

# Vollautomatische Generierung langzeitstabiler Punktmerkmalskarten für die Lokalisierung mit Multikamerasystemen

Marc Sons<sup>\*</sup>, Tobias Strauß<sup>†</sup>, Christoph G. Keller<sup>‡</sup>, Martin Haueis<sup>§</sup>  
und Christoph Stiller<sup>¶</sup>

**Zusammenfassung:** Zentimetergenaue Lokalisierung ist für das hochautomatisierte Fahren fundamental. Videobasierte Ansätze, die eine vorab erzeugte Karte verwenden, ermöglichen eine Lokalisierung in Gebieten, in denen globale Satellitennavigationssysteme aufgrund von Abschattung und Mehrwegeausbreitung unzuverlässig sind. Jedoch hat eine einmalig erzeugte Karte nur eine begrenzte Gültigkeitsspanne, da sich die Umgebung aufgrund äußerer Einflüsse verändert. In dieser Arbeit wird eine Methode zur Erzeugung langzeitstabiler Punktmerkmalskarten für die zentimetergenaue Eigenpositionschätzung in sechs Freiheitsgraden präsentiert. Dazu werden ausschließlich mehrere am Fahrzeug angebrachte Kameras verwendet. Eine existierende Karte wird bei erneuter Befahrung zur Lokalisierung verwendet. Danach wird die Karte mit den aufgezeichneten Bildern der Wiederbefahrung aktualisiert und verbessert. Die iterative Integration neuer Befahrungen ermöglicht eine beliebig lange Gültigkeitsspanne der Karte und steigert die Güte der in ihr gespeicherten Informationen.

**Schlüsselwörter:** Kamerabasierte Lokalisierung, SLAM, Automatisches Fahren

## 1 Einleitung

Die Verwendung von Karten ist für hochautomatisierte Fahrzeuge ein unverzichtbarer Bestandteil. Heutzutage benötigen eine Vielzahl von Anwendungen, beispielsweise die Erkennung und Prädiktion anderer Verkehrsteilnehmer, die Umgebungsinterpretation, die Verhaltensentscheidung oder die Manöverplanung, Kartenwissen, um robust und zuverlässig ihre jeweilige Aufgabe bewältigen zu können [3]. Grundlage dafür ist die hochgenaue Schätzung der eigenen Position und Orientierung innerhalb der Karte. Hierzu werden üblicherweise inertielle Messeinheiten (IMU) und globale Satellitennavigationssysteme (GNSS) kombiniert. Diese Systeme sind besonders in urbanen Gebieten, aufgrund von Abschattung und Mehrwegeausbreitung, nicht zuverlässig. Aktuelle videobasierte Lokalisierungsansätze [1, 13], welche eine vorab erstellte Karte verwenden, haben gezeigt, dass auch in solchen Umgebungen eine hochgenaue Lokalisierung möglich ist. Jedoch führen Veränderungen der Umwelt, beispielsweise Witterung, Baustellen oder dynamische Objekte dazu,

---

<sup>\*</sup>sons@fzi.de, FZI Forschungszentrum Informatik, Karlsruhe

<sup>†</sup>strauss@kit.edu, Institut für Mess- und Regelungstechnik, Karlsruher Institut für Technologie (KIT)

<sup>‡</sup>christoph.g.keller@daimler.com, Daimler AG, Forschung & Entwicklung, Sindelfingen

<sup>§</sup>martin.haueis@daimler.com, Daimler AG, Forschung & Entwicklung, Sindelfingen

<sup>¶</sup>stiller@kit.edu, Institut für Mess- und Regelungstechnik, Karlsruher Institut für Technologie (KIT)

dass eine einmalig erzeugte Karte eine fehlerhafte Umgebungsrepräsentation bereitstellt. Je nach Jahreszeit und Umgebung führt dies bereits innerhalb weniger Wochen zu einer Verschlechterung oder zum Ausfall der Lokalisierung.

Die vorliegende Arbeit präsentiert eine Methode zur Erzeugung langzeitstabiler Punktmerkmalskarten für die zentimetergenaue Lokalisierung in sechs Freiheitsgraden, wobei ausschließlich am Fahrzeug befestigte Kameras verwendet werden. Die Karte speichert rekonstruierte 3D-Landmarken und zugehörige Bildmerkmale. Zur Lokalisierung werden den Kartenlandmarken Beobachtungen in den aktuellen Kamerabildern zugeordnet und darauf basierend die Fahrzeugposition und -orientierung hochgenau geschätzt. Basierend auf der Lokalisierung werden Merkmale aus den Bildern der Wiederbefahrung zu existierenden Kartenmerkmalen assoziiert. Die Selektion häufig wiedererkannter Merkmale sowie von Merkmalen aus aktuelleren Befahrungen, führt zu einer kontinuierlichen Aktualisierung und iterativen Verbesserung der Karte. Zusätzlich wird eine Datenstruktur vorgestellt, welche die kartierten Landmarken in Echtzeit bereitstellt, Landmarken mit höherer Relevanz bevorzugt und die Umgebung möglichst gleichverteilt abtastet.

## 2 Stand der Forschung

Videobasierte Lokalisierung auf Basis vorab erstellter Karten hat in den letzten Jahren zunehmend an Bedeutung gewonnen. Dabei haben sich unterschiedliche Ausprägungen für verschiedene Anwendungszwecke entwickelt. Eine wichtige Klasse sind die topometrischen Methoden [1, 11]. Die Karte speichert hier jeweils Bildsignaturen zu rekonstruierten Fahrzeugpositionen. Zur Lokalisierung wird die am besten übereinstimmende Kartenposition durch Vergleich der Signaturen bestimmt (engl. *place recognition*). Diese Methoden ermöglichen eine grobe Verortung innerhalb großer Karten in Echtzeit.

Für Autobahnen und Überlandstraßen eignen sich fahrbahnmarkierungsbasierte Ansätze [13, 4]. Diese stützen insbesondere die fahrbahnselektive sowie laterale Lokalisierung. In diesen Szenarien hat die Querlokalisierung eine hohe Bedeutung. Des Weiteren kann eine hohe Verfügbarkeit von Spurmarkierungen vorausgesetzt werden.

Diese Arbeit ordnet sich in die Klasse der punktmerkmalsbasierten Methoden zur hochgenauen Kartierung und Lokalisierung ein [1, 18, 20]. Wir erweitern die Ansätze von [1] um die Verwendung mehrerer Kameras zur Steigerung der Lokalisierungsgenauigkeit, -robustheit und -verfügbarkeit sowie um eine iterative Kartengenerierungsmethode, welche sukzessive neue Befahrungen in die Karte integriert. Dies steigert ihre Güte und ermöglicht eine lebenslange Verfügbarkeit (engl. *life-long mapping*).

## 3 Voraussetzungen

Betrachtet wird ein Fahrzeug, an dem mehrere Kameras befestigt sind, welche zeitlich synchrone Bilder aufzeichnen. Durch die vorausgesetzte Kalibrierung sind alle Kameras zueinander referenziert (Abb. 1(a)). Außerdem stellt die Kalibrierung eine Abbildungsfunktion

$$\pi(\mathbf{l}) = \mathbf{z} \tag{1}$$

für jede Kamera bereit, welche einen Objektpunkt  $\mathbf{l} \in R^3$  (im Folgenden als Landmarke bezeichnet) aus der Umgebung auf einen Punkt  $\mathbf{z} \in R^2$  in der Sensorebene abbildet (Abb.

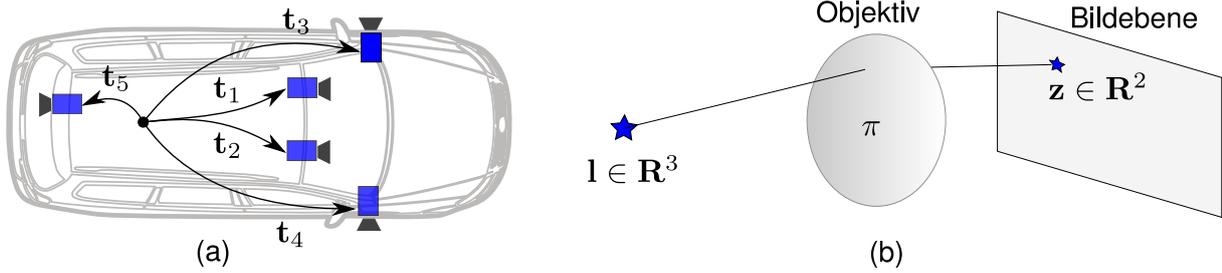


Abbildung 1: (a): Für jede Kamera existiert eine affine 3D-Transformation  $\mathbf{t}_i \in \mathbb{R}^6$ , welche die Pose der Kamera bezüglich eines gemeinsamen Ursprungs beschreibt. (b): Projektion einer Landmarke in die Bildebene.

1(b)). Details zur Kalibrierung von Multikamerasystemen finden sich in [16]. Zu bemerken ist, dass keine Annahmen zur Art der Kameras, Ausprägung der Abbildungsfunktion (1) sowie zur geometrischen Anordnung der Kameras getroffen werden. Jedoch begünstigen überlappende Sichtbereiche zwischen verschiedenen Kameras die Kartierung. Für die Lokalisierung ist es vorteilhaft, wenn die Kameras gemeinsam ein möglichst großes Sichtfeld um das Fahrzeug herum erzeugen (engl. *surround-view*).

### 3.1 Bildverarbeitung

Grundlage der Kartengenerierung ist die Assoziation von redetektierbaren Merkmalen zwischen verschiedenen Bildern (engl. *feature-matching*). Merkmale werden an markanten Bildpunkten erzeugt und dabei mittels eines Deskriptors beschrieben. Zur Extraktion markanter Bildpunkte wird ein feinskaliger Ecken-/Struktur-Detektor (engl. *corner/blob*) verwendet [18]. Als Deskriptor wird der DIRD-Deskriptor [2] verwendet, welcher wegen seiner Robustheit gegen Beleuchtungsveränderungen für die Assoziation von Punkten aus Bildern verschiedener Befahrungen besonders geeignet ist. Zur Kartengenerierung werden Merkmale zwischen zeitlich aufeinanderfolgenden Bildern der gleichen Kamera, zwischen Bildern verschiedener Kameras am Fahrzeug sowie zwischen geeigneten Bildern aus unterschiedlichen Befahrungen assoziiert. Zusätzlich werden alle Assoziationen global verfolgt und zusammengefasst (engl. *feature-tracking*). Während der Lokalisierung werden Kartenmerkmale mit Merkmalen aktueller Kamerabilder assoziiert.

## 4 Kartengenerierung

In diesem Abschnitt wird die iterative Erzeugung der Karte aus mehreren Befahrungen einer Strecke beschrieben. Dafür sind gespeicherte Bilder, Posen, Landmarken sowie ihre zugehörigen Merkmalskorrespondenzen der zuvor bereits gemeinsam kartierten Befahrungen verfügbar. Wurde eine Strecke noch nicht initial befahren, wird eine Karte aus einer einzelnen Befahrung gemäß Abschnitt 4.2 erzeugt. Diese Karte unterscheidet sich strukturell nicht von der Mehrfachbefahrungskarte.

Jede aufgezeichnete Menge von Bildern zu einem Zeitpunkt korrespondiert zu einer Fahrzeugpose  $\mathbf{p} \in \mathbb{R}^6$  in der Karte. Eine Pose ist eine affine 3D-Transformation, welche die Orientierung und Position des Fahrzeugs bezüglich des Kartenursprungs beschreibt. Jede Merkmalskorrespondenz  $Z = \{\mathbf{z}_1, \dots, \mathbf{z}_n\}$  definiert eine Landmarke  $\mathbf{l}_i \in \mathbb{R}^3$  in der Karte.

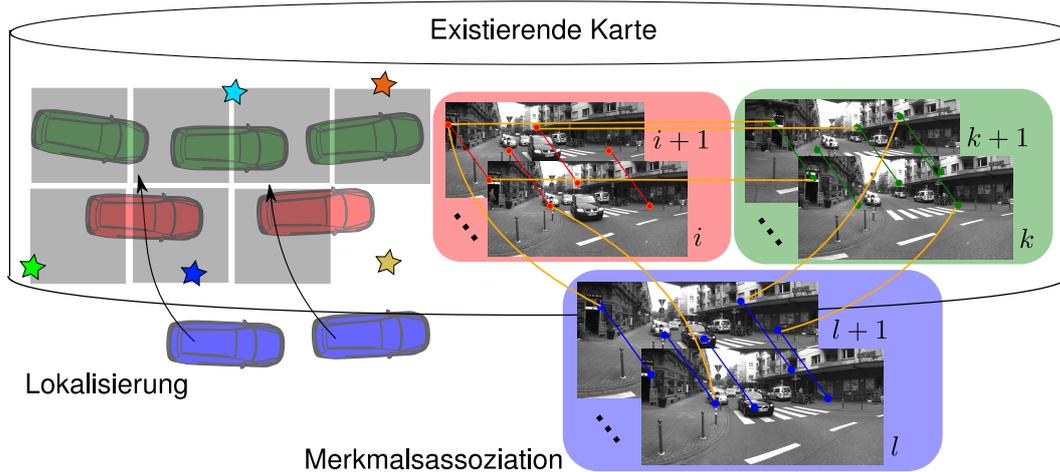


Abbildung 2: Referenzierung einer neuen Befahrung (blau) mit der existierenden Karte. Die neue Befahrung wird initial referenziert und darauf basierend Merkmalskorrespondenzen zwischen Karte (bestehend aus zwei Befahrungen; rot, grün) und neuer Befahrung bestimmt.

Alle Landmarken sind bezüglich des Kartenursprungs verortet. Ein Merkmal  $\mathbf{z}_i \in \mathcal{Z}$  setzt sich zusammen aus einem Bildpunkt  $\mathbf{b}_i$  sowie seinem zugehörigen Deskriptor  $\mathbf{d}_i$ . Durch Gleichung (1) ergibt sich der Zusammenhang einer Fahrzeugpose  $\mathbf{p}$ , einer Landmarke  $\mathbf{l}$ , sowie einem Bildpunkt  $\mathbf{b}$

$$\pi((\mathbf{p}^{-1} \otimes \mathbf{t}) \cdot \mathbf{l}) = \mathbf{b}. \quad (2)$$

Zunächst wird die Landmarke in das kamerafeste Koordinatensystem transformiert und dann in die Bildebene abgebildet (siehe Abb.1). Dabei beschreibt  $\otimes : R^6 \times R^6 \rightarrow R^6$  die Konkatenation zweier Transformationen und  $\cdot : R^6 \times R^3 \rightarrow R^3$  die Basistransformation der Landmarke  $\mathbf{l}$  vom Karten- in das kamerafeste Koordinatensystem.

## 4.1 Referenzierung der neuen Befahrung

Bei erneuter Befahrung der bereits kartierten Strecke lokalisiert sich das Fahrzeug in die bestehende Karte. Zusätzlich werden alle aufgezeichneten Bilder gespeichert. Anhand der Lokalisierungsposen werden örtlich nahe Kartenposen bestimmt und Posendifferenzen zwischen diesen abgeleitet. Mittels visueller Odometrie [14] werden zusätzlich zeitlich aufeinanderfolgende Posendifferenzen der neuen Befahrung geschätzt. Die rekonstruierten Posendifferenzen werden mittels robusten Posenausgleichs [8, 17] initial zur bestehenden Karte referenziert. Somit werden auch bisher unkartierte Gebiete, welche in der erneuten Befahrung durchfahren wurden, in die Karte integriert.

Auf Basis der initialen Referenzierung werden Bilder der neuen Befahrung mit Bildern bereits kartierter Befahrungen assoziiert, welche die gleiche Szene aus einer ähnlichen Perspektive abbilden. Zwischen diesen Bildern sowie zeitlich aufeinanderfolgenden Bildern der neuen Befahrung werden gemäß Abschnitt 3.1 Merkmale assoziiert und diese zu den existierenden Merkmalskorrespondenzen der Karte hinzugefügt. Abb. 2 zeigt dieses Vorgehen schematisch.

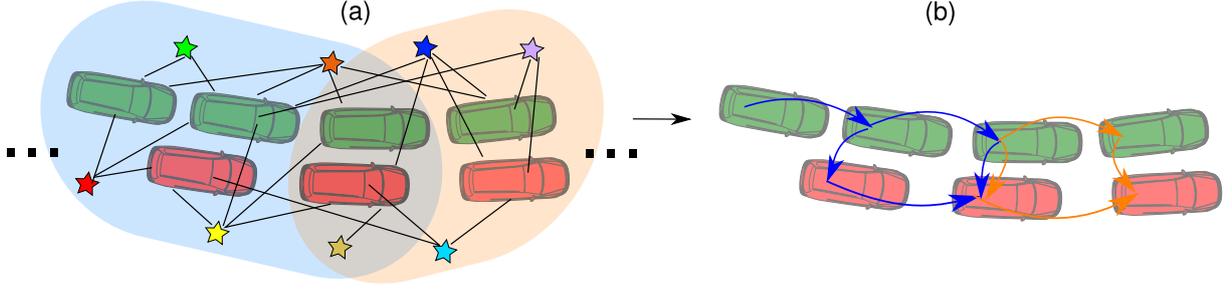


Abbildung 3: Darstellung der Optimierung der Kartenparameter für zwei Befahrungen (rot und grün) (a): Die Kanten zeigen den Zusammenhang zwischen Landmarken (Sterne), Fahrzeugposen und Landmarkenbeobachtungen. Jede Kante erzeugt einen Kostenterm in Problem (3). Durch Aufteilung in Fenster werden zwei kleinere Probleme betrachtet. (b): Die abgeleiteten Posendifferenzen aus den Optimierungsfenstern sind als gerichtete Kanten dargestellt. Im überlappenden Bereich der Fenster entstehen Doppelkanten. Der resultierende Posenausgleich hat deutlich weniger Parameter.

## 4.2 Kartenverbesserung

Dieser Abschnitt erläutert die Verbesserung der Posen und Landmarkenschätzung auf Basis aller Merkmalskorrespondenzen gemäß vorherigem Abschnitt 4.1. Im Folgenden sei  $P \subset R^6$  die Menge aller Fahrzeugposen,  $T$  die Menge aller Merkmalskorrespondenzen sowie  $L \subset R^3$  die Menge aller Landmarken ( $|T| = |L|$ ). Basierend auf Gleichung (2) beschreibt

$$\hat{P}, \hat{L} = \arg \min_{P,L} \left\| \sum_{\mathbf{l}_i \in L} \sum_{\mathbf{z}_{j,k} \in Z_i} \pi_k((p_j^{-1} \otimes T_k) \cdot \mathbf{l}_i) - \mathbf{b}_{j,k} \right\|_{\Omega}^2 \quad (3)$$

die Schätzung der optimalen Fahrzeugosen  $\hat{P}$  sowie Landmarken  $\hat{L}$ , welche den mittleren Rückprojektionsfehler aller Beobachtungen in  $T$  minimiert. Dabei wird der Bildpunkt  $\mathbf{b}_{j,k}$  von Kamera  $(\pi_k, T_k)$  in Fahrzeugpose  $p_j \in P$  erfasst. Des Weiteren korrespondiert Landmarke  $\mathbf{l}_i \in L$  zur Merkmalskorrespondenz  $Z_i \in T$ . Zur Lösung von Problem (3) werden robuste, nichtlineare Bündelausgleichsmethoden verwendet. Details dazu finden sich in [7, 15].

Aufgrund begrenzter Rechen- und Speicherkapazitäten kann  $\hat{P}$  und  $\hat{L}$  in der Praxis nicht gemeinsam gebündelt werden. Daher wird die befahrene Strecke in räumlich überlappende Fenster eingeteilt (siehe Abb. 3(a)). Zur Erzeugung dieser Fenster wird die Karte als Posengraph  $G$  abgebildet (siehe Abb. 3(b)) [8]. Dabei werden nur Kanten zwischen Posen eingefügt, zwischen denen Bildmerkmale assoziiert wurden. Dieser Graph wird daraufhin in disjunkte Partitionen unterteilt. Durch Vergrößerung dieser Partitionen entlang äußerer Kanten entstehen überlappende Bereiche. Die Größe der Fenster legt dabei die Anzahl der gemeinsam zu optimierenden Parameter fest. So können die Posen und Landmarken aus jedem Fenster mittels (3) separat gebündelt werden (Abb. 3(a)). Danach werden Posendifferenzen aus den einzelnen Fenstern abgeleitet und mit diesen gemeinsam die finalen Kartenposen  $\hat{P}$  in einem robusten Posenausgleich geschätzt. Die Posendifferenzen übertragen die Verbesserungen aus der gefensterten Bündeloptimierung in den finalen Posenausgleich. Da sich die Fenster überlappen, entstehen in den Überlappungsbereichen Mehrfachkanten. Abb. 3 zeigt das beschriebene Vorgehen schematisch. Im Gegensatz zu

(3), skaliert der Posenausgleich nur mit der Anzahl der Posendifferenzen, wodurch deutlich mehr Posen zusammen optimiert werden können (Abb. 3(b)). Die optimierten Kartenposen  $\hat{P}$  werden daraufhin als konstant betrachtet, wodurch jede Landmarke  $\mathbf{l} \in L$  separat rekonstruiert werden kann.

### 4.3 Landmarkenselektion

Durch die iterative Hinzunahme neuer Befahrungen in die Karte wächst die Anzahl der Landmarken stetig. Daher wird nur eine Teilmenge aller Landmarken in der finalen Kartenstruktur abgelegt. Dabei werden Landmarken mit hoher Wiedererkennungswahrscheinlichkeit bevorzugt. Ein direktes Maß für diese Wahrscheinlichkeit ist die Anzahl  $R$  der Befahrungen über die sich die Beobachtungskorrespondenz einer Landmarke erstreckt. Mit steigender Anzahl  $N$  an Kartierungsbefahrungen nähert sich die Wiedererkennungsrate  $r = \frac{R}{N}$  der Wahrscheinlichkeit an, dass die Landmarke in der nächsten Befahrung erneut detektiert wird und somit die Lokalisierung stützt. Außerdem haben Landmarken mit Beobachtungen aus jüngeren Befahrungen höhere Wiedererkennungswahrscheinlichkeiten als ältere. Das Alter  $A$  einer Landmarke ist dabei das Alter ihrer jüngsten Beobachtung, wobei  $A \in \{1, \dots, N\}$  der Index der Befahrung ist. Hier ist angenommen, dass die Fahrten zeitlich absteigend sortiert sind. Die resultierende Wiedererkennungswahrscheinlichkeit ergibt sich durch

$$p = \frac{\omega \cdot R + (1 - \omega) \cdot (N - A)}{N}, \quad 0 \leq \omega \leq 1, \quad (4)$$

wobei durch  $\omega$  Wiedererkennungsrate und Alter unterschiedlich gewichtet werden können. Zusätzlich muss sichergestellt werden, dass im gesamten kartierten Gebiet ausreichend Landmarkenbeobachtungen vorhanden sind.

Daher werden die Landmarken in eine zweidimensionale Rasterkarte einsortiert. Dabei wird angenommen, dass sich das Fahrzeug näherungsweise in einer Ebene bewegt. Die Landmarken werden dabei anhand der Position ihrer Beobachtungen in der Befahrungsebene einsortiert (siehe Abb. 4(a)). Jede Zelle der Rasterkarte speichert  $n_c \geq 0$  Landmarken. Um eine lokale Beobachtungsdichte sicherzustellen, wird in jeder Zelle eine Mindestanzahl  $m$  an Landmarken beibehalten. Falls  $n_c > m$  werden die  $n_c - m$  Landmarken mit geringster Wiedererkennungswahrscheinlichkeit  $p$  entfernt. Die Größe der Zellen,  $m$  sowie  $\omega$  sind wählbare Parameter.

## 5 Kartenstruktur

Zur Lokalisierung wird in jedem Schritt zunächst eine grobe Initialschätzung  $\mathbf{p}_i$  der aktuellen Fahrzeugpose bestimmt. Damit können Landmarken aus der Karte ausgewählt werden, welche örtlich nahe Merkmale zu  $\mathbf{p}_i$  haben. Dadurch kann angenommen werden, dass die ausgewählten Landmarken aus einer ähnlichen Perspektive wie zur Kartierung gesehen werden. Dies ermöglicht die robuste Assoziation von Landmarken zu aktuellen Bildmerkmalen. Durch Minimierung des Rückprojektionsfehlers der assoziierten Landmarken wird schließlich die aktuelle Fahrzeugpose hochgenau geschätzt [2]. In diesem Abschnitt wird eine Rasterkartenstruktur vorgestellt, welche die Kartenlandmarken effizient der Lokalisierung bereitstellt.

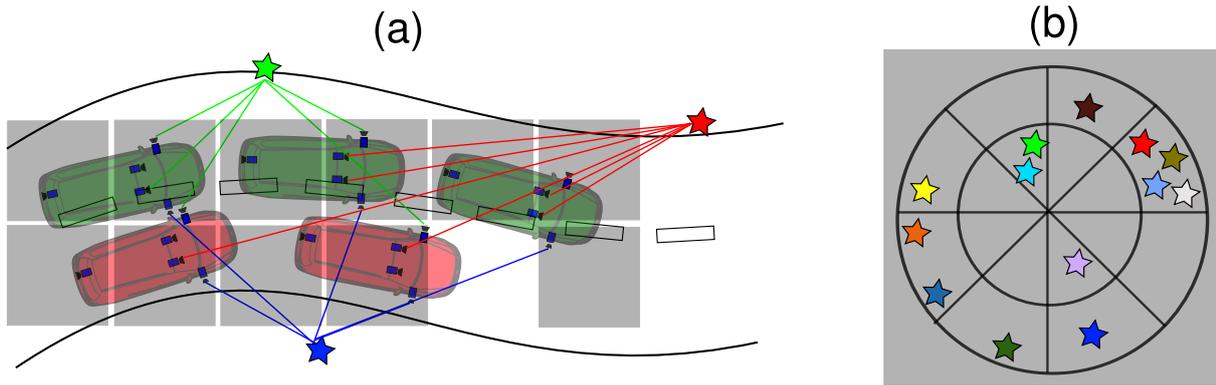


Abbildung 4: Darstellung der Kartenstruktur basierend auf zwei Kartierungsbefahrungen (rot, grün). (a) Landmarken (Sterne) werden ausschließlich anhand ihrer Beobachtungspositionen (unabhängig von Kamera und Befahrung) in Zellen (graue Kacheln) einsortiert. (b) Innerhalb einer Zelle werden alle eingefügten Landmarken und Merkmale anhand der Position der Landmarke in einem polaren Raster abgelegt.

Hierzu werden die gemäß Abschnitt 4.3 selektierten Landmarken in Zellen einer zweidimensionalen Rasterkarte eingefügt. Jede Zelle des Rasters kann separat in den Speicher geladen werden, wodurch die Größe der Karte lediglich durch den verfügbaren Festplattenspeicher begrenzt ist. Die Größe einer Zelle ist ein wählbarer Parameter. Wie in Abb. 4(a) dargestellt, werden i. A. mehrere Beobachtungen einer Landmarke in einer Zelle sowie eine Landmarke in mehreren Zellen gespeichert. Zur Lokalisierung werden die in der Nähe der Initialisierungspose  $\mathbf{p}_i$  liegenden Zellen bereitgestellt.

Innerhalb einer Zelle werden die Landmarken sowie zugehörige Beobachtungen anhand der Position der Landmarke in ein polares Raster einsortiert und innerhalb einer polaren Region anhand ihrer Wiedererkennungswahrscheinlichkeit  $p$  (siehe Abschnitt 4.3) sortiert (Abb. 4(b)). Zur Lokalisierung wird das polare Raster sequentiell durchlaufen. Die Winkel- und Distanzauflösung sind wählbare Parameter. Wurden  $c \leq m$  der  $m$  Landmarken einer Region mit einem Bildmerkmal assoziiert, wird die nächste Region betrachtet. Durch die Sortierung der Landmarken innerhalb einer Region werden Landmarken mit hoher Wiedererkennungswahrscheinlichkeit vorrangig assoziiert. Des Weiteren werden durch das polare Raster alle Raumrichtungen und Distanzen gleichverteilt abgetastet. Eine möglichst gleich verteilte Abtastung von Beobachtungen wirkt sich dabei positiv auf die Robustheit und Genauigkeit der Lokalisierung aus. Zu bemerken ist, dass auf diese Weise eine optimierte Landmarkenauswahl und gleichzeitig eine durch  $c$  sowie der Zellengröße beschränkte Laufzeit garantiert wird.

## 6 Experimente

Zur Evaluierung wurden Bilder aus fünf Kameras mit einer Frequenz von 10 Hz aufgezeichnet. Die Kameras sind ähnlich zu Abb. 1 am Fahrzeug angebracht. Die vorwärts sowie rückwärts gerichteten Kameras haben einen Öffnungswinkel von ca.  $110^\circ$ . Die beiden seitlichen Kameras sind mit Weitwinkel-Objektiven ausgestattet und haben einen Öffnungswinkel von ca.  $175^\circ$ . Nach Entzerrung haben die Bilder eine durchschnittliche Größe von ca.  $1500 \times 500$  Pixel.

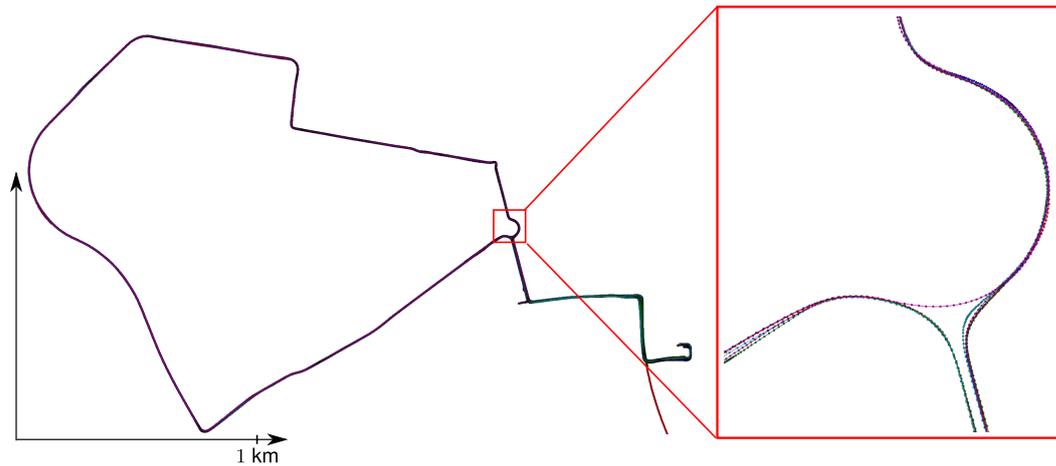


Abbildung 5: Darstellung der ca 6 km langen Evaluierungsstrecke. Es wurden vier Fahrten gebündelt, welche an unterschiedlichen Orten starten und enden. Die Detailansicht des Kreisverkehrs zeigt wie Befahrungen aus unterschiedlichen Richtungen zusammenführen.

Eine ca 6 km lange Strecke wurde insgesamt fünf mal an verschiedenen Tagen, zu unterschiedlichen Tageszeiten und Witterungsbedingungen befahren. Vier Befahrungen wurden zur Kartenerstellung verwendet sowie eine zur Evaluierung der Lokalisierung. Alle Befahrungen wurden innerhalb eines Zeitraumes von ca. zwei Monaten durchgeführt (siehe Abb. 5). Für die gefensterte Bündeloptimierung gemäß Abschnitt 4.2 wurden Fenster mit 25 Posen bestimmt. Der Überlappungsbereich zwischen den Fenstern betrug 3 Kanten.

Abb. 6 zeigt qualitativ die Verteilung von häufig wiedererkannten Landmarken. Gut sichtbare Hausfassaden wurden in vielen Befahrungen wiedererkannt. Hingegen wurden Schattenkanten, Beleuchtungsartefakte und umgeparkte Autos überhaupt nicht wiedererkannt. Aufgrund des relativ kurzen Evaluationszeitraums sind auch Teile der Baumkronen häufig redaktiert worden. Abb. 6 zeigt exemplarisch, dass lediglich durch Auswertung von Wiedererkennungssstatistiken, persistente Strukturen identifiziert werden können, welche zur Lokalisierung besonders geeignet sind.

Das rote Histogramm in Abb. 7 zeigt die Landmarkenverteilung über der Anzahl der Wiedererkennungen in verschiedenen Befahrungen vor Anwendung der in Abschnitt 4.3 beschriebenen Selektionsstrategie. In blau und grün sind die Verteilungen nach der Selektion für unterschiedliche Parameter dargestellt. Die Zellgröße betrug dabei 0.3m und es wurden pro Zelle maximal  $m = 270$  (blaue Verteilung in Fig. 7) beziehungsweise  $m = 90$  (grüne Verteilung in Fig. 7) beibehalten. Der Parameter  $\omega$  wurde aufgrund des kurzen Evaluationszeitraums auf 1 gesetzt, so dass ausschließlich die Redaktionsrate als Auswahlkriterium verwendet wurde. Abb. 7 zeigt deutlich den positiven Effekt der Selektion. Es verbleiben relativ deutlich mehr Landmarken mit hoher Wiedererkennungswahrscheinlichkeit in der Karte. Zusätzlich haben Lokalisierungsexperimente gezeigt, dass in allen Bereichen der Karte eine ausreichende Anzahl an Landmarken erhalten worden ist.

Die fünfte Befahrung wurde zur Evaluierung der Lokalisierung verwendet. Da im kartierten Gebiet keine hochgenauen GPS-Messungen verfügbar sind, wurden lokalisierungsspezifische Kenngrößen, beispielsweise die Anzahl von lokalisierungsstützenden Landmarken-Merkmalsskorrespondenzen (engl. *inlier*) oder das lokale Winkel- und Positionsrauschen ausgewertet. Zusätzlich wurde anhand von Schwellwertvergleichen dieser Kenngrößen so-

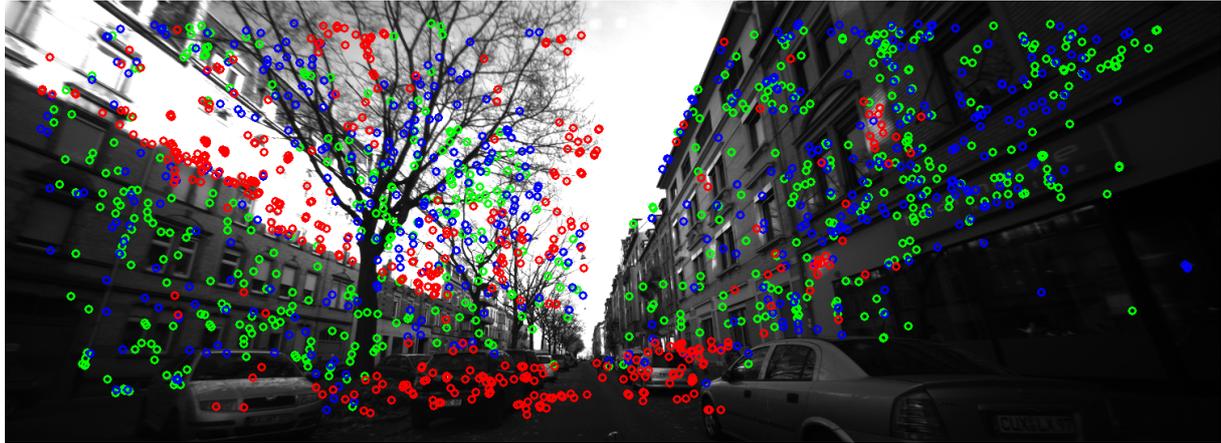


Abbildung 6: Projektion von kartierten Landmarken in das Kamerabild. Die Farbe der Landmarken kodiert die Anzahl der Befahrungen in denen die Landmarke wiedererkannt wurde (grün:  $\geq 3$ , blau 2, rot 1). Parkende Autos, Schattenkanten und andere Beleuchtungsartefakte sowie Teile der Baumkronen wurden nicht wiedererkannt. Im Gegensatz liegen auf Hausfassaden viele Landmarken mit hoher Wiedererkennungswahrscheinlichkeit.

wie weiteren Plausibilitätsüberprüfungen entschieden, ob die Lokalisierung zu einem Zeitschritt erfolgreich war oder nicht. Die Kenngrößen und Schwellwerte basieren dabei auf Erfahrungswerten und Erkenntnissen aus existierenden Arbeiten auf diesem Gebiet [1, 2]. Dabei hat sich gezeigt, dass unabhängig von der Anzahl der Kartierungsbefahrungen eine ähnliche Anzahl erfolgreicher Lokalisierungsschritte erzielt wurde. Außerdem hat sich gezeigt, dass in Gebieten mit überwiegender Anzahl an Landmarken mit hoher Wiedererkennungswahrscheinlichkeit, die Genauigkeit der Lokalisierung gesteigert werden konnte.

## 7 Zusammenfassung und Ausblick

In dieser Arbeit wurde eine Methode zur vollautomatischen Generierung langzeitstabiler Punktmerkmalskarten vorgestellt. Es wurde gezeigt wie die sukzessive Integration neuer Befahrungen und Auswertung von Wiedererkennungsstatistiken die Qualität der Karte verbessert. Zusätzlich wurde eine Datenstruktur präsentiert, welche die Lokalisierung mit mehreren Kameras in Echtzeit ermöglicht und dabei Landmarken mit höherer Detektionswahrscheinlichkeit bevorzugt.

Der Prozess lässt sich ohne weiteres auf die Verwendung von Fahrzeugflotten erweitern. So ist es denkbar, dass mehrere Fahrzeuge die gleiche Kartenbasis zur Lokalisierung verwenden. Nach ihrer Rückkehr können die Befahrungen jedes Fahrzeugs in die Karte integriert werden. Durch die vollautomatische Aktualisierung, Erweiterung und Verbesserung der existierenden Karte würde allen Fahrzeugen der Flotte stets eine robuste und hochgenaue Lokalisierung zur Verfügung stehen.

Insgesamt zeigt sich aber auch schon nach kurzen Evaluationszeiträumen mit relativ wenigen Befahrungen eine qualitative Verbesserung der Karte.

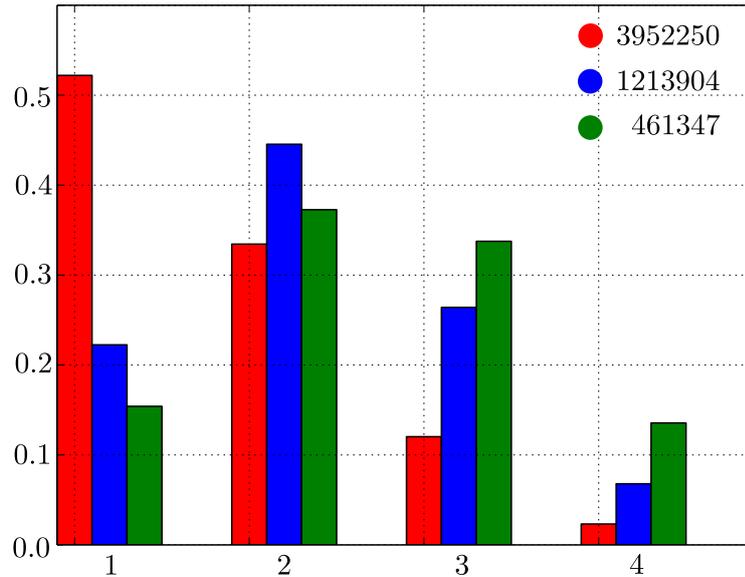


Abbildung 7: Relative Landmarkenverteilung über die Anzahl von Wiedererkennungen in verschiedenen Befahrungen. Die rote Verteilung zeigt die ursprüngliche Verteilung. Die grüne und blaue Verteilung zeigen die Landmarkenverteilung nach der Selektion mit unterschiedlichen Parametern. Größe der Zellen: 0.3m.  $\omega = 1$ . Anzahl erhaltener Landmarken pro Zelle:  $m = 270$  (blau),  $m = 90$  (grün). Bereits nach wenigen Befahrungen zeigt die Karte eine deutlich verbesserte Landmarkenverteilung wobei eine ausreichende räumliche Beobachtungsverteilung erhalten wurde. Zusätzlich wurde die Anzahl an Landmarken deutlich reduziert. So wird zusätzlich die Speicher- und Rechenkomplexität zur Lokalisierung verringert.

## Literatur

- [1] H. Lategahn and C. Stiller, *Vision-only localization*, IEEE Transactions on Intelligent Transportation Systems, 15(3), 1246-1257, 2014
- [2] H. Lategahn, J. Beck and C. Stiller, *Dird is an illumination robust descriptor* In 2014 IEEE Intelligent Vehicles Symposium Proceedings (pp. 756-761), IEEE, 2014
- [3] J. Ziegler, P. Bender, M. Schreiber, H. Lategahn, T. Strauss, C. Stiller, ... and E. Kaus, *Making bertha drive-an autonomous journey on a historic route*, IEEE Intelligent Transportation Systems Magazine, 6(2), 8-20, 2014
- [4] M. Schreiber, A. M. Hellmund and C. Stiller, *Multi-drive feature association for automated map generation using low-cost sensor data*, In Intelligent Vehicles Symposium (IV), IEEE (pp. 1140-1147), 2015
- [5] T. Bailey, and H. Durrant-Whyte, *Simultaneous localization and mapping (SLAM): Part II*, IEEE Robotics Automation Magazine, 13(3), 108-117, 2006
- [6] M. Bosse, P. Newman, J. Leonard and S. Teller, *Simultaneous localization and map building in large-scale cyclic environments using the Atlas framework*, The International Journal of Robotics Research, 23(12), 1113-1139, 2004

- [7] G. Grisetti, H. Strasdat, K. Konolige, W. Burgard, *g2o: A general framework for graph optimization*, In IEEE International Conference on Robotics and Automation, 2011
- [8] M. Sons, H. Lategahn, C. G. Keller and C. Stiller, *Multi trajectory pose adjustment for life-long mapping*, In 2015 IEEE Intelligent Vehicles Symposium (IV) (pp. 901-906), IEEE, 2015
- [9] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge university press, 2003
- [10] B. Kim, M. Kaess, L. Fletcher, J. Leonard, A. Bachrach, N. Roy and S. Teller, *Multiple relative pose graphs for robust cooperative mapping*, In Robotics and Automation (ICRA), 2010 IEEE International Conference on (pp. 3185-3192), IEEE, 2010
- [11] H. Badino, D. Huber and T. Kanade, *Visual topometric localization*, In Intelligent Vehicles Symposium (IV), 2011 IEEE (pp. 794-799), IEEE, 2011
- [12] J. McDonald, M. Kaess, C. Cadena, J. Neira and J. J. Leonard *6-DOF multi-session visual SLAM using anchor nodes*, 2011
- [13] M. Schreiber, C. Knöppel and U. Franke, *Laneloc: Lane marking based localization using highly accurate maps*, In Intelligent Vehicles Symposium (IV), 2013 IEEE (pp. 449-454), IEEE, 2013
- [14] F. Fraundorfer and D. Scaramuzza, *Visual odometry: Part II: Matching, robustness, optimization, and applications*, IEEE Robotics & Automation Magazine, 19(2), 78-90, 2012
- [15] S. Agarwal, N. Snavely, M. S. Seitz and R. Szeliski, *Bundle adjustment in the large*, In European conference on computer vision (pp. 29-42), Springer Berlin Heidelberg, 2010
- [16] T. Strauss, *Kalibrierung von Multi-Kamera-Systemen-Kombinierte Schätzung von intrinsischem Abbildungsverhalten der einzelnen Kameras und deren relativer Lage zueinander ohne Erfordernis sich überlappender Sichtbereiche*, Doctoral dissertation, Karlsruhe, Karlsruher Institut für Technologie (KIT), Diss., 2015
- [17] N. Sünderhauf and P. Protzel, *Switchable constraints for robust pose graph SLAM*, In Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on (pp. 1879-1884), IEEE, 2012
- [18] A. Geiger, J. Ziegler and C. Stiller, *Stereoscan: Dense 3d reconstruction in real-time*, In Intelligent Vehicles Symposium (IV), 2011 IEE (pp. 963-968), IEEE, 2011
- [19] M. Agrawal and K. Konolige, *Real-time localization in outdoor environments using stereo vision and inexpensive gps*, In Pattern Recognition, ICPR 2006, 18th International Conference on, vol.3 (pp. 1063-1068), IEEE, 2006
- [20] H. Chiu, et al. *Sub-meter vehicle navigation using efficient pre-mapped visual landmarks*, In Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on (pp. 505-512), IEEE, 2016