$$\arg \min_{\mathbf{I}} \sum_{i=1}^N \mathbf{d}_i^T \mathbf{d}_i \quad (5.12)$$

schätzen. Da Problem (5.12) quadratisch in \mathbf{I} ist, kann durch die Bedingung

$$\frac{\partial (\sum_{i=1}^N \mathbf{d}_i^T \mathbf{d}_i)}{\partial \mathbf{I}} = \mathbf{0} \quad (5.13)$$

die geschlossene Lösung

$$\hat{\mathbf{I}} = \left(\sum_{i=1}^N \mathbf{I} - \mathbf{r}_i \mathbf{r}_i^T \right)^{-1} \left(\sum_{i=1}^N (\mathbf{I} - \mathbf{r}_i \mathbf{r}_i^T) \mathbf{p}_i \right) \quad (5.14)$$

bestimmt werden, wobei $\mathbf{I} \in \mathbb{R}^{3 \times 3}$ die dreidimensionale Einheitsmatrix bezeichnet.

5.3 Trajektorienvalidierung

Eine unverzerrte und skalentreue Trajektorien-schätzung ist für die Kartengenerierung und Lokalisierung von fundamentaler Bedeutung. Eine Anpassung weiterer Trajektorien an bereits verzerrte Kartentrajektorien kann zu einer falschen und verzerrten Gesamtschätzung der Kartenposen führen. Zur Detektion und Behandlung

solcher Verzerrungen werden die geschätzten Trajektorien mit geometrischen Bewegungsmodellen validiert. Die Validierung basiert auf der Annahme, dass die Kameras auf einem Fahrzeug montiert sind und dessen Bewegung nicht-holonomen Zwangsbedingungen unterliegt. Die folgenden Abschnitte 5.3.1 und 5.3.2 beschreiben die geometrischen Modelle zur Validierung dieser Zwangsbedingungen. Da im Laufe der Kartierung unterschiedliche Schätzungen der gleichen Trajektorie mit unterschiedlichen Nebenbedingungen berechnet werden, können die verschiedenen Trajektorienschätzungen wechselseitig zueinander verglichen und so ebenfalls Verzerrungen identifiziert werden. Dies ist in Abschnitt 5.3.3 beschrieben. Da bis auf die aufgezeichneten Kamerabilder keine weiteren Sensorinformationen vorausgesetzt werden, basieren alle Modelle auf geometrischen Betrachtungen.

5.3.1 Momentanpolbewegung

Bei der Momentanpolbewegung wird angenommen, dass sich das Fahrzeug auf einem Kreis mit konstantem Mittelpunkt und Radius bewegt. Sei Δy die laterale, Δx die longitudinale Positionsänderung und $\Delta \gamma$ die Rotationsänderung um die Fahrzeughochachse zwischen den geschätzten Fahrzeugposen \mathbf{P}_i und \mathbf{P}_{i+n} , so kann bei der Bewegung des Fahrzeugs auf einer Kreisbahn

$$\Delta y = f_{\circ}(\Delta x, \Delta \gamma) = \Delta x \frac{1 - \cos(\Delta \gamma)}{\sin(\Delta \gamma)}. \quad (5.15)$$

als Funktion von Δx und $\Delta \gamma$ bestimmt werden. Abbildung 5.7 (Links) stellt diesen Zusammenhang geometrisch dar. Da die Momentanpolapproximation nur innerhalb eines kurzen Zeitraums gültig ist, muss der Zeitunterschied zwischen dem Aufnahmezeitpunkt t_i der Pose \mathbf{P}_i und dem Aufnahmezeitpunkt t_{i+n} der Pose \mathbf{P}_{i+n} hinreichend klein sein. Durch einen Schwellwertvergleich $|\Delta y - f_{\circ}(\Delta x, \Delta \gamma)| < T_{\circ}$ kann bestimmt werden, ob die Relativbewegung zwischen \mathbf{P}_i und \mathbf{P}_j dieses Modell verletzt und daher eine fehlerhafte Bewegungsschätzung vorliegt. Hierbei ist angenommen, dass sich das Fahrzeug nicht im hochdynamischen Bereich bewegt.

Abbildung 5.7 (Rechts) zeigt den Graphen der Funktion (5.15). Da das Modell (5.15) nur für kurze Zeiträume und daher nur für kleine Winkel $\Delta \gamma$ gültig ist, kann dies durch

$$\Delta y = f_{\circ}(\Delta x, \Delta \gamma) \approx \frac{1}{2} \Delta x. \quad (5.16)$$

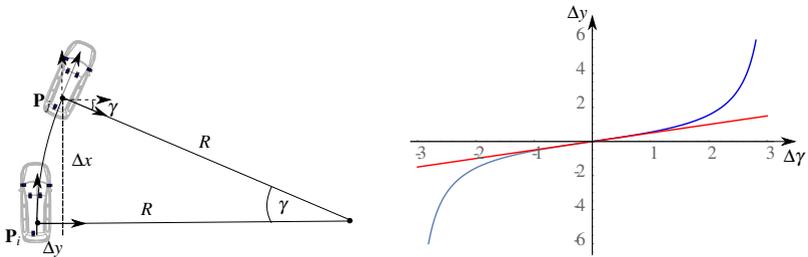


Abbildung 5.7: Links: Geometrie bei einer Bewegung auf einer Kreisbahn mit Radius R .
 Rechts: Graph der Funktion (5.15) (Blau) und der Approximation (5.16) (Rot)
 für $\Delta x = 1$. Für kleine γ kann das Modell (5.15) mit $\frac{1}{2}\Delta x$ approximiert werden.

approximiert werden und ist daher günstig auszuwerten. Ein weiterer Vorteil der Approximation (5.16) ist, dass sie unabhängig von $\Delta \gamma$ ist und daher im Gegensatz zu (5.15) für $\Delta \gamma = 0$ numerisch stabil ist.

5.3.2 Polynomielle Approximation

Eine Möglichkeit, die lokale Glattheit einer Trajektorienabschätzung T zu analysieren, ist der Vergleich eines lokalen Trajektorienabschnitts mit der Approximation dieses Trajektorienabschnitts durch ein Polynom

$$a^m(x) = \sum_{j=0}^m b_j x^j \quad (5.17)$$

m -ten Grades. Mit den in Abschnitt 3.1 getroffenen Annahmen an die maximale Fahrgeschwindigkeit des Fahrzeugs und die Aufnahmefrequenz kann eine glatte und kontinuierliche Trajektorie T angenommen werden. Durch den Vergleich der Trajektorienposen mit einer polynomiellen Approximation können Trajektorien sprünge, die die nicht-holonomen Zwangsbedingungen verletzen, identifiziert werden.

Die Posen $\mathbf{P}_0, \dots, \mathbf{P}_n \in T$ stellen in minimalparameterischer Form $\xi(\mathbf{P}_0), \dots, \xi(\mathbf{P}_n)$ gemäß Abschnitt 5.1.3 eine diskrete Approximation

$$\mathbf{p}(t_i) = \xi(\mathbf{P}_i), 0 \leq i \leq n \quad (5.18)$$

von komponentenweisen Zeitsignalen $\mathbf{p}(t) = [p_1(t), \dots, p_6(t)]^T$ dar. Jedes der sechs Signale bildet dabei eine Rotations- bzw. Translationskomponente ab, wobei t_0, \dots, t_n die Aufnahmezeitpunkte der Bilder zu den Posen $\mathbf{P}_0, \dots, \mathbf{P}_n$ bezeichnen. Zur Überprüfung werden die Signale durch ein Polynom $a_i^m(t)$ innerhalb eines gleitenden Zeitfensters $[t_s, t_e]$ approximiert. Dafür werden die Polynomparameter b_1, \dots, b_m von $a_i^m(t)$ für jedes der sechs Signale durch Lösen eines linearen Ausgleichsproblems geschätzt. Die Residuen sind hierbei die Abweichung $p_i(t) - a_i^m(t)$ des Polynoms $a_i^m(t)$ von den Signalen $p_i(t)$ an den diskreten Abtastzeitpunkten innerhalb des gleitenden Zeitfensters $[t_s, t_e]$. Ist der Grad m des Polynoms dabei kleiner als die Anzahl der in dem Fenster liegenden Abtastzeitpunkte, werden lokale Sprünge bzw. hochfrequente Schwingungen gedämpft.

Auf die gleiche Weise kann das diskrete Posendifferenzsignal

$$\mathbf{d}\left(\frac{t_{i-1} + t_i}{2}\right) = \boldsymbol{\xi}(\mathbf{D}_{i-1,i}), 0 < i \leq n \quad (5.19)$$

bzw. das korrespondierende Geschwindigkeitssignal

$$\mathbf{v}\left(\frac{t_{i-1} + t_i}{2}\right) = \frac{\mathbf{d}\left(\frac{t_{i-1} + t_i}{2}\right)}{\frac{t_k - t_{k-1}}{2}}, 1 < k \leq n \quad (5.20)$$

der Trajektorien schätzung T komponentenweise durch Polynome approximiert und zeitkontinuierlich dargestellt werden. Treten bei der Anpassung des Polynoms an die Messungen ungewöhnlich hohe Residuen auf, deutet dies auf eine Störung hin. Störungen können daher einfach schwellwertbasiert identifiziert werden.

Abbildung 5.8 zeigt beispielhaft das Posendifferenzsignal $d_4(t)$ (Blau) mit einer Störung, die durch den Vergleich zu dem geschätzten Polynom $a_4^1(t)$ (Rot) identifiziert wurde. Die Fensterlänge wurde dabei zu 0.7 Sekunden gewählt, sodass das gleitende Approximationsfenster ~ 7 Posen umfasst. Das Posendifferenzsignal $d_4(t)$ stellt in dem Beispiel die laterale Fahrzeugbewegung dar.

5.3.3 Trajektorienvergleich

Da im Rahmen der Kartengenerierung verschiedene Schätzungen der gleichen Trajektorie mit unterschiedlichen Randbedingungen erfolgen, können diese miteinander

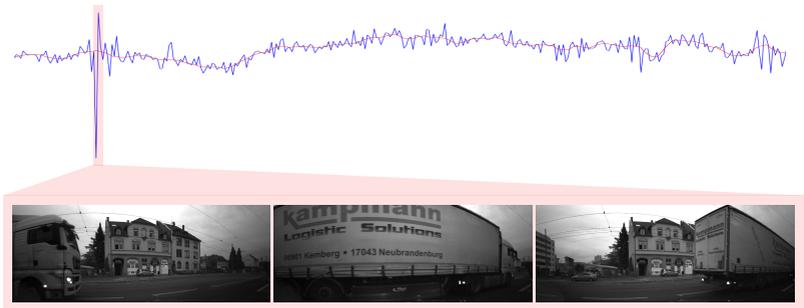


Abbildung 5.8: Oben: Signal $v_4(t)$ (Blau) der lateralen Fahrzeuggeschwindigkeit aus einer Trajektorien-schätzung mittels SVO und der korrespondierenden polynomiellen Approximation $p_4(t)$ (Rot) bei einer Fensterlänge von sieben Aufnahme-punkten und einer Aufnahme-frequenz von $\sim 10\text{Hz}$. Unten: Korrespondierende Bilder zum Zeitpunkt der Störung in der SVO-Trajektorien-schätzung.

verglichen und dadurch Trajektorienverzerrungen identifiziert werden. Dazu wird der Trajektorien-drift [76] zwischen verschiedenen Schätzungen T_a und T_b der gleichen Trajektorie T bestimmt. Sei $\mathbf{D}_i^a = \mathbf{P}_i^{-1} \cdot \mathbf{P}_{i+n}$ die Posendifferenz zwischen Fahrzeugpose $\mathbf{P}_i \in T_a$ und Fahrzeugpose $\mathbf{P}_{i+n} \in T_a$ und \mathbf{D}_i^b die korrespondierende Differenzpose aus Trajektorien-schätzung T_b , so ist $\mathbf{D}_i^{\text{diff}} = (\mathbf{D}_i^a)^{-1} \cdot \mathbf{D}_i^b$ ein Maß für den Drift des betrachteten Trajektorienabschnitts. Durch Iteration des Laufindex i über die gesamte Trajektorie ergibt sich der Drift für jeden Trajektorienabschnitt. Der Index n definiert die zwischen den Posen \mathbf{P}_i und \mathbf{P}_{i+n} zurückgelegte Fahrstrecke. Für kleinere Werte von n wird der lokale Trajektorien-drift bewertet und für größere Werte der globale Drift.

6 Kartengenerierung

In diesem Kapitel wird die iterative Erzeugung einer Kartendatenbank aus mehreren Aufnahmesequenzen beschrieben. Die Kartendatenbank M_n ist die Grundlage zur Erzeugung der Lokalisierungskarte M_n^L , die zur Lokalisierung im Fahrzeug und im Rahmen der Kartengenerierung benötigt wird.

Der grundsätzliche Ablauf der iterativen Kartengenerierung ist in Abbildung 6.1 dargestellt. Es wird angenommen, dass bereits eine Kartendatenbank $M_n = \{P_n, L_n, \Theta_n\}$ besteht, die aus n zuvor aufgezeichneten Aufnahmesequenzen S_1, \dots, S_n erzeugt wurde und die Posen P_n , die Landmarken L_n und die Merkmalskorrespondenzen Θ_n umfasst. Die Kartenposen P_n setzen sich somit aus den Trajektorienposen der zuvor kartierten Sequenzen S_1, \dots, S_n zusammen. Durch die Kartengenerierung liegen die Kartenposen P_n und Landmarken L_n in einer gemeinsamen, konsistenten Repräsentation vor. Das Ziel der in diesem Kapitel beschriebenen iterativen Kartengenerierung ist die Integration einer weiteren Aufnahmesequenz S_{n+1} in die bestehende Kartendatenbank M_n . Durch die Integration der neuen Aufnahmesequenz S_{n+1} werden die Trajektorienposen von S_{n+1} konsistent und präzise an die Kartenposen P_n angepasst und die bestehenden Merkmalskorrespondenzen Θ_n und Landmarken L_n durch Assoziation zu Merkmalen aus S_{n+1} aktualisiert und erweitert. Wurde bei der Aufzeichnungsfahrt von S_{n+1} sowohl kartiertes als auch bisher unkartiertes Gebiet befahren, wird die Karte automatisch um bisher nicht kartierte Gebiete erweitert, da stets die Trajektorienposen der gesamten Aufnahmesequenz S_{n+1} in die Karte übernommen werden. Aus der aktualisierten Kartendatenbank M_{n+1} ergibt sich darauffolgend die verbesserte und aktualisierte Lokalisierungskarte M_{n+1}^L .

In Abschnitt 6.1 wird die Prozesskette zur Aktualisierung und Erzeugung der Kartendatenbank geschildert. Danach werden in Abschnitt 6.2 wichtige Details der Optimierungs- und Assoziationsstrategie im Rahmen der Kartengenerierung ana-

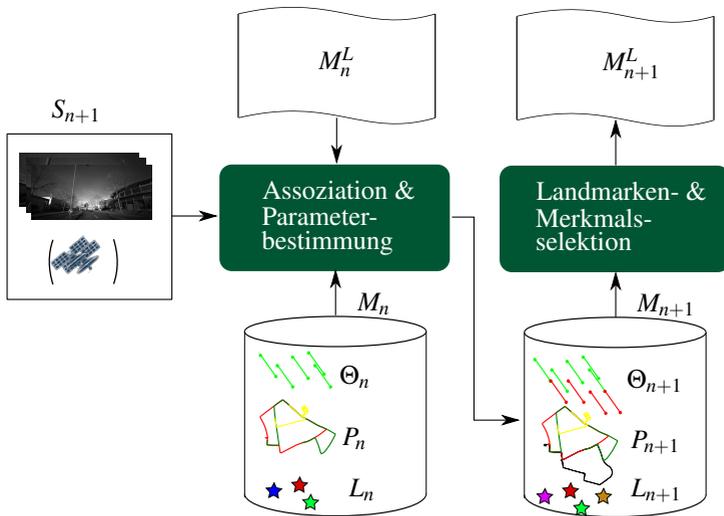


Abbildung 6.1: Ablauf der iterativen Kartengenerierung.

lysiert. Abschließend werden in Abschnitt 6.3 die Ergebnisse der Erstellung einer Kartendatenbank aus 14 Aufnahmesequenzen präsentiert.

6.1 Ablauf

Um die Posen P_n und Landmarken L_n präzise und konsistent schätzen zu können, wird ein hoher Verknüpfungsgrad der Kartenkorrespondenzen Θ_n angestrebt, sodass möglichst viele Merkmale zu jeweils einer Landmarke korrespondieren. Außerdem sollten die Merkmalskorrespondenzen möglichst frei von Ausreißern sein. Wie in Abschnitt 4 gezeigt, kann durch eine genauere Schätzung der Kartenposen P_n wiederum der Verknüpfungsgrad von Θ_n gesteigert und die Wahrscheinlichkeit für Ausreißer verringert werden, da die geometrischen Annahmen zur Merkmalsassoziation besser erfüllt sind. Der in diesem Abschnitt vorgestellte Prozessablauf stellt eine mehrschrittige, aufeinander aufbauende Ausführung von Parameterschätzungs- und Merkmalsassoziationsschritten dar. Die dabei erzielten Zwischenergebnisse der zu schätzenden Fahrzeugposen werden zur geometrischen Merkmalsassoziation ge-

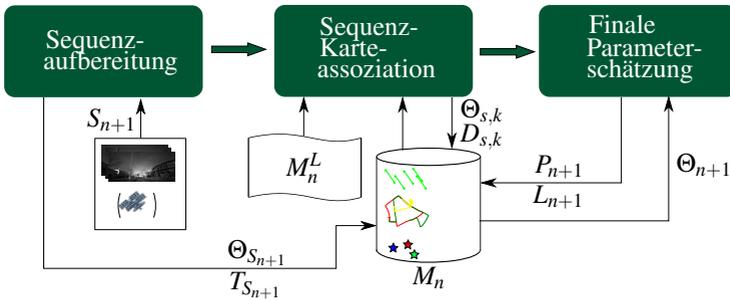


Abbildung 6.2: Ablauf der Aktualisierung der Kartendatenbank M_n durch die Integration von Aufnahmesequenz S_{n+1} .

mäß Abschnitt 4.1.3 verwendet, wodurch der Verknüpfungsgrad der Merkmalskorrespondenzen gesteigert wird. Im nachfolgenden Schritt steigert dies wiederum die Genauigkeit der Parameterschätzung.

Die Generierung der Karte erfolgt vollautomatisch und skaliert für beliebig große Gebiete und beliebig viele Aufnahmesequenzen. Abbildung 6.2 zeigt die drei wesentlichen Schritte zur Integration einer neuen Aufnahmesequenz S_{n+1} in die bestehende Kartendatenbank M_n .

Im ersten Schritt, der in Abschnitt 6.1.1 beschrieben ist, wird die neue Aufnahmesequenz S_{n+1} unabhängig von der bestehenden Datenbank M_n vorverarbeitet, woraus eine Schätzung der gefahrenen Trajektorie $T_{S_{n+1}}$ und Merkmalskorrespondenzen $\Theta_{S_{n+1}}$ resultieren, die ausschließlich Merkmale zwischen aufgezeichneten Bildern aus S_{n+1} verknüpfen.

Im zweiten Schritt wird die Aufnahmesequenz S_{n+1} zur bestehenden Kartendatenbank M_n referenziert. Das Ergebnis dieses Schrittes sind Merkmalskorrespondenzen $\Theta_{s,k}$, welche die Merkmalskorrespondenzen $\Theta_{S_{n+1}}$ zu den bestehenden Kartenmerkmalskorrespondenzen Θ_n assoziieren. Außerdem werden Posendifferenzen $D_{s,k}$ berechnet, welche die Trajektorienposen aus $T_{S_{n+1}}$ präzise zu den Kartenposen P_n referenzieren. Abschnitt 6.1.2 erläutert diesen Prozessschritt.

Im letzten Schritt, der in Abschnitt 6.1.3 beschrieben ist, werden die finalen Kar-

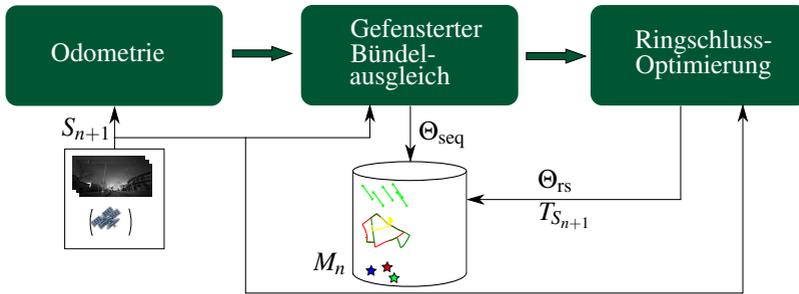


Abbildung 6.3: Prozessschritte der Sequenzaufbereitung.

tenposen P_{n+1} durch Lösen eines Posendifferenzenausgleichs unter Verwendung der zuvor bestimmten Posendifferenzen berechnet. Auf Basis der aktualisierten Kartenposen P_{n+1} werden die finalen Kartenlandmarken L_{n+1} trianguliert.

6.1.1 Sequenzaufbereitung

Abbildung 6.3 zeigt die Prozessschritte der kartenunabhängigen Vorverarbeitung der neuen Aufnahmesequenz S_{n+1} . Durch Odometrie und einen darauffolgenden gleitenden Bündelblockausgleich gemäß Abschnitt 5.1.1 wird eine initiale Schätzung $T_{S_{n+1}}$ der gefahrenen Trajektorie berechnet. Dafür werden Merkmalskorrespondenzen Θ_{seq} zwischen konsekutiven Bildern jeder Kamera des MKS bestimmt. Umfasst das MKS mehrere Kameras mit hinreichender Sichtfeldüberlappung, werden zusätzlich Merkmale zwischen Bildern verschiedener Kameras zum gleichen Zeitpunkt assoziiert. Außerdem wird eine Ringschlusslokalisierung gemäß Abschnitt 5.2.2 durchgeführt und darauf basierend Posendifferenzen zwischen benachbarten Ringschlussposen abgeleitet.

Mit den Ringschlussposendifferenzen und den sequentiellen Posendifferenzen, die sich aus der initialen Trajektorien schätzung $T_{S_{n+1}}$ ergeben, wird ein Posendifferenzenausgleich gelöst. Daraus resultiert eine verbesserte und in sich konsistente Trajektorien schätzung $T_{S_{n+1}}$, die zur robusten Bestimmung der Merkmalskorrespondenzen Θ_{rs} zwischen benachbarten Ringschlussbildern gemäß Abschnitt 4.1.3 verwendet wird. Die globale Verknüpfung aller in diesem Schritt berechneten Merkmals-

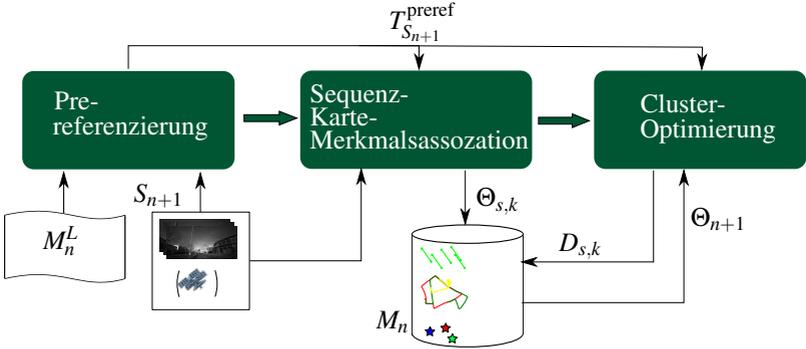


Abbildung 6.4: Ablauf der Sequenz-Karte-Assoziation.

korrespondenzen gemäß Abschnitt 4.3 ergibt die Sequenzmerkmalskorrespondenzen $\Theta_{S_{n+1}}$.

6.1.2 Sequenz-Karte-Assoziation

In diesem Prozessschritt werden die Posen der Sequenztrajektorie $T_{S_{n+1}}$ hoch genau zu den Kartenposen P_n referenziert. Dazu werden die Merkmalskorrespondenzen $\Theta_{s,k}$ berechnet, die einen Teil der Sequenzkorrespondenzen $\Theta_{S_{n+1}}$ zu den Kartenkorrespondenzen Θ_n assoziiert. Abbildung 6.4 zeigt die dazu durchgeführten Prozessschritte, die im Folgenden beschrieben werden:

1. Prereferenzierung

Im ersten Schritt wird die Aufnahmesequenz S_{n+1} in die bestehende Lokalisierungskarte M_n^L gemäß Abschnitt 5.2.2 lokalisiert, wodurch sich die Lokalisierungstrajektorie $T_{S_{n+1}}^{loc}$ ergibt. Die Trajektorie $T_{S_{n+1}}^{loc}$ umfasst somit kartenrelative Lokalisierungsposen zu allen Aufnahmezeitpunkten aus S_{n+1} , die erfolgreich mit der Lokalisierungskarte M_n^L lokalisiert werden konnten. In der Regel sind das Abschnitte der Aufnahmesequenz S_{n+1} , in denen bereits kartiertes Gebiet befahren wurde. In diesen Bereichen werden Posendifferenzen D_{nn}^{loc} zwischen benachbarten Fahrzeugposen aus den Kartenposen P_n und der Lokalisierungstrajektorie $T_{S_{n+1}}^{loc}$ berechnet. Diese werden daraufhin in einem Posendifferenzenausgleich verwendet, dessen Lösung die Trajektorie $T_{S_{n+1}}^{pref}$

ergibt, die initial alle Posen aus S_{n+1} konsistent zu den Kartenposen P_n referenziert.

2. Sequenz-Karte-Merkmalassoziation

Auf Basis der initial referenzierten Trajektorienposen $T_{S_{n+1}}^{\text{pref}}$ werden die Merkmalskorrespondenzen $\Theta_{s,k}$ durch die geometrische Assoziation gemäß Abschnitt 4.1.3 berechnet. Die Merkmalskorrespondenzen $\Theta_{s,k}$ assoziieren Merkmale aus den Kartenmerkmalskorrespondenzen Θ_n zu den Merkmalen aus den zuvor berechneten Sequenzmerkmalskorrespondenzen $\Theta_{S_{n+1}}$, wodurch der Verknüpfungsgrad im Vergleich zur Merkmalsassoziation ausschließlich auf Basis der Lokalisierung gesteigert wird.

3. Cluster-Optimierung

Zur Verbesserung der Konsistenz der Referenzierung $T_{S_{n+1}}^{\text{pref}}$ durch die Lokalisierung werden Trajektoriencluster aus der Sequenztrajektorie $T_{S_{n+1}}$ mit benachbarten Trajektorienclustern aus den Kartenposen P_n gepaart. Durch Lösen lokaler Clusterausgleichsprobleme gemäß Abschnitt 5.1.2 werden die gepaarten Trajektoriencluster präzise aneinander angepasst. Die aktualisierten Kartenkorrespondenzen Θ_{n+1} , die sich aus der Verknüpfung der bestehenden Kartenkorrespondenzen Θ_n mit den Sequenzmerkmalskorrespondenzen $\Theta_{S_{n+1}}$ und den Merkmalskorrespondenzen $\Theta_{s,k}$ ergeben, werden dabei als Messungen verwendet. Das Ergebnis der Clusteroptimierung sind aus den optimierten Clustern abgeleitete Posendifferenzen $D_{s,k}$.

6.1.3 Finale Parameterschätzung

Die letzten beiden Schritte der Kartengenerierung sind in Abbildung 6.5 dargestellt. Die finalen Kartenposen P_{n+1} werden durch Lösen eines Posendifferenzenausgleichs bestimmt. Als Messungen werden die Posendifferenzen aus der Sequenztrajektorie $T_{S_{n+1}}$ und die Posendifferenzen $D_{s,k}$ verwendet, die die Posen der Sequenztrajektorie zu den bestehenden Kartenposen P_n referenzieren. Die Kartenposen P_n werden dem Ausgleichsproblem als fixe Parameter hinzugefügt, sodass sich die Trajektorienposen aus S_{n+1} an die Kartenposen anpassen. Liegen zusätzlich GNSS-Messungen vor, können diese ebenfalls im Posendifferenzenausgleich berücksichtigt werden, wodurch die Kartenposen georeferenziert werden können. Schlussendlich werden die

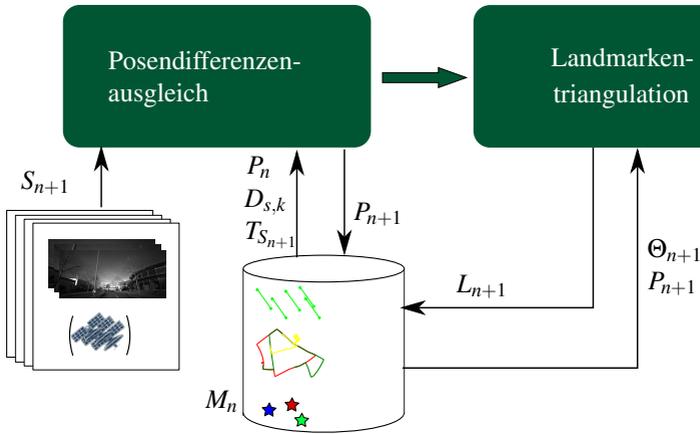


Abbildung 6.5: Ablauf der finalen Parameterschätzung.

Landmarken L_{n+1} auf Basis der Merkmalskorrespondenzen Θ_{n+1} und Kartenposen P_{n+1} gemäß Abschnitt 5.2.4 trianguliert.

6.2 Analyse der Optimierungsstrategie

Durch die fortlaufende Hinzunahme neuer Aufnahmesequenzen wächst die Kartengröße und die Anzahl der Kombinationsmöglichkeiten zur Assoziation von Merkmalen. Eine ganzheitliche gemeinsame Schätzung aller Parameter ist aus diesen Gründen praktisch nicht möglich. Wie zuvor beschrieben, zerlegt die entwickelte Methode die Karte daher in lokale Kartenfenster. Für jedes Fenster wird ein separates Bündelblock- bzw. Clusterausgleichsproblem zur hochgenauen Bestimmung der Landmarken und Posen gelöst. Nach der lokalen Optimierung werden ausschließlich Posendifferenzen aus den Posen jedes Optimierungsfensters abgeleitet und weiterverwendet. Durch das Lösen eines Posendifferenzenausgleichs unter Verwendung der abgeleiteten Posendifferenzen überträgt sich die Genauigkeit, die durch das Lösen der lokalen Ausgleichsprobleme erzielt wurde, in die finale Kartenposenschätzung. Auf diese Weise bleibt die Komplexität jedes Optimierungsproblems, unab-

hängig von der Länge der Aufnahmesequenz S_{n+1} und der Anzahl der bereits kartierten Sequenzen in M_n , begrenzt, da

1. die lokalen Optimierungsfenster immer eine konstante maximale Größe haben. Je länger eine Befahrung ist, desto mehr lokale Bündelblockausgleichsprobleme mit begrenzter Komplexität müssen verarbeitet werden. Des Weiteren wird die Anzahl der Kombinationen von zu lösenden Clusterausgleichsproblemen auf eine maximale Anzahl begrenzt. Dies wird in Abschnitt 6.2.2 im Detail beschrieben.
2. die bestehenden Kartenposen P_n im finalen Posendifferenzenausgleich fixiert werden, sodass ausschließlich die Posen $T_{S_{n+1}}$ der neuen Befahrung optimiert werden. Umfasst $T_{S_{n+1}}$ mehr als eine maximale Posenanzahl, wird die Trajektorie in Abschnitte unterteilt und abschnittsweise an die Kartenposen angepasst, sodass die Größe des finalen Posendifferenzenausgleichsproblems ebenfalls nach oben begrenzt ist.

Da aus jedem lokalen Optimierungsfenster nur Posendifferenzen gespeichert und weiterverwendet werden, wird eine unabhängige parallele Berechnung der Fenster durchgeführt. Die Berechnung der aufwendigsten Schritte des Kartierungsprozesses wird durch diese Parallelisierung effizient an die verfügbaren Hardwareressourcen angepasst.

6.2.1 Gleitender Bündelblockausgleich

Der gleitende Bündelblockausgleich im Rahmen der Sequenzauflistung gemäß Abschnitt 6.1.1 dient zur Verbesserung der vorab durch Odometrie geschätzten Sequenztrajektorie $T_{S_{n+1}}$. Für den gleitenden Bündelblockausgleich werden zunächst Schlüsselposen (engl. *key-frames*) durch Unterabtastung der Trajektorienposen bestimmt. Dabei werden Posen, die zu nah beieinander liegen, entfernt, wodurch sich ein minimaler Positionsunterschied zwischen zwei aufeinanderfolgenden Posen ergibt. Zusätzlich darf die Rotation um die Fahrzeughochachse nicht zu groß sein, da sich sonst die Szene in den Bildern für eine robuste Merkmalsassoziation zu sehr verändert. Im Rahmen der Kartengenerierung wird durch die Schlüsselposenauswahl ein minimaler Positionsunterschied von 0.2m und ein maximaler Winkel von 2° zwischen zwei aufeinanderfolgenden Posen erzielt. Diese Werte haben in Expe-

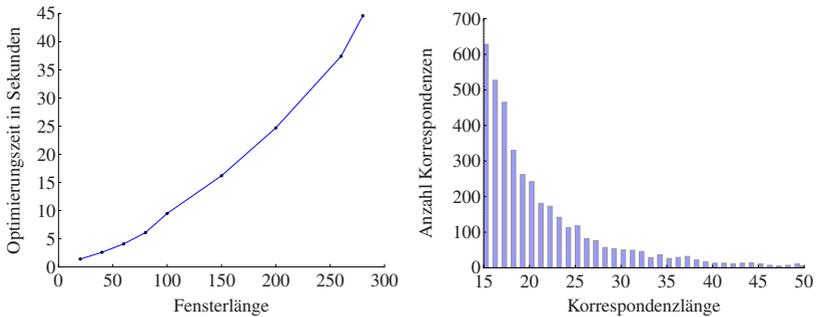


Abbildung 6.6: Links: Typische Optimierungszeiten des Bündelblockausgleichsproblems für verschiedene Fensterlängen (in Fahrzeugposen). Rechts: Exemplarische Verteilung der Korrespondenzen über ihre Korrespondenzlänge (Anzahl verschiedener Fahrzeugposen). Beide Auswertungen stellen gemittelte Werte über mehrere Datenerhebungen mit unterschiedlichen Datensätzen dar.

perimenten und in der Anwendung zufriedenstellende Ergebnisse erzielt. Auf diese Weise werden die Posen in Haltephasen ignoriert, eine ausreichende Parallaxe für die Triangulation sichergestellt und die Menge der zu verarbeitenden Daten ohne signifikanten Informationsverlust reduziert.

Fensterung

Die Länge des gleitenden Optimierungsfensters ist definiert durch die Anzahl n_{window} konsekutiver Schlüsselposen im Optimierungsfenster. Das Optimierungsfenster wird um $n_{window} - n_{overlap}$ Fahrzeugposen weitergeschoben. Die resultierende Überlappung verhindert, dass Posen an den Enden des jeweiligen Optimierungsfensters aufgrund abgeschnittener Merkmalskorrespondenzen schlecht geschätzt werden. Eine größere Fensterlänge kann dabei zu einer höheren Genauigkeit führen, geht jedoch mit einer größeren Rechen- und Speicherkomplexität pro Fenster einher. Bei kleinerer Fensterlänge steigt aufgrund der Fensterüberlappung ebenfalls der Aufwand, da sich insgesamt mehr überlappende Bereiche ergeben. Der linke Graph in Abbildung 6.6 zeigt exemplarisch den Zusammenhang zwischen der Fensterlänge und der benötigten Rechenzeit zur Lösung des Bündelblockausgleichsproblems mit einem durchschnittlichen Desktop-Computer (Intel-i7, 4 Cores, 8GB RAM). Die

Fensterlänge ist proportional (mit großer Proportionalitätskonstante $K > 1000$) zur Anzahl involvierter Landmarkenparameter und Residuen. Wie in Abschnitt 5.1.2 beschrieben, beeinflusst die Dimension der Jacobi-Matrix die Rechenkomplexität des Bündelblockausgleichproblems wesentlich. Da sich die Dimension der Jacobi-Matrix aus dem Produkt der Anzahl der Messungen und der Anzahl der Parameter des Ausgleichsproblems ergibt, zeigt sich ein quadratischer Anstieg der Rechenzeit.

In Experimenten hat sich ergeben, dass Fensterlängen von $n_{window} > 50$ Posen zu keiner nennenswerten Verbesserung der Trajektorienschätzung führen. Das lässt sich anhand der Verteilung der Längen der Merkmalskorrespondenzen Θ_{seq} erklären. Abbildung 6.6 (Rechts) stellt die Verteilung der Korrespondenzlängen von konsekutiven Merkmalskorrespondenzen dar. Zur besseren Darstellung sind nur Korrespondenzlängen > 15 dargestellt. Der überwiegende Anteil der Korrespondenzen ist dabei deutlich kleiner als die Ausdehnung des Optimierungsfensters, sodass nur ein geringer Informationsanteil durch die Fensterung verloren geht. Durch die Optimierungsfensterüberlappung wird dieser Informationsverlust zusätzlich reduziert. Eine Überlappung von $n_{overlap} > 30$ Fahrzeugposen hat ebenfalls keine nennenswerte Verbesserung der Gesamtschätzung erbracht.

Separate Rotationsschätzung

In Übereinstimmung mit aktuellen Arbeiten aus dem Bereich der SVO [77], [78] hat sich gezeigt, dass eine zweischrittige Lösung des gleitenden Bündelblockausgleichproblems von Vorteil ist. Zunächst wird ein Bündelblockausgleich mit Korrespondenzen aus einer einzelnen Kamera durchgeführt, wobei die Translationsparameter der Fahrzeugposen aus der Odometrieschätzung fixiert werden und nur der Rotationsanteil der Fahrzeugposen und die Landmarken geschätzt werden. Im zweiten Schritt wird ein gemeinsamer Bündelblockausgleich mit den Merkmalskorrespondenzen aus allen Kameras des MKS durchgeführt, wobei die Rotationsparameter der Posen festgehalten werden. Dies führt zu einer signifikanten Verbesserung der Rotationsschätzung. In Einklang mit [77] wird angenommen, dass sich die Verbesserung dadurch erklären lässt, dass der Fehler in der extrinsischen Kamerakalibrierung einen signifikanten Einfluss auf die Rotationsschätzung hat. Durch die monoskopische Rotationsschätzung wird diese Fehlerquelle umgangen. Die Schätzung der Translation ist gemäß den durchgeführten Experimenten weniger empfindlich gegenüber dem

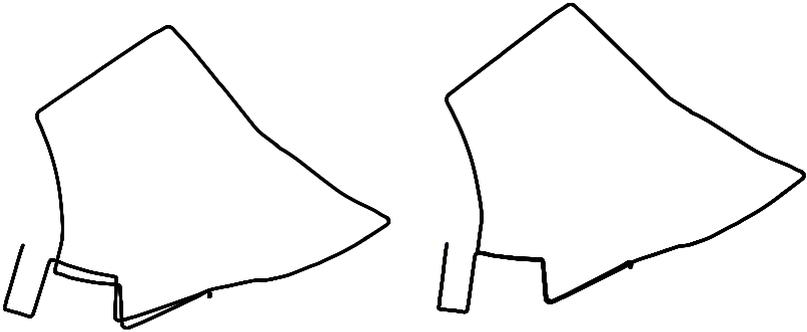


Abbildung 6.7: Schätzung einer Trajektorie mit Ringschluss durch einen direkten gefensternten Bündelblockausgleich (Links) und den zweischrittigen Bündelblockausgleich (Rechts). Beide Resultate zeigen eine Schätzung ohne explizite Ringschluss-optimierung. Die Verbesserung der Rotations-schätzung durch den zweischrittigen Bündelblockausgleich ist deutlich sichtbar.

Fehler der extrinsischen Kamerakalibrierung. Abbildung 6.7 zeigt unterschiedliche Schätzungen einer Ringschlussfahrt. Links dargestellt ist die Schätzung durch einen direkten Bündelblockausgleich und rechts die Schätzung durch den zweischrittigen Bündelblockausgleich. Bei beiden Schätzungen wurden die gleichen Merkmalskorrespondenzen als Messungen und die gleiche Initialisierung der Trajektorien und Landmarken verwendet. Es ist ersichtlich, dass der Endpunkt der Trajektorien-schätzung durch den zweischrittigen Bündelblockausgleich näher und qualitativ richtig am Startpunkt der Trajektorie liegt.

Diese Ergebnisse haben sich mit unterschiedlichen Versuchsträgern in unterschiedlichen Regionen sowie auf KITTI-Odometrie Sequenzen reproduzierbar bestätigt. Abbildung 6.8 zeigt die Verbesserung der VISO2 SVO-Schätzung [79] durch den zweischrittigen Bündelblockausgleich exemplarisch auf unterschiedlichen KITTI-Sequenzen. Die Experimente zeigen, dass der gefensternte Bündelblockausgleich auch eine unpräzise und verzerrte Initialisierungsschätzung der gefahrenen Trajektorien signifikant verbessert.

Unterabtastung der Landmarken

Eine zusätzliche Verringerung der Laufzeit und Verbesserung der Parameterschät-

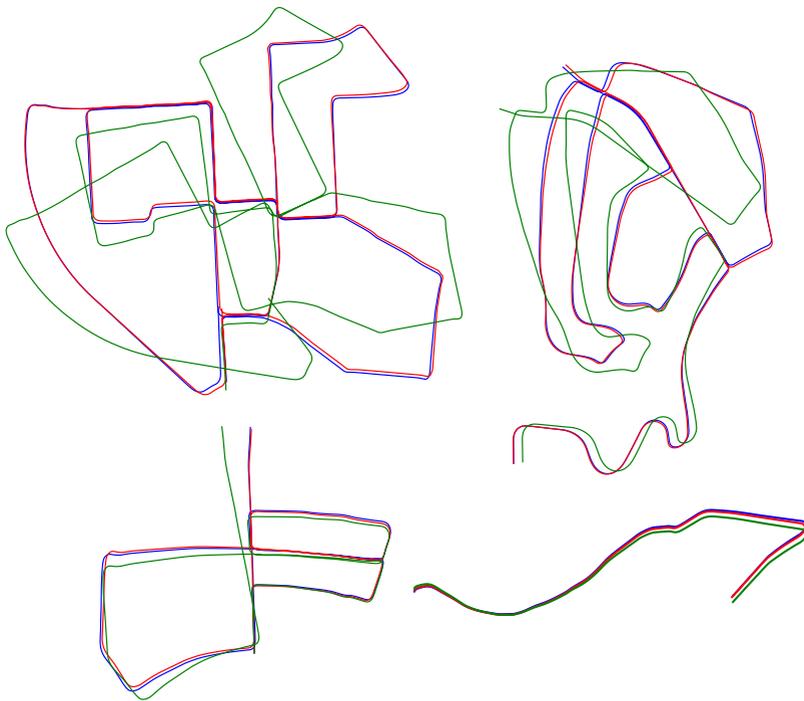


Abbildung 6.8: Verbesserung der initialen SVO-Odometrieschätzung [79] (Grün) durch den zweistufigen Bündelblockausgleich (Blau) ohne Ringschlussoptimierung auf verschiedenen KITTI-Odometrie-Sequenzen. Zum Vergleich ist die Grundwahrheit in rot dargestellt.

zung ergibt sich durch Auswahl einer örtlich gleichverteilten Teilmenge der zur Verfügung stehenden Landmarken und Merkmalskorrespondenzen innerhalb des Optimierungsfensters [73]. Dazu wird eine dünnbesetzte 3D-Rasterkarte definiert, in welche die Landmarken einsortiert werden. Pro Zelle wird die Landmarke beibehalten, die am nächsten zum Zentrum der Zelle liegt. Durch diese räumliche Unterabstufung fallen sowohl nahe als auch entfernte Landmarken ungefähr gleich ins Gewicht. Auf diese Weise wird verhindert, dass weder der Positionsanteil noch der Rotationsanteil der Fahrzeugposen den Gradientenabstieg dominiert. Des Weiteren

reduziert sich die Gesamtanzahl der Landmarken und Residuen und somit die Problemgröße deutlich.

6.2.2 Assoziationstopologie

Die fortlaufende Integration weiterer Aufnahmesequenzen in die bestehende Karte führt zu einer stetig wachsenden Anzahl möglicher Kombinationen zur Merkmalsassoziation. Abbildung 6.9 veranschaulicht diesen Zusammenhang. Dargestellt sind die Trajektorien von mehreren gemeinsam kartierten Aufnahmesequenzen, wobei jeder Punkt eine Fahrzeugpose darstellt, zu der i. A. mehrere aufgezeichnete Kamerabilder existieren. Da viele andere Fahrzeugposen in der unmittelbaren Nähe von \mathbf{P}_j liegen, lassen sich dementsprechend viele Bildpaare finden, zwischen denen Bildmerkmale assoziiert werden können. Die Anzahl der möglichen Paarungen wächst dabei mit jeder neuen Aufnahmesequenz, die das gleiche Gebiet passiert. Daher wird nur eine Teilmenge aller möglichen Paarungen ausgewählt, sodass die Assoziationskomplexität stets unabhängig von der Anzahl der Aufnahmesequenzen ist. Dabei wird zwischen den folgenden drei Varianten unterschieden:

1. Der erste Fall ist die Paarung von Bildern zwischen konsekutiven Posen \mathbf{P}_{j-1} und \mathbf{P}_j aus der gleichen Befahrung. Aus kombinatorischer Sicht ist dieser Fall unkritisch, da die Assoziationskomplexität nur von der Anzahl verwendeter Kameras und der Länge der Aufnahmesequenz abhängig ist.
2. Im Falle der Paarung von Bildern zwischen verschiedenen Aufnahmesequenzen wird sowohl die Anzahl der Kombinationen mit anderen Sequenzen als auch die Anzahl der nächsten Nachbarn pro Sequenz begrenzt. In Abbildung 6.9 ist exemplarisch die Paarung mit zwei der insgesamt sechs anderen Sequenzen und jeweils einer Nächste-Nachbar-Pose pro Sequenz dargestellt ($\{\mathbf{P}_j, \mathbf{P}_m\}, \{\mathbf{P}_j, \mathbf{P}_n\}$). Es werden dabei vorrangig die kartierten Posen und Bilder aus den aktuellsten in diesem Gebiet vorkommenden Aufnahmesequenzen gepaart.
3. Den letzten Fall stellen Ringschlüsse innerhalb der gleichen Befahrung dar. Im Rahmen der Kartengenerierung wird über einen Parameter eine minimale Anzahl konsekutiver Fahrzeugposen vorgegeben, welche zwischen zwei Ringschlussposen liegen müssen, damit ein Ringschluss akzeptiert wird.

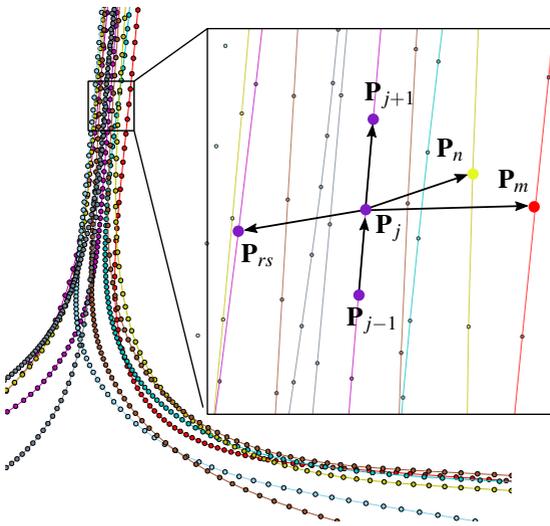


Abbildung 6.9: Trajektorien von sieben kartierten Aufnahmesequenzen an einer Kreuzung. Jede Trajektorie ist in einer anderen Farbe dargestellt. Für jedes Kamerabild, das zu Fahrzeugpose \mathbf{P}_j korrespondiert, lassen sich viele geeignete Paarungen von Bildern aus anderen Befahrungen zur Merkmalsassoziation finden. Die hervorgehobenen Kanten zeigen eine exemplarische Auswahl möglicher Assoziationskombinationen.

So können Mikroringschlüsse zugelassen oder verhindert werden. Bei den durchgeführten Experimenten hat sich jedoch gezeigt, dass Mikroringschlüsse zu keiner nennenswerten Verbesserung der Parameterschätzung führen. Größere Makroringschlüsse steigern hingegen deutlich die globale Konsistenz der Trajektorien-schätzung. Daher werden die Bilder zur Pose \mathbf{P}_j mit den Bildern der nächsten Ringschlusspose \mathbf{P}_{rs} jedes Ringschlusses gepaart.

6.2.3 Lokaler Clusterausgleich

In Abschnitt 6.2.1 wurde gezeigt, dass die Schätzung der Trajektorie $T_{S_{n+1}}$ durch den gleitenden Bündelblockausgleich eine hohe Genauigkeit erzielt. Diese Trajektorien-schätzung wird zusätzlich mit den in Abschnitt 5.3 beschriebenen Modellannahmen validiert. Daher wird angenommen, dass ein gemeinsamer Bündelblockausgleich von Trajektorienabschnitten aus unterschiedlichen Aufnahmesequenzen keine si-

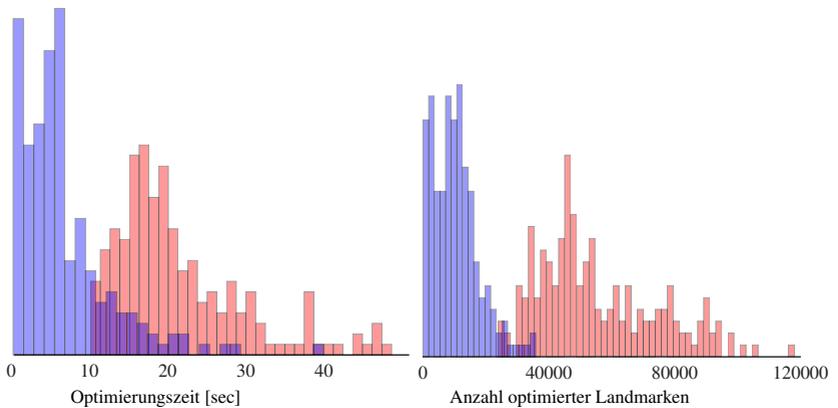


Abbildung 6.10: Vergleich des Berechnungsaufwands des Bündelblockausgleichs und des Clusterausgleichs. Links: Verteilung der Optimierungszeiten von Bündelblockausgleich (Rot) und Clusterausgleich (Blau). Rechts: Verteilung der Anzahl involvierter Landmarken von Bündelblockausgleich (Rot) und Clusterausgleich (Blau).

gnifikante Verbesserung der Posen innerhalb jedes einzelnen Trajektorienabschnitts erzielt. Aufgrund dessen verbleibt lediglich die Anpassung der Trajektoriensegmente zueinander, was durch einen paarweisen Clusterausgleich gemäß Abschnitt 5.1.2 mit deutlich verringerten Berechnungsaufwand im Vergleich zum Bündelblockausgleich möglich ist.

Um die getroffenen Annahmen empirisch zu untermauern, wurden Experimente durchgeführt, bei denen der Clusterausgleich mit einem Bündelblockausgleich verglichen wurde. Dazu wurden überlappende Trajektorienabschnitte aus mehreren, unterschiedlichen Aufnahmesequenzen sowie Trajektorienabschnitte aus der gleichen Sequenz in Ringschlussbereichen zueinander optimiert. Das Experiment wurde in verschiedenen Umgebungen, wie beispielsweise Häuserschluchten oder wäldlichen Regionen, wiederholt. Die gefahrene Distanz pro Cluster beträgt im Mittel $\sim 12\text{m}$ und es werden, je nach Konfiguration und Posendichte, zwischen 80 und 300 Kamerabilder verwendet. Die gleichen Paarungen wurden zum Vergleich durch Bündelblockausgleichsprobleme optimiert. Abbildung 6.10 zeigt Resultate dieses

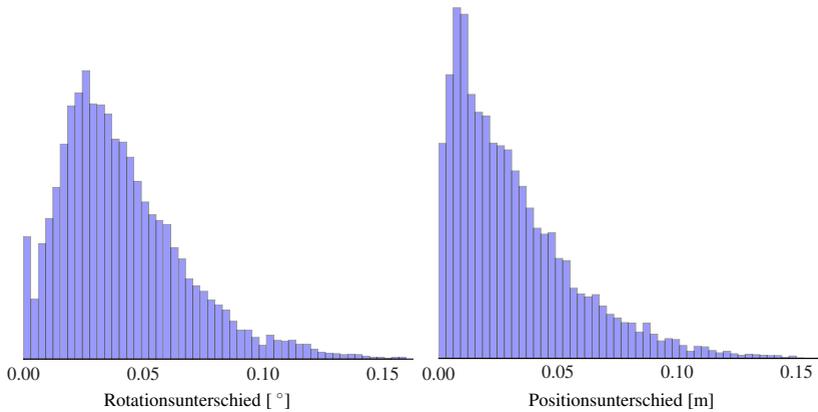


Abbildung 6.11: Vergleich der Schätzergebnisse des Bündelblockausgleichs und des Clusterausgleichs. Links: Verteilung des absoluten Rotationsunterschieds der Posenschätzungen. Rechts: Verteilung des absoluten Positionsunterschieds.

Experiments. Links ist die Verteilung der Optimierungszeiten und rechts die Verteilung der Anzahl optimierter Landmarken dargestellt. Die Verteilungen zeigen eine klare Abhängigkeit zwischen der Landmarkenanzahl und der benötigten Optimierungszeit. Durch die verringerte Landmarkenanzahl hat die Clusteroptimierung einen geringeren Speicherbedarf, was insbesondere vorteilhaft für die Parallelisierung ist, da mehr Cluster parallel gelöst werden können. Des Weiteren zeigt sich, dass der Clusterausgleich die mittlere Laufzeit mehr als halbiert.

Zum Vergleich der Schätzergebnisse des Bündelblockausgleichs und des Clusterausgleichs wurden beide Schätzungen in ein gemeinsames Koordinatensystem transformiert und für korrespondierende Posen die absolute Posendifferenz gebildet. Abbildung 6.11 zeigt links die Verteilung des Rotationsunterschieds und rechts die Verteilung des Positionsunterschieds der geschätzten Fahrzeugposen. Insgesamt umfasst die Verteilung die Unterschiede von $\sim 23\text{Tsd}$ optimierten Posen aus allen gelösten Ausgleichsproblemen. Die Verteilungen zeigen, dass der Unterschied der Schätzungen des Clusterausgleichs zum Bündelblockausgleich in einer vernachlässigbaren Größenordnung ist und die signifikante Reduktion der Laufzeit die approximati-

ve Schätzung durch den Clusterausgleich rechtfertigt. Der geringe Unterschied der beiden Schätzungen zeigt des Weiteren, dass die zuvor berechnete separate Trajektorienschätzung durch das gleitende Bündelblockausgleichsfenster bereits eine hohe Schätzungsgüte erzielt, da der gemeinsame Bündelblockausgleich beider Cluster keine wesentliche Verbesserung gegenüber dem Clusterausgleich erzielt.

6.2.4 Robuste Kartenposenschätzung

Im letzten Schritt der Kartengenerierung werden die Posen der neuen Aufnahmesequenz S_{n+1} durch Lösen eines Posendifferenzenausgleich an die bestehenden Kartenposen P_n angepasst. Die folgende Übersicht beschreibt die getroffenen Maßnahmen für eine möglichst robuste Kartenposenschätzung und zur Validierung der resultierenden Kartenposen P_{n+1} :

- **Wahl der Gewichtungsmatrizen $\Omega_{D_{i,j}}$**

Empirische Untersuchungen und die praktische Anwendung haben gezeigt, dass eine heuristische Wahl der Gewichtungsmatrizen $\Omega_{D_{i,j}}$ in der überwiegenden Mehrzahl der Fälle eine akkurate und konsistente Posenschätzung erzielt. Um zu vermeiden, dass der Translationsfehler den Rotationsfehler unter den gegebenen Rahmenbedingungen dominiert, werden Diagonalmatrizen verwendet, bei denen die Rotationskomponenten mit Faktor 10 gegenüber den Translationskomponenten gewichtet sind.

- **Ausreißerbehandlung**

Zur Ausreißerbehandlung werden alle Ringschlussposendifferenzen und Posendifferenzen zwischen Posen verschiedener Befahrungen als unsicher angenommen und gemäß Abschnitt 5.1.4 als deaktivierbare Residuen modelliert. Für die Posendifferenzen zwischen konsekutiven Fahrzeugposen wird aufgrund des gleitenden Bündelblockausgleichs und der Trajektorienvalidierung durch geometrische Modelle angenommen, dass keine Ausreißer vorhanden sind.

- **Analyse des Trajektorien-drifts**

Die an die Kartenposen angepasste Trajektorie der neuen Aufnahmesequenz wird durch einen Trajektorienvergleich gemäß Abschnitt 5.3.3 automatisch mit der Trajektorienschätzung $T_{S_{n+1}}$ aus der Sequenzaufbereitung verglichen.

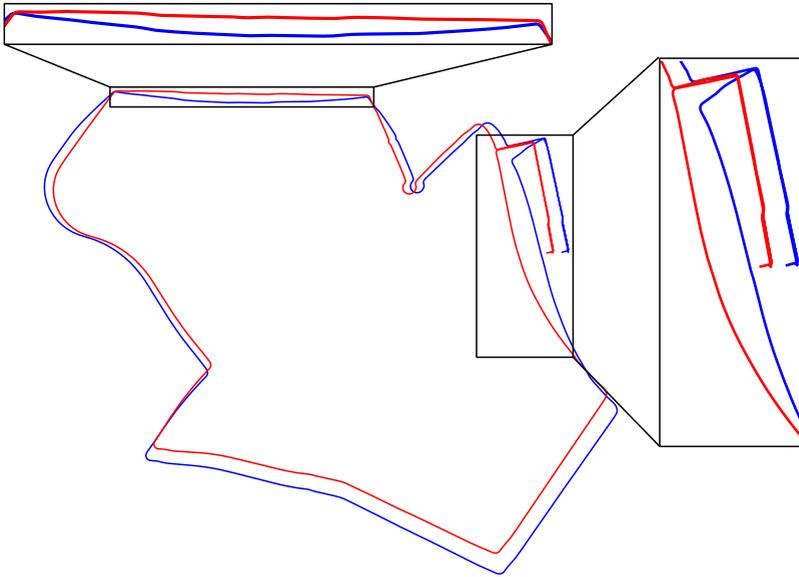


Abbildung 6.12: Identifizierte Trajektorienverzerrungen durch den Trajektorienvergleich. Dargestellt ist die geschätzte Trajektorie einer Aufnahmesequenz nach der Sequenzaufbereitung (Rot) und nach dem gemeinsamen Posendifferenzenausgleich mit zuvor kartierten Trajektorien (Blau). Die hervorgehobenen Bereiche zeigen unterschiedliche Krümmungen von verschiedenen Trajektorienabschnitten.

Abbildung 6.12 zeigt exemplarisch identifizierte Trajektorienverzerrungen. Es zeigen sich Verkrümmungen der Trajektorie in den hervorgehobenen Bereichen, die aus einem Skalfehler in der verwendeten Kamerakalibrierung zu dieser Aufnahmesequenz resultieren. Der Skalfehler konnte in diesem Beispiel anhand der zurückgelegten Distanz der Trajektorien-schätzung, die deutlich von der zurückgelegten Distanz der korrespondierenden GNSS-Trajektorie abweicht, nachgewiesen werden. Bei der Anpassung der falsch skalierten Trajektorie an die bestehende Karte durch den Posendifferenzenausgleich wurde die Skaleninkonsistenz durch Verkrümmungen von



Abbildung 6.13: Abschnitte der geschätzten Trajektorien (Blau) und korrespondierender GNSS-Trajektorien (Rot) aus dem seriennahen GNSS-Empfänger. Von links nach rechts sind unterschiedliche Fehlerarten in den GNSS-Daten sichtbar. Links: Positionssprung. Mitte: Ausreißermessungen. Rechts: Fehlerhafter systematischer Positionsversatz in einem Ringschlusszenario.

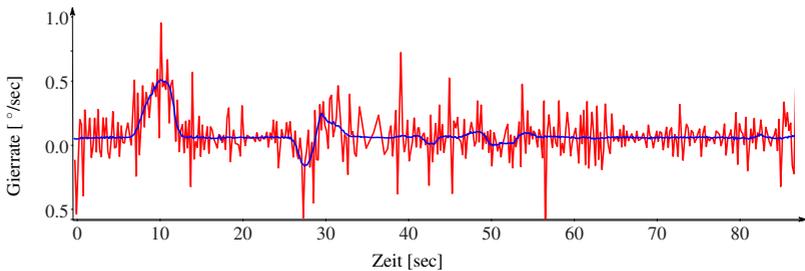


Abbildung 6.14: Ausschnitt der Gierratensignale einer kartierten Trajektorie (Blau) und der korrespondierenden GNSS-Trajektorie (Rot).

Trajektorienabschnitten, an denen sich die Trajektorie nicht mit den Kartentrajektorien überlappt, kompensiert.

- **Georeferenzierung**

Die Hinzunahme von GNSS-Messungen in den finalen Posendifferenzenausgleich der Kartengenerierung ermöglicht zum einen eine Georeferenzierung und gewährleistet zum anderen eine globale Konsistenz der Kartenposenschätzung. Auch wenn die GNSS-Daten häufig lokale Störungen aufweisen, lässt sich durch gemeinsame Optimierung der Trajektorien unter Einbeziehung

ung der GNSS-Messungen im Mittel eine akkurate Referenzierung erzielen. Abbildung 6.13 zeigt exemplarisch drei typische Fehlerklassen in den aufgezzeichneten GNSS-Trajektorien. Solche Fehler sowie ein hohes Signalrauschen, dass exemplarisch in Abbildung 6.14 dargestellt ist, treten bei Befahrungen im innerstädtischen Bereich aufgrund von Signalabschattung und Mehrwegeausbreitung durch Gebäude und Bäume zahlreich auf. Bei der Lösung des Posendifferenzenausgleichs werden die GNSS-Messungen daher mit einem Faktor von 10^{-3} gegenüber den Posendifferenzen gewichtet, sodass die gezeigten Störungen der GNSS-Signale nahezu keinen Einfluss auf die lokale Genauigkeit der Kartenposen haben, sondern nur zur globalen Konsistenz der Kartentrajektorien beitragen.

6.3 Experimente

Zur Evaluation der Kartengenerierung wird die Erstellung einer Kartendatenbank aus $n = 14$ Aufnahmesequenzen von Befahrungen im Umkreis von Karlsruhe präsentiert. Die Befahrungen wurden innerhalb eines Zeitraums von November bis April in zeitlich unterschiedlichen Abständen durchgeführt. Sechs der 14 Befahrungen befinden sich, abgesehen von Spurwechseln, auf demselben Rundkurs und umfassen innerstädtische Bereiche, Walddurchfahrten und größere Zubringerstraßen. Die anderen Befahrungen erweitern diesen Rundkurs um weitere Strecken. Das gesamte kartierte Gebiet ist in Abbildung 6.15 dargestellt. Die Aufzeichnungen umfassen verschiedene Witterungsbedingungen und wurden zu unterschiedlichen Tages- und Jahreszeiten mit Sensorkonfiguration MKS-A (siehe Abschnitt 3.1) aufgezeichnet. Im Laufe der Experimente wurden Objektive gewechselt und die Einbaupositionen verändert, sodass die aufgezeichneten Daten insgesamt eine breite Variation von Bildverzeichnungen umfassen. Zusätzlich zu den Bildern wurden GNSS-Daten von einem seriennahen GPS-Empfänger aufgezeichnet. In Summe wurden 194Tsd Bilder und 18Tsd GNSS-Messungen aufgezeichnet und bei der Kartengenerierung verarbeitet.

Die kartierten Trajektorien wurden durch Hinzunahme der aufgezeichneten GNSS-Messungen im finalen Posendifferenzenausgleich georeferenziert. Für eine quantitative Evaluation der Genauigkeit und Konsistenz der geschätzten Kartentrajek-

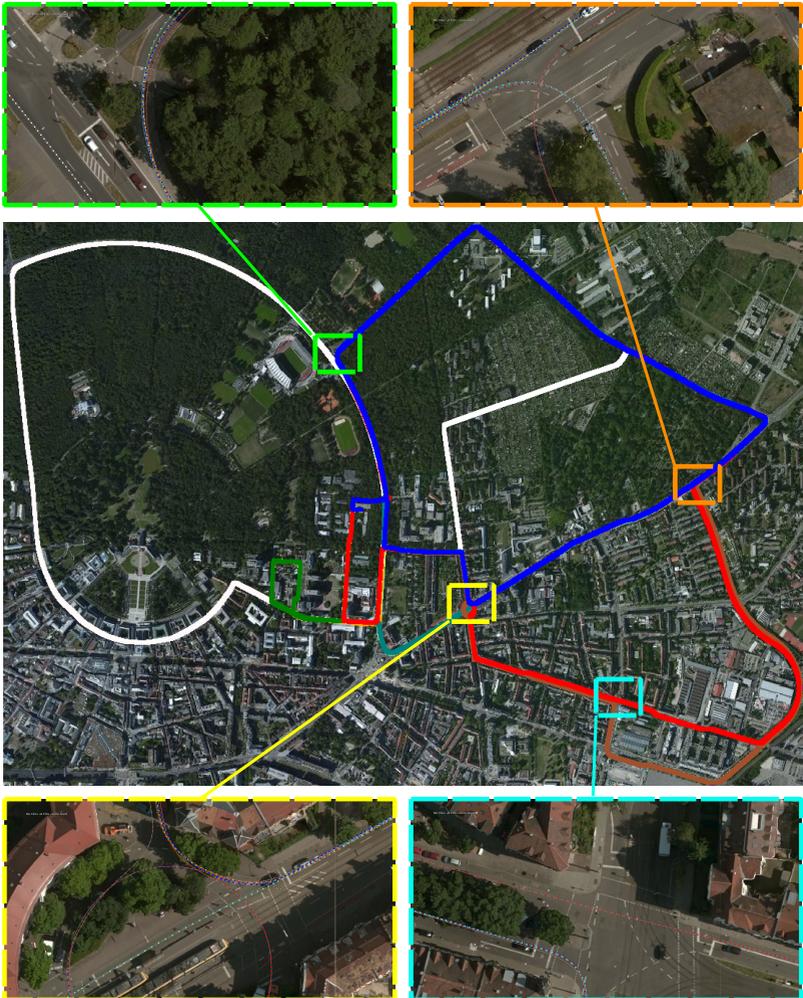


Abbildung 6.15: Überlagerung der geschätzten Trajektorien mit einer Luftbild-Referenz, wobei jede kartierte Trajektorie in einer anderen Farbe dargestellt ist. Die Detailansichten lassen sich am besten digital betrachten. Luftbilder: ©Microsoft/Here

torien wurde ein absoluter Vergleich der georeferenzierten Kartentrajektorien mit korrespondierenden Trajektorien eines hochgenauen GNSS-Empfängers in Betracht gezogen. Jedoch stellen die Trajektorien eines hochgenauen GNSS-Empfängers, insbesondere bei der gemeinsamen Referenzierung mehrerer Trajektorien der gleichen Strecke zu verschiedenen Tages- und Jahreszeiten, keine zuverlässige Grundwahrheit dar. Der Hauptgrund dafür ist, dass der systematische Fehler durch Mehrwegeausbreitung nicht nur von der umgebenden Struktur, sondern auch von der geometrischen Anordnung der Satelliten abhängt. Wird das gleiche Gebiet an verschiedenen Tagen und zu verschiedenen Zeiten befahren, kann keine gleiche Anordnung der Satelliten angenommen werden. Daher ergeben sich unterschiedliche systematische Fehler bei den verschiedenen Befahrungen. Im Einklang mit anderen Arbeiten [80] wurde dies in eigenen Arbeiten empirisch nachgewiesen [67]. Als Alternative zur Demonstration der globalen Konsistenz werden die Trajektorien mit georeferenzierten Luftbildern überlagert und visuell analysiert. Abbildung 6.15 zeigt Ausschnitte aus dieser Überlagerung, wobei die Trajektorien der verschiedenen Aufnahmesequenzen in unterschiedlichen Farben dargestellt sind. Aus den höheren Zoomstufen ist ersichtlich, dass die Trajektorien sich akkurat und erwartungstreu mit der Straßengeometrie in den Luftbildern überlagern.

Zur Demonstration der lokalen Genauigkeit der Kartenschätzung werden manuell annotierte Strukturen der Umgebung in die Bilder verschiedener Aufnahmesequenzen projiziert. Dazu wurden markante Strukturen in jeweils zwei synchron aufgezzeichneten Bildern der beiden nach vorne ausgerichteten Kameras von Sensorconfiguration MKS-A pixelgenau annotiert, trianguliert und mit der korrespondierenden Kartenpose in das Kartenkoordinatensystem transformiert. Darauf basierend wurden die Annotationen in die Bilder aus der gleichen Aufnahmesequenz und anderer Aufnahmesequenzen projiziert. Die Abbildungen 6.16, 6.17 und 6.18 zeigen exemplarisch Projektionen von unterschiedlichen Objekten der Umwelt. In Grün dargestellt ist die Projektion in Bilder der gleichen Aufnahmesequenz und in Gelb die Projektion in Bilder aus den anderen Aufnahmesequenzen. Die dargestellten Bilder entspringen aus verschiedenen Kameras zu verschiedenen Tages- und Jahreszeiten und zeigen die Projektionen aus unterschiedlichen Perspektiven. Es ist ersichtlich, dass die Projektionen unabhängig von der Beobachtungspose mit hoher Genauigkeit an der erwarteten Position im Bild liegen, was die lokale Genauigkeit der Trajekto-

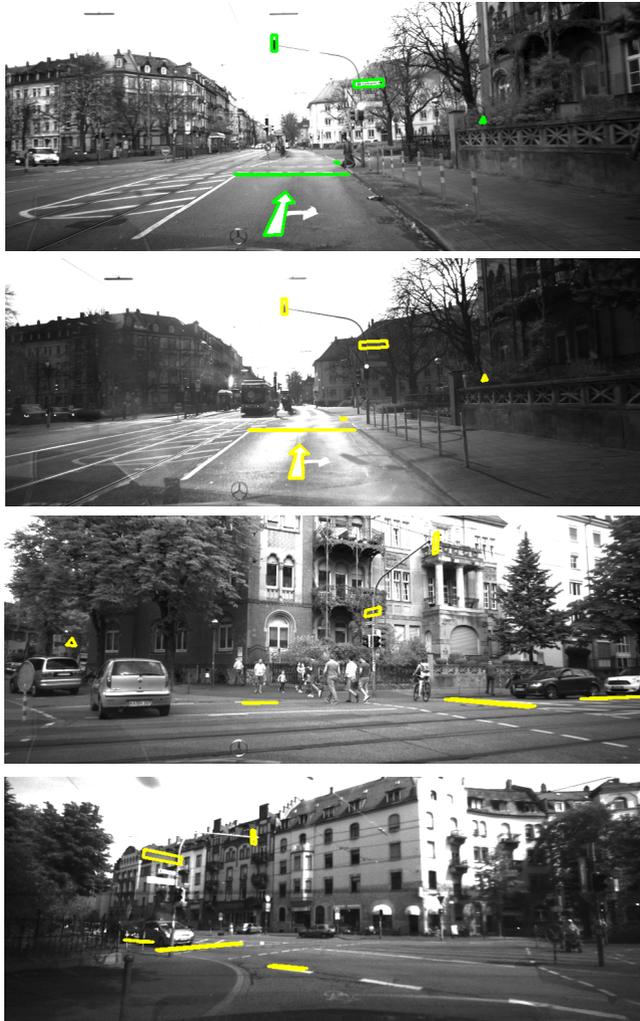


Abbildung 6.16: Projektion (Gelb) manuell annotierter, markanter Strukturen in die aufgezeichneten Bilder verschiedener Aufnahmesequenzen. Die Strukturen wurden in zwei synchron aufgezeichneten Bildern aus einer einzelnen Befahrung annotiert (Grün), trianguliert und auf Basis der geschätzten Kartenposen in die Bilder anderer Aufnahmesequenzen projiziert.

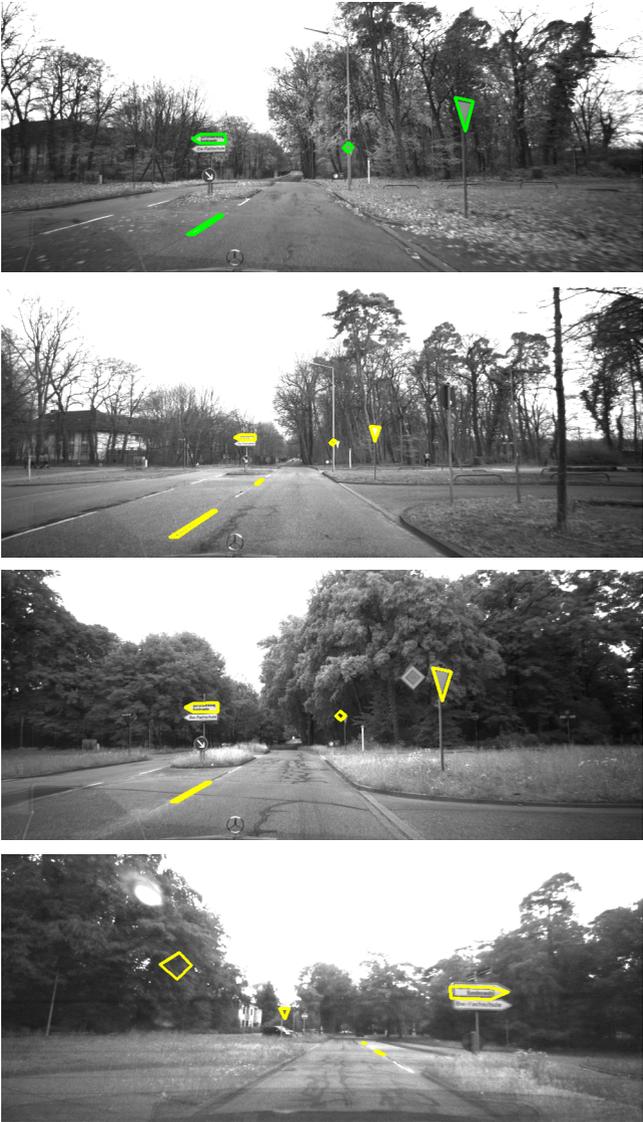


Abbildung 6.17: Projektion von markanten Strukturen in wäldlichen Gebieten.



Abbildung 6.18: Projektion von markanten Strukturen in Gebieten mit wenig Umgebungsstruktur.

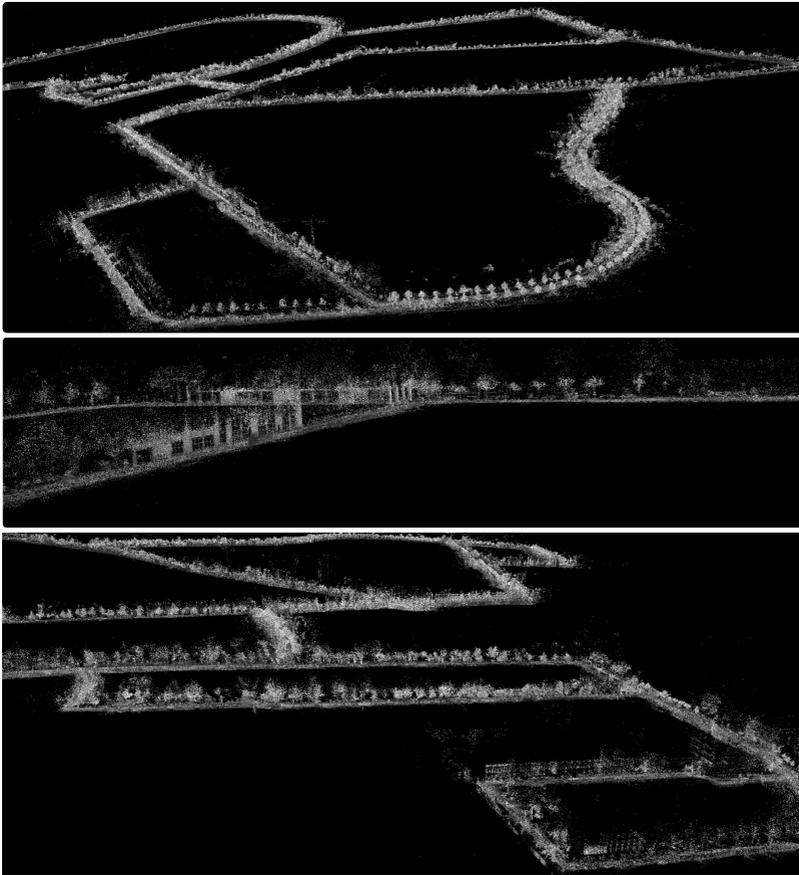


Abbildung 6.19: Darstellungen der Punktwolke der kartierten Landmarken aus den 14 Aufnahme sequenzen. Jede Landmarke ist mit dem korrespondieren Grauwert eines ihrer zugehörigen Merkmale eingefärbt.

rienreferenzierung untermauert.

Zur weiteren Demonstration der Genauigkeit wird die Punktwolke der Kartenlandmarken dargestellt. Abbildung 6.19 zeigt verschiedene Ansichten auf die Punktwolke, in denen sich eine konsistente Umgebungsrepräsentation erkennen lässt. Die

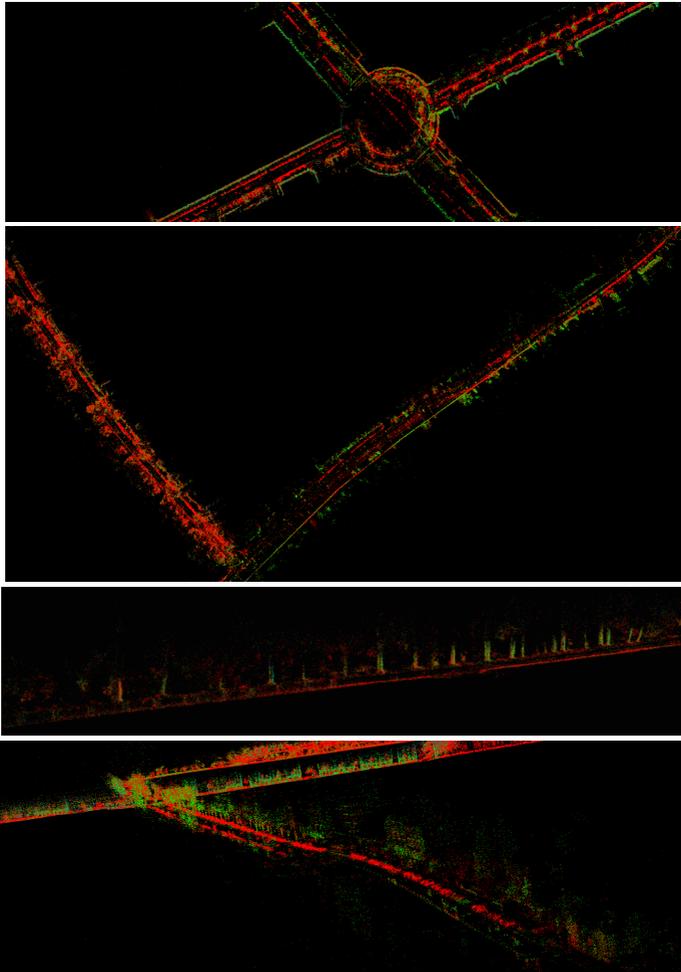


Abbildung 6.20: Darstellungen der Landmarkenpunktvolke hinsichtlich der Häufigkeit der Wiedererkennung der Landmarken in verschiedenen Aufnahme­sequenzen. Eher rötlich dargestellte Landmarken wurden nur in wenigen Befahrungen wiedererkannt. Eher grünlich dargestellte Landmarken wurden demhingegen häufiger wiedererkannt.

Landmarken sind gemäß des Grauwerts des ersten zugeordneten Merkmals eingefärbt. Es lassen sich die präzise rekonstruierten Konturen von Bäumen und Hauswänden erkennen. Abbildung 6.20 zeigt weitere Ansichten der Punktwolke, wobei die Landmarken hier in Abhängigkeit ihrer Anzahl von Wiedererkennungen in verschiedenen Aufnahmesequenzen dargestellt sind. Grün dargestellte Landmarken wurden in vielen Befahrungen wiedererkannt, wohingegen eher rötlich eingefärbte Landmarken nur in wenigen Aufnahmesequenzen wiedererkannt wurden. Insbesondere aus der Vogelperspektive zeigen sich eindeutige Strukturen von Hausfassaden und Bäumen, die gemäß der Erwartung in vielen Aufnahmesequenzen wiedererkannt wurden. Zur Fahrbahnmitte häufen sich hingegen viele Landmarken auf parkenden Fahrzeugen, die in den meisten Fällen nur in einer einzelnen Aufnahmesequenz detektiert wurden. Die Darstellung der Wiedererkennungshäufigkeit in Abbildung 6.20 deutet insbesondere darauf hin, dass Landmarken auf persistenten Strukturen, wie beispielsweise Häusern und Baumstämmen, gemäß der Erwartung besonders nutzvoll für eine zuverlässige Lokalisierung sind.

Zusammenfassend zeigen die qualitativen Analysen eine hohe Schätzgenauigkeit der Kartenposen und -landmarken. Insbesondere die Projektion von Umgebungsstrukturen ist empfindlich gegenüber falsch geschätzten Winkeln und Positionen der Kartenposen. Die pixelgenauen Projektionen der Annotationen an ihren erwarteten Positionen belegen daher eindeutig die hohe lokale Konsistenz der Kartenparameter. Durch die Hinzunahme von GNSS-Messungen in die Optimierung wird zusätzlich eine globale Konsistenz der Kartentrajektorien erzielt, auch wenn die verfügbaren GNSS-Messungen häufig fehlerbehaftet und stark verrauscht sind.

7 Multi-Kamera-Lokalisierung

Dieses Kapitel befasst sich mit der in dieser Arbeit entwickelten Multi-Kamera-Lokalisierung. Das Ziel der Lokalisierung ist eine möglichst genaue Schätzung der kartenrelativen Fahrzeugpose zum Aufnahmezeitpunkt t der Bilder Σ_t . Die zum Aufnahmezeitpunkt t geschätzte Lokalisierungspose $\mathbf{P}_t^{\text{loc}} \in SE(3)$ beschreibt die Transformation zwischen dem Fahrzeugreferenzsystem R und dem Ursprung der Lokalisierungskarte M_n^L . Zur Lokalisierung werden ausschließlich die mit dem MKS aufgezeichneten Bilder Σ_t verwendet. Es ist sowohl eine Lokalisierung mit einer als auch mit mehreren Kameras möglich.

Die Evaluationen in diesem Kapitel stellen immer die Ergebnisse der Schätzung einer einzelnen Pose pro Lokalisierungsschritt dar (engl. *single-shot localization*). Es wird keine zeitliche Filterung durchgeführt und keine weitere Sensorik benötigt. Die Bestimmung einer zeitlich gefilterten Lokalisierungsschätzung durch Lösen eines Ausgleichsproblems innerhalb eines gleitenden Fensters unter Echtzeitbedingungen wurde im Rahmen einer eigenen Arbeit [51] und einer studentischen Arbeit [81] unter verschiedenen Gesichtspunkten analysiert. Die in diesem Kapitel vorgestellten Methoden und Auswertungen fokussieren hingegen die Verwendung eines MKS für die Lokalisierung.

Die Verwendung eines MKS zur Lokalisierung steigert die Genauigkeit, Robustheit und Verfügbarkeit der Lokalisierung gegenüber der Verwendung eines monoskopischen Kamerasystems. Das MKS stellt i. A. Bilder aus verschiedenen Blickrichtungen bereit, sodass Merkmale aus unterschiedlichen Orientierungen beobachtet werden. Dadurch ist es möglich, dass beispielsweise Kartenmerkmale aus Bildern, die mit einer rückwärts gerichteten Kamera aufgezeichnet wurden, zu detektierten Merkmalen aus Bildern aus einer vorwärts gerichteten Kamera assoziiert werden. Mit diesen Korrespondenzen kann somit auch eine Lokalisierungspose geschätzt werden, wenn die kartierte Strecke in entgegengesetzter Richtung befahren wird.

Dies ist insbesondere bei punktförmigen Merkmalen ein großer Vorteil, da eine robuste Merkmalsassoziation ähnliche Beobachtungsposen erfordert. Die in diesem Kapitel beschriebenen Konzepte und Kartenstrukturen stellen eine kamerabasierte Lokalisierungsmethode dar, die diese vorteilhaften Eigenschaften eines MKS effizient nutzt.

Die Lokalisierung umfasst zwei Teilsysteme. Das erste Teilsystem ist eine Methode zur Ortswiedererkennung und wird zur erstmaligen Bestimmung des ungefähren Aufenthaltsortes und der ungefähren Orientierung des Fahrzeugs in der Karte benötigt. Das zweite Teilsystem stellt die metrische Lokalisierung dar, welches eine Initialisierungspose benötigt, um damit Kartenlandmarken aus der lokalen Umgebung bereitzustellen. Mit den bereitgestellten Kartenlandmarken und -merkmalen wird schließlich die hochpräzise kartenrelative Lokalisierungspose $\mathbf{P}_t^{\text{loc}}$ durch Lösen eines reduzierten Bündelblockausgleichsproblems geschätzt. Zunächst wird in Abschnitt 7.1 der Ablauf der Lokalisierung erläutert sowie ein Überblick über das Gesamtsystem gegeben. Im darauffolgenden Abschnitt 7.2 ist die Ortswiedererkennung beschrieben. Der letzte Abschnitt 7.3 befasst sich mit der metrischen Lokalisierung.

7.1 Überblick

Abbildung 7.1 zeigt den grundsätzlichen Ablauf der Lokalisierung. Die resultierende Lokalisierungspose $\mathbf{P}_t^{\text{loc}}$ wird durch die metrische Lokalisierung geschätzt. Dafür benötigt die metrische Lokalisierung in jedem Schritt eine Initialisierungspose $\mathbf{P}_t^{\text{init}} \in SE(3)$, die den ungefähren Aufenthaltsort und die ungefähre Fahrzeugorientierung relativ zum Kartenursprung beschreibt. Das Initialisierungsmodul stellt die Initialisierungspose $\mathbf{P}_t^{\text{init}}$ für die metrische Lokalisierung bereit. Für die erstmalige Initialisierung wird die Ortswiedererkennung benötigt.

Initialisierungsmodul

Die folgenden zwei Fälle beschreiben die Funktionsweise des Initialisierungsmoduls, welches die beiden Teilsysteme der entwickelten Lokalisierungsmethode verknüpft:

1. Erstmalige Initialisierung

Da vor Eintritt des Fahrzeugs in den kartierten Bereich noch keine Infor-

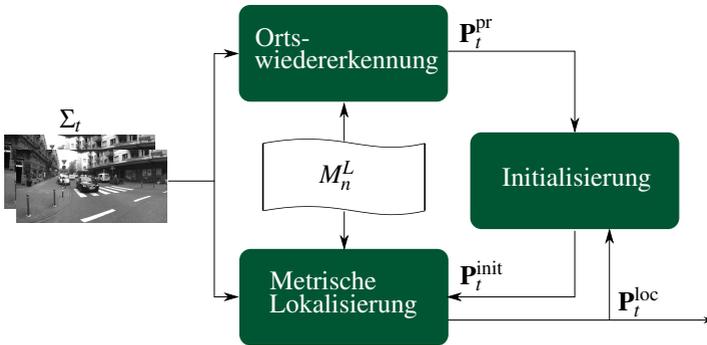


Abbildung 7.1: Zusammenspiel der Lokalisierungs-komponenten.

Information über den aktuellen Aufenthaltsort bekannt ist, wird die Ortswiedererkennung benötigt. Diese schätzt bildbasiert die Kartenpose $\mathbf{P}_t^{\text{pr}} \in SE(3)$, die der aktuellen Fahrzeugpose am nächsten ist, wobei Bildmerkmale mit allen in der Karte gespeicherten Merkmalen verglichen werden. Die geschätzte nächste Kartenpose \mathbf{P}_t^{pr} repräsentiert in diesem Fall den ungefähren Aufenthaltsort und wird unabhängig vom Ort des Eintritts in den kartierten Bereich bestimmt. In diesem Fall wird daher die aktuelle Schätzung \mathbf{P}_t^{pr} der Ortswiedererkennung direkt an die metrische Lokalisierung weitergeleitet $\mathbf{P}_t^{\text{init}} = \mathbf{P}_t^{\text{pr}}$. Nach der ersten erfolgreichen Initialisierung durch die Ortswiedererkennung wird diese erst wieder benötigt, wenn die fortlaufende Lokalisierung unterbrochen und die Information über den aktuellen Aufenthaltsort verloren wurde. Das ist beispielsweise der Fall, wenn der kartierte Bereich verlassen und zu einem späteren Zeitpunkt erneut betreten wird.

2. Fortlaufende Initialisierung

Bei der fortlaufenden Initialisierung ist angenommen, dass eine Historie an vorhergegangenen Lokalisierungsposenschätzungen $\mathbf{P}_{t-N}^{\text{loc}}, \dots, \mathbf{P}_{t-1}^{\text{loc}}$ durch die metrische Lokalisierung existiert. Basierend auf diesem Wissen wird die Initialisierungspose $\mathbf{P}_t^{\text{loc}}$ prädiziert. Die fortlaufende Initialisierung hat stets Priorität vor der Initialisierung durch die Ortswiedererkennung, da der unge-

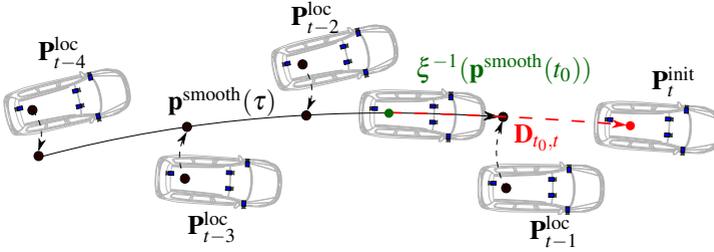


Abbildung 7.2: Fortlaufende Initialisierung durch polynomielle Approximation der Lokalisierungsposenhistorie.

fähre Aufenthaltsort in der Karte bereits bekannt ist und die Initialisierungspose robust und präzise durch die Prädiktion auf Basis der Lokalisierungsposenhistorie bestimmt werden kann.

Zur Prädiktion von $\mathbf{P}_t^{\text{loc}}$ wird eine polynomielle Approximation $\mathbf{p}^{\text{smooth}}(\tau)$ der Lokalisierungsposenhistorie $\mathbf{P}_{t-N}^{\text{loc}}, \dots, \mathbf{P}_{t-1}^{\text{loc}}$ gemäß Abschnitt 5.3.2 bestimmt. Die Initialisierungspose $\mathbf{P}_t^{\text{init}}$ kann darauf basierend durch Extrapolation $\xi^{-1}(\mathbf{p}^{\text{smooth}}(t))$ der polynomiellen Approximation $\mathbf{p}^{\text{smooth}}(\tau)$ bestimmt werden.

Alternativ kann eine externe Odometrie verwendet werden, die eine Posendifferenz $\mathbf{D}_{t_0,t}$, $t-N \leq t_0 < t$ bereitstellt, welche die Fahrzeugbewegung zwischen dem Zeitschritt t_0 und dem aktuellen Zeitschritt t beschreibt. Die Initialisierungspose $\mathbf{P}_t^{\text{init}}$ kann so durch Komposition $\mathbf{P}_t^{\text{init}} = \xi^{-1}(\mathbf{p}^{\text{smooth}}(t_0)) \cdot \mathbf{D}_{t_0,t}$ der geglätteten Lokalisierungstrajektorie und der Odometrieschätzung $\mathbf{D}_{t_0,t}$ bestimmt werden. Dadurch wird die Extrapolation der polynomiellen Approximation vermieden und die Information über die Bewegung des Fahrzeugs verwendet, wodurch in der Regel eine genauere Initialisierung erzielt werden kann. Abbildung 7.2 stellt dies schematisch dar.

Eine Lokalisierungsposenhistorie von $\sim 0.5\text{s}$ Länge bei 10Hz Aufnahme Frequenz und einem Polynomgrad von 1 hat in der Anwendung gute Ergebnisse erzielt.

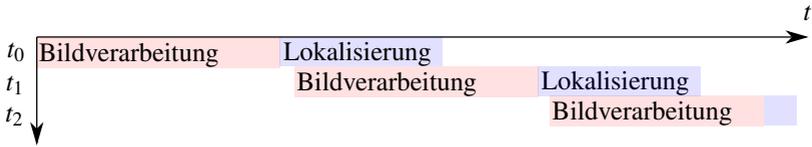


Abbildung 7.3: Prozessstaffelung der Lokalisierung.

Prozessstaffelung

Um eine möglichst hohe Lokalisierungsfrequenz zu erzielen, werden, wie in Abbildung 7.3 dargestellt, konsekutive Prozessschritte zeitlich gestaffelt durchgeführt (engl. *pipelining*). In der praktischen Anwendung hat sich gezeigt, dass dadurch die Lokalisierungsfrequenz bei Verwendung von 4 Kameras von 8 Hz auf bis zu 13 Hz gesteigert werden kann. Eine höhere Lokalisierungsfrequenz begünstigt u. a. die Schätzung der Initialisierungspose $\mathbf{P}_t^{\text{init}}$ durch die fortlaufende Initialisierung, da die Stützstellen der polynomiellen Approximation zeitlich dichter beieinander liegen.

Struktur der Lokalisierungskarte

Die Lokalisierungskarte M_n^L stellt, wie in Abbildung 7.1 dargestellt, Kartenmerkmale für die metrische Lokalisierung und für die Ortswiedererkennung bereit. Da die beiden Lokisierungsmodule methodisch unterschiedlich sind und verschiedene Merkmalstypen benötigen, umfasst die Lokalisierungskarte M_n^L zwei deckungsgleiche Kartenschichten:

1. Die erste Kartenschicht stellt eine Selektion holistischer Kartenmerkmale und ihre korrespondierenden Kartenkameraposen aus der Kartendatenbank M_n für die Ortswiedererkennung bereit. Ein holistisches Merkmal ist durch einen einzelnen Deskriptorvektor repräsentiert, der ein ganzes Bild beschreibt. Die Kartendatenbank M_n stellt daher für jedes Bild, das zur Generierung der Karte verwendet wurde, ein holistisches Kartenmerkmal bereit.
2. Die zweite Kartenschicht stellt eine Teilmenge der Landmarken L_n und ihre korrespondierenden Merkmale aus der Kartendatenbank M_n für die metrische Lokalisierung bereit.

Die Verwendung und Struktur der jeweiligen Kartenschicht wird in den folgenden Abschnitten 7.2 und 7.3 detailliert beschrieben. Für beide Kartenschichten wurde vorab eine Selektion durchgeführt, die eine Teilmenge der kartierten Merkmale und Landmarken aus M_n auswählt und diese in die jeweilige Schicht der Lokalisierungs-karte M_n^L überführt. Die Merkmalsselektion wird im folgenden Kapitel 8 detailliert erläutert.

7.2 Ortswiedererkennung

Die Ortswiedererkennung schätzt die nächste zum aktuellen Aufenthaltsort des Fahrzeugs liegende Kartenpose durch Vergleich von holistischen Kartenmerkmalen und holistischen Merkmalen, die aus den aktuellen Kamerabildern Σ_t berechnet werden. Hierfür werden holistische DIRD-Merkmale $\mathbf{h}_i \in \mathbb{R}^M$ verwendet [82], die eine robuste bildbasierte Ortswiedererkennung auch bei starken Beleuchtungsunterschieden ermöglichen.

7.2.1 Ablauf

Die Ortswiedererkennungskartenschicht speichert holistische Kartenmerkmale $H^L = \{\mathbf{h}_1^L, \dots, \mathbf{h}_n^L\}$, die zu den Kartenkameraposen $C^L = \{\mathbf{C}_1^L, \dots, \mathbf{C}_n^L\}$ aus der Kartendatenbank M_n korrespondieren. Die Kartenstruktur ist ein M -dimensionaler k - d -Suchbaum, in dem die Kartenmerkmale H^L abgelegt sind. Zur Bestimmung des wahrscheinlichsten Aufenthaltsortes werden die folgenden drei Schritte durchgeführt:

1. Merkmalsassoziation

Für jedes Bild $I_k \in \Sigma_t$ wird ein holistisches Merkmal \mathbf{h}_k berechnet und mit diesem eine Radiussuche im k - d -Suchbaum mit Radius f_{\max} durchgeführt. Auf diese Weise werden die n Kartenmerkmale $\mathbf{h}_1^k, \dots, \mathbf{h}_n^k \in H^L$ gefunden, deren Assoziationsdistanz $d(\mathbf{h}_k, \mathbf{h}_i^k)$ zu \mathbf{h}_k kleiner als der Schwellwert f_{\max} ist. Als Distanzfunktion $d(\mathbf{h}_k, \mathbf{h}_i^k) = \|\mathbf{h}_k - \mathbf{h}_i^k\|_1$ wird die $L1$ -Norm der Deskriptordifferenz verwendet [82]. Ist die Anzahl n von assoziierten Kartenmerkmalen größer als eine Maximalanzahl n_{\max} , werden nur die n_{\max} Kartenmerkmale mit den kleinsten Assoziationsdistanzen in den nachfolgenden Schritten weiter verwendet. Pro Lokalisierungsschritt werden daher o. B. d. A.

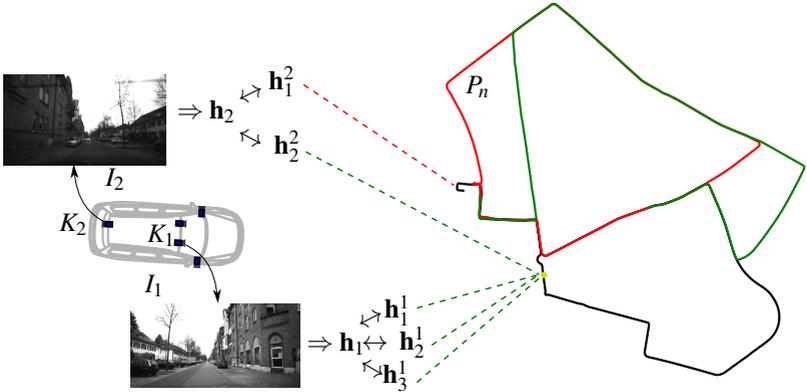


Abbildung 7.4: Bestimmung von Aufenthaltsortshypothesen im Rahmen der Ortswiedererkennung. Für das Bild I_1 wurden drei korrekte Hypothesen (Grüne Kanten) und für das Bild I_2 eine richtige und eine Ausreißerhypothese (Rote Kante) durch die Radiussuche im k-d-Suchbaum bestimmt. Der gelbe Punkt stellt den wahren Aufenthaltsort des Fahrzeugs zum Aufnahmezeitpunkt dar.

$m, 0 \leq m \leq |\Sigma_t| \cdot n_{\max}$ Kartenmerkmale zu den Merkmalen der zum Zeitpunkt t aufgezeichneten Bilder Σ_t assoziiert.

2. Bestimmung von Aufenthaltsortshypothesen

Jedes Kartenmerkmal h_i^k korrespondiert zu einer Kartenkamerapose $C_i^k \in C^L$ und jedes aktuell detektierte Merkmal h_k zu einer am Fahrzeug montierten Kamera K_k . Wurde h_k zu h_i^k assoziiert, stellt C_i^k eine Aufenthaltsortshypothese der Kamera K_k in der Karte dar. Damit kann auf die Kartenpose $P_i^k = C_i^k \cdot K_k^{-1} \in P_n$ unter Verwendung der extrinsischen Kamerakalibrierung K_k der Kamera K_k geschlossen werden, wobei P_n die kartierten Posen aus der Kartendatenbank M_n bezeichnet. Die Kartenpose P_i^k stellt somit die zum Merkmal h_i^k korrespondierende Aufenthaltsortshypothese des Fahrzeugs dar und ist im Idealfall die nächste zum wahren Aufenthaltsort des Fahrzeugs liegende Kartenpose. Da o. B. d. A. m Kartenmerkmale assoziiert wurden, ergeben sich auf diese Weise m Aufenthaltsortshypothesen P_1^k, \dots, P_m^k des Fahr-

zeugs. Abbildung 7.4 stellt die Bestimmung der Aufenthaltsortshypothesen schematisch dar.

3. Verfolgung von Aufenthaltsortshypothesen

In der praktischen Anwendung stellt in der Regel nur ein Teil der m geschätzten Aufenthaltsortshypothesen $\mathbf{P}_1^k, \dots, \mathbf{P}_m^k$ eine richtige Hypothese dar, wie in Abbildung 7.4 exemplarisch dargestellt ist. Das liegt zum einen daran, dass der verwendete Deskriptor empfindlich gegenüber Veränderungen in der Perspektive ist und sich das Fahrzeug i. A. nicht exakt auf einer der Kartenposen befindet. Zum anderen besteht eine hohe Wahrscheinlichkeit, dass durch periodische Umgebungsstrukturen oder durch ähnliche Szenen an verschiedenen Orten eine falsche Aufenthaltsortshypothese bestimmt wird. Daher werden im zweiten Schritt Hypothesenspuren verwaltet, durch die die Aufenthaltsortshypothesen über konsekutive Zeitschritte verfolgt werden. Eine Hypothesenspur $S = \{\mathbf{P}_{t-N}, \dots, \mathbf{P}_{t-J}\}_l$ ist dabei eine Sequenz von konsekutiven Aufenthaltsortshypothesen, wobei $t-N, \dots, t-J < t$ die Bestimmungszeitpunkte der einzelnen Hypothesen sind. Der Zeitpunkt $t-J$, der jüngsten in der Spur S vorkommenden Hypothese \mathbf{P}_{t-J} , stellt das Alter der Spur dar. Die Anzahl $|S|$ der Spurlypothesen der Spur S wird im Folgenden als Länge der Hypothesenspur bezeichnet.

Zur Erzeugung der Hypothesenspuren werden die m bestimmten Aufenthaltsortshypothesen $\mathbf{P}_1^k, \dots, \mathbf{P}_m^k$ zu den bereits bestehenden Spuren assoziiert. Dafür wird eine aktuelle Hypothese \mathbf{P}_i^k mit der jüngsten Hypothese \mathbf{P}_{t-J} einer bestehenden Spur S verglichen. Sind die Winkel- und Positionsdifferenz der Posendifferenz $\mathbf{D}_{t-J,i}^k = \mathbf{P}_{t-J}^{-1} \cdot \mathbf{P}_i^k$ kleiner als vorab definierte Schwellwerte, unterstützt die Hypothese \mathbf{P}_{t-J} die Spur S . In diesem Fall wird die Hypothese \mathbf{P}_i^k zu dieser Spur hinzugefügt. Dadurch verlängert sich zum einen die Spur S und zum anderen verringert sich ihr Alter. Unterstützt \mathbf{P}_{t-J} keine der bestehenden Spuren, wird eine neue Spur initialisiert. Spuren, die zu alt sind, werden aus der Verfolgung entfernt.

Längere Spuren von korrekt assoziierten Merkmalen repräsentieren somit eine Sequenz von Fahrzeugaufenthaltshypothesen, die näherungsweise eine valide Fahrzeugbewegung beschreiben. Ist dabei eine der verwalteten Spuren im

Vergleich zur Mehrheit aller Spuren deutlich länger, besteht daher eine hohe Wahrscheinlichkeit, dass die jüngste Hypothese dieser Spur den aktuellen ungefähren Aufenthaltsort des Fahrzeugs in der Karte darstellt. Die Entstehung von Spuren mit vergleichbarer Länge, die sich aus einer validen Folge von Fehlassoziationen zusammensetzen, ist demhingegen eher unwahrscheinlich.

Die beschriebene Methode zur Ortswiedererkennung durch zeitliche Verfolgung von Aufenthaltsortshypothesen hat insbesondere für MKSe mehrere Vorteile:

- Die Methode und Kartenstruktur abstrahiert von der Anzahl der verwendeten Kameras im MKS und von der Anzahl der Aufnahmesequenzen, aus denen die Karte besteht.
- Je mehr Bilder aus Σ_t verwendet werden, desto robuster ist die Hypothesenverfolgung gegenüber Ausreißern und desto schneller entwickeln sich längere richtige Hypothesenspuren. Die Wahrscheinlichkeit, dass sich lange falsche Hypothesenspuren ergeben, verringert sich durch Verwendung mehrerer Kameras, da es unwahrscheinlich ist, dass die Assoziation der Merkmale aller Bilder aus Σ_t in Übereinstimmung eine gleiche, jedoch falsche Fahrzeugaufenthaltshypothese ergeben.
- Da die Aufenthaltsortshypothesen der Kameras mittels der gegebenen Kalibrierung auf eine Aufenthaltsortshypothese des Fahrzeugs abgebildet werden, ergibt sich eine implizite Orientierungshypothese des Fahrzeugs. Wird beispielsweise das Merkmal \mathbf{h}_k einer rückwärts ausgerichteten Kamera mit einem Kartenmerkmal assoziiert, welches aus dem Bild einer in Kartenrichtung vorwärts ausgerichteten Kamera generiert wurde, ergibt sich durch die Transformation ins Fahrzeugreferenzsystem eine Aufenthaltsortshypothese mit entgegengesetzter Orientierung relativ zur korrespondierenden Kartenpose.

7.2.2 Experimente

Dieser Abschnitt beschreibt die durchgeführten Experimente zur Analyse der Ortswiedererkennung, die mit der Kamerakonfiguration MKS-A (siehe Abschnitt 3.1) durchgeführt wurden. Die folgende Auflistung beschreibt die Parameterisierung der Ortswiedererkennung bei den durchgeführten Experimenten:

- Es werden pro Kamerabild maximal $n_{\max} = 3$ Aufenthaltsortshypothesen assoziiert.
- Für die Assoziation von Aufenthaltsortshypothesen zu den bestehenden Spuren wird eine maximale Positionsdistanz von 3m und eine maximale Winkel-differenz von 10° verwendet. Durch diese Wahl toleriert die Methode Lokalisierungssprünge von bis zu 3m. Es wird daher auch dann eine Spur weiter verfolgt, wenn nicht die bestmögliche Kartenpose assoziiert wurde, sondern eine in der unmittelbaren Umgebung. Der Toleranzbereich ist dabei so gewählt, dass die metrische Lokalisierung auf Basis dieser Initialisierung zuverlässig eine Lokalisierungspose berechnen kann.
- Hypothesenspuren, die älter als 10 Zeitschritte sind, werden verworfen. Die Hypothesenverfolgung toleriert auf diese Weise kurze zeitliche Lücken in den Spuren.

Als Vergleichsverfahren wurde die topologische Lokalisierung [82] von Lategahn et al. verwendet. Das Verfahren verwendet den gleichen Deskriptor und führt eine Verfolgung von Bildhypothesen mittels dynamischer Programmierung durch. Es unterstützt jedoch lediglich die Verwendung einer einzelnen Kamera. Bei dem Vergleichsverfahren wird ebenfalls über einen einstellbaren Parameter entschieden, ob eine Schätzung des Aufenthaltsortes akzeptiert wird. Dieser wird in den Experimenten stets korrespondierend zur Länge $|S^{\max}|$ der längsten Hypothesenspur S^{\max} gewählt.

In den Experimenten werden die folgenden drei Konfigurationen verglichen:

1. **Hypothesenverfolgung mit zwei Kameras (HV-2)**

Die erste Versuchsreihe zeigt die Ortswiedererkennung mit dem vorgestellten Verfahren durch Aufenthaltsortshypothesenverfolgung mit einer vorwärts und einer rückwärts gerichteten Kamera.

2. **Hypothesenverfolgung mit einer Kamera (HV-1)**

Im zweiten Fall wird die vorgestellte Methode mit einer einzelnen Kamera evaluiert.

3. **Bildverfolgung mit dynamischer Programmierung (DYNP)**

Die dritte Versuchsreihe stellt zum Vergleich die Ergebnisse des Ansatzes von [82] dar.

Bei den Experimenten hat sich gezeigt, dass die Verwendung der seitlich ausgerichteten Kameras keine Verbesserungen erbracht und in manchen Fällen zu einer Verschlechterung der Ergebnisse geführt haben. Der Grund dafür ist, dass die Bilder der seitlichen Kameras häufig zu einem großen Teil parkende, entgegenkommende oder parallel fahrende Fahrzeuge abbilden.

Grundwahrheit

Die Evaluation der Ortswiedererkennung wird mit verschiedenen Evaluationsaufnahmesequenzen durchgeführt. Die Evaluationssequenzen wurden dafür zuvor mit der in Kapitel 6 beschriebenen Kartierungsmethode unter Hinzunahme von GNSS-Messungen an die Karte angeglichen. Alle Merkmale aus Bildern der Evaluationsaufnahmesequenzen wurden nach der Kartengenerierung wieder entfernt. Auf diese Weise verbleiben von den Evaluationssequenzen nur die kartierten Trajektorien in der Kartendatenbank, die als Grundwahrheit für die Evaluation verwendet werden. Bei der Kartengenerierung wird die Trajektorie einer Aufnahmesequenz zeitlich gefiltert und unter Hinzunahme von GNSS-Messungen an die Kartenposen angepasst. Zusätzlich werden die Kartentrajektorien automatisch auf Störungen und Verzerrungen untersucht. Daher wird angenommen, dass die Referenzierungsgenauigkeit der Kartentrajektorien ausreichend ist, um die Kartentrajektorien als Grundwahrheit für die Evaluation der Lokalisierung, die nur die aktuellen Bilder auswertet und keine zeitliche Filterung durchführt, zu verwenden.

Evaluationsmetrik

Zu jeder Kandidatenbildmenge Σ_t zu einem Zeitpunkt t aus einer Evaluationsaufnahmesequenz existiert eine Grundwahrheitspose \mathbf{P}_t^{gt} , die den Aufenthaltsort und die Orientierung des Fahrzeugs in der Karte zum Zeitpunkt t mit hoher Genauigkeit beschreibt. Die zu \mathbf{P}_t^{gt} korrespondierende nächste Kartenpose $\mathbf{P}_{\text{nn}}^{\text{gt}}$ ist die bestmögliche Schätzung, welche die Ortswiedererkennung erzielen kann. Da sich zu Evaluationszwecken nicht alle Abschnitte der Evaluationsaufnahmesequenzen mit der Karte überlappen, wird anhand des euklidischen Abstands zwischen \mathbf{P}_t^{gt} und $\mathbf{P}_{\text{nn}}^{\text{gt}}$ unterschieden, ob eine Ortswiedererkennung möglich ist. Ist der Abstand zwischen den Posen dabei größer als $d_{\text{max}} = 4\text{m}$, wird angenommen, dass keine Ortswiedererkennung möglich ist. Es wird immer die jüngste Hypothese der aktuell längsten Hypothesenspur S_t^{max} als Schätzung der Ortswiedererkennung \mathbf{P}_t^{pr} verwendet. Als

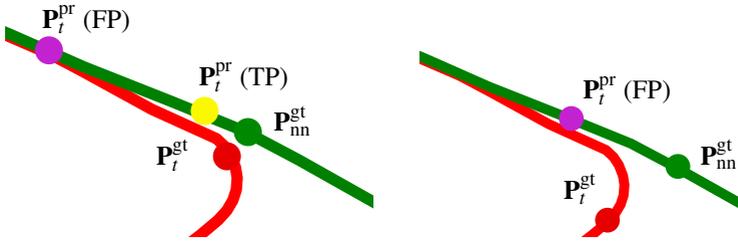


Abbildung 7.5: Klassifizierung der Schätzung \mathbf{P}_t^{pr} der Ortswiedererkennung in den Experimenten. Die in Rot dargestellte Trajektorie stellt die Grundwahrheit der Evaluationssequenz dar und die in Grün dargestellte Trajektorie zeigt die nächste Kartentrajektorie. Links: Der Abstand zwischen $\mathbf{P}_{\text{nn}}^{\text{gt}}$ und \mathbf{P}_t^{pr} ist kleiner als d_{max} , d. h. eine Ortswiedererkennung ist prinzipiell möglich. Rechts: Der Abstand zwischen $\mathbf{P}_{\text{nn}}^{\text{gt}}$ und \mathbf{P}_t^{pr} ist größer als d_{max} . Das kartierte Gebiet wird verlassen und es ist keine Ortswiedererkennung möglich.

Entscheidungskriterium, ob eine Schätzung herausgegeben wird, wurde im Rahmen der Experimente $|S_l^{\text{max}}| > 30$ gewählt. Darauf basierend ist eine Klassifizierung einer Schätzung \mathbf{P}_t^{pr} der Ortswiedererkennung möglich:

- **Korrekte Ortswiedererkennung (TP; engl. *true-positive*)**

Eine Schätzung \mathbf{P}_t^{pr} ist richtig, sofern eine Ortswiedererkennung möglich ist und die geschätzte Aufenthaltsortshypothese \mathbf{P}_t^{pr} in der Nähe von $\mathbf{P}_{\text{nn}}^{\text{gt}}$ liegt. Die maximale Distanz d_{max} zwischen \mathbf{P}_t^{pr} und $\mathbf{P}_{\text{nn}}^{\text{gt}}$ muss dafür kleiner als 2m sein. Der in Gelb dargestellte Punkt in Abbildung 7.5 (Links) stellt diesen Fall dar. Als zusätzliche Überprüfung wird mit dieser Schätzung die metrische Lokalisierung initialisiert. Nur wenn die metrische Lokalisierung erfolgreich eine Lokalisierung bestimmen konnte, wird die Schätzung \mathbf{P}_t^{pr} als korrekt gewertet.

- **Falsche Ortswiedererkennung (FP; engl. *false-positive*)**

Eine herausgegebene Schätzung \mathbf{P}_t^{pr} wird in zwei Fällen als falsch klassifiziert. Den ersten Fall stellt der in Lila dargestellte Punkt in Abbildung 7.5 (Links) dar. Hier ist eine Ortswiedererkennung möglich, jedoch ist die Distanz zwischen \mathbf{P}_t^{pr} und $\mathbf{P}_{\text{nn}}^{\text{gt}}$ größer als der Schwellwert $d_{\text{max}}^{\text{pr}}$. Den zweiten Fall zeigt der in Lila dargestellte Punkt in Abbildung 7.5 (Rechts). In diesem Fall

wurde eine Schätzung herausgegeben. Es ist jedoch keine Ortswiedererkennung möglich, da die nächste Kartenpose $\mathbf{P}_{nn}^{\text{gl}}$ zu weit von der Evaluationsstrajektorie entfernt ist.

- **Fehlende Ortswiedererkennung (FN; engl. *false-negative*)**

Wird keine Schätzung herausgegeben, obwohl eine Ortswiedererkennung möglich ist, wird dies als fehlende Schätzung gewertet.

- **Keine Ortswiedererkennung möglich (TN; engl. *true-negative*)**

Wenn sich das Fahrzeug in einem nicht kartierten Bereich bewegt, ist keine Ortswiedererkennung möglich. Wird in diesem Fall keine Schätzung herausgegeben, wird dies als korrekt zurückgewiesene Schätzung bewertet.

Zur Auswertung der verschiedenen Verfahren und Evaluationssequenzen wird die Ortswiedererkennung an 100 äquidistant ausgewählten Startpositionen entlang der jeweiligen Evaluationsstrecke gestartet und fortgeführt. Die in den folgenden Evaluationen dargestellten Kurven der TP-Rate und der FP-Rate [83] zeigen die entsprechende Rate nach n durchgeführten Verfolgungsschritten. Jeder Datenpunkt ergibt sich dabei stets aus der Mittelung über die 100 Ergebnisse für jede Startposition.

Städtische Fahrten

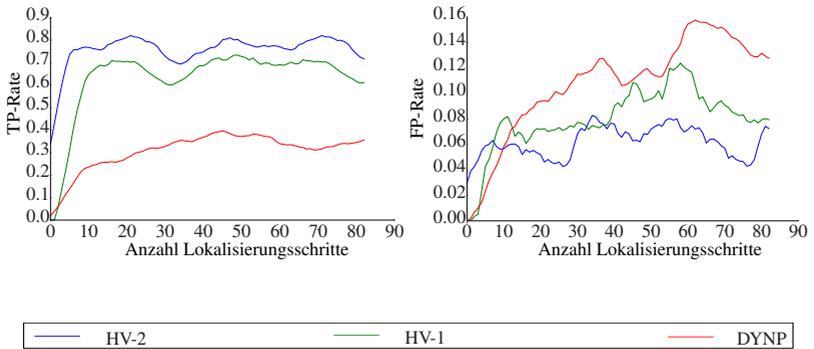
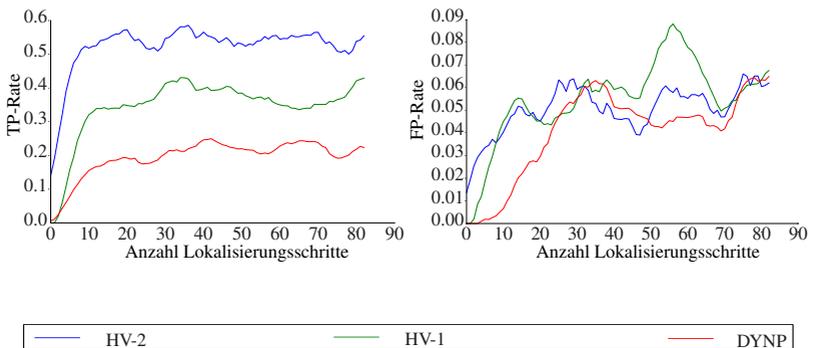
Für dieses Experiment werden drei Aufnahmesequenzen durch die Karlsruher Oststadt betrachtet, wobei die erste Sequenz M_{Stadt} als Karte verwendet wird und die anderen beiden Sequenzen E_{Stadt}^1 und E_{Stadt}^2 zur Evaluation genutzt werden. Die Kartensequenz M_{Stadt} wurde im November 2017, die erste Evaluationssequenz E_{Stadt}^1 im März und die zweite Evaluationssequenz E_{Stadt}^2 im Juni des Folgejahres aufgezeichnet. Die Umweltbedingungen bei der Evaluationssequenz E_{Stadt}^1 und der Kartensequenz M_{Stadt} sind ähnlich, jedoch ist die Sicht in der Evaluationssequenz E_{Stadt}^1 teilweise durch Regen und teilweise durch starke Sonneneinstrahlung gestört. Die Umweltbedingungen bei der Aufnahme der Evaluationssequenz E_{Stadt}^2 im Juni unterscheiden sich signifikant von den Bedingungen zum Aufnahmezeitpunkt der Kartensequenz M_{Stadt} aufgrund der jahreszeitbedingten Veränderung der Vegetation. Außerdem ist ein Teil der Strecke in der Sequenz E_{Stadt}^2 durch Baustellen belegt, sodass die kartierte Strecke in diesen Bereichen auf einer parallelen Umleitungsspur befahren wurde. Abbildung 7.6 zeigt Beispielbilder gleicher Orte in den verschiedenen Sequenzen.



Abbildung 7.6: Beispielhafte Bilder gleicher Orte aus der Kartensequenz M_{Stadt} (Links), der Evaluationssequenz E_{Stadt}^1 (Mitte) und der Evaluationssequenz E_{Stadt}^2 (Rechts).

Die Abbildungen 7.7 und 7.8 zeigen die erzielten TP- und FP-Raten der drei evaluierten Konfigurationen für die beiden Evaluationssequenzen E_{Stadt}^1 bzw. E_{Stadt}^2 . Es ist ersichtlich, dass die Verwendung mehrerer Kameras auf beiden Sequenzen deutlich schneller eine größere TP-Rate erreicht als die Vergleichsverfahren, bei denen nur eine Kamera verwendet wurde. Demhingegen ist die FP-Rate von Konfiguration HV-2 etwas geringer als die der Vergleichsverfahren. Der Grund für die größere TP-Rate von HV-2 nach weniger Lokalisierungsschritten ist, dass im Falle von richtig assoziierten Merkmalen mehr Hypothesen pro Lokalisierungsschritt zur richtigen Spur hinzugefügt werden und sich diese daher schneller verlängert. Die Konfiguration HV-1 erzielt ebenfalls höhere TP-Raten als das Vergleichsverfahren DYNP. Der Grund dafür ist zum einen die höhere Toleranz der Hypothesenverfolgung gegenüber sporadischen Verfolgungsaussetzern durch das Beibehalten von Spuren, die kurzzeitig nicht aktualisiert wurden, und zum anderen die höhere Toleranz gegenüber kleinen Lokalisierungssprüngen.

Beide Evaluationssequenzen stellen herausfordernde Szenarien dar. In der Anwendung wäre insbesondere durch die vegetationsbedingten und die baulichen Veränderungen eine höhere Aktualisierungsfrequenz der Karte vorteilhaft. Jedoch zeigen die

Abbildung 7.7: TP-Raten (Links) und FP-Raten (Rechts) für Evaluationssequenz E^1_{Stadt} .Abbildung 7.8: TP-Raten (Links) und FP-Raten (Rechts) für Evaluationssequenz E^2_{Stadt} .

Experimente, dass auch unter schwierigen Bedingungen noch hohe Verfügbarkeiten erreicht und durch die Hinzunahme weiterer Kameras deutliche Verbesserungen erzielt werden.

Campusfahrten

Im zweiten Experiment werden drei Fahrten auf dem Campus Süd des Karlsruher Instituts für Technologie evaluiert. Die Fahrten auf dem Campus wurden jeweils im



Abbildung 7.9: Beispielhafte Bilder gleicher Orte aus der Kartensequenz M_{Campus} (Links), der Evaluationssequenz E_{Campus}^1 (Mitte) und der Evaluationssequenz E_{Campus}^2 (Rechts). Für die Evaluationssequenzen sind Bilder aus der rückwärts gerichteten Kamera dargestellt.

März, Mai und Juni 2018 durchgeführt. Die erste Sequenz M_{Campus} stellt die Kartierungsfahrt dar. Bei der Aufnahme der beiden Evaluationssequenzen E_{Campus}^1 und E_{Campus}^2 wurde die Strecke in entgegengesetzter Richtung befahren. Dies steigert die Wahrscheinlichkeit für Fehlassoziationen, da sich die Projektionseigenschaften der rückwärts ausgerichteten Kamera von denen der nach vorne orientierten Kameras unterscheiden. Außerdem haben die Trajektorien der Evaluationssequenzen an vielen Stellen einen lateralen Versatz zur Kartentrajektorie, wodurch die korrekte Assoziation holistischer Merkmale zusätzlich erschwert wird. Abbildung 7.9 zeigt Beispielbilder aus den aufgezeichneten Sequenzen und Abbildung 7.10 die gefahrenen Trajektorien.

Prinzipiell ist die Ortswiedererkennung mit nur einer am Fahrzeug montierten Kamera bei entgegengesetzter Befahrungsrichtung nicht möglich, da die Kartenbilder nie aus der gleichen Perspektive gesehen werden. Daher wurden die holistischen Deskriptoren der vorwärtsschauenden Kamerabilder zur Kartengenerierung und die Bilder der rückwärts schauenden Kameras zur Evaluation der Konfigurationen HV-1

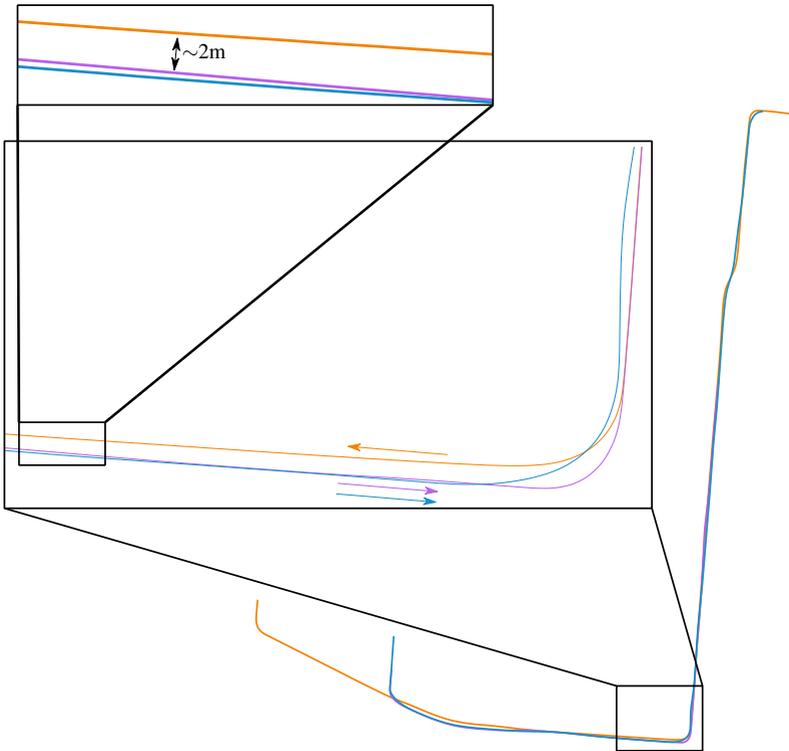


Abbildung 7.10: Darstellung der gefahrenen Trajektorien der Campussequenzen M_{Campus} (Orange), E_{Campus}^1 (Blau) und E_{Campus}^2 (Lila) in verschiedenen Detailstufen. Die Pfeile zeigen die Befahrungsrichtungen der einzelnen Trajektorien.

und DYNP verwendet. In den Experimenten mit Konfiguration HV-2 wurden sowohl die Bilder der rückwärts als auch die der vorwärts schauenden Kameras zur Kartengenerierung und zur Lokalisierung verwendet.

Die Abbildungen 7.11 und 7.12 zeigen die erzielten TP- und FP-Raten bei der Evaluation von Sequenz E_{Campus}^1 bzw. E_{Campus}^2 . Es zeigt sich eine verringerte Effizienz im Vergleich zu den Ergebnissen der Experimente der innerstädtischen Fahrten. Erklären lässt sich dies, wie zuvor beschrieben, durch die erschwerten Bedingungen bei

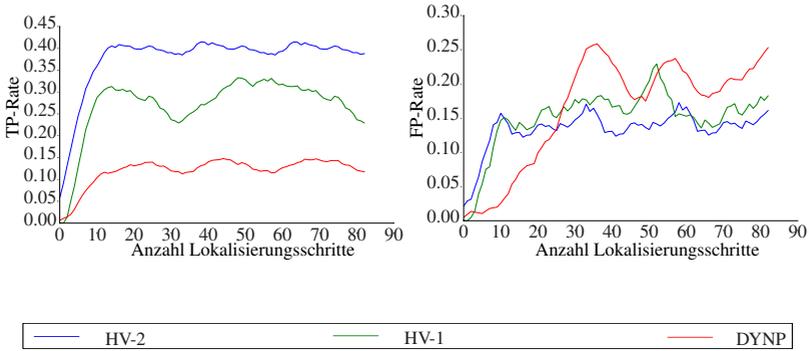


Abbildung 7.11: TP-Raten (Links) und FP-Raten (Rechts) für Evaluationssequenz E_{Campus}^1 .

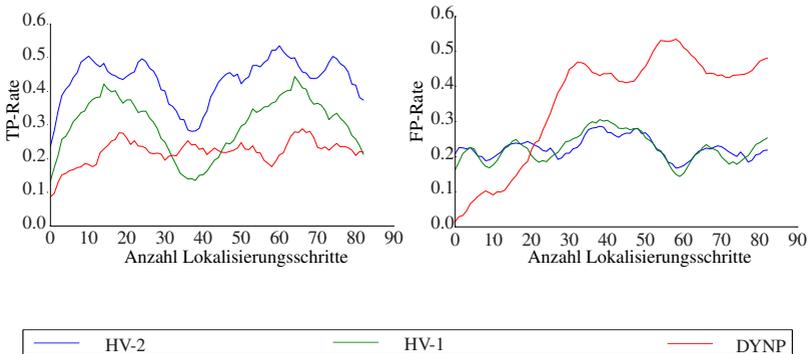


Abbildung 7.12: TP-Raten (Links) und FP-Raten (Rechts) für Evaluationssequenz E_{Campus}^2 .

der Merkmalsassoziation. Hier zeigt sich eine Beschränkung des DIRD-Deskriptors, der empfindlich gegenüber veränderten Szenendarstellungen durch Paralaxe und unterschiedlichen Abbildungsverhalten ist. Die Experimente zeigen dennoch eine deutliche Verbesserung der TP-Rate durch die Verwendung mehrerer Kameras. Außerdem belegen die Ergebnisse, dass eine effizientere und flexiblere Verwendung der kartierten Informationen durch ein MKS möglich ist. Durch die Hypothesenverfolgung abstrahiert das Verfahren von der Zuordnung der Kartenbilder zu den aktuell aufgezeichneten Bildern und ermöglicht so eine initiale Fahrzeugorientierungsschätzung, wofür Bilder aus mehreren Kameras mit unterschiedlicher Blickrichtung benötigt werden.

7.3 Metrische Lokalisierung

Die metrische Lokalisierung schätzt eine möglichst exakte kartenrelative Lokalisierungspose $\mathbf{P}_t^{\text{loc}}$ in sechs Freiheitsgraden. Voraussetzung dafür ist eine ungefähre Schätzung $\mathbf{P}_t^{\text{init}}$ der Fahrzeugpose, welche die metrische Lokalisierung zur Initialisierung benötigt. Darauf basierend werden die folgenden vier Schritte durchgeführt:

1. Auswahl von Kartenlandmarken und -merkmalen

Im ersten Schritt werden in Abhängigkeit der Initialisierungspose $\mathbf{P}_t^{\text{init}}$ Landmarken und Merkmale aus der Karte bereitgestellt, die in den Kartierungsfahrten aus einer ähnlichen Perspektive gesehen wurden wie die Merkmale aus den aktuellen Bildern Σ_t . Die entwickelte Kartenstruktur stellt für die Bereitstellung der Kartenlandmarken eine effiziente Zugriffsmethode bereit, die in Abschnitt 7.3.1 detailliert beschrieben wird.

2. Landmarkenassoziation

Die Assoziation von Kartenmerkmalen zu detektierten Merkmalen basiert auf dem Deskriptorvergleich gemäß Abschnitt 4.2. Der Suchraum wird dabei in Abhängigkeit der Initialisierungspose $\mathbf{P}_t^{\text{init}}$ gemäß Abschnitt 4.1.2 eingeschränkt.

3. Optimierung

Nach der Assoziation von Kartenlandmarken zu detektierten Merkmalen wird die finale Lokalisierungspose $\mathbf{P}_t^{\text{loc}}$ durch eine robuste Optimierung des Bündelblockausgleichsproblems 5.3 bestimmt. Das Bündelblockausgleichspro-

blem umfasst in diesem Fall nur die Lokalisierungspose $\mathbf{P}_t^{\text{loc}}$ als Posenparameter und die im vorherigen Schritt assoziierten Kartenlandmarken als Landmarkenparameter. Bei der Lösung des Problems werden alle Landmarkenparameter fixiert, sodass nur die Lokalisierungspose $\mathbf{P}_t^{\text{loc}}$ geschätzt wird, die mit $\mathbf{P}_t^{\text{init}}$ initialisiert wurde. Zur robusten Schätzung wird eine RANSAC-basierte Optimierung durchgeführt. Da die Lokalisierung unter Echtzeitbedingungen im Fahrzeug verwendet wird, endet die numerische Optimierung im Falle der Konvergenz oder nachdem eine maximale Berechnungszeit überschritten wurde. Auf diese Weise ist die Laufzeit der Optimierung stets konstant und nach oben begrenzt.

4. Validierung der Posenschätzung

Mittels einer Heuristik wird entschieden, ob der Lokalisierungsschätzung $\mathbf{P}_t^{\text{loc}}$ vertraut werden kann. Die Heuristik bewertet das Verhältnis r_{i0} von Unterstützerlandmarken zu Ausreißerlandmarken und die absolute Anzahl n_i von Unterstützerlandmarken, die im Rahmen des RANSAC-Schrittes identifiziert wurden. Eine Landmarke wird immer dann als Ausreißer gewertet, wenn der Rückprojektionsfehler im Kamerabild größer als ein maximaler Fehler e_{max} ist. Ist die Anzahl n_i der Unterstützer größer als eine untere Grenze n_i^{min} , wird die Lokalisierungsschätzung $\mathbf{P}_t^{\text{loc}}$ als zuverlässig eingestuft und das Verhältnis r_{i0} ignoriert. Das Kriterium basiert auf der Annahme, dass die Wahrscheinlichkeit für eine falsche Lokalisierung sinkt, je mehr Unterstützer diese Schätzung hat. In Fällen, in denen weniger als n_i^{min} Kartenlandmarken assoziiert werden, wird das Verhältnis r_{i0} als Entscheidungskriterium verwendet. Ist r_{i0} größer als ein Schwellwert $r_{i0}^{\text{min}} \in [0, 1]$, wird die Schätzung unabhängig von der absoluten Anzahl n_i der Unterstützer akzeptiert. In diesem Fall ist die Anzahl assoziierter Kartenlandmarken in der Regel kleiner als n_i^{min} , sodass die Anzahl n_i der Unterstützer diesen Schwellwert nicht erreichen kann. Die geringe Anzahl assoziierter Kartenlandmarken kann jedoch einer schwach strukturierten Umgebung oder einer gestörten Sicht geschuldet sein und ist daher kein eindeutiger Indikator für eine falsche Initialisierung. Empirisch hat sich gezeigt, dass falsche Lokalisierungsschätzungen durch diese Kriterien in der überwiegenden Mehrheit der Fälle identifiziert und zurückgewiesen werden.

7.3.1 Kartengestützte Merkmalsassoziation

Durch die kartengestützte Merkmalsassoziation werden die kartierten Landmarken effizient für die Assoziation zu den aktuell detektierten Merkmalen bereitgestellt. Die Kartenstruktur ermöglicht in Abhängigkeit der Initialisierungspose P_t^{init} eine flexible Bereitstellung jener Kartenlandmarken, die potenziell von den Kameras des MKS gesehen werden können. Zum einen werden vorrangig Landmarken bereitgestellt, die eine hohe Wiedererkennungswahrscheinlichkeit haben. Zum anderen werden die Landmarken unabhängig von der umgebenden Struktur räumlich gleichverteilt abgetastet. Eine möglichst gleichverteilte Abtastung von Beobachtungen wirkt sich positiv auf die Robustheit und Genauigkeit der Lokalisierung aus [5]. Außerdem wird so die Anzahl der zu assoziierenden Landmarken pro Ort reduziert und dadurch die Laufzeit zur Merkmalsassoziation reduziert. Die entwickelte Struktur der metrischen Lokalisierungskarte ermöglicht die beschriebene Auswahl nahezu ohne zusätzlichen Rechenaufwand während der Lokalisierung.

Struktur der metrischen Lokalisierungskarte

Die metrische Lokalisierungskarte ist eine zweidimensionale euklidische Rasterkarte, wobei angenommen ist, dass sich das Fahrzeug in einer Ebene bewegt. Die Rasterkarte speichert nur Zellen an Orten, an denen Kartenmerkmale beobachtet wurden. Wie in Kapitel 4 gezeigt wurde, ist die Assoziation zwischen einem Kartenmerkmal und einem detektierten Merkmal nur dann zuverlässig, wenn die beiden Merkmale aus einer ähnlichen Perspektive beobachtet werden. Daher sind die Landmarken sowie ihre zugehörigen Merkmale, wie in Abbildung 7.13 dargestellt, anhand der Beobachtungsposen ihrer zugehörigen Merkmale in den Zellen der Rasterkarte abgelegt. Die Beobachtungsposen sind hierbei durch die Kartenposen P_n aus der Kartendatenbank M_n bestimmt. Eine Landmarke wird i. A. von mehreren Zellen bereitgestellt, sofern sich die Beobachtungsposen ihrer zugehörigen Landmarken über mehrere Zellen erstrecken. Jede Zelle des Rasters kann separat von der Festplatte in den Speicher geladen werden.

Landmarkenbereitstellung

Zur Bereitstellung und Assoziation der Kartenlandmarken zu den in den aufgezeich-

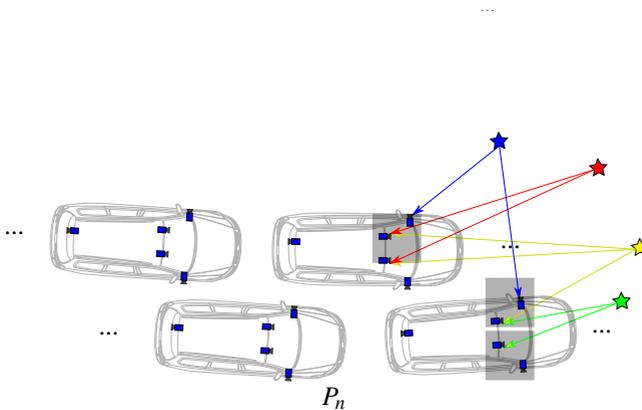


Abbildung 7.13: Landmarken (Sterne) werden anhand der Aufnahme-
posen ihrer zugehörigen Merkmale in die zweidimensionale Rasterkarte
einsortiert. Die Aufnahme-
posen ergeben sich durch die Kartenposen P_n . Die Pfeile zeigen, in welche
Zellen der Rasterkarte die Landmarken einsortiert werden.

neten Bildern Σ_t detektierten Merkmalen werden die folgenden Schritte durch-
geführt:

- **Auswahl geeigneter Zellen**

Durch die extrinsische Kalibrierung kann für jedes Bild $I_k \in \Sigma_t$ eine initiale Kameraposenschätzung $\mathbf{C}_k^{\text{init}} = \mathbf{P}_t^{\text{init}} \cdot \mathbf{K}_k$ in Abhängigkeit der Initialisierungs-
pose $\mathbf{P}_t^{\text{init}}$ berechnet werden. Zur Assoziation von den in Bild I_k detektierten
Merkmalen können mit der Kameraposenschätzung $\mathbf{C}_k^{\text{init}}$ geeignete Karten-
landmarken und Kartenmerkmale aus der Lokalisierungskarte ausgewählt
werden. Für das Bild I_k wird daher die Zelle aus der Lokalisierungskarte gela-
den, in der die zu I_k korrespondierende Initialisierungskamerapose $\mathbf{C}_k^{\text{init}}$ liegt.
Liegt $\mathbf{C}_k^{\text{init}}$ in der Nähe des Zellenrandes, werden zusätzlich die benachbarten
Zellen geladen. Abbildung 7.14 (Links) stellt diesen Schritt schematisch dar.

- **Landmarkenabtastung**

Wie in Abbildung 7.14 (Rechts) gezeigt, sind die kartierten Landmarken und
ihre zugehörigen Merkmale einer Zelle der Lokalisierungskarte anhand der
Landmarkenposition in ein zweidimensionales polares Raster einsortiert. Das
Einfügen der Landmarken in die polare Rasterstruktur pro Zelle ist zuvor bei

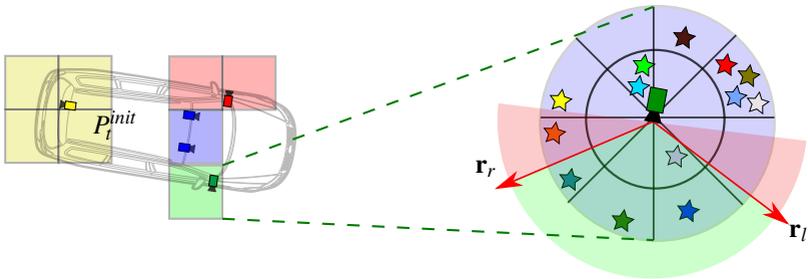


Abbildung 7.14: Links: Ausgewählte Zellen der metrischen Lokalisierungskarte für verschiedene Kameras. Die Farbe kodiert die Zuordnung der Kameras zu den aus der Lokalisierungskarte geladenen Zellen in Abhängigkeit der Initialisierungspose $\mathbf{P}_t^{\text{init}}$. Rechts: Zweidimensionale polare Rasterkartenstruktur innerhalb einer Zelle. In Abhängigkeit des Sichtkegels (Grün), der durch die Sichtstrahlen \mathbf{r}_l und \mathbf{r}_r definiert ist, werden polare Regionen für die Merkmalsassoziation ausgewählt. Da $\mathbf{P}_t^{\text{init}}$ unsicher ist, wird der Sichtkegel um einen Toleranzbereich (Rot) erweitert.

der Erstellung der Lokalisierungskarte durchgeführt worden. Die Winkel- und Distanzauflösung des polaren Rasters sind wählbare Parameter. Die Landmarken innerhalb einer Region des polaren Rasters sind anhand eines Gütekriteriums sortiert abgelegt. Dieses Gütekriterium bewertet die Wiedererkennungswahrscheinlichkeit einer Landmarke und wird im folgenden Kapitel 8 näher beschrieben.

Unter Verwendung der Initialisierungskamerapose $\mathbf{C}_k^{\text{init}}$ und der intrinschen Kamerakalibrierung kann der Sichtkegel der Kamera K_k im Kartenkoordinatensystem bestimmt werden. Dazu werden die Sichtstrahlen \mathbf{r}_l , \mathbf{r}_r des linken und rechten Randpixels von I_k ins Kartenkoordinatensystem transformiert. Darauf basierend werden nur polare Regionen innerhalb des geschätzten Sichtkegels für die nachfolgende Merkmalsassoziation ausgewählt. Da von einer ungenauen Initialposenschätzung $\mathbf{P}_t^{\text{init}}$ ausgegangen wird, wird der Sichtkegel zusätzlich um einen Toleranzbereich erweitert. Der grüne Sichtkegel in Abbildung 7.14 (Rechts) stellt dies exemplarisch dar. Der Toleranzbereich ist zusätzlich in Rot dargestellt.

- **Landmarkenassoziation**

Zur Landmarkenassoziation werden die in Abhängigkeit der Sichtkegelschätzung ausgewählten polaren Regionen sequentiell durchlaufen. Die Landmarken aus einer einzelnen polaren Region werden zur Merkmalsassoziation sequentiell gemäß ihrer Einsortierungsreihenfolge durchlaufen und zu den in Bild I_k detektierten Merkmalen assoziiert. Wurden n_{\max} Landmarken einer polaren Region erfolgreich zu den in I_k detektierten Merkmalen assoziiert oder wird eine maximale Anzahl nicht erfolgreicher Assoziationsversuche überschritten, wird die nächste Region betrachtet. Aufgrund der Diskretisierung des polaren Rasters werden dabei alle Raumrichtungen und Distanzen innerhalb des Sichtkegels gleichverteilt abgetastet und die Landmarken mit der höchsten Wiedererkennungswahrscheinlichkeit pro Region zuerst ausgewählt. Durch die zusätzliche Beschränkung der Assoziationsversuche pro polarer Region ergibt sich außerdem eine obere Grenze des Assoziationsaufwands pro bereitgestellter Zelle aus der Lokalisierungsrasterkarte.

7.3.2 Experimente

Dieser Abschnitt beschreibt die durchgeführten Experimente zur Evaluation der vorgestellten Lokalisierungsmethode.

Parameterisierung

In der praktischen Anwendung und in den Experimenten hat eine Zellgröße der Lokalisierungsrasterkarte von $1\text{m} \times 1\text{m}$ und eine Winkel- und Distanzauflösung des polaren Rasters von 9° bzw. 0.5m zufriedenstellende Ergebnisse erzielt. Abbildung 7.15 zeigt beispielhaft die polare Struktur einer einzelnen Zelle bei dieser Parameterisierung.

Als Entscheidungsschwellwert des RANSAC-Algorithmus zur robusten Schätzung der Lokalisierungspose hat sich ein maximaler Rückprojektionsfehler von 6 Pixel als geeignet erwiesen. Der Parameter r_{i0} der beschriebenen Heuristik zur Einschätzung der Zuverlässigkeit einer Lokalisierungsposenschätzung wurde in allen Experimenten zu 0.4 gewählt. Da der zweite Parameter n_i^{\min} der Entscheidungsheuristik eine untere Schranke für die absolute Anzahl der Unterstützerlandmarken darstellt, wird dieser in Abhängigkeit der Anzahl und der Art der verwendeten Kameras ge-

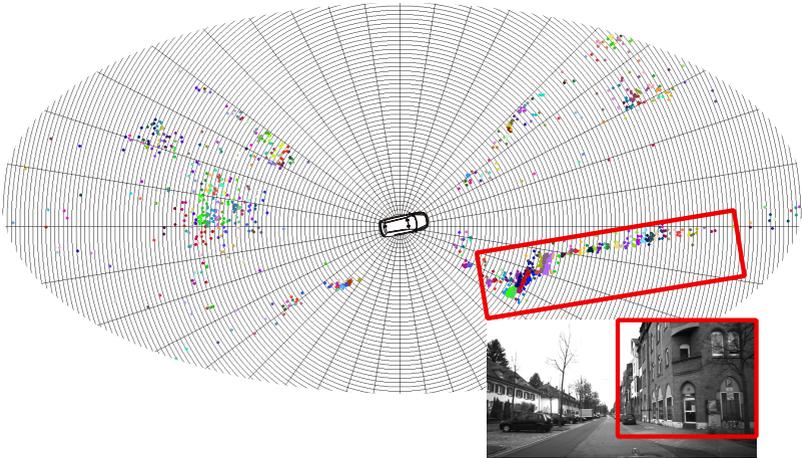


Abbildung 7.15: Darstellung der Zellenstruktur. Die unterschiedlichen eingefärbten Punkte stellen die Landmarken dar. Die Farben der Landmarken kodieren die polare Region, in der die Landmarken liegen. Im rot hervorgehoben Bereich lässt sich die Struktur einer kartierten Hausfassade gut erkennen.

wählt. In den Experimenten wurde $n_i^{\min} = |\Sigma| \times 70$ gewählt, wobei $|\Sigma|$ die Anzahl der verwendeten Kameras in dem entsprechenden Experiment darstellt. Mit dieser Parameterisierung wird je nach Anzahl der verwendeten Kameras eine echtzeitfähige Lokalisierung mit 8-12Hz erzielt.

Evaluationskonfigurationen

Zur Demonstration der Verbesserung der Lokalisierung durch die Verwendung mehrerer Kameras werden drei verschiedene Kamerakonfigurationen basierend auf der Sensorkonfiguration MKS-A (siehe Abschnitt 3.1) verglichen:

- **MKS-A1:** Bei der Konfiguration MKS-A1 wird nur die linke der beiden vorwärts gerichteten Kameras verwendet.
- **MKS-A2:** Bei der Konfiguration MKS-A2 werden die linke der beiden vorwärts gerichteten Kameras und die rückwärts gerichtete Kamera verwendet.



Abbildung 7.16: Projektionen von assoziierten Landmarken in das Bild einer vorwärts gerichteten Kamera (Links) und einer Seitenkamera (Rechts).

- **MKS-A3:** Bei der Konfiguration MKS-A3 werden die beiden vorwärts gerichteten Kameras und die rückwärts gerichtete Kamera verwendet.

Die Hinzunahme der seitlichen Kameras hat in allen Fällen keine nennenswerte Veränderung der Ergebnisse erzielt. Die niedrige Einbauposition und die zur Fahrtrichtung orthogonale Orientierung der Seitenkameras führen bereits bei geringen Fahrgeschwindigkeiten zu einem großen optischen Fluss und Bewegungsunschärfe in den aufgezeichneten Bildern. Daher ist der Anteil der kartierten Landmarken mit Merkmalskorrespondenzen aus den Bildern der seitlichen Kameras in der Lokalisierungskarte verhältnismäßig gering. Der überwiegende Anteil der kartierten Landmarken ist zu Merkmalen aus den Bildern der vorwärts bzw. rückwärts gerichteten Kameras assoziiert. Abbildung 7.16 zeigt beispielhaft Projektionen von Landmarken in die Bilder aus einer der vorwärts gerichteten Kameras und einer seitwärts gerichteten Kamera. Die Abbildung illustriert das unausgeglichene Verhältnis von assoziierten Landmarken in den vorderen Kameras und den Seitenkameras.

Evaluationsmetrik

Wie bei den Experimenten zur Ortswiedererkennung in Abschnitt 7.2.2 wurde eine Karte erstellt, die sowohl Evaluationssequenzen als auch Kartensequenzen umfasst. Nach der Erstellung der Karte wurden alle Merkmale, die aus den Evaluationssequenzen entspringen, aus der Kartendatenbank entfernt, sodass von den Evaluationssequenzen ausschließlich die kartierten Trajektorien in der Kartendatenbank verbleiben.

Darauf basierend kann eine geschätzte Lokalisierungspose $\mathbf{P}_t^{\text{loc}}$ gegen ihre korre-

spondierende Grundwahrheitspose \mathbf{P}_i^{gt} verglichen werden. Die Genauigkeit der Lokalisierung wird durch den Positionsfehler t_{err} und den Orientierungsfehler r_{err} quantifiziert. Der Positionsfehler t_{err} ergibt sich durch die Norm der Translation $\mathbf{t}_{\mathbf{D}_{\text{err}}}$ der Posendifferenz $\mathbf{D}_{\text{err}} = (\mathbf{P}_i^{\text{loc}})^{-1} \cdot \mathbf{P}_i^{\text{gt}}$ und der Orientierungsfehler r_{err} durch den Betrag des Rotationswinkels von \mathbf{D}_{err} .

Für die Analyse der Lokalisierungsverfügbarkeit wird die Anzahl erfolgreicher Lokalisierungsschritte mit der Anzahl der möglichen Lokalisierungen verglichen. Dazu wird geprüft, ob der Positionsfehler $t_{\text{err}} > 1.5\text{m}$ und der Orientierungsfehler $r_{\text{err}} > 3^\circ$ ist. In den durchgeführten Experimenten hat sich gezeigt, dass die Entscheidungsheuristik mit der gewählten Parameterisierung in über 95% der Fälle mit diesen Kriterien übereinstimmt. Die Anzahl der möglichen Lokalisierungen ergibt sich wie bei der Evaluation der Ortswiedererkennung durch den Abstand von \mathbf{P}_i^{gt} zur nächsten Kartenpose $\mathbf{P}_{\text{gt}}^{\text{nn}}$. Ist der Abstand der Posen kleiner als 3m , wird angenommen, dass sich das Fahrzeug im kartierten Bereich bewegt und eine Lokalisierung möglich ist.

Städtische Fahrten

Zur Evaluation der Lokalisierung in vor- und innerstädtischen Regionen werden die gleichen Sequenzen verwendet, die auch bei den Experimenten zur Ortswiedererkennung in Abschnitt 7.2.2 verwendet wurden. Aus der Sequenz M_{Stadt} wurde die Karte erstellt und die Sequenzen E_{Stadt}^1 und E_{Stadt}^2 werden zur Evaluation verwendet. Tabelle 7.1 zeigt den mittleren und den maximalen Lokalisierungsfehler sowie die Lokalisierungsverfügbarkeit der drei Konfigurationen auf den Evaluationssequenzen. Die Tabelle zeigt, dass die Unterschiede der Fehler der drei Konfigurationen gering sind, wobei die Ergebnisse der Konfigurationen MKS-A2 und MKS-A3 etwas geringere Fehlerwerte zeigen. Dies entspricht der Erwartung, dass die Lokalisierung robuster und genauer bei der Verwendung mehrerer Kameras und der damit verbundenen Vergrößerung des Sichtfeldes wird. Die Tabelle zeigt außerdem den Effekt der Kartenalterung, da die Genauigkeit und insbesondere die Verfügbarkeit der Lokalisierung auf der Evaluationssequenz E_{Stadt}^2 deutlich geringer ist als auf der Sequenz E_{Stadt}^1 . Die Verschlechterung der Verfügbarkeit ist deutlich stärker als der Anstieg des Rotations- bzw. Positionsfehlers. Das zeigt, dass die Entscheidungsheuristik zuverlässig fehlerhafte Lokalisierungsschätzungen detektiert. Die Ergebnisse zeigen außerdem, dass die Verwendung von mindestens zwei Kameras zu einer deutlichen

Sequenz / Konfiguration	$r_{\text{err}}^{\text{avg}} / r_{\text{err}}^{\text{max}} [^\circ]$	$t_{\text{err}}^{\text{avg}} / t_{\text{err}}^{\text{max}} [\text{m}]$	Verfügbarkeit
E_{Stadt}^1 / MKS-A1	0.09 / 1.24	0.11 / 0.98	71%
E_{Stadt}^1 / MKS-A2	0.06 / 1.24	0.08 / 1.14	94%
E_{Stadt}^1 / MKS-A3	0.06 / 1.32	0.07 / 0.92	98%
E_{Stadt}^2 / MKS-A1	0.15 / 2.09	0.14 / 1.04	48%
E_{Stadt}^2 / MKS-A2	0.16 / 1.49	0.11 / 0.91	77%
E_{Stadt}^2 / MKS-A3	0.15 / 1.81	0.1 / 0.87	86%

Tabelle 7.1: Tabellarische Darstellung des mittleren / maximalen Rotationsfehlers $r_{\text{err}}^{\text{avg}} / r_{\text{err}}^{\text{max}}$, des mittleren / maximalen Positionsfehlers $t_{\text{err}}^{\text{avg}} / t_{\text{err}}^{\text{max}}$ und die erzielte Lokalisierungsverfügbarkeit der drei Kamerakonfiguration MKS-A1, MKS-A2, MKS-A3 auf den beiden Testsequenzen E_{Stadt}^1 und E_{Stadt}^2 .

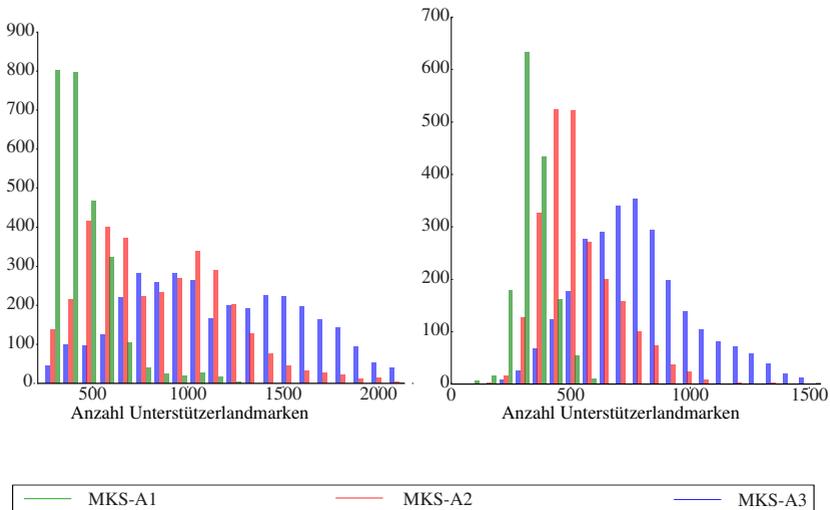


Abbildung 7.17: Histogramme der Unterstützerlandmarkenanzahlen für die drei Testkonfigurationen MKS-A1 (Grün), MKS-A2 (Rot) und MKS-A3 (Blau) bei der Evaluation auf den beiden Evaluationssequenzen E_{Stadt}^1 (Links) und E_{Stadt}^2 (Rechts).

Steigerung der Verfügbarkeit führt. Die Verteilungen in Abbildung 7.17 zeigen eine höhere Anzahl der Unterstützerlandmarken bei den Kamerakonfigurationen MKS-A2 und MKS-A3 gegenüber der Konfiguration MKS-A1, wodurch sich die Steigerung der Verfügbarkeit hauptsächlich erklären lässt.

Eine Verbesserung der Lokalisierung durch die Hinzunahme zusätzlicher Kameras zeigt sich insbesondere bei der Durchfahrt durch schwächer strukturierte Gebiete und bei Befahrung von nicht kartierten parallel verlaufenden Nachbarfahrspuren. Die Abbildungen 7.18 und 7.19 zeigen jeweils einen Streckenabschnitt aus der Evaluationssequenz E_{Stadt}^1 bzw. E_{Stadt}^2 , an der sich das Fahrzeug auf der Parallelspur zur kartierten Fahrspur bewegt. Die im oberen Bereich dargestellten Trajektorien zeigen die Verläufe der gefahrenen Trajektorien und im unteren Bereich sind die evaluierten Lokalisierungsfehler über die gefahrene Distanz aufgetragen. Die in Grün hervorgehobenen Zeitbereiche zeigen Abschnitte, in denen die Lokalisierung mit Konfiguration MKS-A1 nicht mehr verfügbar ist, wohingegen mit den anderen beiden Konfigurationen weiterhin eine gemäß der Entscheidungsheuristik zuverlässige Lokalisierungsschätzung erzielt wird. Der Lokalisierungsfehler steigt bei den Konfigurationen MKS-A2 und MKS-A3 in diesem Bereich an, da in beiden Fällen weniger Landmarken assoziiert werden als bei der Fahrt auf der kartierten Spur, wobei mit Konfiguration MKS-A3 der Lokalisierungsfehler im Mittel geringer ist. Insbesondere bei dem dargestellten Szenario in Abbildung 7.19 aus der Evaluationssequenz E_{Stadt}^2 zeigt sich stellenweise ein deutlicher Unterschied der Fehlerwerte zwischen Konfiguration MKS-A2 und MKS-A3. Auch wenn sich die Perspektive der linken und rechten Kamera nur wenig voneinander unterscheiden, wirkt sich die Hinzunahme der rechten vorderen Kamera positiv auf die Genauigkeit der Lokalisierung aus.

Campusfahrten

Zur Analyse der Lokalisierungsgenauigkeit und -verfügbarkeit bei entgegengesetzter Befahrungsrichtung wird die metrische Lokalisierung mit den Campussequenzen durchgeführt, die ebenfalls in den Experimenten zur Ortswiedererkennung in Abschnitt 7.2.2 verwendet wurden. Tabelle 7.1 zeigt den mittleren und den maximalen Lokalisierungsfehler sowie die Lokalisierungsverfügbarkeit der drei Konfigurationen auf den Sequenzen E_{Campus}^1 und E_{Campus}^2 . Da die Trajektorie der Kartensequenz M_{Campus} bei der Aufnahme der Evaluationssequenzen E_{Campus}^1 und E_{Campus}^2 in ent-

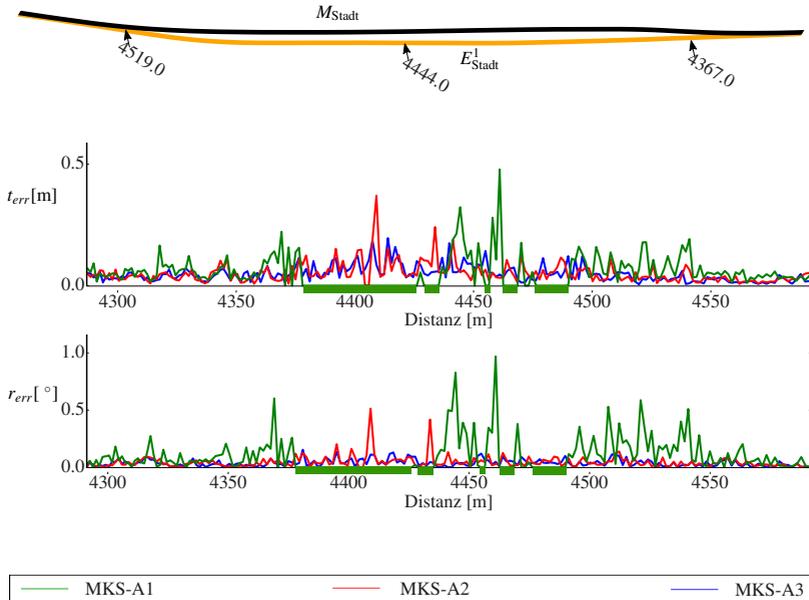


Abbildung 7.18: Oben: Kartentrajektorie (Schwarz) und Grundwahrheitstrajektorie (Gelb) der Evaluationsequenz E_{Stadt}^1 in einem Bereich, in dem die zur Kartentrajektorie parallele Nebenspur befahren wird. Mitte: Positionsfehler t_{err} aufgetragen über die gefahrene Distanz in diesem Bereich. Unten: Rotationsfehler r_{err} aufgetragen über die gefahrene Distanz in diesem Bereich. Hervorgehoben sind die Bereiche, in denen mit Konfiguration MKS-A1 keine Lokalisierung herausgegeben wurde, da keine zuverlässige Lokalisierungspose geschätzt werden konnte.

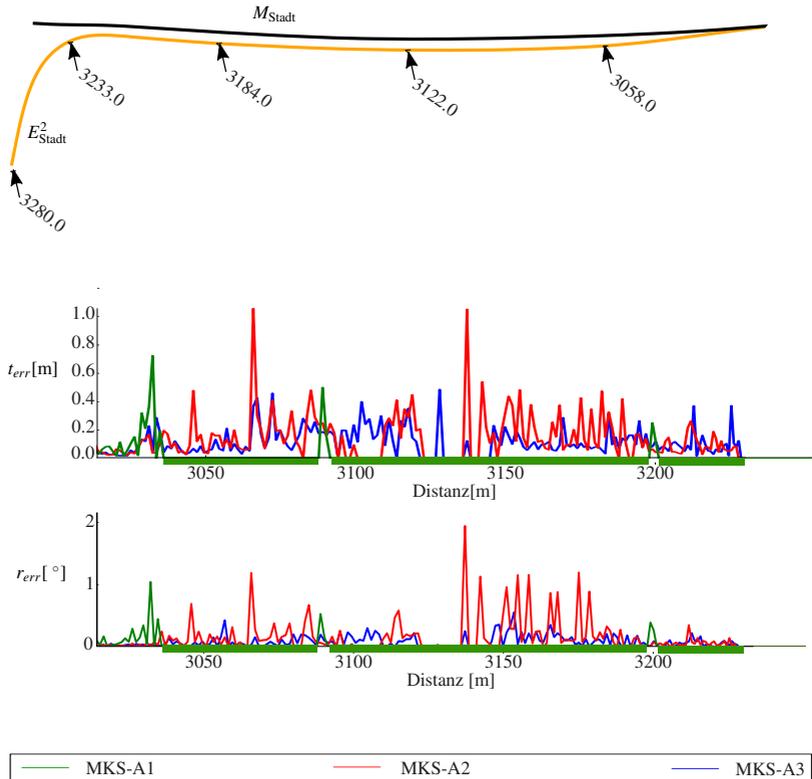


Abbildung 7.19: Oben: Kartentrajektorie (Schwarz) und Grundwahrheitstrajektorie (Gelb) der Evaluationsequenz E_{Stadt}^2 in einem Bereich, in dem eine zur Kartentrajektorie parallele Nebenspur befahren wird und danach der kartierte Bereich verlassen wird. Mitte: Positionsfehler t_{err} aufgetragen über die gefahrene Distanz in diesem Bereich. Unten: Rotationsfehler r_{err} aufgetragen über die gefahrene Distanz in diesem Bereich. Hervorgehoben sind die Bereiche, in denen mit Konfiguration MKS-A1 keine Lokalisierung herausgegeben wurde, da keine zuverlässige Lokalisierungspose geschätzt werden konnte.

Sequenz / Konfiguration	$r_{\text{err}}^{\text{avg}} / r_{\text{err}}^{\text{max}}$ [°]	$t_{\text{err}}^{\text{avg}} / t_{\text{err}}^{\text{max}}$ [m]	Verfügbarkeit
E_{Campus}^1 / MKS-A1	0.39 / 2.27	0.19 / 1.24	15%
E_{Campus}^1 / MKS-A2	0.23 / 1.89	0.11 / 0.92	41%
E_{Campus}^1 / MKS-A3	0.27 / 1.92	0.16 / 1.43	58%
E_{Campus}^2 / MKS-A1	0.39 / 2.52	0.16 / 0.98	11%
E_{Campus}^2 / MKS-A2	0.31 / 2.29	0.14 / 1.43	29%
E_{Campus}^2 / MKS-A3	0.32 / 2.35	0.16 / 1.45	42%

Tabelle 7.2: Tabellarische Darstellung des mittleren / maximalen Rotationsfehlers $r_{\text{err}}^{\text{avg}} / r_{\text{err}}^{\text{max}}$, des mittleren / maximalen Positionsfehlers $t_{\text{err}}^{\text{avg}} / t_{\text{err}}^{\text{max}}$ und die erzielte Lokalisierungsverfügbarkeit der drei Kamerakonfiguration MKS-A1, MKS-A2, MKS-A3 auf den beiden Evaluationsequenzen E_{Campus}^1 und E_{Campus}^2 .

gegengesetzter Richtung befahren wurde, sind die Bedingungen für die Merkmalsassoziation, wie bereits in Abschnitt 7.2.2 beschrieben, wesentlich erschwert. Dies spiegelt sich in den größeren Lokalisierungsfehlern und insbesondere auch in den wesentlich geringen Lokalisierungsverfügbarkeiten im Vergleich zu den städtischen Sequenzen wieder. Die Lokalisierung mit nur einer Kamera (MKS-A1) ermöglicht prinzipiell keine robuste Lokalisierung mehr. Bei der Hinzunahme weiterer Kameras (MKS-A2, MKS-A3) zeigt sich eine deutliche Steigerung der Verfügbarkeit. In Übereinstimmung mit den Experimenten der städtischen Sequenzen belegen die Experimente, dass die Verwendung eines MKS insbesondere unter schwierigen Bedingungen eine Verbesserung der Lokalisierungsgüte erbringt.

Experimente zur Befahrungsrichtung

Zur Demonstration der Lokalisierung bei einer Befahrung der kartierten Strecke in unterschiedliche Richtungen wurde eine meanderförmige Aufnahmesequenz M auf einer freien Fläche in einem verkehrsberuhigten Bereich aufgezeichnet und damit eine Lokalisierungskarte berechnet. Die rote Trajektorie in Abbildung 7.20 stellt die Kartentrajektorie dar. Danach wurden in einem zeitlichen Abstand von mehreren Tagen zwei Evaluationsequenzen E_m und E_r aufgezeichnet. Wie in Abbildung 7.20 dargestellt, wurde die Karte bei der Aufnahme von Evaluationssequenz E_m meanderförmig und orthogonal zur Kartentrajektorie befahren. Bei der Aufnahme von

Sequenz	$t_{\text{err}}^{\text{avg}} [^\circ]$	$t_{\text{err}}^{\text{avg}} [\text{m}]$	Verfügbarkeit
E_m	0.8	0.23	72%
E_r	0.92	0.28	70%

Tabelle 7.3: Mittlerer Rotations- und Positionsfehler sowie die erzielte Lokalisierungsverfügbarkeit der Evaluationssequenzen M und R .

Evaluationssequenz E_r wurde das kartierte Gebiet willkürlich und zufällig aus verschiedenen Richtungen befahren. Bei diesen Experimenten wurden alle Kameras der Kamerakonfiguration MKS-A verwendet, wodurch die Lokalisierungskarte Merkmale bereitstellt, welche die kartierten Landmarken aus nahezu allen Raumrichtungen abbilden. Dadurch ist eine Assoziation von in den Seitenkameras detektierten Merkmalen zu Landmarken möglich, die bei der Kartengenerierung aus den nach vorne bzw. nach hinten gerichteten Kameras beobachtet wurden. Somit kann eine Lokalisierung auch bei orthogonaler Befahrung der Karte ermöglicht werden.

Die gelb markierten Bereiche in Abbildung 7.20 zeigen die Orte, an denen sich die Evaluationssequenzen E_m und E_r mit der Kartentrajektorie M überschneiden. Bei allen Überschneidungspunkten konnte sich das Fahrzeug mehrfach aufeinanderfolgend gemäß der vorgestellten Entscheidungsheuristik erfolgreich lokalisieren. Die fortlaufende Initialisierung gemäß Abschnitt 7.1 wurde bei diesen Experimenten durch eine Odometrieschätzung durch SVO unterstützt. Auf diese Weise konnte eine zuverlässige Initialisierung bis zum nächsten Überkreuzungsbereich der Kartentrajektorie ermöglicht werden. Das Verhältnis der Unterstützer zu Ausreißern lag im Mittel bei $r_{i0} = 0.32$ bei einer durchschnittlichen Anzahl von $n_i = 120$ Unterstützerlandmarken. Die geringe Unterstützerlandmarkenanzahl erklärt sich hauptsächlich durch die unterschiedlichen Einbaupositionen und das unterschiedliche Projektionsverhalten der Seitenkameras im Vergleich zu den vorderen Kameras bzw. der hinteren Kamera.

Tabelle 7.3 zeigt den mittleren Positions- und Rotationsfehler aller herausgegebenen Lokalisierungsschätzungen. Die Ergebnisse zeigen, dass auch bei einer beliebig orientierten Durchfahrt durch die Karte eine hohe Verfügbarkeit und Genauigkeit mit der vorgestellten Lokalisierungsmethode erzielt wird. Die Experimente demonstrieren

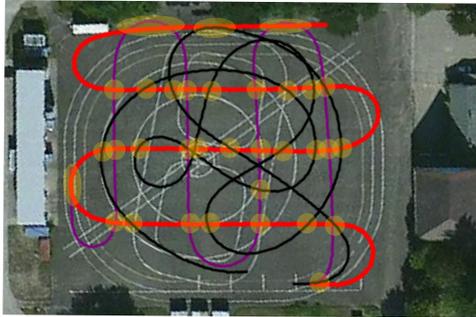


Abbildung 7.20: Überlagerung der Kartentrajektorie M (Rot) und der Trajektorien der Evaluationssequenzen E_m (Lila) und E_r (Schwarz) mit dem Luftbild der Freifläche. Luftbilder: ©Microsoft/Here

ren außerdem die Flexibilität der vorgestellten Kartenstruktur, wodurch das umfassende Sichtfeld eines MKS effizient ausgenutzt wird.

8 Merkmalsselektion

Ein wesentlicher Prozessschritt zur Erzeugung der Lokalisierungskarte M_n^L ist die Selektion von Landmarken und Merkmalen aus der Kartendatenbank M_n , die in der Lokalisierungskarte M_n^L gespeichert werden sollen und somit der Lokalisierung zur Verfügung stehen. Wie in Abschnitt 7.1 beschrieben, umfasst die Lokalisierungskarte M_n^L zwei Kartenschichten, wobei jeweils eine für die Ortswiedererkennung und eine für die metrische Lokalisierung benötigt wird. Für beide Kartenschichten wird eine Merkmalsselektion durchgeführt. Da beide Kartenschichten unterschiedlich verwendet werden und unterschiedliche Merkmalstypen bereitstellen, unterscheiden sich dementsprechend die Auswahlkriterien voneinander.

Die Palette der Möglichkeiten zur Auswahl von Merkmalen ist breit gefächert. Eine Selektion kann anhand verschiedener Gesichtspunkte vorgenommen werden. In Betracht dafür kommen

- **geometrische Kriterien**, wie beispielsweise die Sichtbarkeit der Landmarken aus verschiedenen Perspektiven oder die Ungenauigkeit der Landmarkenschätzung.
- **merkmalsbasierte Kriterien**, wie beispielsweise die lokale Eindeutigkeit einer Landmarke anhand ihrer Merkmalsdeskriptoren.
- **statistische Kriterien**, wie beispielsweise die Wiedererkennungswahrscheinlichkeit von Landmarken in zukünftigen Befahrungen des kartierten Gebiets.
- **semantische Kriterien**, wie beispielsweise die Objektklassen, welche die Landmarken repräsentieren.

Verschiedene merkmalsbasierte und statistische Selektionsstrategien von Landmarken und Merkmalen wurden im Rahmen einer studentischen Arbeit [84] hinsichtlich verschiedener Gesichtspunkte detailliert untersucht und evaluiert.

Wie bereits in Abschnitt 1.2 erwähnt, soll die Merkmalsselektion nicht im Fokus

dieser Arbeit stehen. Da dennoch ein vollumfängliches Gesamtsystem präsentiert werden soll, wird in diesem Kapitel ein kurzer Einblick in diesen Prozessschritt gegeben.

Im folgenden Abschnitt 8.1 wird das grundsätzliche Vorgehen zur Merkmals- bzw. Landmarkenselektion beschrieben. Darauffolgend wird in Abschnitt 8.2 exemplarisch eine mögliche Variante zur statistischen Selektion kartierter Landmarken anhand einer Schätzung ihrer Wiedererkennungswahrscheinlichkeit beschrieben. Im letzten Abschnitt 8.3 wird diese Variante empirisch analysiert.

8.1 Selektionsmethode

Die entwickelte Selektion zielt darauf ab, die Anzahl der Landmarken und Merkmale in der Lokalisierungskarte zu minimieren und gleichzeitig den Informationsgehalt der Lokalisierungskarte hinsichtlich einer robusten und präzisen Lokalisierung zu maximieren. Grundlegende Anforderungen für die Erzeugung der Lokalisierungskarte sind:

- Die Selektion soll von der Anzahl der Aufnahmesequenzen abstrahieren und eine Auswahlstrategie unter Einbeziehung der Merkmale aus allen kartierten Aufnahmesequenzen ermöglichen.
- Es sollen überall im kartierten Bereich ausreichend viele Landmarken vorhanden sein, die eine möglichst hohe Wahrscheinlichkeit haben, erneut detektiert und assoziiert werden zu können. Da das Spektrum der Umgebungsstruktur in städtischen Bereichen sehr vielfältig ist, variiert dementsprechend die Anzahl, Wiedererkennungswahrscheinlichkeit und Eindeutigkeit der Merkmale. So ist beispielsweise die Merkmalsausprägung in Häuserschluchten deutlich dichter und eindeutiger als in wäldlichen Gebieten. Es sollte außerdem möglichst in allen Gebieten eine ausreichende Anzahl von Merkmalen vorhanden sein, sodass an jedem kartierten Ort eine robuste Lokalisierung möglich ist. Des Weiteren ist eine möglichst gleichmäßige Landmarkendichte erstrebenswert, sodass die Menge an zu speichernder und zu verarbeitender Karteninformation im gesamten kartierten Gebiet gleich und nach oben begrenzt ist. Auf diese Weise kann garantiert werden, dass die Ladezeit der Lokalisierungskarte in allen kartierten Bereichen ähnlich ist, wodurch eine näherungsweise

konstante Laufzeit der Lokalisierung unabhängig vom Aufenthaltsort erreicht wird.

Diese beiden Anforderungen können durch die Anwendung einer Selektion pro Ort erfüllt werden. Dafür wird eine Selektionsrasterkarte verwendet, in welche die zu selektierenden Landmarken aus der Kartendatenbank eingefügt werden. Zur Selektion werden die Landmarken nach den gleichen Kriterien wie bei der Erstellung der metrischen Lokalisierungskarte gemäß Abschnitt 7.3.1 in die Selektionsrasterkarte eingefügt, wodurch gemäß rein geometrischer Kriterien die Sichtbarkeit und Wiedererkennbarkeit der Landmarken sichergestellt wird. Darauffolgend wird pro Zelle eine Selektion durchgeführt, wobei in jeder Zelle maximal n_{\max} Landmarken ausgewählt werden, die in die Lokalisierungskarte M_n^L überführt werden. Enthält eine Zelle weniger als n_{\max} Landmarken, werden alle in der Zelle liegenden Landmarken ausgewählt. So verbleibt auch in schwach strukturierten oder wenig befahrenen Gebieten eine minimale Anzahl von Merkmalen und Landmarken. Die Selektionsrasterkarte ist dabei zweidimensional und erstreckt sich über die angenommene Befahrungsebene, die durch die Kartenposen aufgespannt wird.

8.2 Landmarkenselektion

Die Selektion zur Erstellung der metrischen Lokalisierungskarte bestimmt eine Teilmenge der Landmarken L_n aus der Kartendatenbank M_n , die zusammen mit ihren korrespondierenden Merkmalskorrespondenzen in die Lokalisierungskarte M_n^L überführt werden. Bei den Betrachtungen ist angenommen, dass die Landmarken in jeder der n kartierten Aufnahmesequenzen prinzipiell sichtbar und wiedererkennbar sind.

Zur Selektion wird jeder Landmarke $\mathbf{l}_i \in L_n$ ein statistischer Wiedererkennungswert $d_i \in [0, 1]$ und ein altersbedingter Wiedererkennungswert $a_i \in \mathbb{R}$ zugeordnet. Die folgende Auflistung beschreibt die Interpretation und Bestimmung dieser beiden Kennwerte:

- **Statistischer Wiedererkennungswert d_i**

Eine Landmarke, die in den bisher kartierten Aufnahmesequenzen oft redetektiert wurde, hat eine hohe Wahrscheinlichkeit, während der Lokalisierung wiedererkannt zu werden. Wurde eine Landmarke $\mathbf{l}_i \in L_n$ in $r \leq n$ Aufnahmesequenzen mindestens einmal wiedererkannt – also umfasst die zu \mathbf{l}_i zuge-

hörige Merkmalskorrespondenz $\Theta_i \in \Theta_n$ mindestens k Merkmale aus k verschiedenen Befahrungen – so ist ihr statistischer Wiedererkennungswert

$$d_i = \frac{k}{n}. \quad (8.1)$$

- **Altersbedingter Wiedererkennungswert a_i**

Eine Annahme, die insbesondere für längere Betrachtungszeiträume gilt, ist, dass jüngere Landmarken eine höhere Wiedererkennungswahrscheinlichkeit haben. Das Alter einer Landmarke \mathbf{l}_i ergibt sich aus dem Alter des jüngsten Merkmals aus ihrer zugeordneten Merkmalskorrespondenz Θ_i . Das Alter eines Merkmals ergibt sich wiederum aus dem Alter der Aufnahmesequenz S_i , in der das Merkmal detektiert wurde. Der altersbedingte Wiedererkennungswert a_i ergibt sich daher direkt aus dem absoluten Aufnahmezeitpunkt der Sequenz S_i . Je jünger das Aufnahmedatum und der Aufnahmezeitpunkt einer Sequenz ist, desto höher ist der Wiedererkennungswert a_i der Landmarken, die zu mindestens einem Merkmal aus dieser Sequenz assoziiert wurden.

Zur Selektion wird angenommen, dass der statistische Wiedererkennungswert d_i für die Bewertung der Wiedererkennungswahrscheinlichkeit einer Landmarke eine höhere Relevanz als der altersbedingte Wiedererkennungswert a_i hat, sofern die Landmarke \mathbf{l}_i nicht zu alt ist. Für die Selektion werden die Landmarken sortiert und die ersten n_{\max} Landmarken pro Zelle der Selektionsrasterkarte ausgewählt und in der Lokalisierungskarte gespeichert. Die Landmarken werden dabei gemäß der folgenden Kriterien sortiert:

- Sind zwei Landmarken \mathbf{l}_i und \mathbf{l}_j beide jünger oder älter als ein maximales Alter a_{\min} , d.h. gilt $a_i, a_j \geq a_{\min} \vee a_i, a_j < a_{\min}$, werden die Landmarken ausschließlich anhand ihres statistischen Wiedererkennungswerts d_i bewertet. Landmarken mit einer höheren Bewertung d_i werden bevorzugt. Bei gleicher Bewertung $d_i = d_j$ wird die Landmarke \mathbf{l}_i mit dem höheren altersbedingten Wiedererkennungswert $a_i > a_j$ bevorzugt. Sind ihre Altersbewertungen ebenfalls gleich, wird zufällig eine der beiden Landmarken bevorzugt.
- Gilt für die ältere Landmarke \mathbf{l}_i , $a_i < a_{\min}$ und für die jüngere Landmarke \mathbf{p}_j , $a_j \geq a_{\min}$, hat \mathbf{l}_j Vorrang vor \mathbf{l}_i , unabhängig von den statistischen Wiedererkennungswerten. Auf diese Weise erhöht sich für jüngere Landmarken,

die beispielsweise aufgrund ihres jungen Alters noch nicht so häufig wiedererkannt werden konnten, die Wahrscheinlichkeit, in die Lokalisierungskarte überführt zu werden. Sehr alte Landmarken werden demhingegen häufiger aussortiert.

Durch diese beiden Kriterien ergibt sich eine Auswahl aktueller und häufig in verschiedenen Befahrungen wiedererkannter Landmarken, von denen angenommen wird, dass sie für eine zuverlässige Lokalisierung von Vorteil sind. Die Lokalisierungskarte M_n^L adaptiert sich somit stets an die aktuellen Umgebungs- und Umweltbedingungen und präferiert Landmarken auf persistenten Strukturen, da diese sich auch über lange Zeiträume nicht verändern und daher fortlaufend wiedererkannt werden.

8.3 Experimente

Zur Evaluation der Selektionsstrategie werden zehn Aufnahmesequenzen, die im Zeitraum von November 2017 bis Januar 2019 mit Konfiguration MKS-A aufgezeichnet wurden, betrachtet. Die Sequenzen wurden in zeitlich aufsteigender Reihenfolge kartiert und nach jeder Iteration eine Selektion durchgeführt. Die kartierte Strecke entspricht im Wesentlichen der in rot dargestellten Trajektorie in Abbildung 6.15 und umfasst Häuserschluchten, wäldliche Gebiete und größere vorstädtische Straßen. Die kartierte Strecke hat eine Länge von $\sim 7.5\text{km}$ und die Summe der gefahrenen Distanzen zur Aufzeichnung der zehn Aufnahmesequenzen beträgt $\sim 83\text{km}$. Des Weiteren sind größere Teile der Strecke im kartierten Zeitraum durch Baustellen verändert worden, sodass teilweise auf einer entfernten Parallelspur oder eine Umleitung gefahren wurde. Tabelle 8.1 gibt eine Übersicht über die Aufnahmezeitpunkte und Bedingungen der einzelnen Befahrungen. Des Weiteren stellt Abbildung 8.1 die zeitlichen Abstände zwischen den Aufnahmesequenzen dar. Der Altersschwellwert a_{\min} beträgt sechs Monate.

Zur Analyse der Selektionsstrategie wurde die Selektion mit den Merkmalen und Landmarken aus der Kartendatenbank M_9 , welche die gemeinsam kartierten Landmarken und Merkmale aus allen Sequenzen S_0 bis S_9 umfasst, für verschiedene Kompressionsstufen untersucht. Dazu wurde die Selektion mit einer konstanten Zellengröße von $0.4\text{m} \times 0.4\text{m}$ mit unterschiedlichen Mindestanzahlen n_{\max} von

Sequenz	Datum / Zeit	Bedingungen
S_0	2017-11-20 15:11	bewölkt, wenig Vegetation
S_1	2018-03-01 14:53	bewölkt, wenig Vegetation
S_2	2018-03-12 10:40	bewölkt, wenig Vegetation, regnerisch, Veränderungen am Kamerasystem
S_3	2018-03-31 15:19	sonnig, wenig Vegetation
S_4	2018-06-25 10:54	wechselhaft, viel Vegetation, Baustellen
S_5	2018-07-20 14:27	sonnig, viel Vegetation, Baustellen
S_6	2018-08-29 13:47	sonnig, viel Vegetation, teilweise Straßenerneuerung
S_7	2018-10-22 14:37	sonnig, leichte Vegetation, Veränderungen am Kamerasystem
S_8	2018-11-29 15:50	bewölkt, wenig Vegetation
S_9	2019-01-15 10:31	bewölkt, wenig Vegetation, regenerisch

Tabelle 8.1: Aufnahmezeitpunkte und Bedingungen der evaluierten Aufnahmesequenzen für die Merkmalsselektion.

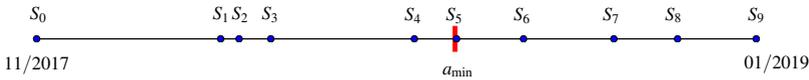


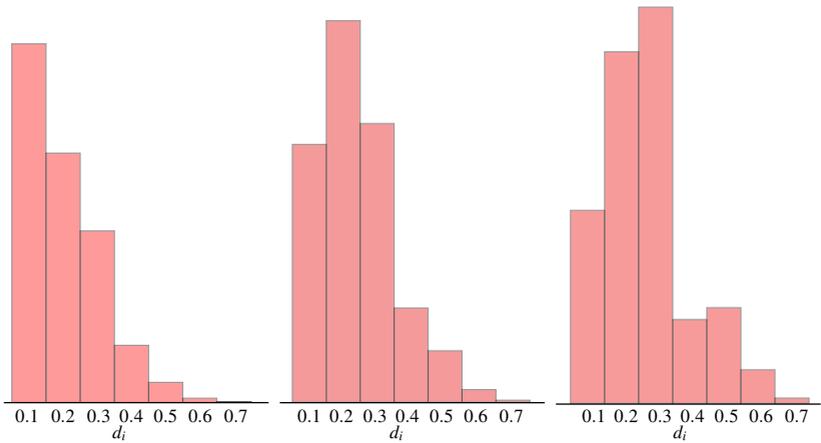
Abbildung 8.1: Aufnahmezeitpunkte der Evaluationssequenzen. Die rote Grenze stellt den Altersschwellewert $a_{\min} = 6$ Monate dar.

Landmarken pro Zelle durchgeführt, wobei n_{\max} den Grad der Kompression beschreibt. Es werden exemplarisch drei Kompressionsstufen $n_{\max} = 350$, $n_{\max} = 150$ und $n_{\max} = 50$ verglichen. Tabelle 8.2 zeigt die Gesamtanzahl ausgewählter Landmarken für die verschiedenen Kompressionsstufen. Bei der höchsten Kompressionsstufe $n_{\max} = 50$ wird im Verhältnis zur niedrigsten evaluierten Kompressionsstufe $n_{\max} = 350$ nur ungefähr ein Sechstel der Landmarken in der Lokalisierungskarte gespeichert. Dies ist insbesondere vorteilhaft für die Laufzeit der Lokalisierung, da die Bereitstellungszeit der Zellen der metrischen Lokalisierungskarte zur Lokalisierung näherungsweise linear mit der Anzahl der Landmarken pro Zelle einhergeht.

Abbildung 8.2 zeigt die Verteilung der Landmarken aufgetragen über dem statistischen Wiedererkennungswert d_i für die verschiedenen Kompressionsstufen. Der Anteil der Landmarken, die in mehr als sieben Aufnahmesequenzen wiedererkannt wurden, ist vergleichsweise klein und nicht dargestellt. Die Abbildung zeigt gemäß der Erwartung eine Verschiebung der Verteilung zu einem größeren Anteil von Landmarken mit höherem statistischen Wiedererkennungswert für steigende Kom-

n_{\max}	Anzahl Landmarken in Millionen / Anteil zur niedrigsten Kompressionsstufe
350	$\sim 7,3 / 100\%$
150	$\sim 3 / 42\%$
50	$\sim 1,2 / 16\%$

Tabelle 8.2: Landmarkenanzahlen für verschiedene Kompressionsstufen.

Abbildung 8.2: Landmarkenverteilungen aufgetragen über dem statistischen Wiedererkennungswert d_i für verschiedene Kompressionsstufen n_{\max} (Links: 350, Mitte: 150, Rechts: 50).

pressionsstufen. Des Weiteren stellt Abbildung 8.3 die Verteilungen der Landmarken über ihren altersbedingten Wiedererkennungswert a_i für die evaluierten Kompressionsstufen dar. Hier zeigt sich ebenfalls der erwartete Effekt, dass der Anteil der Landmarken aus jüngeren Befahrungen mit steigender Kompression steigt und die Selektion veraltete Merkmale vergisst. Es ist ersichtlich, dass Landmarken mit hohem statistischen Wiedererkennungswert bevorzugt werden, da der Anteil der Landmarken mit Merkmalen aus den Sequenzen S_1 , S_2 und S_3 sich deutlich weniger verringert als Landmarken mit Merkmalen aus den jüngeren Befahrungen S_4 , S_5 , S_6 und S_7 . Da diese Fahrten in den Sommermonaten durchgeführt wurden, wurde ein Großteil der persistenten Struktur auf Hauswänden und der Baumverästelung

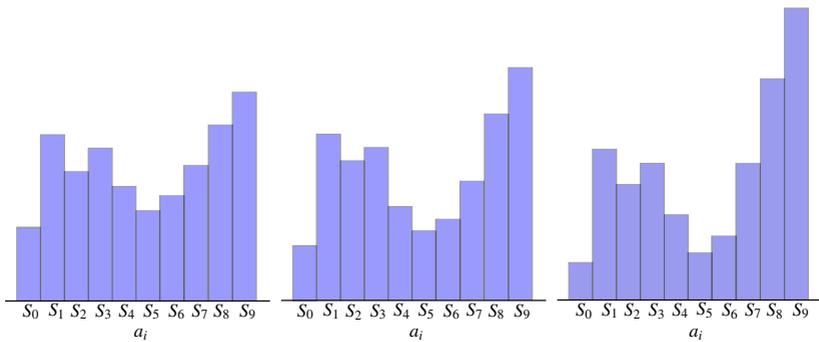


Abbildung 8.3: Landmarkenverteilungen aufgetragen über dem altersbedingten Wiedererkennungswert a_i für verschiedene Kompressionsstufen n_{\max} (Links: 350, Mitte: 150, Rechts: 50). Der altersbedingte Wiedererkennungswert a_i korrespondiert direkt zum Alter der Aufnahmesequenzen.

durch Blätter und Grünwuchs verdeckt. In den späteren Winterfahrten S_8 und S_9 mit weniger Vegetation wurden viele dieser Merkmale wieder sichtbar und zu Merkmalen aus den Fahrten S_1 , S_2 und S_3 im Winter des Vorjahres assoziiert. Da sich durch eine erneute Assoziation einer Landmarke der statistische Wiedererkennungswert und der altersbedingte Wiedererkennungswert dieser Landmarke erhöht, sind die Landmarken mit Merkmalen aus den Fahrten S_1 , S_2 und S_3 weniger von der altersbedingten Wiedererkennungswahrscheinlichkeit betroffen als die Landmarken aus den Sequenzen S_4 , S_5 , S_6 und S_7 . Da in den Sommermonaten außerdem größere Teile der Strecke durch Baustellen belegt waren, verstärkt sich der Effekt der Landmarkenalterung in den Sommersequenzen $S_4 - S_7$. Auch wenn die gezeigten Statistiken keine eindeutigen Schlussfolgerungen ermöglichen, da die vorgestellte Selektionsstrategie durch viele äußeren Einflussfaktoren beeinflusst wird, zeigt sich dennoch eine für die Lokalisierung günstige Entwicklung der Landmarkenverteilung.

Die Verbesserung der Lokalisierungskarte hinsichtlich der Lokalisierungsgüte zeigt sich ebenfalls bei der Analyse der Verfügbarkeit erfolgreicher Lokalisierungsschritte. Dafür wurde die Lokalisierungsverfügbarkeit, welche die Robustheit und Genau-

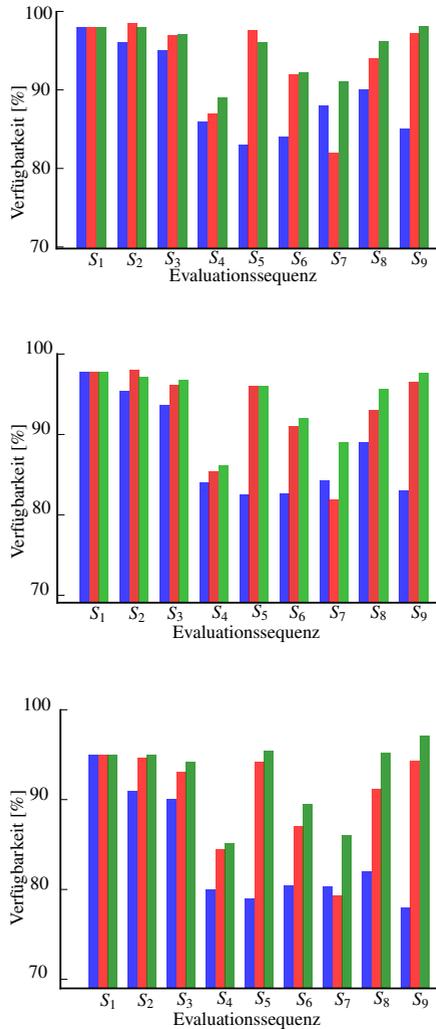


Abbildung 8.4: Erzielte Lokalisierungsverfügbarkeiten für die jeweilige Evaluationssequenz S_{i+1} für die drei Kompressionsstufen $n_{\max} = 350$ (Oben), $n_{\max} = 150$ (Mitte) und $n_{\max} = 50$ (Unten). Grüne Balken: Ergebnisse mit der fortlaufend aktualisierten Lokalisierungskarte M_i^L . Blaue Balken: Ergebnisse mit der einmalig erstellten Lokalisierungskarte M_0^L . Rote Balken: Ergebnisse mit der Lokalisierungskarte M_i^L , die nur Merkmale und Landmarken aus der jeweils vorgehenden Sequenz S_i umfasst.

igkeit der Lokalisierung maßgeblich repräsentiert, nach jeder Kartierungsiteration evaluiert. Zu diesem Zweck wurde mit den selektierten Landmarken nach jeder Kartierungsiteration eine Lokalisierungskarte berechnet. Für jede der drei Kompressionsstufen $n_{\max} = 350$, $n_{\max} = 150$ und $n_{\max} = 50$ wurden somit neun Lokalisierungskarten M_0^L, \dots, M_8^L berechnet, wobei die Sequenzen S_0, \dots, S_8 in chronologischer Reihenfolge kartiert wurden. Die zeitlich nachfolgende Aufnahmesequenz S_{i+1} wurde zur Evaluation der Lokalisierungskarte M_i^L , $0 \leq i < 9$ verwendet. Dies entspricht dem praxisnahen Anwendungsfall, in dem die Lokalisierungskarte fortlaufend aktualisiert wird. In der praktischen Anwendung wird in der Regel jedoch nicht jede weitere Befahrung kartiert.

Die Lokalisierungsverfügbarkeiten wurden auf die gleiche Weise wie bei der Evaluation der metrischen Lokalisierung in Abschnitt 7.3.2 bestimmt. Die Ergebnisse sind in Abbildung 8.4 dargestellt. Die grünen Balken zeigen die erzielten Lokalisierungsverfügbarkeiten bei der Lokalisierung der Evaluationssequenz S_{i+1} in die fortlaufend aktualisierte Lokalisierungskarte M_i^L für die verschiedenen Kompressionsstufen $n_{\max} = 350$ (Oben), $n_{\max} = 150$ (Mitte) und $n_{\max} = 50$ (Unten). Zum Vergleich wurde jede Evaluationssequenz S_{i+1} in die Lokalisierungskarte M_0^L , die nur Merkmale und Landmarken aus der ersten Sequenz S_0 umfasst, lokalisiert. Die blauen Balken in Abbildung 8.4 zeigen die erzielten Verfügbarkeiten dieses Experiments. Zusätzlich wurde für jede Evaluationssequenz S_{i+1} eine Lokalisierung mit der Lokalisierungskarte \hat{M}_i^L durchgeführt, die nur Merkmale und Landmarken aus der jeweils vorherigen Aufnahmesequenz S_i bereitstellt. In diesem Fall zeigen die roten Balken in Abbildung 8.4 die erzielten Ergebnisse.

Die in Abbildung 8.4 dargestellten Lokalisierungsverfügbarkeiten untermauern die Schlussfolgerungen aus den Statistiken der altersbedingten und statistischen Wiedererkennungswerte. Bei der Lokalisierung mit der Lokalisierungskarte M_0^L wird eine höhere Verfügbarkeit auf den Sequenzen $S_7 - S_9$ als auf den Sequenzen $S_4 - S_6$ erzielt, was sich durch die veränderte Vegetation im Winter bzw. Sommer erklären lässt.

Bei Verwendung der aktuellen Lokalisierungskarte \hat{M}_i^L wird erwartungsgemäß eine höhere Verfügbarkeit als bei der Lokalisierung mit der fortlaufend älter werdenden Karte M_0^L erzielt. Es zeigt sich jedoch insbesondere bei der Lokalisierung von Se-

quenz S_4 mit der Karte \hat{M}_5^L und der Lokalisierung von Sequenz S_7 mit der Karte \hat{M}_6^L ein verringerte Verfügbarkeit als bei den anderen Iterationen. Dies lässt sich durch den großen zeitlichen Abstand und die damit zusammenhängende Umweltveränderung zwischen den Aufnahmesequenzen erklären.

Die Lokalisierung mit der fortlaufend aktualisierten Lokalisierungskarte M_i^L zeigt bei höherer Kompressionsstufe vergleichsweise bessere Verfügbarkeiten auf allen Evaluationssequenzen. Durch das Beibehalten von Landmarken auf persistenten Strukturen und der Beimischung eines Anteils aktueller Landmarken wird über den gesamten Evaluationszeitraum die beste Verfügbarkeit erzielt. Die Landmarkenselektion kompensiert u. a. die Verschlechterung der Lokalisierungsverfügbarkeit bei den signifikanten Umweltveränderungen durch die verschiedenen Jahreszeiten. Je weniger Landmarken in die Lokalisierungskarte überführt werden, desto stärker zeigt sich dieser Effekt, was die Effizienz dieser exemplarisch vorgestellten statistischen Selektionsmethode belegt.

9 Anwendungsfälle

In diesem Kapitel werden Anwendungsfälle vorgestellt, in denen die präsentierte Kartierungs- und Lokalisierungsmethode im Rahmen verschiedener Projekte zum kartenbasierten automatischen Fahren erfolgreich eingesetzt wird. Die ursprüngliche Motivation dieser Arbeit entspringt aus dem A2B-Projekt, welches am Forschungszentrum Informatik in Kooperation mit der Daimler AG durchgeführt wurde. Kapitel 9.1 gibt eine Übersicht über die Rahmenbedingungen, Zielsetzung und Durchführung des A2B-Projekts. Ein weiterer Meilenstein aus Sicht der Anwendung ist das Mercedes-Benz Future Bus-Projekt, welches in Abschnitt 9.2 vorgestellt wird. Zuletzt wird in Abschnitt 9.3 die Verwendung der Verfahren im Adenauer-Projekt vorgestellt, bei dem seit Ende 2016 bis zum aktuellen Tag das automatisierte Fahren in vor- und innerstädtischen Szenarien in der Kalsruher Oststadt erprobt wird.

9.1 A2B

Das zugrunde liegende Konzept für diese Arbeit wurde im Rahmen des A2B-Projekts, das zusammen mit der Daimler AG im Zeitraum zwischen Januar 2014 und Dezember 2015 durchgeführt wurde, entwickelt. Das Projekt ist ein Nachfolgeprojekt der Bertha-Benz-Fahrt (BBF) [4] aus dem Jahr 2013. Für das A2B-Projekt wurden verschiedene Probleme der Kartengenerierung und der Lokalisierungsmethode, die im BBF-Projekt festgestellt wurden, aufgegriffen:

- **Automatisierung der Kartengenerierung**

Ein Problem im BBF-Projekt war die Generierung der Karte, die im Wesentlichen aus einer Lokalisierungskartenschicht und einer Planungskartenschicht bestand. Insbesondere die Planungskartenschicht konnte nur teilautomatisiert erstellt werden, wodurch viele Arbeitsstunden in die manuelle Bearbeitung dieser Karte geflossen sind. Dieser Prozess soll durch die automatische Extraktion von Fahrspurmarkierungen aus den aufgezeichneten Bildern auto-



Abbildung 9.1: Das verwendete seriennahe MKS, welches im Rahmen des A2B-Projekts für die Lokalisierung verwendet wurde.

matisiert werden [85]. Um die Detektionsrate der Fahrspurmarkierungen zu steigern, sollen mehrere Befahrungen der gleichen Strecke automatisch gemeinsam kartiert werden.

- **Aktualisierung der Lokalisierungskarte**

Die Lokalisierungskarte soll iterativ aktualisiert und um neue Gebiete erweitert werden können.

- **Verwendung mehrerer seriennaher Kameras zur Lokalisierung**

Das zweigeteilte Lokalisierungskonzept, welches im BBF-Projekt (Kapitel 1) eingesetzt wurde, soll durch eine einheitliche Lokalisierungsmethode unter

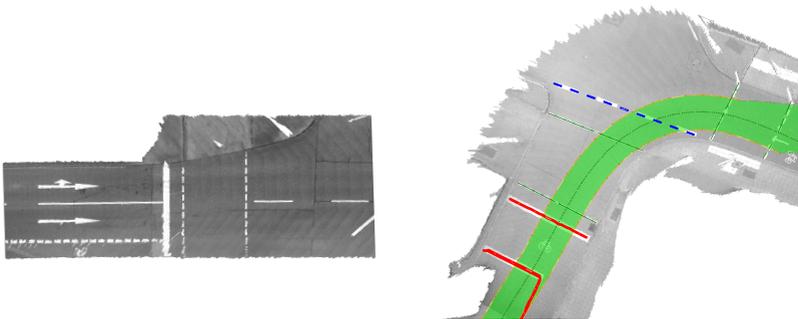


Abbildung 9.2: Virtuelle Luftbilder, die durch Stereorekonstruktion und Homographie erzeugt wurden. Die Bilder wurden mit den kartierten Fahrzeugposen aus mehreren aufeinanderfolgenden Einzelbildern zusammengefügt. Im rechten Bild sind zusätzlich detektierte und klassifizierte Fahrspurmarkierungen und der eigene Fahrkorridor hervorgehoben.

Verwendung eines seriennahen MKS ersetzt werden. Abbildung 9.1 zeigt das MKS und exemplarische Rohbilder, die zur Durchführung und Demonstration des Projektfortschritts verwendet wurden. Robuste Assoziationsverfahren, wie sie in Kapitel 4 vorgestellt wurden, waren für eine erfolgreiche Kartengenerierung mit einem solchen MKS erforderlich.

Das Projekt wurde im Januar 2016 im Rahmen einer Live-Demonstration auf einer 4 km langen Strecke in Sindelfingen, bei der u. a. die Robustheit und Genauigkeit der Lokalisierung unter Echtzeitbedingungen im Fahrzeug demonstriert wurde, erfolgreich abgeschlossen.

Abbildung 9.2 zeigt virtuelle Luftbilder, die durch Stereorekonstruktion und Homographie aus den aufgezeichneten Rohbildern berechnet wurden. Die virtuellen Luftbilder wurden aus Einzelbildern aus mehreren aufeinanderfolgenden Zeitschritten auf Basis der mit der in Kapitel 6 vorgestellten Kartengenerierung zuvor erstellten Kartenposen zusammengefügt (engl. *stitched*). Die dargestellte Fahrbahnoberfläche in Abbildung 9.2 ist präzise rekonstruiert und frei von Verknüpfungsartefakten, was die Genauigkeit der Kartenposen demonstriert. Außerdem sind detektierte und klassifizierte Fahrspurmarkierungen dargestellt, die im Rahmen der Live-Demonstration in die aktuellen Kamerabilder projiziert wurden. Hiermit wurde die Genauigkeit der



Abbildung 9.3: Typische Umgebung der Strecke entlang der Bustrasse zwischen Amsterdam und Haarlem.

Lokalisierung im Realbetrieb demonstriert.

Eine Besonderheit der Demonstrationsstrecke ist ein ~ 1 km langer Landstraßenabschnitt mit wenig Umgebungsstruktur. Auf diesem Streckenabschnitt konnte sich der Versuchsträger reproduzierbar und in Echtzeit erfolgreich lokalisieren. Dies zeigt, dass mit einem seriennahen MKS eine Lokalisierung mit Punktmerkmalen auch in merkmalschwachen Umgebungen möglich ist und somit ein Problem, welches im Rahmen des BBF-Projekts identifiziert wurde, gelöst ist.

9.2 Future Bus

Das Mercedes-Benz Future Bus Projekt [86] ist in Kooperation der Daimler AG mit dem Forschungszentrum Informatik und weiteren Institutionen im Zeitraum von Oktober 2015 bis Juli 2016 durchgeführt worden. Das Ziel dieses Projektes war die Inbetriebnahme und Demonstration eines vollautomatisiert fahrenden Linienbusses auf einer 20 km langen Bustrasse, die den Stadtteil Amsterdam Schiphol mit der Stadt Haarlem in den Niederlanden verbindet. Die Strecke führt weitestgehend durch ländliche und strukturschwache Gebiete und umfasst mehrere Tunneldurchfahrten von bis zu 1,2 km Länge. Abbildung 9.3 zeigt die typische Landschaft entlang der Strecke. Auf der Strecke fahren ausschließlich Linienbusse. Zur Lokalisierung wurde im Rahmen des Projektes eine Fusionslösung zwischen einem GNSS-System und der in dieser Arbeit vorgestellten Lokalisierung umgesetzt. Dafür wurde der verwendete Bus u. a. mit einem MKS ausgerüstet. Abbildung 9.4 zeigt schematisch die Sicht-



Abbildung 9.4: Sichtbereiche des verwendeten MKS im Future Bus.

kegel des verwendeten MKS. Die Lokalisierungsfusion wurde mit einem Unscented Kalman Filter (UKF) realisiert. Sowohl das GNSS-System als auch die visuelle Lokalisierung übermitteln neben der Schätzung der EgoPose eine Schätzung der Unsicherheit in Form einer Kovarianzmatrix an den UKF. Insbesondere in Tunneldurchfahrten und bei Durchfahrten zahlreicher Unterführungen wurde die Lokalisierung ausschließlich durch das visuelle System sichergestellt, da das GNSS-System hier keine verlässliche Messung liefern kann. Die wesentlichen Herausforderungen in diesem Projekt waren

- die Lokalisierung in einer Umgebung mit wenig Umgebungsstruktur. Insbesondere im Nahbereich finden sich auf weiten Teilen der Strecke nur Fahrbahnstruktur, Masten und Laternen.
- die Lokalisierung in den langen Tunnelpassagen, in denen nicht nur wenig visuelle Struktur vorhanden ist, sondern die zusätzlich auch nur schwach und inhomogen beleuchtet ist.
- die hochgenaue Anpassung der Kartentrajektorie an die GNSS-Messungen insbesondere in den Bereichen der Tunnelausfahrten. Da das GNSS-Signal bei Tunnelausfahrten nicht-deterministisch wiedereinsetzt, müssen beide Lokalisierungsschätzungen in Position und Orientierung mit geringer Unsicherheit übereinstimmen, da sonst Lokalisierungssprünge auftreten.

Dem Problem mangelnder Umgebungsstruktur wurde durch mehrere Maßnahmen entgegengewirkt. Zum einen wurde bei der Aufnahme der Kartierungssequenzen mit verringerter Geschwindigkeit gefahren, wodurch die subsequente Merkmalsassoziation insbesondere in den zur Seite schauenden Kameras verbessert werden konnte. Zum anderen wurden verschiedene Parameter der Bildaufzeichnung und der Merkmalsdetektion angepasst und die Lokalisierungsfrequenz durch zeitlich überlappen-

de Prozessausführung erhöht. Auf diese Weise konnte die Initialisierung der metrischen Lokalisierung verbessert werden.

Zur Anpassung der Trajektorie an die GNSS-Trajektorie wurde die extrinsische Kalibrierung zwischen dem GNSS-System und dem MKS auf Basis gefahrener Trajektorien verbessert. Dazu wurden die Signale der aus der Kartengenerierung entspringenden hochgenauen Trajektorien mit den Signalen der GNSS-Trajektorien verglichen. Neben der extrinsischen Kalibrierung wurde zusätzlich der zeitliche Versatz der Signale geschätzt. Außerdem wurde der Ausgangsbereich der Tunnel in der Karte annotiert und das Wiedereinsetzen des GNSS-Signals über eine Rampenfunktion gedämpft.

Diese Maßnahmen haben dazu geführt, dass sich der Bus kontinuierlich, störungsfrei und hochgenau auf der gesamten Strecke lokalisieren konnte. Im Juli 2016 wurde die Strecke im Rahmen einer mehrtägigen Abschlussdemonstration viele Male und vollautomatisch in beiden Richtungen befahren. Hierbei waren keine Eingriffe des Sicherheitsfahrers erforderlich. Der Bus fuhr während der Demonstrationsfahrten kartentbasiert Haltestellen an, sodass stets ein behindertengerechtes Ein- und Aussteigen möglich war. Das stellt u. a. die hohe Präzision der entwickelten Lokalisierung dar.

9.3 Adenauer

Im Adenauer-Projekt wird das automatisierte Fahren in vor- und innerstädtischen Szenarien in Karlsruhe erprobt. Das Projekt wird in Zusammenarbeit der Forschungsgruppe “Mobile Perception Systems” des Forschungszentrums Informatik Karlsruhe und dem Institut für Mess- und Regelungstechnik des Karlsruher Instituts für Technologie weiterentwickelt. Die blau und rot dargestellten Strecken in Abbildung 6.15 zeigen die beiden Hauptstrecken, auf denen seit Anfang des Jahres 2017 in wöchentlichen Zyklen das vollautomatisierte Fahren erprobt wird.

Die in dieser Arbeit vorgestellte Lokalisierung und Kartengenerierung ermöglicht seit Beginn des Projekts bis zur Veröffentlichung dieser Arbeit und voraussichtlich noch lange darüber hinaus die Bereitstellung der präzisen Karteninformationen für das Verhaltensgenerierungs- und Planungsmodul. Abbildung 9.5 zeigt beispielhaft die Projektion einer kartierten Ampel, die unter Verwendung der Lokalisierungsschät-



Abbildung 9.5: Klassifikation des Ampelstatus durch ein CNN anhand der Projektion einer kartierten Ampel in die aktuellen Kamerabilder auf Basis der Lokalisierung.

zung in die aktuellen Kamerabilder projiziert wurde. Auf Basis dieser Projektion wird ein Bildausschnitt bestimmt, der als rote Box in Abbildung 9.5 dargestellt ist. Dieser Bildausschnitt ist das Eingangsbild für einen gelernten Klassifikator, der das Lichtsignal der Ampel interpretiert. Das dafür verwendete CNN ist besonders empfindlich gegenüber Verschiebungen des Bildausschnitts. Eine Verschiebung von wenigen Pixeln führt in den meisten Fällen bereits zu einer falschen Klassifikation. Dieser Anwendungsfall zeigt beispielhaft die Notwendigkeit einer hochgenauen 3D-Lokalisierung. Die vorgestellte Methode wird diesen Anforderungen reproduzierbar und über lange Zeiträume im Realbetrieb gerecht und ermöglicht so eine robuste und zuverlässige Verhaltensgenerierung an den zahlreichen Ampelszenarien entlang der gefahrenen Strecke.

Auch die Trajektorienplanung benötigt eine möglichst präzise Lokalisierung in der Planungskarte, die Fahrspurbegrenzungen und Fahrkorridore bereitstellt. Abbildung 9.6 zeigt beispielhaft die geplante Trajektorie innerhalb eines Fahrkorridors, der auf Basis der Lokalisierung aus der Karte bereitgestellt wurde. Idealerweise findet die Planung der Trajektorie in einem Koordinatensystem statt, welches glatt und sprunfrei zwischen aufeinanderfolgenden Planungsschritten transformiert wird. Da Kartenwissen benötigt wird, wirkt sich die Lokalisierungsungenauigkeit auf die Trajektorienplanung aus, wobei bereits Lokalisierungssprünge von 0,3 - 0,5m zu ungewünschten und unkomfortablen Fahrmanövern führen. In den wöchentlichen Probefahrten zeigt sich, dass die vorgestellte Lokalisierungsmethode auch diesen hohen

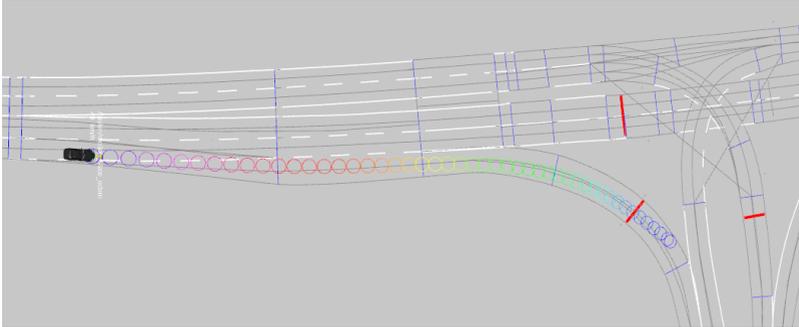


Abbildung 9.6: Trajektorienplanung innerhalb der Planungskarte.

Genauigkeitsanforderungen gerecht wird und somit auf allen Streckenabschnitten ein glattes und natürliches Fahrverhalten erzielt wird.

10 Zusammenfassung

Automatisches Fahren im öffentlichen Straßenverkehr ist eine der großen soziotechnischen Herausforderungen der nächsten Jahre. Die Bereitstellung von geometrisch verorteten Informationen aus Karten ist gemäß des aktuellen Stands der Technik ein notwendiger Bestandteil zur Lösung dieses anspruchsvollen Problems. Je feiner und exakter die kartierten Informationen sind, desto höher sind die Anforderungen an die Lokalisierung.

Aufgrund der Abschirmung der Satellitensignale von GNSS-Systemen durch Bäume und Häuser ermöglichen selbst hochgenaue GNSS-Empfänger keine verlässliche Lokalisierung in vor- und innerstädtischen Bereichen. Kamerabasierte Lokalisierungsansätze unter Verwendung von Punktmerkmalen sind eine vielversprechende Alternative, benötigen jedoch aktuell eine zuvor erstellte Karte, die aufgrund einer dynamischen Umwelt altert und schnell unbrauchbar wird. Daraus resultiert die Notwendigkeit, weitere Gebiete in die Karte zu integrieren und bereits kartierte Gebiete bei Bedarf zu aktualisieren. Ein weiteres Problem stellen wenig oder uneindeutig strukturierte Umgebungen dar, da eine Mindestanzahl von lokal eindeutigen und re-detektierbaren Merkmalen benötigt wird.

In dieser Arbeit wurden bestehende Konzepte eines vielversprechenden Ansatzes aufgegriffen und sukzessive in ein robustes, skalierbares und langzeitverfügbares Gesamtsystem integriert. Im Fokus standen dabei die folgenden Aspekte:

- **Skalierbarkeit**

Die Kartengenerierung ist unabhängig von der Größe des zu kartierenden Gebietes und von der Anzahl der Kartierungsiterationen. Die Erstellung, Aktualisierung und Erweiterung der Karte ist stets mit ressourcenbegrenzter Hardware möglich und kann an diese angepasst werden.

- **Robustheit**

Durch die sorgfältige Auswahl und Validierung der gefundenen Merkmals-

korrespondenzen, durch Verwendung von Bewegungsmodellen und durch Kreuzvalidierung der geschätzten Trajektorien wird eine hohe Schätzgüte sichergestellt.

- **Effizienz**

Durch Aufteilung des Gesamtproblems in mehrere unabhängige Teilprobleme mit beschränkter Komplexität wird eine parallele Verarbeitung der Bildverarbeitung und Optimierung ermöglicht.

Die Methode setzt eine bestehende Kalibrierung des Multi-Kamera-Systems und synchron aufgezeichnete Bilder der verwendeten Kameras voraus. Darauf basierend wird die Aufzeichnung vorverarbeitet, wobei eine möglichst genaue Trajektorienrekonstruktion mit Ringschlusserkennung und -optimierung durchgeführt wird. Im Anschluss werden die extrahierten Merkmalskorrespondenzen der neuen Befahrung mit denen der Karte durch eine effiziente, bildbasierte Merkmalsassoziation verknüpft. Darauf basierend wird die Trajektorie der neuen Befahrung in einer geteilten und mehrschrittigen Optimierung an die bestehende Karte angepasst. Durch Hinzunahme von ungenauen GNSS-Messungen aus einem seriennahen GNSS-Empfänger in die Kartenoptimierung kann eine georeferenzierte Kartenbasis erstellt werden. Die verschiedenen Trajektorien-schätzungen werden durch Modelle und geometrische Annahmen in sich und zueinander validiert, wodurch die Wahrscheinlichkeit für eine Fehlerpropagierung verringert wird.

Der zweite Schwerpunkt der Arbeit ist die Lokalisierung mit einem MKS. Die entwickelte Lokalisierung umfasst zwei Komponenten:

- **Eine Ortswiedererkennung**, die durch einen Vergleich von holistischen Bildmerkmalen der aktuellen Bilder mit allen Kartenmerkmalen Aufenthaltshypothesen bestimmt und diese zeitlich verfolgt. Dadurch ergeben sich Hypothesenverfolgungen, die eine valide Fahrzeugbewegung in der Karte aufweisen und somit mit hoher Wahrscheinlichkeit den richtigen Aufenthaltsort in der Karte darstellen.
- **Eine metrische Lokalisierung**, die eine initiale Posenschätzung entgegennimmt und darauf basierend Kartenlandmarken auswählt und diese mit den aktuell detektierten Bildmerkmalen aus allen aufgezeichneten Bildern assoziiert. Mit diesen Korrespondenzen wird durch nichtlineare Optimierung ei-

ne exakte Relativpose zum Kartenursprung bestimmt. Es wurde dafür eine Kartenstruktur entwickelt, die eine effiziente und flexible Abtastung der kartierten Landmarken ohne zusätzlichen Rechenaufwand während der Lokalisierung ermöglicht. Die Kartenstruktur gewährleistet zum einen eine räumlich möglichst gleichverteilte Abtastung der Kartenlandmarken und zum anderen, dass Landmarken mit höherer Wiedererkennungswahrscheinlichkeit bevorzugt werden.

Es wurde in Experimenten gezeigt, dass die vorgestellte Ortswiedererkennung unter Verwendung mehrerer Kamerabilder eine deutliche Steigerung der Anzahl richtiger Aufenthaltshypothesen im Vergleich zu bestehenden Verfahren unter Verwendung einer einzelnen Kamera erzielt. Außerdem wird durch das vorgestellte Verfahren eine Schätzung der relativen Fahrzeugorientierung ermöglicht, wodurch eine Lokalisierung bei Befahrung in entgegengesetzter Kartenrichtung bereitgestellt werden kann.

Die Experimente zur metrischen Lokalisierung haben ebenfalls eine Steigerung der Lokalisierungsverfügbarkeit und -genauigkeit im Vergleich zur Verwendung von nur einer einzelnen Kamera offenbart. Insbesondere wenn die Karte durch saisonale Veränderung der Umwelt veraltet ist oder bei Befahrung einer nicht kartierten Nebenspur zeigt die Multi-Kamera-Lokalisierung eine deutliche Robustheits- und Verfügbarkeitssteigerung gegenüber der Verwendung von nur einer einzelnen Kamera. In weiteren Experimenten wurde gezeigt, dass eine genaue Lokalisierung bei Befahrung der Karte mit beliebiger Fahrzeugorientierung mit der vorgestellten Methode möglich ist.

Zusätzlich wurde eine statistikbasierte Merkmalsselektion vorgeschlagen, die kartierte Landmarken und Merkmale auswählt und in die Lokalisierungskarte überführt. Die Auswahl der Landmarken soll eine robuste und langzeitverfügbare Lokalisierung im kartierten Gebiet gewährleisten und gleichzeitig die Informationsmenge in der Lokalisierungskarte minimieren. Die vorgestellte Auswahlmethode bevorzugt Landmarken, die eine hohe Wahrscheinlichkeit haben, in der Zukunft erneut detektiert zu werden und somit potenziell die Lokalisierung unterstützen. In Experimenten wurde gezeigt, dass die Selektionsstrategie stets eine Auswahl aus aktuellen und langzeitverfügbaren Merkmalen bereitstellt, mit der die Robustheit und Genauigkeit

der Lokalisierung über lange Zeiträume aufrecht erhalten werden kann.

Zusammenfassend wurde ein vollumfassendes langzeitverfügbares Kartierungs- und Lokalisierungssystem vorgestellt, das in mehreren praktischen Anwendungen und Projekten reproduzierbar und robust kartenbasiertes automatisches Fahren in städtischen und ländlichen Regionen sowie bei kilometerlangen Tunneldurchfahrten mit wenig Umgebungsstruktur ermöglicht hat.

Literatur

- [1] Statistisches Bundesamt. (2018). Polizeilich erfasste Unfälle, Adresse: https://www.destatis.de/DE/ZahlenFakten%20/Wirtschaftsbereiche/TransportVerkehr/Verkehr%20sunfaelle/Tabellen_/Strassenverkehrsunfaelle.%20html (besucht am 05.12.2018).
- [2] C.-Y. Chan, “Advancements, prospects, and impacts of automated driving systems”, *International Journal of Transportation Science and Technology*, Bd. 6, Nr. 3, S. 208–216, 2017, Elsevier.
- [3] F. Poggenhans, J.-H. Pauls, J. Janosovits, S. Orf u. a., “Lanelet2: A high-definition map framework for the future of automated driving”, in *International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, Island of Maui, Hawaii, USA, 2018, S. 1672–1679.
- [4] J. Ziegler, P. Bender, M. Schreiber, H. Lategahn u. a., “Making Bertha Drive - An Autonomous Journey on a Historic Route”, *Intelligent Transportation Systems Magazine*, Bd. 6, Nr. 2, S. 8–20, 2014, ISSN: 1939-1390.
- [6] H. Lategahn und C. Stiller, “Vision-Only Localization”, *Transactions on Intelligent Transportation Systems (TITS)*, Bd. 15, Nr. 3, S. 1246–1257, Juni 2014, ISSN: 1524-9050.
- [7] T. Bailey und H. Durrant-Whyte, “Simultaneous localization and mapping (SLAM): Part II”, *Robotics & Automation Magazine*, Bd. 13, Nr. 3, S. 108–117, 2006, IEEE, ISSN: 1558-223X.

- [8] J. Engel, T. Schöps und D. Cremers, “Lsd-slam: Large-Scale Direct Monocular SLAM”, in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele und T. Tuytelaars, Hrsg., Cham: Springer International Publishing, 2014, S. 834–849, ISBN: 978-3-319-10605-2.
- [9] R. Mur-Artal und J. D. Tardós, “ORB-SLAM2: An open-source SLAM system for monocular, stereo, and rgb-d cameras”, *Transactions on Robotics*, Bd. 33, Nr. 5, S. 1255–1262, 2017, IEEE.
- [10] A. Rituerto, L. Puig und J. J. Guerrero, “Visual SLAM with an omnidirectional camera”, in *International Conference on Pattern Recognition (ICPR)*, Istanbul, Turkey, 2010, S. 348–351.
- [11] J. McDonald, M. Kaess, C. Cadena, J. Neira u. a., “Real-time 6-Dof multi-session visual SLAM over large-scale environments”, *Robotics and Autonomous Systems*, Bd. 61, Nr. 10, S. 1144–1158, 2013, Elsevier.
- [12] P. Muehlfellner, P. Furgale, W. Derendarz und R. Philippsen, “Evaluation of fisheye-camera based visual multi-session localization in a real-world scenario”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Gold Coast, Australia, 2013, S. 57–62.
- [13] A. Geiger, J. Ziegler und C. Stiller, “Stereoscan: Dense 3d reconstruction in real-time”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Baden-Baden, Germany, 2011, S. 963–968.
- [14] H. Lategahn, J. Beck und C. Stiller, “DIRD is an illumination robust descriptor”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Ypsilanti, Michigan, USA, 2014, S. 756–761.
- [15] M. Bürki, M. Dymczyk, I. Gilitschenski, C. Cadena u. a., “Map Management for Efficient Long-Term Visual Localization in Outdoor Environments”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Changshu, Suzhou, China, 2018, S. 682–688.

-
- [16] M. Bürki, I. Gilitschenski, E. Stumm, R. Siegwart u. a., “Appearance-based landmark selection for efficient long-term visual localization”, in *International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Daejeon, Korea, 2016, S. 4137–4143.
- [18] P. Mühlfellner, M. Bürki, M. Bosse, W. Derendarz u. a., “Summary maps for lifelong visual localization”, *Journal of Field Robotics*, Bd. 33, Nr. 5, S. 561–590, 2016, Wiley Online Library.
- [19] H. Durrant-Whyte und T. Bailey, “Simultaneous localization and mapping: Part I”, *Robotics and Automation Magazine*, Bd. 13, Nr. 2, S. 99–110, Juni 2006, ISSN: 1070-9932.
- [20] G. Grisetti, R. Kummerle, C. Stachniss und W. Burgard, “A tutorial on graph-based SLAM”, *Intelligent Transportation Systems Magazine*, Bd. 2, Nr. 4, S. 31–43, 2010, IEEE.
- [21] S. Agarwal, N. Snavely, S. M. Seitz und R. Szeliski, “Bundle Adjustment in the Large”, in *Computer Vision – ECCV 2010*, K. Daniilidis, P. Maragos und N. Paragios, Hrsg., Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, S. 29–42, ISBN: 978-3-642-15552-9.
- [22] C. Cadena, L. Carlone, H. Carrillo, Y. Latif u. a., “Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age”, *Transactions on Robotics*, Bd. 32, Nr. 6, S. 1309–1332, 2016, IEEE.
- [23] T. Schneider, M. Dymczyk, M. Fehr, K. Egger u. a., “Maplab: An open framework for research in visual-inertial mapping and localization”, *Robotics and Automation Letters*, Bd. 3, Nr. 3, S. 1418–1425, 2018, IEEE.
- [24] M. Dymczyk, I. Gilitschenski, R. Siegwart und E. Stumm, “Map summarization for tractable lifelong mapping”, in *Robotics: Science and Systems (RSS) Workshop*, Cambridge MA, USA, 2016.

- [25] M. Dymczyk, T. Schneider, I. Gilitschenski, R. Siegwart u. a., “Erasing bad memories: Agent-side summarization for long-term mapping”, in *International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Daejeon, Korea, 2016, S. 4572–4579.
- [26] M. Dymczyk, S. Lynen, M. Bosse und R. Siegwart, “Keep it brief: Scalable creation of compressed localization maps”, in *International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Hamburg, Germany, 2015, S. 2536–2542.
- [27] M. Dymczyk, E. Stumm, J. Nieto, R. Siegwart u. a., “Will it last? learning stable features for long-term visual localization”, in *International Conference on 3D Vision (3DV)*, IEEE, Stanford, CA, USA, 2016, S. 572–581.
- [28] D. M. Rosen, J. Mason und J. J. Leonard, “Towards lifelong feature-based mapping in semi-static environments”, in *International Conference on Robotics and Automation (ICRA)*, IEEE, Stockholm, Sweden, 2016, S. 1063–1070.
- [31] B. Triggs, P. F. McLauchlan, R. I. Hartley und A. W. Fitzgibbon, “Bundle Adjustment — A Modern Synthesis”, in *Vision Algorithms: Theory and Practice*, B. Triggs, A. Zisserman und R. Szeliski, Hrsg., Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, S. 298–372, ISBN: 978-3-540-44480-0.
- [32] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige u. a., “G 2 o: A general framework for graph optimization”, in *International Conference on Robotics and Automation (ICRA)*, IEEE, Shanghai, China, 2011, S. 3607–3613.
- [33] K. Ni, D. Steedly und F. Dellaert, “Tectonic Sam: Exact, out-of-core, submap-based SLAM”, in *International Conference on Robotics and Automation (ICRA)*, IEEE, Roma, Italy, 2007, S. 1678–1685.

-
- [35] M. Kaess, H. Johannsson, R. Roberts, V. Ila u. a., “Isam2: Incremental smoothing and mapping using the Bayes tree”, *The International Journal of Robotics Research*, Bd. 31, Nr. 2, S. 216–235, 2012, Sage Publications Sage UK: London, England.
- [36] F. Dellaert und M. Kaess, “Square Root Sam: Simultaneous localization and mapping via square root information smoothing”, *The International Journal of Robotics Research*, Bd. 25, Nr. 12, S. 1181–1203, 2006, Sage Publications Sage CA: Thousand Oaks, CA.
- [37] A. Cunningham, V. Indelman und F. Dellaert, “Ddf-Sam 2.0: Consistent distributed smoothing and mapping”, in *International Conference on Robotics and Automation (ICRA)*, IEEE, Karlsruhe, Germany, 2013, S. 5220–5227.
- [38] H. Johannsson, M. Kaess, M. Fallon und J. J. Leonard, “Temporally scalable visual SLAM using a reduced pose graph”, in *International Conference on Robotics and Automation (ICRA)*, IEEE, Karlsruhe, Germany, 2013, S. 54–61.
- [39] T. Cieslewski, S. Lynen, M. Dymczyk, S. Magnenat u. a., “Map api-scalable decentralized map building for robots”, in *International Conference on Robotics and Automation (ICRA)*, IEEE, Seattle, Washington, USA, 2015, S. 6241–6247.
- [40] S. Lowry, N. Sünderhauf, P. Newman, J. J. Leonard u. a., “Visual place recognition: A survey”, *Transactions on Robotics*, Bd. 32, Nr. 1, S. 1–19, 2016, IEEE, ISSN: 1941-0468.
- [41] H. Badino, D. Huber und T. Kanade, “Real-time topometric localization”, in *International Conference on Robotics and Automation (ICRA)*, IEEE, St. Paul, Minnesota, USA, 2012, S. 1635–1642.
- [42] T. Naseer, L. Spinello, W. Burgard und C. Stachniss, “Robust Visual Robot Localization Across Seasons Using Network Flows.”, in *Asso-*

- ciation for the Advancement of Artificial Intelligence (AAAI)*, Québec City, Québec, Canada, 2014, S. 2564–2570.
- [43] S. Lynen, M. Bosse, P. Furgale und R. Siegwart, “Placeless place-recognition”, in *International Conference on 3D Vision (3DV)*, IEEE, Bd. 1, Tokyo, Japan, 2014, S. 303–310.
- [44] J. Sivic und A. Zisserman, “Video Google: A text retrieval approach to object matching in videos”, in *International Conference on Computer Vision (ICCV)*, IEEE, Nice, France, 2003, S. 1470.
- [45] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla u. a., “Netvlad: Cnn architecture for weakly supervised place recognition”, in *Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Las Vegas, Nevada, USA, 2016, S. 5297–5307.
- [46] N. Sünderhauf, S. Shirazi, A. Jacobson, F. Dayoub u. a., “Place recognition with convnet landmarks: Viewpoint-robust, condition-robust, training-free”, *Proceedings of Robotics: Science and Systems XII*, 2015.
- [47] R. Girshick, “Fast R-CNN”, in *International Conference on Computer Vision (ICCV)*, IEEE, Santiago, Chile, 2015, S. 1440–1448.
- [48] H. Lategahn, J. Beck, B. Kitt und C. Stiller, “How to learn an illumination robust image feature for place recognition”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Gold Coast, Australia, 2013, S. 285–291.
- [49] M. Schreiber, C. Knöppel und U. Franke, “Laneloc: Lane marking based localization using highly accurate maps”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Gold Coast City, Australia, 2013, S. 449–454.
- [50] F. Poggenhans, N. O. Salscheider und C. Stiller, “Precise Localization in High-Definition Road Maps for Urban Regions”, in *International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Madrid, Spain, 2018, S. 2167–2174.

-
- [52] C. Häne, L. Heng, G. H. Lee, F. Fraundorfer u. a., “3d visual perception for self-driving cars using a multi-camera system: Calibration, mapping, localization, and obstacle detection”, *Image and Vision Computing*, Bd. 68, S. 14–27, 2017, Elsevier.
- [53] E. Stenborg, C. Toft und L. Hammarstrand, “Long-term Visual Localization using Semantically Segmented Images”, *ArXiv preprint arXiv:1801.05269*, 2018.
- [54] A. Kendall, M. Grimes und R. Cipolla, “Posenet: A convolutional network for real-time 6-dof camera relocalization”, in *International Conference on Computer Vision (ICCV)*, IEEE, Santiago, Chile, 2015, S. 2938–2946.
- [55] A. Kendall und R. Cipolla, “Modelling uncertainty in deep learning for camera relocalization”, *ArXiv preprint arXiv:1509.05909*, 2015.
- [56] A. Kendall, R. Cipolla u. a., “Geometric loss functions for camera pose regression with deep learning”, in *Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Bd. 3, Honolulu, Hawaii, USA, 2017, S. 8.
- [57] W. Maddern, G. Pascoe, C. Linegar und P. Newman, “1 Year, 1000km: The Oxford Robotcar Dataset”, *The International Journal of Robotics Research (IJRR)*, Bd. 36, Nr. 1, S. 3–15, 2017. eprint: <http://ijr.sagepub.com/content/early/2016/11/28/0278364916679498.full.pdf+html>.
- [58] J. N. Cederberg, *A course in modern geometries*. New York, NY: Springer, 2013, Springer Science & Business Media, ISBN: 978-0-387-98972-3.
- [59] T. Strauß, J. Ziegler und J. Beck, “Calibrating multiple cameras with non-overlapping views using coded checkerboard targets”, in *International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, Qingdao, China, 2014, S. 2623–2628.

- [60] E. Rublee, V. Rabaud, K. Konolige und G. Bradski, “Orb: An efficient alternative to SIFT or SURF”, in *International Conference on Computer Vision (ICCV)*, IEEE, Barcelona, Spain, 2011, S. 2564–2571.
- [61] D. G. Lowe, “Distinctive image features from scale-invariant keypoints”, *International Journal of Computer Vision*, Bd. 60, Nr. 2, S. 91–110, 2004, Springer, ISSN: 1573-1405.
- [62] M. Calonder, V. Lepetit, C. Strecha und P. Fua, “Brief: Binary Robust Independent Elementary Features”, in *Computer Vision – ECCV 2010*, K. Daniilidis, P. Maragos und N. Paragios, Hrsg., Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, S. 778–792.
- [63] S. Leutenegger, M. Chli und R. Y. Siegwart, “Brisk: Binary robust invariant scalable keypoints”, in *2011 International Conference on Computer Vision (ICCV)*, IEEE, Barcelona, Spain, 2011, S. 2548–2555.
- [64] R. I. Hartley und A. Zisserman, *Multiple View Geometry in Computer Vision*, Second. Cambridge University Press, ISBN: 0521540518, 2004, Cambridge University Press.
- [65] J. L. Bentley, “Multidimensional binary search trees used for associative searching”, *Communications of the ACM*, Bd. 18, Nr. 9, S. 509–517, 1975, ACM.
- [66] S. Agarwal, K. Mierle u. a. (2018). Ceres Solver, Adresse: <http://ceres-solver.org> (besucht am 05. 12. 2018).
- [68] J. Diebel, “Representing attitude: Euler angles, unit quaternions, and rotation vectors”, *Matrix*, Bd. 58, Nr. 15-16, S. 1–35, 2006.
- [69] S. Kreitlow, A. Brettschneider, C.-H. Jahn und U. Feldmann-Westendorff, “Etrs89/Utm–Der Bezugssystemwechsel und die Auswirkungen auf die Geodatenutzung”, *Kartographische Nachrichten*, Bd. 60, Nr. 4, S. 179–188, 2010.

-
- [70] M. A. Fischler und R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography”, *Communications of the ACM*, Bd. 24, Nr. 6, S. 381–395, 1981, ACM.
- [71] N. Sünderhauf und P. Protzel, “Switchable constraints for robust pose graph SLAM”, in *International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Vilamoura, Algarve, Portugal, 2012, S. 1879–1884.
- [72] J. Gräter, T. Strauss und M. Lauer, “Momo: Monocular motion estimation on manifolds”, in *International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, Yokohama, Japan, 2017, S. 1–6.
- [73] J. Gräter, A. Wilczynski und M. Lauer, “Limo: Lidar-Monocular Visual Odometry”, in *International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Madrid, Spain, 2018, S. 7872–7879.
- [74] T. D. Gillespie, *Fundamentals of vehicle dynamics*. Society of automotive Engineers Warrendale, PA, 1992, Bd. 400.
- [75] S. S. Skiena, *The algorithm design manual: Text*. London: Springer, 1998, Bd. 1, Springer Science & Business Media, ISBN: 978-1-84800-069-8.
- [76] R. Kümmerle, B. Steder, C. Dornhege, M. Ruhnke u. a., “On measuring the accuracy of SLAM algorithms”, *Autonomous Robots*, Bd. 27, Nr. 4, S. 387, 2009, Springer.
- [77] I. Cvišić, J. Česić, I. Marković und I. Petrović, “Soft-SLAM: Computationally efficient stereo visual simultaneous localization and mapping for autonomous unmanned aerial vehicles”, *Journal of field robotics*, Bd. 35, Nr. 4, S. 578–595, 2018.

- [78] K. Lenac, J. Ćesić, I. Marković und I. Petrović, “Exactly sparse delayed state filter on Lie groups for long-term pose graph SLAM”, *The International Journal of Robotics Research*, Bd. 37, Nr. 6, S. 585–610, 2018, SAGE Publications Sage UK: London, England.
- [79] A. Geiger, J. Ziegler und C. Stiller, “Stereoscan: Dense 3d reconstruction in real-time”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Baden-Baden, Germany, 2011, S. 963–968.
- [80] T. Sattler, W. Maddern, C. Toft, A. Torii u. a., “Benchmarking 6dof outdoor visual localization in changing conditions”, in *Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Bd. 1, Salt Lake City, Utah, USA, 2018.
- [81] R. A. Zia, “Tightly-Coupled Localization Using Point Features from Surround View”, MA. Thesis, Institut für Mess- und Regelungstechnik, Karlsruher Institut für Technologie (KIT), Karlsruhe, 2017.
- [82] H. Latégahn, *Mapping and Localization in Urban Environments Using Cameras*, Ser. Schriftenreihe / Institut für Mess- und Regelungstechnik, Karlsruher Institut für Technologie. Karlsruhe: KIT Scientific Publishing, 2014, ISBN: 9783731501350.
- [83] C. M. Bishop, *Pattern recognition and machine learning*, Ser. Information Science and Statistics. New York, NY: Springer, 2006, Softcover published in 2016.
- [84] J. Janosovits, “Entwicklung eines Verfahrens zur intelligenten Landmarkenauswahl anhand Befahrungsstatistiken und Bildmerkmalen”, MA. Thesis, Institut für Mess- und Regelungstechnik, Karlsruher Institut für Technologie (KIT), Karlsruhe, 2017.
- [85] F. Poggenhans, M. Schreiber und C. Stiller, “A universal approach to detect and classify road surface markings”, in *International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, Las Palmas de Gran Canaria, Island of Gran Canaria, 2015, S. 1915–1921.

- [86] Daimler. (2016). Mercedes-Benz Future Bus, Adresse: <https://www.daimler.com/innovation/autonomes-fahren/future-bus.html> (besucht am 09.05.2019).

Veröffentlichungen des Autors

- [5] M. Sons, M. Lauer, C. G. Keller und C. Stiller, “Mapping and localization using surround view”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Los Angeles, CA, USA, 2017, S. 1158–1163.
- [17] M. Sons, T. Strauss, C. G. Keller, M. Haueis u. a., “Vollautomatische Generierung langzeitstabiler Punktmerkmalskarten für die Lokalisierung mit Multikamerasystemen”, in *Workshop for Driving Assistance Systems and Autonomous Driving*, UniDAS, Walting, Germany, 2017.
- [29] L. P. N. Matias, M. Sons, J. R. Souza, D. F. Wolf u. a., “Veigan: Vectorial Inpainting Generative Adversarial Network for Depth Maps Object Removal”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Paris, France, Juni 2019, S. 310–316.
- [30] H. Hu, M. Sons und C. Stiller, “Accurate Global Trajectory Alignment using Poles and Road Markings”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Paris, France, Juni 2019.
- [34] M. Sons und C. Stiller, “Efficient Multi-Drive Map Optimization towards Life-long Localization using Surround View”, in *International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, Island of Maui, Hawaii, USA, 2018, S. 2671–2677.
- [51] J. Kümmerle, M. Sons, F. Poggenhans, T. Kühner u. a., “Accurate and Efficient Self-Localization on Roads using Basic Geometric Primitives”, in *International Conference on Robotics and Automation (ICRA)*, IEEE, Montreal, Canada, 2019, S. 5965–5971.

- [67] M. Sons, H. Lategahn, C. G. Keller und C. Stiller, “Multi trajectory pose adjustment for life-long mapping”, in *Intelligent Vehicles Symposium (IV)*, IEEE, Seoul, Korea, 2015, S. 901–906.