

Bauhaus-Universität Weimar
Fakultät Medien
Studiengang Medieninformatik

Echtzeit Hinderniserkennung für unbemannte Flugsysteme unter Benutzung eines Stereo Kamera Systems

Bachelorarbeit

Hagen Hiller
Geboren am 04.06.1992 in Berlin

Matrikelnummer 110514

1. Gutachter: Prof. Dr. Volker Rodehorst
2. Gutachter: TBA

Datum der Abgabe: 25.01.2016

Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Weimar, den 25.01.2016

.....
Hagen Hiller

Zusammenfassung

Lorem ipsum dolor sit Amet.

Inhaltsverzeichnis

1	Einführung	3
1.1	Motivation	3
1.2	Setup	4
1.3	Ziel der Arbeit	4
2	Anforderungsanalyse	5
2.1	Bildaufnahme und Preprocessing	5
2.2	Erfassung von Hindernissen	5
2.3	Performanz	5
2.4	Positionsinformationen	5
3	State of the Art Algorithmen	6
3.1	Kamerabasierte Hinderniserkennung	6
3.2	Sensorbasierte Hinderniserkennung	6
4	Zugrunde liegende Konzepte und Algorithmen	7
4.1	Epipolargeometrie	7
4.2	Stereo Matching	8
4.2.1	Klassifikation	9
4.2.2	Block Matching	10
4.2.3	Semi Global Block Matching	10
4.3	mvStereoVision Framework	11
5	Subimage Detection	13
5.1	Konzept	13
5.2	Subimages	13
5.3	foo	13
6	Samplepoint Detection	14
6.1	Konzept	14
6.2	Erkennung	14
6.3	foo	14

7 Konflikte in der Erkennung und Lösungsansätze	15
7.1 Frame Skipping	15
7.2 Hindernisgröße	15
7.3 Position Mapping	15
8 Evaluierung	16
8.1 Aufbau des Testsetups	16
8.2 Evaluierung Subimage Detection	16
8.2.1 Robustheit	16
8.2.2 Performanz	16
8.3 Evaluierung Samplepoint Detection	16
8.3.1 Robustheit	16
8.3.2 Performanz	16
9 Diskussion	17
9.1 Anforderungsevaluierung	17
9.2 Gegenüberstellung beider Algorithmen	17
10 Fazit und zukünftige Arbeiten	18
A Anhang 1	19
Abbildungsverzeichnis	20
Tabellenverzeichnis	21
Literaturverzeichnis	22

Kapitel 1

Einführung

1.1 Motivation

Die Forschung im Bereich der Robotik ist ein immer weiter fortschreitender Prozess. Eine besondere Domäne innerhalb dessen ist die autonome Navigation sowie Steuerung von unbemannten Flugsystemen. Dies findet Anwendung in der Erkundung von unbekannten oder gefahrenträchtigen Gebieten. Dabei müssen nicht nur die physikalischen Eigenschaften der Drohne betrachtet werden sondern auch die Fusion verschiedenster Sensoren.

Ein wichtiges Kriterium in der Entwicklung autonomer Roboter ist die Erkennung von Hindernissen in Echtzeit. Dabei muss jedoch zuerst definiert werden was vom System als potentiell Hindernis erkannt werden soll. Einerseits sind alle Objekte welche sich in der unmittelbaren Nähe des Systems befinden eine Gefahrenquelle, andererseits wird im Falle Kamera basierter Navigation von einer Bewegung in Blickrichtung ausgegangen. Dies schließt Objekte welche sich außerhalb des Sichtfeldes, also neben oder hinter der Kamera befinden, für die Erfassung durch die Algorithmen aus. Auch sehr weit entfernte Objekte sind prinzipiell nicht als Hindernis anzusehen, wobei sich die maximale Gefahrendistanz relativ zur aktuellen Geschwindigkeit verhält. Unter Betrachtung dieser Gesichtspunkte wird ein Hindernis innerhalb dieser Arbeit als ein Objekt definiert welches sich innerhalb eines definierten Distanzbereichs und innerhalb des Sichtfeldes der Kamera befindet.

Die hauptsächliche Anwendung des im Rahmen dieser Arbeit entwickelten Systems zielt auf die autonome Navigation von Unbemannten Flugsystemen ab. Weitere Anwendungsbereiche können sowohl komplexerer als auch einfacherer Natur sein. Prinzipiell ist es möglich die entwickelten Algorithmen im Automobil Bereich zu verwenden um beispielsweise Objekte vor oder Hinter

dem Kraftfahrzeug zu erkennen und die Distanz zu berechnen. Weiterhin ist es vorstellbar grobe kartographische Höhen Klassifikationen vorzunehmen, um Bild basiert eine Höhenkarte geographischer Areale zu erstellen. Das jeweilige Anwendungsszenario richtet sich jedoch dabei nach der verwendeten Hardware.

1.2 Setup

1.3 Ziel der Arbeit

Die Verwendung Stereobild basierter Daten ermöglicht eine simple Berechnung räumlicher Tiefe.

Mithilfe photogrammetrischer Verfahren ist es mögliche räumliche Tiefe aus zweidimensionalen Daten zu berechnen. Dabei gilt die Verwendung Stereobild basierter Daten als ein einfachste Möglichkeit Disparitäten zwischen korrespondierenden Pixeln zu bestimmen um damit einen räumlichen Eindruck der abgebildeten Szenerie zu erhalten. Unter Betrachtung dieser Techniken wurde im Rahmen dieser Arbeit

Vor diesem Hintergrund werden diverse Anforderungen definiert, welche in Kapitel 2 näher beschrieben werden. Die zugrunde liegende Berechnung von *Disparity Maps* basiert auf dem Semi Global Block Matching Algorithmus [[Hirschmuller, 2005]] sowie dem Konzept der Epipolargeometrie welche in Kapitel 4 näher erläutert werden. Kapitel 3 gibt einen Einblick in State of the Art Algorithmen zur Hinderniserkennung. Dabei wird sowohl auf Kamera basierte als auch auf Sensor basierte Algorithmen eingegangen. In den Kapiteln 5 und 6 werden die beiden im Rahmen dieser Arbeit entwickelten Algorithmen zur Hinderniserkennung vorgestellt. Um die Echtheit gefundener Hindernisse zu validieren wurde das Verfahren *Frame Skipping* implementiert welches in Kapitel 7 erläutert wird. Ebenfalls werden bestehende Konflikte in der Erkennung sowie praktische Lösungen als auch hypothetische Ansätze zur Bewältigung dieser erläutert. Die Evaluierung der beiden implementierten Methoden erfolgt in Kapitel 8. Hierbei werden diese hinsichtlich ihrer Robustheit und Performanz statistisch ausgewertet. Kapitel 9 diskutiert die Erfüllung der durch die Anforderungsanalyse definierten Vorgaben. Des Weiteren erfolgt eine Gegenüberstellung der beiden entwickelten Algorithmen. Abschließend fasst Kapitel 10 die Beiträge dieser Abschlussarbeit zusammen und gibt einen Ausblick auf mögliche zukünftige Arbeiten in diesem Bereich.

Kapitel 2

Anforderungsanalyse

Lorem ipsum dolor sit amet.

2.1 Bildaufnahme und Preprocessing

Anforderung 1 *Die zur Erkennung von Hindernissen verwendeten Bilder müssen kalibriert, entzerrt sowie stereo-rektifiziert werden.*

Um 1 zu erfüllen müssen die aufgenommenen Bilder vor der Berechnung

2.2 Erfassung von Hindernissen

Anforderung 2 *Die Erfassung von Hindernissen sollte eine signifikant hohe Wahrscheinlichkeit haben um etwaige Fehlentscheidungen zu vermeiden.*

2.3 Performanz

Anforderung 3 *Das System sollte Hindernisse in Echtzeit erfassen und erkennen können.*

2.4 Positionsinformationen

Anforderung 4

Kapitel 3

State of the Art Algorithmen

3.1 Kamerabasierte Hinderniserkennung

3.2 Sensorbasierte Hinderniserkennung

Kapitel 4

Zugrunde liegende Konzepte und Algorithmen

Dieses Kapitel beleuchtet dieser Arbeit zugrunde liegende Konzepte und Algorithmen. Zunächst wird das Prinzip der Epipolargeometrie beschrieben welches ein wesentlicher Bestandteil photogrammetrischer Verfahren ist. Daraufhin folgt eine Erläuterung des Terms Stereo Matching sowie eine Klassifizierung in lokale und globale Algorithmen. Im Anschluss daran wird das Prinzip des Block Matching Algorithmus grundlegend beschrieben, welcher die Grundlage für den im Rahmen dieser Arbeit verwendeten Semi Global Block Matching Algorithmus bildet. Anschließend erfolgt eine Erläuterung des verwendeten Frameworks sowie Details zur Implementierung dessen.

4.1 Epipolargeometrie

Bei der Verwendung der meisten photogrammetrischer Verfahren spielt die Epipolargeometrie eine entscheidende Rolle. Das mathematische Modell beschreibt die geometrische Beziehung zwischen verschiedenen Kamerabildern desselben Objektes, sowie die Abhängigkeit zwischen korrespondierenden Punkten. Generell gesehen ist die Epipolargeometrie durch das Lochkamera-Modell beschrieben. Dabei liegt jeder Punkt des aufgenommenen Objektes mit dem Projektionszentrum sowie dem Bildpunkt auf einer Geraden. Unter Zuhilfenahme der extrinsischen Orientierung sowie der intrinsischen Parameter der Kamera, ist es möglich den Schnittpunkt zweier Geraden zu berechnen um die dreidimensionalen Koordinaten des Objektpunktes zu erhalten. Dabei gilt generell, wenn ein Punkt P im linken Bild gegeben ist, so wird die Suche des korrespondierenden Punktes P' auf die Epipolarlinie des rechten Bildes reduziert.

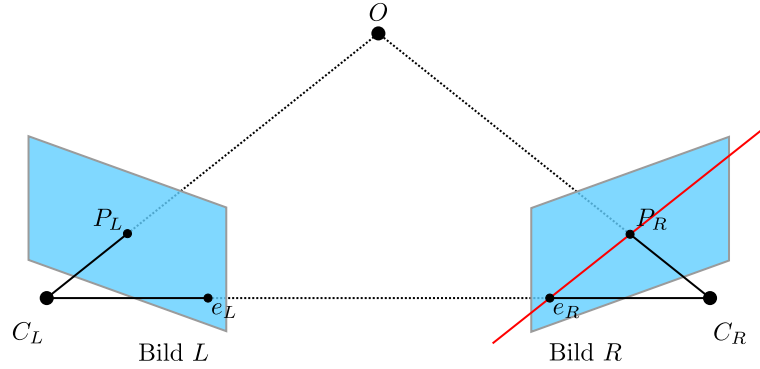


Abbildung 1: Darstellung der Epipolargeometrie.

Abbildung 1 visualisiert diesen Prozess. Gegeben sind die beiden Projektionszentren C_L und C_R der beiden Bildebenen sowie die Bildpunkte P_L und P_R . Der Objektpunkt O bildet im linken Kamerabild auf P_{x_1, y_1} ab. Zunächst ist es nur möglich den Strahl auf welchem sich O befindet zu bestimmen. Aufgrund des Wissens der extrinsischen Orientierung der Kameras, können mithilfe der Basislinie zwischen den beiden Kameras die beiden Epipole in L und R ermittelt werden. Dabei sind diese durch den Schnittpunkt der Basislinie mit den beiden Bildebenen definiert¹. Somit ist es möglich mithilfe der vorhandenen Epipole e_L , e_R die Epipolarlinie e_R, P_R zu bestimmen. Mithilfe dieser kann im Anschluss die dreidimensionale Position des Objektpunktes O bestimmt werden.

4.2 Stereo Matching

Das grundlegende Konzept des Stereo Matchings beschreibt das Finden korrespondierender Punkte in zwei simultan aufgenommenen Bildern. Die Position der beiden Kameras ist dabei leicht versetzt, um einen jeweils anderen Blickwinkel auf die Szene zu erhalten. Mithilfe der verschiedenen Perspektiven können Disparitäten (Differenz der Projektion eines Objektes vom linken zum rechten Bild) zwischen korrespondierenden Pixeln berechnet werden. Die dabei erhaltenen Tiefeninformationen eines Objektes sind mit der errechneten Disparität sowie der relativen Position der Kameras, und deren intrinsischen Parametern REF, verbunden. Im Laufe dieses Prozesses kommt es zu

¹Im Falle eines Stereonormalfalles ist ein Schnittpunkt der Basislinie mit den Bildebenen nicht möglich, sodass die Epipole in der Unendlichkeit und parallel zur x-Achse liegen. Dies hat unter anderem den Vorteil, dass die Epipolargeometrie bereits bekannt ist, und korrespondierende Bildpunkte nur noch innerhalb einer Pixelreihe gesucht werden müssen.

zwei wesentlichen Problemstellungen, der Berechnung der Disparität (Stereo Correspondence) sowie die Invertierung der Projektiven Geometrie um dreidimensionale Informationen aus der errechneten Disparität zu erhalten. Sofern eine Lösung beider Probleme vorhanden ist können diese Informationen durch einfache Triangulierung errechnet werden.

4.2.1 Klassifikation

Lokale Methoden:

Zur Berechnung der Disparität in lokalen Stereo Matching Algorithmen gilt grundlegendes Prinzip: "Finde Pixel P_2 korrespondierend zu P_1 im Referenzbild". Dabei wird die Korrelation von P_1 und P_2 durch deren lokale Umgebung bestimmt. Der Referenzpunkt ist dabei das Zentrum eines Bereiches in welchem das Matching ausgeführt wird. Geläufige Methoden dafür sind Sum of Absolute Differences (SAD) (Hirschmüller 2011), Zero-mean Normalized Cross-Correlation (ZNCC) (Chen & Medioni, 1999), (Sára, 2002), Sum of Squared Differences (SSD) (Cox et al., 1996), etc. Jedoch können aufgrund fehlender Beschränkungen verzerrte Tiefenberechnungen auftreten, da benachbarte Pixel verschiedene Disparitäten aufweisen können (beispielsweise an horizontalen Kanten wie Türrahmen etc.). Strukturell gesehen sind lokale Methoden einfacher gehalten als Globale Methoden was einen hohen Grad an Optimierung in der Implementierung zulässt.

Globale Methoden:

Bei globalen Stereo Matching Algorithmen wird ein globales Modell der zu betrachtenden Szene erstellt, sowie eine ebenfalls globale Kostenfunktion definiert, welche minimal gehalten werden soll. Dabei werden im Gegensatz zu lokalen Methoden Matches in einer Reihe und des gesamten Bildes verglichen. Zusätzlich zu den benachbarten Pixeln werden hier ebenfalls die Matches der angrenzenden Pixel betrachtet. Zur Vereinfachung dieses Vorgangs betrachten einige Algorithmen nur die Epipolargeometrie, wobei ein zweidimensionales Problem auf ein eindimensionales reduziert wird. Resultierend daraus liegt die Stärke globaler Methoden in der Bewältigung schwacher Texturen sowie auftretende Okklusionen und unterschiedlicher Lichteinfälle, was, aufgrund der höheren Komplexität in einem höheren Rechen- und Speicheraufwand mündet. Weitere Verbesserungen der Ergebnisse können durch Techniken wie Dynamische Programmierung und Graph Cut erreicht werden.

Abbildung 2: Visualisierung des Block Matching Algorithmus

4.2.2 Block Matching

Einen der einfachsten Ansätze zur Berechnung von Korrespondenz zwischen zwei Bildern bietet der Block Matching Algorithmus. Bei diesem lokalen Matching Algorithmus werden Blöcke bestimmter Größe mithilfe mathematischer Berechnungen (NCC, SAD, etc.) auf Korrespondenz untersucht. Dabei reduziert die Epipolargeometrie diesen Prozess auf ein eindimensionales Problem indem ein Block im linken Bild mit einem Block im rechten Bild auf der Epipolarlinie des rechten Bildes pixelweise verglichen wird. Das eigentliche Matching erfolgt über die Berechnung eines Blockes mithilfe einer Energiefunktion. Dabei gilt es bei der Korrespondenzanalyse mit dem möglichst gleichen Wert zu erhalten.

Diesen Prozess veranschaulicht Abbildung 2. Durch den Fokus auf einzelne Pixelreihen und deren unmittelbare Umgebung ist dieses Verfahren wesentlich schneller aber auch ungenauer als globale Matching Algorithmen.

4.2.3 Semi Global Block Matching

Das im Rahmen dieser Arbeit verwendete Verfahren zur Berechnung von Disparity Maps ist der in der freien Computer Vision Library implementierte Semi Global Block Matching Algorithmus. Diese leicht abgewandelte Implementierung des Semi Global Matching Algorithmus von Hirschmüller et. al [Hirschmüller, 2005] zeichnet sich sowohl durch seine Genauigkeit, als auch durch seine performante Berechnung aus. Unter Zuhilfenahme dieses Verfahren ist es möglich Disparity Maps in Echtzeit, mit sehr guter Framerate, zu berechnen. Im folgenden wird die grundlegende Funktionsweise des SGM erläutert. Im Anschluss daran werden die Unterschiede zum SGBM herausgearbeitet, sowie eine kurze Erklärung der Parameter dessen vorgenommen.

SGM:

Die grundlegende Idee des Semi-Global Matching Algorithmus besteht aus dem pixelweisen Matching von Mutual Information ². Die globale zweidimensionale Smoothness Einschränkung wird dabei mithilfe mehrerer eindimensionaler Einschränkungen erzeugt. Ausgangsvoraussetzung dafür sind verschiedene Bilder derselben Szene mit vorhandener und bekannter Epipolargeometrie. Zunächst

²Mutual Information(MI), zu deutsch Transinformationen, beschreiben den statistischen Zusammenhang zweier Zufallsgrößen...

wird für jeden Pixel P , aus der Intensität I_{bp} sowie der vermuteten Korrespondenz I_{mq} auf der Epipolarlinie $q = e_{bm}(pd)$ die Matching Kosten berechnet. Die eigentliche Berechnung dieser erfolgt dann mit Hilfe von Mutual Information.

$$MI_{I_1, I_2} = \sum_p mi_{I_1, I_2}(I_{1p}, I_{2p}) \quad (4.1)$$

$$mi_{I_1, I_2}(i, k) = h_{I_1}(i) + h_{I_2}(k) - h_{I_1, I_2}(i, k) \quad (4.2)$$

Zur Berechnung der MI mithilfe der in 4.1 und 4.2 dargestellten Formeln, wird zunächst eine Initial Disparity Map benötigt. Diese wird nach Kims [Zureiki et al., 2008] Ansatz zufällig gewählt um die Kosten berechnen zu können. Dies ist insbesondere für iterative verfahren geeignet, da andernfalls die Berechnung verlangsamt werden könnte. Zur Steigerung der Performanz wird die Disparity Map zunächst nur auf halber Auflösung berechnet, was den Rechenaufwand um den Faktor $2 \cdot 3$ reduziert. Zur Vermeidung falscher Kostenberechnung durch auftretende Noise innerhalb des Bildes nimmt der Algorithmus weitere Einschränkungen vor. Dabei werden die benachbarten Disparitäten in die Berechnung mit einbezogen (siehe Abbildung 4.3).

$$foo = foo \quad (4.3)$$

Das letzte Problem besteht in der Berechnung der Korrespondenz sowie der daraus resultierenden Disparitäten. Dabei wird nach der Disparität D mit der geringsten berechneten Energiefunktion gesucht. Anstatt nun einfach den minimalen Pfad der Kosten zu summieren werden alle Richtungen zur aktuellen Disparität mit einbezogen (siehe Abbildung 3)

Zur Berechnung valider Werte, sollten dabei mindestens 8 Richtungen vorliegen, eine vermehrte Anzahl, beispielsweise 16, sind bei vorteilhaft. Die berechnete Disparität ergibt sich dabei aus den minimalen Kosten anhand dieser Pfade. Dabei gilt: "Je mehr Übereinstimmungen in den Kosten, desto wahrscheinlicher ist es das D zur Intensität I korrespondiert".

SGBM:

4.3 mvStereoVision Framework

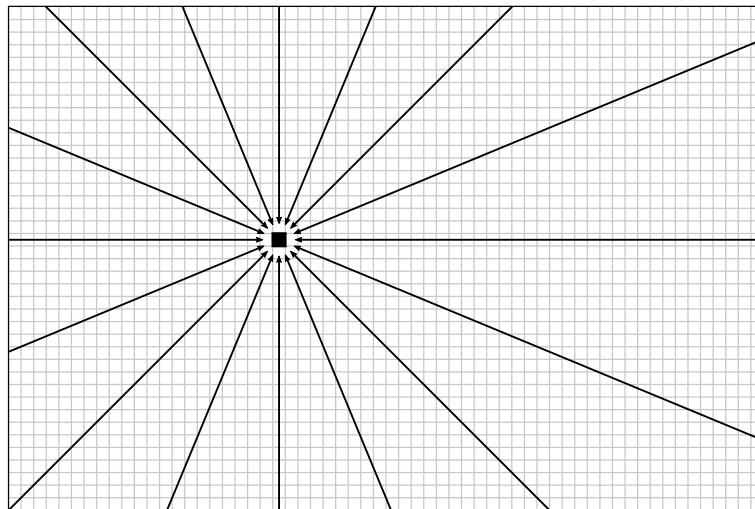


Abbildung 3: Darstellung der verschiedenen Richtungen

Kapitel 5

Subimage Detection

5.1 Konzept

5.2 Subimages

5.3 foo

Kapitel 6

Samplepoint Detection

6.1 Konzept

6.2 Erkennung

6.3 foo

Kapitel 7

Konflikte in der Erkennung und Lösungsansätze

7.1 Frame Skipping

7.2 Hindernisgröße

7.3 Position Mapping

Kapitel 8

Evaluierung

8.1 Aufbau des Testsetups

8.2 Evaluierung Subimage Detection

8.2.1 Robustheit

8.2.2 Performanz

8.3 Evaluierung Samplepoint Detection

8.3.1 Robustheit

8.3.2 Performanz

Kapitel 9

Diskussion

9.1 Anforderungsevaluierung

9.2 Gegenüberstellung beider Algorithmen

Kapitel 10

Fazit und zukünftige Arbeiten

Anhang A

Anhang 1

Abbildungsverzeichnis

1	Darstellung der Epipolargeometrie.	8
2	Visualisierung des Block Mathching Algorithmus	10
3	Darstellung der verschiedenen Richtungen	12

Tabellenverzeichnis

Literaturverzeichnis

[Hirschmuller, 2005] Hirschmuller, H. (2005). Accurate and efficient stereo processing by semi-global matching and mutual information. 2:807–814 vol. 2.

[Zureiki et al., 2008] Zureiki, A., Devy, M., and Chatila, R. (2008). *Stereo Matching and Graph Cuts*. INTECH Open Access Publisher.