

Filtern bewegter Objekte aus mobilen LiDAR-Daten auf Grundlage volumetrischer Modelle

JOACHIM GEHRUNG^{1,2}, MARCUS HEBEL¹, MICHAEL ARENS¹ & UWE STILLA²

Zusammenfassung: Die automatische Änderungsdetektion in mobil durchgeführten Lasermessungen hat das Potential, ein wichtiges Werkzeug für die Aktualisierung von Stadtmodellen zu werden. Eine Voraussetzung ist das Entfernen von mobilen Objekten, welche zusammen mit der Bausubstanz in den Messdaten enthalten ist. Im Rahmen dieser Arbeit wird ein Ansatz vorgestellt, der sich eine Volumenrepräsentation zunutze macht, um auf Basis von Einzelmessungen bewegte mobile Objekte zu erkennen und vom statischen Hintergrund zu separieren. Die Methode wurde anhand zweier Datensätze einer qualitativen Untersuchung unterzogen. Die Resultate zeigen, dass eine hohe Anzahl der Oberflächenpunkte aller bewegten Objekte sowie des Hintergrunds korrekt erfasst wird.

1 Einleitung

Heutige Mobile-Mapping-Systeme verfügen neben Kameras in der Regel auch über aktive Lasersensoren bzw. Laserscanner zur direkten 3D-Erfassung der Umgebung. Mit derartigen Systemen durchgeführte Messungen verfügen über eine hohe Genauigkeit und 3D-Punktdichte und eignen sich daher hervorragend für die Änderungsdetektion urbaner Strukturen, wie sie beispielsweise für die automatische Aktualisierung von Stadtmodellen benötigt wird. Die mit mobilen Laserscannern (MLS) in urbanen Gebieten durchgeführten Aufnahmen enthalten neben der zu erfassenden Bausubstanz jedoch auch eine Reihe von mobilen Objekten. Hierbei handelt es sich beispielsweise um Fußgänger, Fahrräder, aber auch abgestellte PKWs, die alle nicht für die Stadtmmodellierung von Interesse sind und daher separiert werden müssen. Ein erster Schritt in diese Richtung ist die Entfernung bewegter Objekte, die eine vergleichsweise einfach zu detektierende Sonderklasse der mobilen Objekte darstellen. Gerade diese haben beim Einsatz von Laserscannern einen großen negativen Einfluss auf die Messergebnisse, da sie zu Messartefakten und Mehrfachbeobachtungen führen, die sich über die Zeit hinweg akkumulieren und so einen Schweif entlang der Bewegungstrajektorie des Objekts bilden (vergleiche „rot“ in Abb. 1). In belebten Innenstädten ist dieser Sachverhalt aufgrund von Fußgängern und sich bewegendem Fahrzeuge besonders ausgeprägt und stellt einen großen Störfaktor für die automatische Änderungsdetektion bzgl. der Bausubstanz dar.

Aktuelle Ansätze zur Identifikation mobiler Objekte unterscheiden entweder sehr grob zwischen diesen und dem Hintergrund oder aber arbeiten mit einem kurzen Zeithorizont, welcher zu einer fehlerhaften Segmentierung führen kann, sobald ein Objekt für eine längere Zeitdauer seinen Standort hält. Im Rahmen dieses Papers wird ein Ansatz vorgestellt, welcher eine volumenbasierte

¹ Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB, Gutleuthausstraße 1, D-76275 Ettlingen, E-Mail: [joachim.gehrung, marcus.hebel, michael.arens]@iosb.fraunhofer.de

² Photogrammetrie und Fernerkundung, Technische Universität München, Arcisstr. 21, 80333 München, E-Mail: stilla@tum.de

Hintergrundrepräsentation der Umgebung dazu nutzt, die in deren Erstellung eingegangenen Einzelmessungen möglichst präzise in statischen Hintergrund und bewegte Objekte zu segmentieren. Der hierbei berücksichtigte Zeithorizont umfasst die komplette Messung, sodass auch über längere Zeitabschnitte hinweg stehende Objekte nach wie vor als bewegte Objekte erkannt werden.

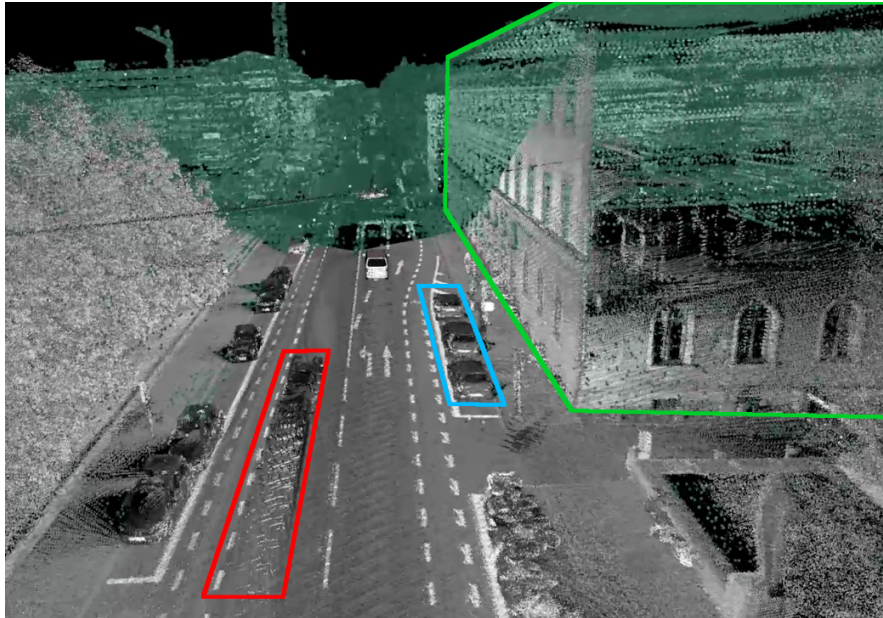


Abb. 1: Eine Beispielszene mit statischen Bausubstanzen (grün), statischen mobilen Objekten (blau) und bewegten mobilen Objekten (rot). Die Mehrfachbeobachtung (rot) ist durch das Akkumulieren mehrerer Einzelmessungen entstanden.

2 Stand der Forschung

2.1 Aufbereitung von Punktwolken

Die Unterscheidung zwischen dynamischen und statischen Strukturen ist in vergleichbarer Form in verschiedenen Bereichen der Forschung relevant. Viele vom Sinn her thematisch ähnliche Arbeiten sind im maschinellen Sehen angesiedelt, konkret im Bereich der Objektverfolgung auf zweidimensionalen Bildern. Im Hinblick auf 3D-Daten sind Parallelen zur Detektion und Verfolgung von mobilen Objekten (engl.: Detection and Tracking of Mobile Objects, DATMO) zu erkennen. Hierbei handelt es sich um einen Themenbereich, der vor allem in der Robotik und dem autonomen Fahren untersucht wird.

Im Kontext von letzterem identifizieren (VALLET et al. 2015) mit einem auf der Dempster-Shafer-Evidenztheorie basierten Ansatz mobile Objekte durch die schritthaltende Suche von Konflikten über mehrere Einzelmessungen hinweg. Die Anzahl der für den Vergleich gewählten Epochen ist dabei relativ gering, sodass nur ein relativ kurzer Zeithorizont abgedeckt wird. Da im Rahmen dieser Arbeit in erster Linie langfristig statische Strukturen gesucht werden, ist das Verfahren daher für die angestrebte Anwendung nur bedingt geeignet.

AZIM & AYCARD (2012) erzeugen eine Repräsentation der Umgebung auf Basis eines octree-basierten 3D-Belegungsgitters. Anhand dieser werden Einzelmessungen im Hinblick auf Konflikte

überprüft. Anhand dieses Vorgehens wird sukzessiv eine Liste von Voxeln aktualisiert, welche in weiterführenden Verarbeitungsschritten zu Objekthypothesen zusammengefasst werden. Das Verfahren ist laut den Autoren anfällig für Rauschen, welches durch ein Schwellwertverfahren auf Ebene der Objekthypothesen oder in späteren Schritten zu beheben ist. Da sich auch dieses Verfahren schnell an Veränderungen in der Umgebung anpasst, ist es ebenfalls nur unzureichend zur Extraktion langfristig stabiler Strukturen geeignet.

Bei der simultanen Lokalisierung und Umgebungskartierung (engl.: Simultaneous Localization and Mapping, SLAM) kann das Entfernen bewegter Objekte dazu beitragen, die Konsistenz der Umgebungserfassung zu erhöhen. LITOMISKY & BHANU (2013) entfernen bewegte Objekte aus Punktwolken, indem Cluster in Punktwolken zweier Epochen gebildet und Korrespondenzen dazwischen gesucht werden. Anhand dieser Korrespondenzen werden bewegte Objekte ermittelt und entfernt. Hierbei bleibt ein hoher Prozentsatz der statischen Objekte vorhanden, bei weitem jedoch nicht alle. Da das Verfahren primär für SLAM konzipiert ist, wird mehr Wert auf Konsistenz als auf exakte Repräsentation der Umgebung gelegt.

Alle hier erläuterten Ansätze haben zum Ziel, entweder eine konsistente Repräsentation unter Billigung von Informationsverlusten zu generieren oder aber sich schnell auf Änderungen in der statischen Umgebung einstellen zu können. Beide Kriterien sind im Hinblick auf die möglichst exakte Änderungsdetektion in urbanen Umgebungen nicht zielführend, da hier der Fokus auf langfristig stabilen Strukturen liegt, die möglichst exakt und vollständig zu erfassen sind.

2.2 Volumenmodelle

Zur Repräsentation des Hintergrunds ist eine punktuelle Oberflächenbeschreibung, wie sie durch Laserscannerdaten gegeben ist, unzureichend. Stattdessen wird ein auf Belegungsgittern basierender Ansatz herangezogen, da dieser zur Beschreibung ganzer Raumbereiche in der Lage ist. MORAVEC & ELFES (1985) erzeugen anhand von ultraschall-basierten Entfernungsmessungen ein probabilistisches Belegungsgitter, welches dem Zustand nach *freie* und *belegte* Raumbereiche in der Umgebung auf Höhe des Sensors beschreibt. Obwohl das kegelförmige Schallsignal eine eher unpräzise Form der Abstandsmessung darstellt, sind die auf diese Art und Weise erzeugten Umgebungsrepräsentationen überraschend exakt.

Eine geometrische Modellierungstechnik von beliebigen 3D-Objekten wird von MAEGHER (1982) vorgeschlagen. Diese umfasst eine auf einem Octree basierende Datenstruktur, welche binäre Belegungsinformationen enthält. PAYEUR et al. (1997) ergänzen diesen Ansatz um probabilistische Informationen. Um eine schnellere Anpassung an eine sich ändernde Umgebung zu gewährleisten, führen HORNING et al. (2013) eine als „Clamping“ bezeichnete Begrenzung der Wahrscheinlichkeiten ein, welche eine Kompression der Datenstruktur durch das Zusammenfassen von Kinds-knoten erlaubt.

Auf dieser Basis wurde in einer vorangegangenen Arbeit ein Modellierungskonzept vorgestellt, welches die globale Verwaltung von probabilistischen Volumenrepräsentationen ermöglicht und ein Konzept zu deren effizienter Abspeicherung vorstellt (GEHRUNG et al. 2016).

3 Filtern bewegter Objekte

Bei der Auswertung von LiDAR-Daten eines mobilen Sensors im Kontext der automatischen Änderungsdetektion stellen bewegte Objekte unerwünschte Störungen dar, die zu einer fehlerhaft diagnostizierten Änderung führen können. Die im Rahmen dieser Veröffentlichung vorgestellte Methode macht sich die in (GEHRUNG et al. 2016) vorgestellte volumetrische Umgebungsrepräsentation zunutze, um eine Hypothese des nicht-beweglichen Hintergrunds aus Sensoreinzelmessungen zu erstellen. Letztere liegen in Form von Entfernungsmessungen vor. Diese Hintergrundrepräsentation wird in einem weiteren Verarbeitungsschritt dazu eingesetzt, die darin eingeflossenen Einzelmessungen in bewegte Objekte und statische Strukturen zu segmentieren. Erstere umfassen alles, was sich zum Zeitpunkt der Messung zumindest eine Zeit lang in Bewegung befunden hat. Eine Annahme hinsichtlich Struktur oder Klasse bewegter Objekte wird nicht getroffen. Die statischen Strukturen umfassen in erster Linie urbane Strukturen wie beispielsweise Gebäude, Schilder und Straßenlaternen sowie in eingeschränktem Maß den Pflanzenbestand. Ebenfalls betroffen sind jedoch auch sonst mobile Objekte wie stillstehende Personen und abgestellte Fahrzeuge, die sich jedoch zum Zeitpunkt der Messung nicht in Bewegung befinden.

3.1 Volumenbasierte Repräsentation

Zum besseren Verständnis des hier vorgestellten Ansatzes wird kurz die in (GEHRUNG et al. 2016) verwendete volumenbasierte Repräsentation von Umgebungen umrissen (Abb. 2). Der zu repräsentierende Raum wird in würfelförmige Kacheln zerlegt, welche je einen Octree enthalten. In diesem sind Voxel in unterschiedlichen Auflösungsstufen abgelegt, welche den Grad der Belegtheit des davon abgedeckten Raumbereichs beschreiben. Dies erfolgt durch eine im weiteren Verlauf als Belegungswahrscheinlichkeit bezeichnete Fließkommazahl, welche das Resultat eines Zwei-Klassen-Entscheidungsproblems darstellt. Es handelt sich hier um die Division der Wahrscheinlichkeit für eine Belegung durch diejenige für eine Nicht-Belegung, auf welche der Logarithmus angewandt wurde. Das logarithmierte Verhältnis zweier Wahrscheinlichkeiten wird auch als Log-Odd bezeichnet.

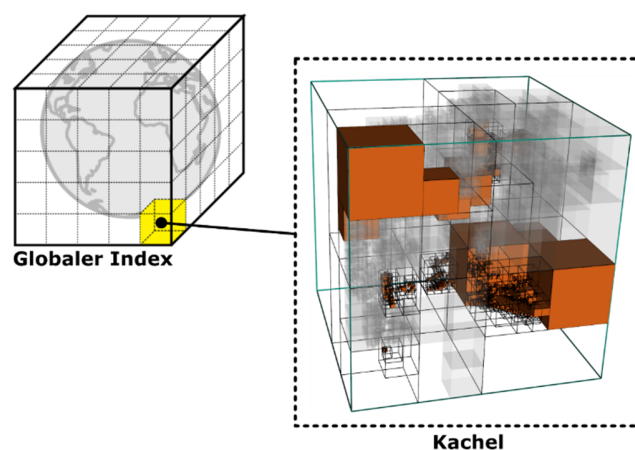


Abb. 2: In einem globalen Referenzgitter werden einzelne Kacheln abgelegt. Jede dieser Kacheln enthält eine probabilistische Umgebungsrepräsentation in Form eines Octree.

Für jede einzelne Entfernungsmessung werden die traversierten Voxel bestimmt und die Belegungswahrscheinlichkeit darin entsprechend verringert. Weiterhin wird das Voxel bestimmt, in welchem der Strahl endet und darin die Belegungswahrscheinlichkeit erhöht. Die Werte für das Absenken und Erhöhen werden im Rahmen des Sensormodells festgelegt. Die Belegungswahrscheinlichkeit ist auf ein festes Intervall begrenzt. Dieser Prozess wird als Klammerung (engl.: Clamping) bezeichnet. Wird die untere Grenze des Intervalls erreicht, so gilt das Voxel als frei; beim Erreichen der oberen Grenze gilt es als belegt.

3.2 Konvergenz der volumetrischen Umgebungsrepräsentation

Dem Ansatz liegt die Annahme zugrunde, dass die volumetrische Umgebungsrepräsentation für längere Beobachtungszeiträume zu einem Punkt hin konvergiert, an dem nur noch die statischen Elemente der Umgebung darin enthalten sind. Dieses erwartete Verhalten lässt sich dadurch begründen, dass ein in Bewegung befindliches Objekt jedes Voxel auf seiner Bahn nur für einen relativ kurzen Augenblick durchläuft. Für jedes der besagten Voxel liegen im Schnitt betrachtet also nur relativ wenige Messungen vor, die für deren Belegung sprechen. Der Großteil der Messungen impliziert hingegen, dass das Voxel Freiraum umfasst.

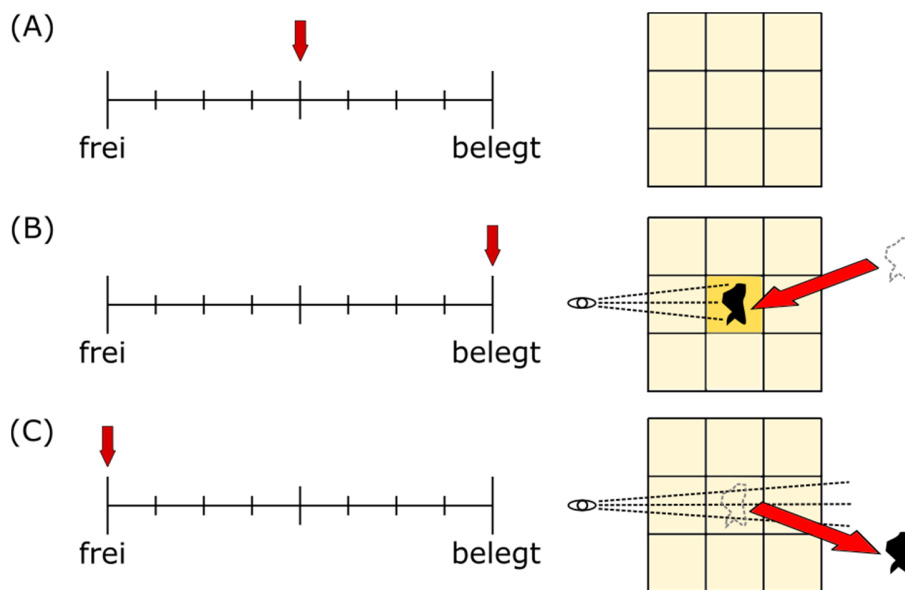


Abb. 3: (A) Ein Voxel wird mit Unwissen instanziiert. (B) Ein Objekt bewegt sich in das Voxel, die Belegungswahrscheinlichkeit wird erhöht, bis die Obergrenze des Intervalls erreicht ist. (C) Das Objekt hat das Voxel verlassen, weitere Messungen senken die die Belegungswahrscheinlichkeit sukzessive ab.

In Bezug auf das im letzten Abschnitt beschriebene Repräsentationskonzept findet folgender Prozess statt (vergleiche auch Abb. 3). Einige Messungen implizieren die Belegung eines Voxels, worauf der entsprechende Wert des Sensormodells auf die Belegungswahrscheinlichkeit aufaddiert wird. Hierbei kann es vorkommen, dass die obere Intervallgrenze erreicht wird und das Voxel als belegt angesehen werden kann. Hat das belegte Objekt besagtes Voxel wieder verlassen, so sorgen

alle durchlaufenden Messstrahlen dafür, dass dessen Belegungswahrscheinlichkeit um den im Sensormodell definierten Wert dekrementiert wird. Traversieren ausreichend viele Messungen das Voxel, so wird der untere Schwellwert erreicht, woraufhin das Voxel wieder als frei betrachtet wird. Die Annahme ist unter einer Reihe von Randbedingungen gültig. So muss eine ausreichend lange Messung für den von einem Objekt durchlaufenen Raumbereich vorhanden sein. Diese hängt von mehreren Faktoren ab; so etwa von der Anzahl der Einzelmessungen und der Parametrisierung des Sensormodells. Auch die Bewegungsgeschwindigkeit des Objekts spielt eine Rolle. Ist das Objekt sehr langsam unterwegs, so kann eine zuverlässige Erkennung bewegter Objekte aus zeitlicher Sicht letzten Teils der Entfernungsmessungen nur erfolgen, wenn es den Messbereich des Laserscanners verlassen hat. Auch kann es problematisch sein, wenn sich sehr viele Objekte sehr dicht zueinander bewegen, da bestimmte Voxel immer wieder von Objekten verlassen und kurz darauf durch andere Objekte betreten werden.

3.3 Filtern der Einzelmessungen

Zunächst wird aus den Einzelmessungen eine volumetrische Hintergrundrepräsentation erstellt. Erstere liegen in Form von Punktwolken mit Messwegen von Laserpulsen vor, die hierfür wie oben beschrieben als Sichtstrahlen interpretiert werden. Unter Voraussetzung der Korrektheit der obigen Annahme konvergiert die volumetrische Repräsentation zu einem stabilen Zustand, welcher die Eigenschaft der Umgebung hinsichtlich des Grades der Belegtheit zutreffend wiedergibt. Kurzzeitige Störungen durch bewegte Objekte werden durch Akkumulation der Messungen beseitigt. Das Resultat stellt eine Hypothese über den statischen Hintergrund der gemessenen Szene dar.

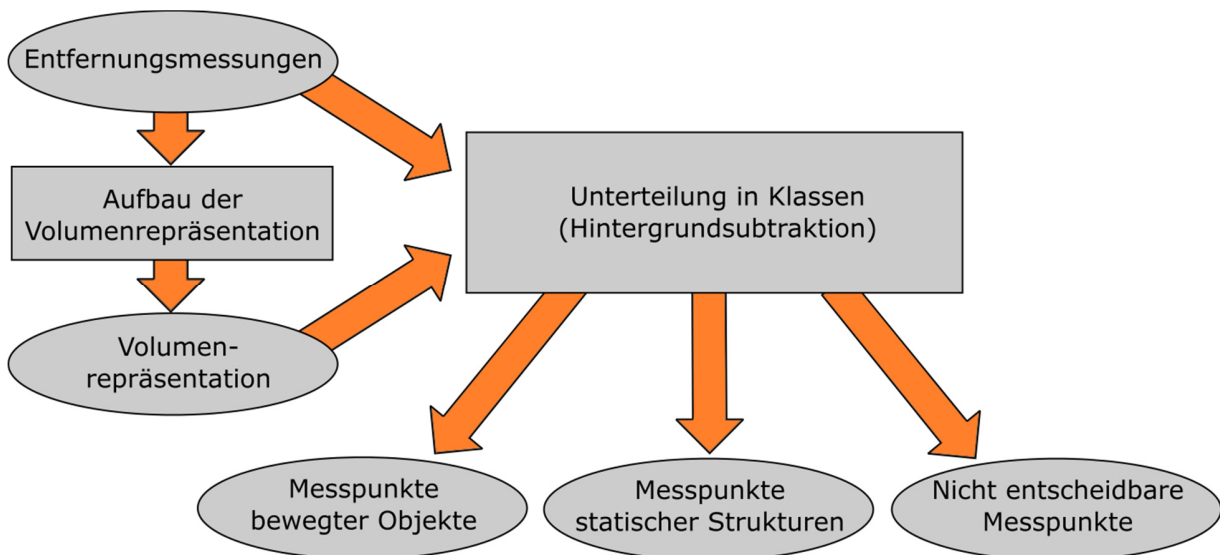


Abb. 4 Ablaufdiagramm der Hintergrundsubtraktion. Aus den Entfernungsmessungen wird zunächst die Volumenrepräsentation hergeleitet. Anhand dieser werden die Einzelmessungen in verschiedene Klassen aufgeteilt.

In einem weiteren Schritt werden die ursprünglich in die Erstellung der Volumenrepräsentation eingeflossenen Einzelmessungen in die Klassen bewegte Objekte, Hintergrund sowie in eine Restklasse für nicht entscheidbare Messungen unterteilt. Ein Ablaufdiagramm ist Abb. 4 zu entnehmen. Der Prozess ähnelt im weitesten Sinn der im Bereich des Maschinensehens angewandten Hintergrundsubtraktion. Für jede Einzelmessung, d.h. jeden Punkt der Punktwolke, wird das Voxel bestimmt, in dem es liegt. Anhand der darin abgelegten Belegungswahrscheinlichkeit wird die Zugehörigkeit des Messpunktes zu einer der oben genannten Klassen bestimmt. Liegt die Belegungswahrscheinlichkeit am oberen Rand des Intervalls, so ist der Messpunkt Teil einer statischen Struktur. Befindet sich die Belegungswahrscheinlichkeit hingegen am unteren Ende, d.h. das Voxel wird gegenwärtig als frei betrachtet, so handelt es sich um einen Messpunkt, der von einem bewegten Objekt stammt. Messpunkte, deren Belegungswahrscheinlichkeit an keiner der Intervallgrenzen liegt, werden der Restklasse zugeordnet.

4 Experimente

4.1 Mobile-Mapping-System

Der vorgestellte Ansatz wurde mit Messdaten mobiler Laserscanner evaluiert, die auf dem Stammgelände der TU München und entlang öffentlicher Gebäude im Umkreis aufgenommen wurden. Zur Erfassung wurde das Mobile-Mapping-System MODISSA des Fraunhofer IOSB eingesetzt (siehe Abb. 5). Das Fahrzeug verfügt insgesamt über vier Laserscanner. Zwei Velodyne HDL-64E sind in einem Winkel von 35° auf dem vorderen Dach angebracht. Diese Konstellation erlaubt es, sowohl einen Teil der Straße vor dem Fahrzeug als auch die Fassaden der umliegenden Gebäude zu erfassen. Zwei weitere Velodyne VLP-16 Puck befinden sich auf dem hinteren Teil des Dachs und sind ebenfalls geneigt angebracht. Acht an den Ecken des Dachs angebrachte Rundumkameras dienen unter anderem zur Texturierung der Lasermessungen. Eine Georeferenzierung der Messdaten erfolgt anhand der Navigationsdaten, die ein Applanix POS LV Navigationssystem mit zwei GNSS-Antennen, einer inertialen Messeinheit und einem Raddrehgeber liefert. Die Navigationsdaten wurden nachprozessiert, um die Genauigkeit der Georeferenzierung zu erhöhen.



Abb. 5: Das Messfahrzeug MODISSA, mit dem die im Rahmen dieser Arbeit verwendeten Datensätze aufgenommen wurden.

4.2 Qualitative Bewertung

In diesem Abschnitt wird eine qualitative Untersuchung der vorgestellten Methode beschrieben. Die Lasermessungen werden wie im vorherigen Abschnitt dargestellt durch Hintergrundsubtraktion in verschiedene Klassen unterteilt. Das Resultat wird anhand folgender Kriterien durch einen Betrachter beurteilt:

- Sind für jedes bewegte Objekt zumindest einige Messpunkte als zu dieser Klasse zugehörend erfasst?
- Zu welchem Prozentsatz sind die Oberflächenpunkte eines bewegten Objektes erfasst?
- Wie verhält sich der Ansatz bei feinen beweglichen Strukturen wie z.B. Vegetation?
- Wie hoch ist die Fehlerkennungsrate; sprich die Anzahl der Messpunkte, die als zu bewegten Objekten gehörend gezählt werden?
- Wie hoch ist der Anteil derjenigen Messpunkte, für die keine Aussage über ihre Zugehörigkeit zu einer der beiden Klassen *bewegt* oder *statisch* getroffen werden kann?
- In welchem Ausmaß wurde der statische Hintergrund erfasst?
- Treten sonstige Artefakte auf?

Die Beantwortung der oben genannten Fragen erfolgt anhand eines Datensatzes, der mit dem Mobile-Mapping-System MODISSA im Bereich des Stammgeländes der TU München erfasst wurde. Es handelt sich um einen belebten Innenhof auf dem Campus. Die Messung hat eine Dauer von knapp 4 Minuten und umfasst mehrere Personengruppen sowie sitzende und umhergehende Einzelpersonen. Ebenfalls in den Aufnahmen enthalten sind mehrere Bäume, die zum Zeitpunkt der Aufnahme über ein ausgeprägtes Blattwerk verfügten. Abb. 6 zeigt einen exemplarischen Ausschnitt aus dem Datensatz.

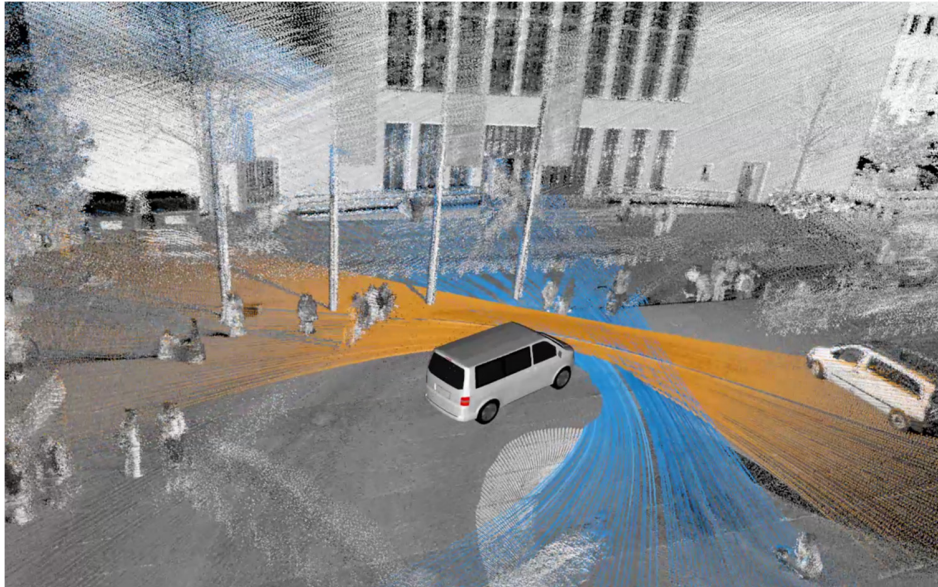


Abb. 6: Akkumulierte Lasermessungen aus dem Innenhof-Datensatz. Die zum Zeitpunkt der Aufnahme aktuellen Messungen der vorderen Laserscanner sind farblich hervorgehoben.

Neben der qualitativen Betrachtung wird auch die Laufzeit des Ansatzes für unterschiedliche Auflösungsstufen untersucht. Hierbei wird die Hintergrundsubtraktion an beiden Datensätzen durchgeführt und die mittlere Dauer pro Umdrehung des Laserscanners bestimmt, welche ca. 100.000 Einzelmessungen umfasst.

5 Diskussion der Ergebnisse

Die hier vorgestellten Ergebnisse basieren auf einer Volumenrepräsentation mit 40 cm Auflösung. Die maximale Länge der berücksichtigten Messstrahlen wurde auf 30 m festgelegt, um die Effekte von Ungenauigkeiten in der Geolokalisierung zu minimieren. Um die Anzahl der Messpunkte ohne Zuordnung zu reduzieren und das Gesamtergebnis zu verbessern wurde das Intervall der Belegungswahrscheinlichkeit für die Einordnung in bewegte Objekte und Freiraum auf $[-0.5, 0.5]$ gesetzt.

5.1 Grad der Erfassung

Alle bewegten Objekte wurden im Rahmen einer manuellen Betrachtung identifiziert und einer genauen Betrachtung unterzogen. Hierzu wurden die Ergebnisse der Segmentierung übereinandergelegt und farblich so kodiert, dass sich anhand der Farbe die Zugehörigkeit zur entsprechenden Klasse erkennen lässt. Alle so aufbereiteten Einzelmessungen wurden durch einen menschlichen Betrachter ausgewertet. Zusammenfassend lässt sich festhalten, dass auf allen bewegten Objekten Oberflächenpunkte detektiert wurden.

Der Prozentsatz korrekt zugeordneter Messpunkte pro bewegtem Objekt liegt schätzungsweise bei 90-95%. Gelegentlich kommt es zu Fehlzuordnungen, wenn eine Person ein belegtes Voxel oder ein Voxel, das nicht zugeordnet werden kann, durchläuft. In einem solchen Fall ist für eine kurze

Zeitspanne von ca. 1-2 Sekunden ein größerer Anteil der Messpunkte falsch zugeordnet. Im vorliegenden Datensatz wurde dies allerdings nur an sehr wenigen Stellen beobachtet. Häufiger fehlen Messpunkte in Grenzbereichen zu anderen Flächen, wie etwa der Bodenebene. Dies ist auf Diskretisierungsfehler zurückzuführen. In der vorliegenden Konfiguration der Volumenrepräsentation liegt die Größe eines Voxel bei $40 \times 40 \times 40 \text{ cm}^3$. Befindet sich die Bodenebene nun mittig in einem Voxel, so fallen die Füße einer Person in ein als solide angesehenes Volumen und werden daher falsch klassifiziert. Beispiele für eine farblich kodierte Einzelmessung sowie die hier beschriebenen Beobachtungen sind Abb. 7 zu entnehmen.

5.2 Kurzzeitig stationäre Objekte

An mehreren Stellen im Datensatz wurde beobachtet, wie Personen kurzfristig an einem Ort verweilen, um dann ihren Weg fortzusetzen. Die Person ist also stationär, ohne dass in den Lasermessungen eine größere Bewegung erkennbar ist. Die zugehörigen Messdaten wurden in besagten Fällen ohne zeitliche Unterbrechung als zu einem bewegten Objekt gehörend eingestuft. Dieser Sachverhalt ist darauf zurückzuführen, dass die Messdaten im Kontext der volumenbasierten Hintergrundrepräsentation betrachtet werden. Der Raumbereich, in den die Messungen fallen, wurde größtenteils als frei beobachtet und war nur kurz durch besagte Personen belegt.

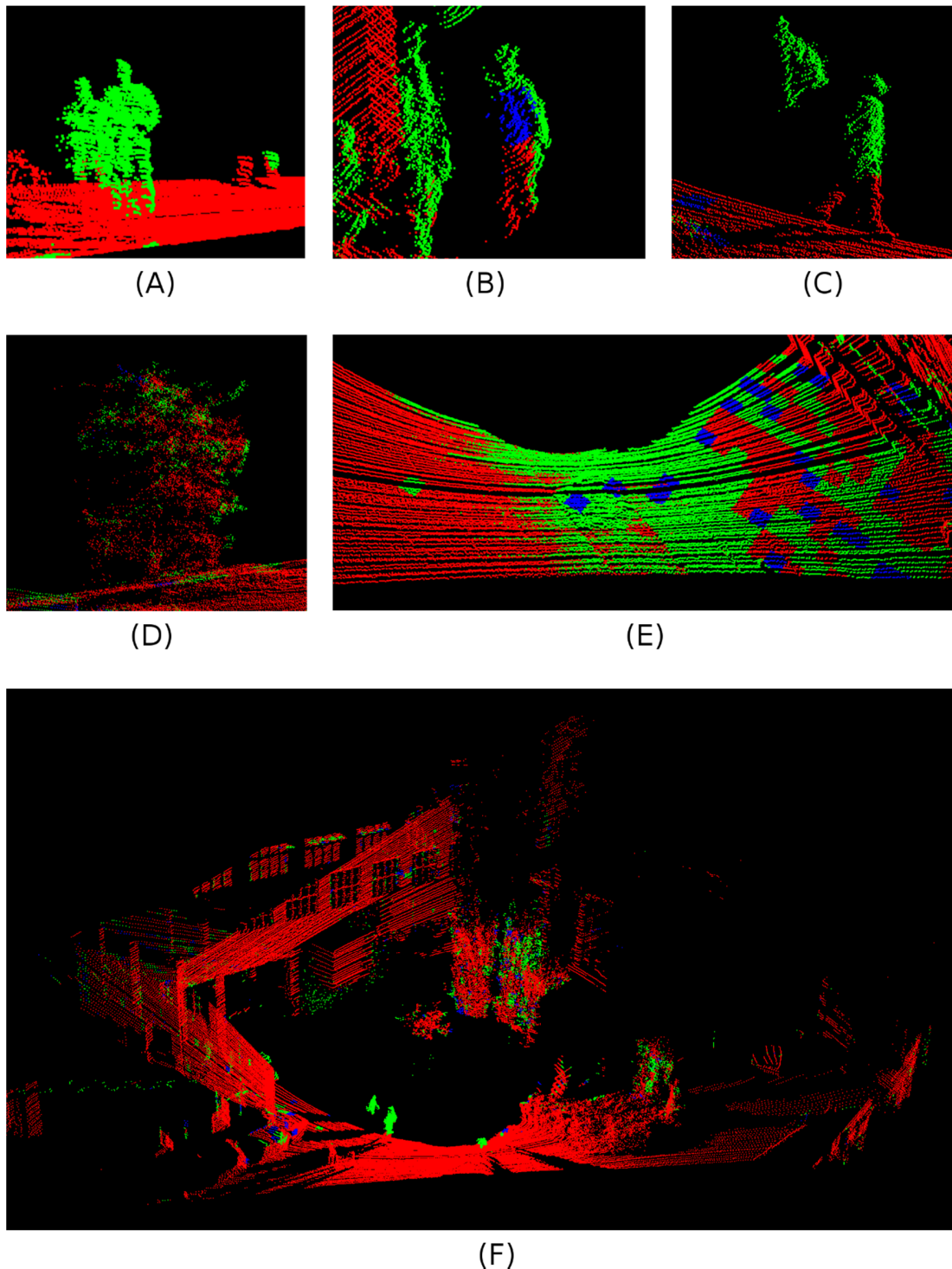


Abb. 7: Die Messpunkte sind in die Kategorien *bewegtes Objekt* (grün), *statische Struktur* (rot) und *keine Zuordnung möglich* (blau) unterteilt. (A) Segmentierung einer bewegten Person. (B) Segmentierungsfehler aufgrund falscher Zuordnung eines Voxels. (C) Diskretisierungsfehler. (D) Vegetation. (E) Systematischer Fehler, voraussichtlich ausgelöst durch flach auftreffende Messstrahlen. (F) Segmentierter Einzelscan eines Velodyne-Laserscanners.

5.3 Fehlererkennung bewegter Objekte

Dieser Abschnitt behandelt mögliche Gründe für die Fehlererkennung bewegter Objekte und der dazugehörigen Objektpunkte beim beschriebenen Verfahren. Wird die Auflösung des Volumenmodells zu hoch gewählt, so kommt es aufgrund von Mess- und Registrierungsungenauigkeiten zu einer hohen Fehldetektionsrate. Eine Auflösung von 40 cm hat sich beim betrachteten Messsystem als guter Kompromiss zwischen Rauschunterdrückung und Anfälligkeit bzgl. Segmentierungsfehlern erwiesen.

Neben dem eher zufällig verteilten Hintergrundrauschen wurden eine Reihe von Artefakten beobachtet (vergleiche auch Abb. 7, D bis E). So werden teilweise durchlässige Strukturen wie Vegetation oder gitterartige Objekte zum Teil als bewegt erkannt, da hier die Messstrahlen gelegentlich das Objekt durchdringen. Ebenfalls beobachtet wurden mehrere als bewegt eingestufte Flecken an statischen Strukturen, vorwiegend an waagrechten oder senkrechten Ebenen. Deren Kantenlänge variiert von ca. einem halben Meter bis hin zu mehreren Metern. Als Ursache hierfür werden Messstrahlen angenommen, die in einem sehr flachen Winkel auf die Ebene treffen. Im Zuge dessen werden andere zu der Ebene gehörende Voxel traversiert, die als belegt eingestuft sind. Messung für Messung wird so der Grad der Belegtheit umgekehrt, bis die traversierten Voxel als frei angesehen werden. Eine grafische Veranschaulichung des hier beschriebenen Effekts ist Abb. 8 zu entnehmen.

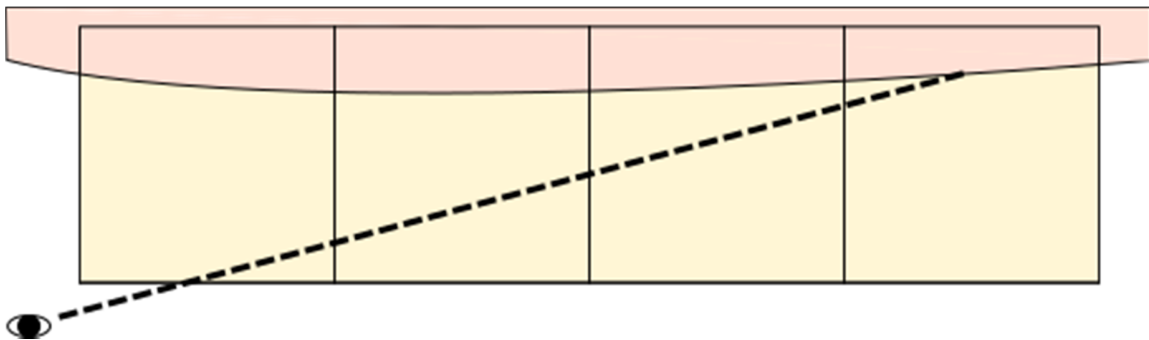


Abb. 8: Flach auf eine Oberfläche (rot) auftreffende Messstrahlen führen dazu, dass eigentlich als statisch angesehene Voxel sukzessive als frei angesehen werden.

5.4 Erfassung des Hintergrunds

Das Resultat der Hintergrundsubtraktion ist eine Menge von Lasermessungen, die sich in Bezug auf die Dichte mit einer unbearbeiteten Messung vergleichen lässt. Vegetation ist ebenfalls in Teilen noch enthalten, auch wenn diese wie oben beschrieben zu einem Großteil den bewegten Objekten zugeordnet wird. Personen sind keine mehr in den Messdaten vorhanden, lediglich deren Schatten. Gelegentlich kann es vorkommen, dass die Füße einer oder mehrerer Personen sichtbar bleiben.

5.5 Messpunkte ohne eindeutige Zuordnung

Messpunkte lassen sich nicht eindeutig einer der beiden oben diskutierten Klassen zuordnen, wenn die Belegungswahrscheinlichkeit des Voxels, in welches ein Messpunkt fällt, an keiner der beiden Intervallgrenzen liegt. Wie oben beschrieben wurden die Intervallgrenzen für die Segmentierung angepasst, sodass auch nicht eindeutige Messungen einer der beiden Klassen zugeordnet werden. Das Resultat ist, dass die Anzahl der Messpunkte ohne Zuordnung bei unter 5% liegt. Es handelt sich meist um einzelne Messpunkte oder aber um Flecken in Form und Größe eines Voxels.

5.6 Laufzeitbetrachtung

Die Laufzeitmessung wurde auf einem Rechner mit 32 GiB an RAM und einem Intel i7 mit 3.5 GHz und 12 Kernen durchgeführt. Die Implementierung wurde mittels OpenMP parallelisiert. Nach Initialisierung des Programms werden schrittweise die benötigten Elemente der volumetrischen Hintergrundrepräsentation geladen. Sobald diese komplett geladen ist beträgt die Laufzeit für jedes Bündel aus ca. 100.000 Einzelmessungen ca. 450 ms. Experimente für mehrere Auflösungsstufen von grob nach fein haben ergeben, dass die Tiefe im Octree keinen nennenswerten Einfluss auf die Laufzeit hat.

6 Fazit & Ausblick

Das im Rahmen dieser Arbeit vorgestellte Verfahren ermöglicht es, einzelne Lasermessungen in die Klassen *bewegtes Objekt* und *statischer Hintergrund* zu unterteilen. Hierfür wird ein einfacher Ansatz ohne Modellannahmen auf Basis einer Volumenrepräsentation verwendet, der zuerst aus Lasermessungen aufgebaut und dann zu deren Unterteilung in die beiden Klassen eingesetzt wird. Die Methode wurde anhand eines durch ein Messfahrzeug aufgenommenen Datensatzes qualitativ untersucht. Für jedes bewegte Objekt wird eine gewisse Mindestanzahl von Messpunkten richtig zugeordnet. Der prozentuale Anteil an korrekt zugeordneten Messpunkten pro Objekt liegt bei ca. 90-95%, wobei es an Übergängen zu statischen Strukturen wie etwa der Bodenebene zu Diskretisierungsfehlern kommen kann. Weiterhin treten Artefakte an Vegetation, teildurchsichtigen Strukturen sowie aufgrund von flach auftreffenden Messstrahlen auf. Zum statischen Hintergrund gehörende Messungen werden abgesehen von wenigen Diskretisierungsfehlern korrekt zugeordnet. Der Anteil der Messpunkte, die nicht zugeordnet werden können, ist mit weniger als 5% relativ gering. In weiterführenden Arbeiten sind Ansätze zu untersuchen, um das durch flach auftreffende Messstrahlen ausgelöste Artefakt sowie durch Vegetation bedingte Fehlzuordnungen zu beheben. Weiterhin ist es denkbar, die den bewegten Objekten zugeordneten Messpunkte zu gruppieren und damit einen Klassifikator zu trainieren, um auch unbewegte Instanzen der Objektklassen automatisiert aus den Lasermessungen entfernen zu können.

7 Literaturverzeichnis

- AZIM, A. & AYCARD, O., 2012: Detection, classification and tracking of moving objects in a 3D environment. Proceedings of the IEEE Intelligent Vehicles Symposium (IV), 802-807.
- GEHRUNG, J., HEBEL, M., ARENS, M. & STILLA, U., 2016: A Framework for Voxel-based Global Scale Modeling of Urban Environments. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, **XLII-2/W1**, 45-51.
- HORNUNG, A., WURM, K. M., BENNEWITH, M., STACHNISS, C. & BURGARD, W., 2013: OctoMap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees. Autonomous Robots **34**(3), 189-206.
- LITOMISKY, K. & BHANU, B., 2013: Removing Moving Objects from Point Cloud Scenes. Advances in Depth Image Analysis and Applications, Springer Berlin Heidelberg, 189-206.
- MEAGHER, D., 1982: Geometric modeling using octree encoding. Computer Graphics and Image Processing **19**, 129-147.
- MORAVEC, H. & ELFES, A., 1985: High Resolution Maps from Wide Angle Sonar. Proceedings of the 1985 IEEE International Conference on Robotics and Automation, 116-121.
- PAYEUR, P., HEBERT, P.; LAURENDEAU, D. & GOSSELIN, C. M., 1997: Probabilistic octree modeling of a 3D dynamic environment. Proceedings of the IEEE International Conference on Robotics and Automation, 1289-1296.
- VALLETT, B., XIAO, W. & BRÉDIF, M., 2015: Extracting Mobile Objects in Images Using a Velodyne LIDAR Point Cloud. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, **II-3/W4**, 247-253.