Hochschule für Technik und Wirtschaft Berlin

Internationaler Studiengang Medieninformatik

Masterarbeit

von

Daniel Schneider

Photogrammetrie zur Platzierung von standortbezogenen dynamischen Inhalten in AR

Photogrammetry for placement of location-based dynamic content in AR

Hochschule für Technik und Wirtschaft Berlin Fachbereich Informatik, Kommunikation und Wirtschaft

Studiengang Internationaler Studiengang Medieninformatik

Masterarbeit

von

Daniel Schneider

Photogrammetrie zur Platzierung von standortbezogenen dynamischen Inhalten in AR

Photogrammetry for placement of location-based dynamic content in AR

Bearbeitungszeitraum: von 13.05.2019

bis 16.09.2019

1. Prüfer: Prof. Dr. Tobias Lenz

2. Prüfer: Prof. Dr. Klaus Jung

| Fachbereich Informatik, Kommunikation und Wirtschaft | | | | | | |
|--|--|--|--|--|--|--|
| Eigenständigkeitserklärung | | | | | | |
| Eigenstandigkenserklarung | | | | | | |
| Name und Vorname der Studentin/des Studenten: | Schneider, Daniel | | | | | |
| Studiengang: | Internationaler Studiengang Medieninformatik | | | | | |
| Ich bestätige, dass ich die Master | rarbeit mit dem Titel: | | | | | |
| Photogrammetrie zur Platzieru | ing von standortbezogenen dynamischen Inhalten in AR | | | | | |
| | t anderweitig für Prüfungszwecke vorgelegt, keine uellen oder Hilfsmittel benutzt sowie wörtliche und kennzeichnet habe. | | | | | |
| Datum: 10. Juli 2019 | | | | | | |
| Unterschrift: | | | | | | |

Hochschule für Technik und Wirtschaft Berlin

Hochschule für Technik und Wirtschaft Berlin Fachbereich Informatik, Kommunikation und Wirtschaft

Masterarbeit Zusammenfassung

Studentin/Student (Name, Vorname): Schneider, Daniel

Studiengang: Internationaler Studiengang Me-

dieninformatik

Aufgabensteller, Professor: Prof. Dr. Tobias Lenz

Durchgeführt in (Firma/Behörde/Hochschule): HTW Berlin

Betreuer in Firma/Behörde:

Ausgabedatum: 13.05.2019 Abgabedatum: 16.09.2019

Titel:

Photogrammetrie zur Platzierung von standortbezogenen dynamischen Inhalten in AR

Zusammenfassung:

"Zusammenfassung"

Schlüsselwörter: Photogrammetrie, Augmented Reality, standortbezogene Daten, Android, Java

Inhaltsverzeichnis

| 1 | Einführung und Motivation | | | | | |
|----|--|-----------------------|--|--|--|--|
| 2 | Photogrammetrie | | | | | |
| 3 | Verfahren zur Generierung von Mapping Daten der Umwelt 3.1 Simultaneous Localisation and Mapping | 3 6 7 7 | | | | |
| 4 | Implementation einer AR Anwendung für Android4.1Verwendete Hard und Software4.1.1Ar Core4.1.2Sceneform4.1.3Google Location Service4.1.4Dexter4.1.5Volley | 8 8 8 8 8 | | | | |
| 5 | Praxistest5.1 Anwendungsbeispiel | 9 9 | | | | |
| 6 | | 10 10 | | | | |
| 7 | Zusammenfassung und Ausblick | 11 | | | | |
| Li | iteraturverzeichnis | | | | | |
| Αl | Abbildungsverzeichnis | | | | | |

1 Einführung und Motivation

Photogrammetrie ist "die Wissenschaft und Technologie der Gewinnung von Informationen über die physische Umwelt aus Bildern, mit einem Schwerpunkt auf Vermessung, Kartierung und hochgenauer Messtechnik". (Heipke, 2017, S.5 [1]) Die Photogrammetrie beschäftigt sich mit der Rekonstruktion von dreidimensionalen Daten aus zweidimensionalen Informationsträgern, wie Bildern oder Laserscan Daten. Dabei gehen diese Daten alle auf das Prinzip der Aufnahme der elektromagnetischen Strahlung zurück. Bei Bildern ist das die Helligkeits und Farbverteilung, bei Laserscans, Entfernungsbilder, beziehungsweise Punktwolken. Die Disziplin der Photogrammetrie ist dabei dem Bereich der Fernerkundung zuzuordnen, die sich mit der Auswertung von geometrischen oder semantischen Informationen beschäftigt. Beides sind Fachbereiche, die sich über die Jahrzehnte entwickelt haben und sich dem Gebiet der Geodäsie zuordnen lassen. Die Geodäsie erfasst Geoinformationen über die Erde, die dann beispielsweise mit Kartographie visualisiert werden können. Das Gebiet der Computer Vision, das sich großteils parallel mit der Photogrammetrie entwickelt hat, verfolgt den gleichen Ansatz. Auch hier ist die Auswertung digitaler Bilder das zentrale Element. Nachdem sich Photogrammetrie und Computer Vision lange unabhängig voneinander entwickelt haben, ist heute Photogrammetrie als Grundlage von Computer Vision anerkannt. (vgl. [1] S. 5-7)

Durch die rasante Entwicklung der Technik und der damit eingehenden Steigerung der Rechenpower, sind die Bereiche, in denen Photogrammetrie eingesetzt werden kann gestiegen. Im Rahmen dieser Arbeit soll evaluiert werden, ob photogrammetrische Verfahren bei Smarthphones eingesetzt werden können, um in Echtzeit aus Videodaten die dreidimensionalen Beschaffenheit der gefilmten Objekte zu rekonstruieren. Weiterhin soll dieses Verfahren genutzt werden um anschließend ein dreidimensionales Koordinatensystem mit Tiefeninformationen zu erzeugen, in welches dann dynamische Standortinformationen, im Stil von Augumented Reality, eingeblendet werden können. Dazu ist im Rahmen dieser Arbeit eine auf dem Android Betriebsystem basierende Anwendung erstellt worden.

2 Photogrammetrie

3 Verfahren zur Generierung von Mapping Daten der Umwelt

3.1 Simultaneous Localisation and Mapping

Simultaneous Localisation and Mapping, kurz SLAM, ist das Problem der Auswertung einer unbekannten Umgebung und Erstellung einer Map, während gleichzeitig die lokale Position innerhalb dieser Map bestimmt wird. Die Lösung dieses SLAM Problems war vorallem in der Robotik eine fundamentale Aufgabe der letzten zwei Jahrzehnte. Dabei ist SLAM ein Alltagsproblem: Das Problem der räumlichen Erkundung. Jeder Mensch und jedes Tier hat dieses Verfahren gemeistert und benutzt es unterbewusst zur Navigation in unserer Realität. Die Lösung für dieses Problems, wenn es für einen Roboter automatisiert ausgeführt werden soll, ist dagegen sehr komplex. Durch das Meistern dieser Technik kann man Roboter wirklich autonom steuern. Bei SLAM wird die Bewegung des Objekts an sich durch den Raum und die Position aller zur positionsbestimmung notwendigen Merkmale berechnet, ohne auf vorheriges Wissen, über Position oder Lage im Raum, Kenntniss zu haben. (vgl. [2] S. 1-2)

Dabei benötigt der Roboter mindestens einen exterozeptiven Sensor um äußere Informationen zu sammeln. SLAM besteht aus drei grundlegenden Operationen, die iterativ pro Zeitintervall ausgeführt werden.

Der Roboter bewegt sich und erreicht eine neue Position in der Umwelt. Diese Bewegung erzeugt, durch unvermeidbares Rauschen und Fehler, Ungewissheit über die wirkliche Position des Roboters. Eine automatisierte Lösung benötigt ein mathematisches Modell für diese Bewegung. Dies ist das "*Motion Model*"

Der Roboter entdeckt neue Features in seiner Umgebung, welche in die Umgebungskarte aufgenommen werden müssen. Diese Features heißen "Landmarks". Da die Position der Landmarks, durch Fehler in den exterozeptiven Sensoren und die Position des Roboters ungewiss ist, müssen diese beiden Faktoren passend arrangiert werden.

Eine automatisierte Lösung benötigt ein mathematisches Modell, das die Position der Landmarks anhand er Sensordaten bestimmt. Dies ist das "textitInverse Oberservation Model."

Der Roboter entdeckt Landmarks, die schon gemappt wurden und verwendet diese um seine eigene Position, sowie die aller Landmarks zu korrigieren. Diese Operation reduziert die Unsicherheit über den Standort des Roboters, sowie der Landmarks. Die automatisierte Lösung erfordert ein mathemathisches Modell, um die Werte der Messungen aus den prognostizierten Positionen der Landmarks und der Position des Roboters zu berechnen. Dies ist das "Direct Observation Model"

Mit diesen drei Modellen ist es möglich eine automatisierte Lösung für SLAM zu entwerfen. Diese Lösung muss diese drei Modelle verbinden und alle Daten korrekt und oganisiert halten, sowie die korrekten Entscheidungen bei jedem Schritt machen. (vgl. [4] S.2-3)

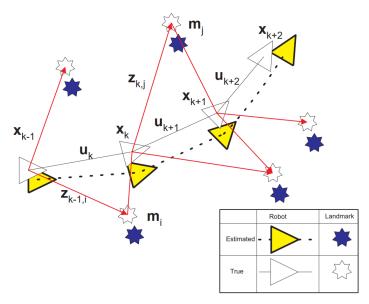


Abbildung 3.1: Das SLAM Problem: Die wahren absoluten Positionen der extrahierten Features sind nie wirklich bekannt. Bildquelle [2]

Wie in Abbildung 3.1. erkennbar ist, bewegt sich ein Roboter durch eine unbekannte Umgebung und nimmt mit seinem Sensor Features der näheren Objekte (Landmarks) auf. Wobei $\mathbf{x}k$ der Vektor des Roboters, $\mathbf{u}k$ der Bewegungsvektor, $\mathbf{m}i$ der Vektor des Landmarks und $\mathbf{z}ik$ die Oberservation eines Landmarks durch den Roboter zur Zeit k sind. Wie man sehen kann, ist der Fehler zwischen echten und geschätzten Landmarks, bei allen geschätzten Landmarks ähnlich, was an der initialen Betrachtung der Umgebung liegt. Zu diesem Zeitpunk wird nur das erste Feature erkannt. Daraus kann man

schließen, dass die Fehler in der Schätzung der Landmarkpositionen korrelieren. Praktisch bedeutet dies, dass die relative Position zweier Landmarks, $\mathbf{m}i$ - $\mathbf{m}j$ zueinander sehr genau sein kann, auch wenn die absolute Position sehr ungenau ist.

Je mehr Landmarks in das Modell aufgenommen werden, desto gleichbleibend besser wird das Modell der relativen Positionen, egal wie sich der Roboter bewegt. Dieser Prozess wird in Abbildung 3.2. veranschaulicht.

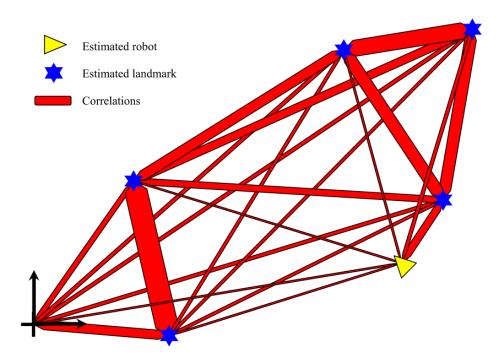


Abbildung 3.2: Die Landmarks sind durch Federn verbunden, welche die Korrelation zwischen ihnen darstellen. Bildquelle [2]

Während sich der Roboter durch die Umgebung bewegt, werden die Korrelationen stetig aktualisiert. Je mehr Beobachtungen über die Umwelt gemacht werden, desto steifer werden die Federn in diesem Modell. Im Nachhinein werden neue Beobachtungen von Landmarks durch das ganze Netzwerk propagiert und je nach Input, kleinere oder größere Anpassungen vorgenommen.

Lösungen für das SLAM Problem benötigen eine angemessene Repräsentation für die Observierungen der Landmarks, welche eine konsistente und schnelle Berechnung ermöglichen. Die geläufigste Repräsentation besteht in der Form einer Zustandsraumdarstellung mit Gaußschen Rauschen, was zur Verwendung des "Extended Kalman Filter"(EKF) führt. Eine weitere alternative Repräsentation ist die Beschreibung der Features als Datenset aus Stichproben in einer nicht gaußschen Wahrscheinlichkeitsverteilung. Diese Methodik benutzt den "Rao-Blackwellised particle filter"oder den

Fast-SLAM Algorithmus. (vgl. [2] S. 2-4)

3.1.1 Extended Kalman Filter - SLAM

Der Kalman Filter ist eine Schätzfunktion für das "linear-quadratic-problem ", was das Problem der Schätzung des augenblicklichen Zustands eines linearen dynamischen Systems, gestört durch weißes Rauschen darstellt. Der Kalman Filter wird auch dazu benutzt um die mögliche Zukunft von dynamischen Systemen vorherzusagen, die von Menschen nicht kontrolliert werden können, wie zum Beispiel die Flugbahn von Himmelskörpern, oder der Kurs von gehandelten Rohstoffen. (vgl. [?] S.1)

Bei EKF-SLAM ist die Map ein großer Stapel an Vektor und Sensordaten, sowie Zuständen von Landmarks und wird durch eine gaußsche Variable modelliert.

$$x = \begin{bmatrix} R \\ M \end{bmatrix} = \begin{bmatrix} R \\ L_1 \\ \dots \\ L_n \end{bmatrix}$$
 (3.1)

R ist der Zustand des Roboters und $M=(L_1,...,L_n)$ ist das Set an Zuständen der Landmarks. Bei EKF wird die Map durch eine gaußsche Variable modelliert, die den Mittelwert und die Kovarianzmatrix des Zustandsvektors verwendet, die jeweils durch \overline{x} und P beschrieben werden. Das Ziel ist es die Map $\{\overline{x},P\}$ zu allen Zeiten auf dem aktuellsten Stand zu halten.

$$\overline{x} = \begin{bmatrix} \overline{R} \\ \overline{M} \end{bmatrix} = \begin{bmatrix} \overline{R} \\ \overline{L_1} \\ \dots \\ \overline{L_n} \end{bmatrix} \qquad P = \begin{bmatrix} P_{RR} & P_{RM} \\ P_{MR} & P_{MM} \end{bmatrix} = \begin{bmatrix} P_{RR} & P_{RL1} & \dots & P_{RLn} \\ P_{L1R} & P_{L1L1} & \dots & P_{L1Ln} \\ \dots & \dots & \dots & \dots \\ P_{LnR} & P_{LnL1} & \dots & P_{LnLn} \end{bmatrix}$$
(3.2)

Diese Map, die als stochastische Map bezeichnet wird, wird durch den EKF in Stand gehalten, durch die Vorhersage- und Korrekturprozesse. Um eine echte Erkundung der Umgebung zu erreichen, wird der EKF Algorithmus mit einem extra Schritt der Landmark Erkennung und Initialisierung gestartet, bei dem neue Landmarks der Map hinzugefügt werden. Die Landmark Initialisierung erfolgt durch eine Umkehrung der Bewertungsfunktion und der Verwendung dieser und der Ableitungsmatrix, um die beobachteten Landmarks und die benötigten Co- und Crossvarianzen für den Rest der

Map zu berechnen. Diese Beziehungen werden dann an den Zustandsvektor und die Kovarianzmatrix angehängt. (vgl. [4] S.6-7)

http://www.iri.upc.edu/people/jsola/JoanSola/objectes/curs_SLAM/SLAM2D/SLAM% 20course.pdf

3.1.2 FAST-SLAM

http://www.cs.cmu.edu/~mmde/mmdeaaai2002.pdf http://srl.informatik.uni-freiburg.de/publicationsdir/grisettiRAS07.pdf

3.1.3 SLAM für mobiles Augumented Reality

Das Ziel von Augumented Reality ist es virtuelle Objekte oder Informationen in die echte Welt zu integrieren, um den Benutzer zusätzliche Informationen in die betrachtete Szene zu liefern. Dazu ist es notwendig, die echte und die virtuelle Welt präzise aneinander auszurichten. Dann kann für jedes Frame aus der Sequenz des Videobildes die genaue Position des mobilen Gerätes bestimmt werden. Um dieses Ziel des exakten Matchings von Realität und generierter Virtueller Realität zu erreichen, ist "Camera Localization", also die Lokalisierung der Kamera im dreidimensionalen Raum, anhand von aufgenommenen zweidimensionalen Daten, die Schlüsseltechnologie für alle Augumented Reality Anwendungen. (vgl. [3] S.1)

https://hal.inria.fr/hal-00994756/document https://pdfs.semanticscholar.org/00f4/41387f04f40aad6491ce23bdeb0ece17d12e.pdf

3.1.4 SLAM als Core für viele AR APIs

4 Implementation einer AR Anwendung für Android

- 4.1 Verwendete Hard und Software
- **4.1.1** Ar Core
- 4.1.2 Sceneform
- 4.1.3 Google Location Service
- **4.1.4** Dexter
- **4.1.5 Volley**

5 Praxistest

- 5.1 Anwendungsbeispiel
- 5.2 Benchmarks

6 Weitere Verfahren

6.1 Depth Map by Dual Camera

7 Zusammenfassung und Ausblick

Literaturverzeichnis

- [1] Heipke, C. (2017), "Photogrammetrie und Fernerkundung", 1.Auflage, Berlin: Springer Verlag, S. 5-7.
- [2] Hugh Durrant-Whyte, Tim Bailey (2006), Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms. URL: https://people.eecs.berkeley.edu/~pabbeel/cs287-fa09/readings/Durrant-Whyte_Bailey_SLAM-tutorial-I.pdf (Zuletzt abgerufen am 09.07.2019)
- [3] Pierre Martin, Eric Marchand, Pascal Houlier, Isabelle Marchal. Mapping and re-localization for mobile augmented reality. IEEE Int. Conf. on Image Processing, Oct 2014, Paris, France. URL: https://hal.inria.fr/hal-00994756/document (Zuletzt abgerufen am 09.07.2019)
- [4] Joan Sol'a, (2014), Simulataneous localization and mapping with the extended Kalman filter. URL: http://www.iri.upc.edu/people/jsola/JoanSola/objectes/curs_SLAM/SLAM2D/SLAM%20course.pdf (Zuletzt aufgerufen am 10.07.2019)
- [5] Mohinder S. Grewal, Angus P. Andrews (2001), Kalman Filtering: Theory and Practice with MATLAB. URL: http://staff.ulsu.ru/semoushin/_index/_pilocus/_gist/docs/mycourseware/13-stochmod/2-reading/grewal.pdf (Zuletzt aufgerufen am 10.07.2019)

Abbildungsverzeichnis

| 3.1 | Das SLAM Problem: Die wahren absoluten Positionen der extrahierten | |
|-----|--|---|
| | Features sind nie wirklich bekannt. Bildquelle [2] | 4 |
| 3.2 | Die Landmarks sind durch Federn verbunden, welche die Korrelation | |
| | zwischen ihnen darstellen Bildquelle [2] | _ |