



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Bachelorarbeit

Christoph Schmiedecke

Tiefenbilderzeugung mit Hilfe von
skalierungsinvarianten Merkmalen für ein
Stereokameramodul

Christoph Schmiedecke
Tiefenbilderzeugung mit Hilfe von
skalierungsinvarianten Merkmalen für ein
Stereokameramodul

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung
im Studiengang Technische Informatik
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer : Prof. Dr. -Ing. Andreas Meisel
Zweitgutachter : Prof. Dr. rer. nat. Bettina Buth

Abgegeben am 23. August 2009

Christoph Schmiedecke

Thema der Bachelorarbeit

Tiefenbilderzeugung mit Hilfe von skalierungsinvarianten Merkmalen für ein Stereokameramodul

Stichworte

Koordinatensysteme, Koordinatentransformation, 2D-/3D-Transformation, Distanzbe-
rechnung, SIFT, Impresario, LTI-Bibliothek, Stereokamera, Bumblebee

Kurzzusammenfassung

Gegenstand dieser Bachelorarbeit ist die Entwicklung eines Tiefenbildes unter Verwendung des scale-invariant feature transform (SIFT) Algorithmus. Hierbei werden mit der Bumblebee-Stereokamera Bildpaare einer Szene aufgenommen und aus den gewonnenen SIFT-Punkten die entsprechenden Raumpunkte berechnet. Diese sollen dann für die Erstellung eines Tiefenbildes genutzt werden. Zusätzlich findet eine Analyse bezüglich der Genauigkeit und Verwendbarkeit des Verfahrens für einen praktischen Einsatz, beispielsweise im Bereich der Kollisionserkennung und -vermeidung, statt.

Christoph Schmiedecke

Title of the paper

Annotating images from a stereo camera with depth information using scale-invariant features

Keywords

coordinate systems, coordinate transformation, 2D/3D transformation, distance calculation, SIFT, Impresario, LTI-Library, stereo camera, Bumblebee

Abstract

The object of this bachelor thesis is the implementation of the scale-invariant feature transform (SIFT) algorithm for finding depth information contained in pairs of pictures. The pairs of pictures are taken with the Bumblebee stereo camera system and three-dimensional coordinates are calculated for the SIFT points. The method is analysed for its use in practical applications.

Inhaltsverzeichnis

Tabellenverzeichnis	7
Abbildungsverzeichnis	8
1. Einführung	10
1.1. Motivation	10
1.2. Zielsetzung	12
1.3. Gliederung der Arbeit	13
2. Grundlagen	14
2.1. Tiefenbild	14
2.1.1. Passive Verfahren	14
2.1.2. Aktive Verfahren	17
2.2. Scale-Invariant Feature Transform (SIFT)	18
2.3. Koordinatensysteme	20
2.3.1. Weltkoordinatensystem	21
2.3.2. Projektionszentrumskoordinatensystem	22
2.3.3. Bildhauptpunktkoordinatensystem	22
2.3.4. Bildpunktkoordinatensystem	22
2.4. Kamera	22
2.4.1. Kameraparameter	23
2.4.2. Linsenverzeichnung	24
2.4.3. Kamerakalibrierung	26
2.5. Herleitung des vollständig projektiven Kameramodells	26
2.5.1. Abbildung eines Raumpunktes auf Bildkoordinaten	26
2.5.2. Bildhauptpunktverschiebung	28
2.5.3. Weltkoordinatensystem zu Projektionszentrumskoordinatensystem	30
2.5.4. Vollständig projektives Kameramodell	31
2.5.5. Raumpunktvermessung mit Stereobildpaaren	31
3. Anforderungsanalyse	34
3.1. Anforderungen an ein Tiefenbild und das Verfahren zur Bestimmung von korrespondierenden Bildpunkten	34

3.1.1. Große Menge an Tiefeninformationen	34
3.1.2. Homogene Verteilung der Tiefeninformationen	35
3.1.3. Distanzgenauigkeit	35
3.1.4. Verlässlichkeit des Verfahrens und Verfügbarkeit der Tiefeninformationen	35
3.1.5. Unabhängigkeit des Verfahrens gegenüber äußeren Einflüssen	36
3.2. Anforderungen an das Kamerasystem	36
3.2.1. Gleichheit der verwendeten Kameras und der Kameraparameter	36
3.2.2. Stabilität und Unveränderlichkeit des Kamerasystems	36
3.3. Anforderungen an das Programm zur Berechnung der Raumpunkte und Erstellung des Tiefenbildes	37
3.3.1. Bedienbarkeit	37
3.3.2. Geschwindigkeit	37
4. Konzeption	38
4.1. Bestimmung des Verfahren zur Auffindung von korrespondierenden Bildpunkten	38
4.1.1. Analyse der Punktemenge und der Punkteverteilung in verschiedenen Anwendungsfällen	38
4.1.2. Unabhängig gegenüber äußeren Einflüssen	39
4.2. Bestimmung der zu verwendenden Software	39
4.3. Bestimmung des Stereokamerasystems	40
4.4. Bestimmung der Kalibrierungssoftware	40
4.5. Genauigkeitsanalyse der berechneten Raumpunkte	40
5. Versuchsdurchführung	42
5.1. Entwicklungsumgebung des Projekts	42
5.1.1. Entwicklung des Stereomessmoduls	43
5.1.2. Kameraanbindung	43
5.1.3. Kamerakalibrierung	45
5.1.4. 3D-Modell der Kalibrierungsumgebung	48
5.2. Merkmalsextraktion	50
5.2.1. Bildaufnahme	50
5.2.2. Bestimmung der korrespondierenden Bildpunktepaare	51
5.3. Ergebnisfilterung und Raumpunktvermessung	53
5.3.1. Raumpunkte	55
5.4. Erstellung des Tiefenbildes	57
6. Analyse	59
6.1. Genauigkeitsanalyse der Raumpunkte	59
6.2. Analyse zur SIFT-Punkte-Verteilung	62
6.2.1. Flur	63
6.2.2. Labor	67

6.2.3. Arbeitsraum	69
6.2.4. Auswertung der SIFT-Punktemenge	71
6.2.5. Analyse des Verfahrens am Menschen	71
6.3. Qualitätseinflüsse	74
6.3.1. Downsampling	74
6.3.2. Bildauflösung	77
6.3.3. Helligkeit und Kontrast	78
6.3.4. Basisbreite des Stereokamerasystems	78
6.3.5. Auffälligkeiten der SIFT-Punkteverteilung	80
7. Diskussion und Ausblick	83
Literaturverzeichnis	86
A. Homogene Koordinaten	89
B. Impresario	91
C. LTI-Library	92
D. Inhalt der CD-ROM	93

Tabellenverzeichnis

5.1. Berechnete Raumpunkte	56
6.1. Auswertung der Genauigkeitsanalyse	61
6.2. Auswertung der Punkteverteilung im leeren Flur	66
6.3. Auswertung der Punkteverteilung im Flur mit Hindernissen	67
6.4. Auswertung der Punkteverteilung im Labor	69
6.5. Auswertung der Punkteverteilung im Arbeitsraum	71

Abbildungsverzeichnis

1.1. Entfernungsmessung mit Lasersystem	11
1.2. Panoramabilderzeugung: Vorstufe	12
1.3. Panoramabilderzeugung: Endstufe	12
2.1. Theoretischer Aufbau des Stereoverfahrens	15
2.2. Tiefenbild aus Stereobildern	16
2.3. Bildinformationen aus Schattierungen	16
2.4. Theorie des Laufzeitverfahrens	17
2.5. Triangulation mit Lichtstrahl	18
2.6. Triangulation mit Lichtebenen	19
2.7. Übersicht Koordinatensysteme	21
2.8. Effekte der Linsenverzeichnung	24
2.9. Aufbau der Lochkamera	27
2.10. Mathematisches Modell der Lochkamera	27
2.11. Bildhauptpunktverschiebung	29
2.12. Grundprinzip räumlicher Vorwärtsschnitt	32
5.1. Ablaufdiagramm des C++ Projekts	44
5.2. Bumblebee-Stereokamera	45
5.3. Kamerakalibrierungsmuster	46
5.4. Innere Kameraparameter	47
5.5. Kalibrierungsmodell aus Lego	47
5.6. Kalibrierungsbilder mit Referenzpunkten	48
5.7. Äußere Kameraparameter	49
5.8. 3D-Modell der Kalibrierungsumgebung	49
5.9. Erstellung der SIFT-Punktlisten	50
5.10. Impresario-Projekt	51
5.11. Gefundene SIFT-Punkte	52
5.12. Gefundene SIFT-Punktepaare	52
5.13. Liste mit SIFT-Punktepaaren	53
5.14. Berechnung der Raumpunkte	54
5.15. SIFT-Punktepaare mit mehreren Zielpunkten	55
5.16. SIFT-Punktepaare mit horizontaler Abweichung	56

5.17. Ergebnisbild mit markierten Raumpunkten	58
6.1. Versuchsaufbau Genauigkeitsanalyse	60
6.2. Fotoserie der Genauigkeitsanalyse	60
6.3. Mittlere Abweichung	62
6.4. Mittlere relative Abweichung	63
6.5. Szenario: Leerer Flur	64
6.6. Leerer Flur mit SIFT-Punktpaaren (Distanz 5m)	64
6.7. Leerer Flur mit SIFT-Punktpaaren (Distanz 20m)	65
6.8. Szenario: Flur mit Hindernissen	66
6.9. Szenario: Labor	68
6.10. Labor mit SIFT-Punktpaaren (Distanz 5m)	68
6.11. Labor mit SIFT-Punktpaaren (Distanz 15m)	69
6.12. Szenario: Reinfahrt in einen Arbeitsraum	70
6.13. Szenario: Arbeitsraum aus statischer Position	70
6.14. Auswertung der verwertbaren SIFT-Punkte	72
6.15. SIFT-Punkte an einer Person (Teil 1)	73
6.16. SIFT-Punkte an einer Person (Teil 2)	74
6.17. Impresario-Projekt mit Downsampling	75
6.18. SIFT-Punkte: Mit und ohne Downsampling	76
6.19. SIFT-Punktpaare: Mit und ohne Downsampling, Teil 1	76
6.20. SIFT-Punktpaare: Mit und ohne Downsampling, Teil 2	77
6.21. Quellbild mit Weißabgleich	79
6.22. SIFT-Punktpaare nach einem Weißabgleich	79
6.23. SIFT-Punkte an einer Person nach einem Weißabgleich	80
6.24. Schlechte Anordnung der Stereokameras	81
6.25. Gute Anordnung der Stereokameras	81

1. Einführung

Der Mensch verfügt über eine hervorragende visuelle Wahrnehmungsfähigkeit. In Sekundenbruchteilen verarbeitet das menschliche Gehirn die Eindrücke, welche von seinen Augen wahrgenommen werden und bildet daraus ein dreidimensionales Abbild seiner Umgebung. Zusätzlich spielen auch menschliche Erfahrungswerte eine wichtige Rolle, die es dem Menschen ermöglichen auch in unbekanntem oder ungewohnten Situationen eine Einschätzung seiner Umgebung zu bekommen (Erhardt 2008). Hierbei werden sowohl bekannte als auch unbekannte Objekte visuell erfasst und entsprechend den vorhandenen Erfahrungen in Größe, Entfernung und Beschaffenheit eingeordnet.

Für Maschinen, die mit der visuellen Erkennung ihrer Umgebung arbeiten, auch Vision-Systems genannt, stellt diese Aufgabe eine bisher ungelöste Herausforderung dar (Davies 2005, Hornberg 2006). Auch wenn die Entwicklungen in diesem Bereich immer weiter fortschreiten, ist der Mensch der Technik in diesem Bereich noch weit überlegen. Es gibt aber verschiedene Techniken, die den Maschinen das autonome Arbeiten bereits heute ermöglichen. So werden Vision-Systems in Bereichen wie der Positionsbestimmung, der Lokalisierung und Sortierung von Objekten, in der Qualitätskontrolle oder der Erkennung von Fehlstücken auf einem Fließband sehr erfolgreich eingesetzt.

Die meisten dieser Systeme basieren auf dem Einsatz von Infrarot, Ultra-Schall, Laser, Radar und GPS, aber auch der Einsatz von optischen Geräten wie Kameras nimmt immer weiter zu.

In diesem Sinne beschäftigt sich auch diese Arbeit mit der Abschätzung von Entfernungen, die aus den Bildern einer Stereokamera bestimmt werden.

1.1. Motivation

Sollen in einem Arbeitsbereich mehrere Maschinen autonom agieren, so bedarf es hierbei einiger Anforderungen an das Umfeld und die Maschinen. Befinden sich zusätzlich Menschen im Arbeitsbereich, so steigen diese Anforderungen noch weiter. Der Schwerpunkt liegt dabei im Bereich der Sicherheit, denn ein Menschenleben bedeutet mehr als der Verlust einer

Maschine. Ein mögliches Einsatzgebiet von Vision-Systemen in einer solchen Arbeitsumgebung bildet der Kollisionsschutz oder die Kollisionsvermeidung. Besonders in einem Arbeitsumfeld, wo große Maschinen und schwere Ladungen bewegt werden, z.B. in Häfen oder in Hochregallagern, ist es wichtig Kollisionen zu vermeiden. Vorhandene Systeme basieren hier häufig auf dem Einsatz von Sensoren und Schaltern oder auf Lasertechnik. Im Bereich der Lasertechnik werden für die Abstandsmessung bzw. Positionsbestimmung so genannte Reflexionsmarken an den beteiligten Objekten benötigt. Bild 1.1 zeigt den Einsatz eines solchen Lasersystems im Hafen von Shanghai.



Abbildung 1.1.: Bockkrane im Hafen von Shanghai mit Lasersystem zur Entfernungsmessung (Reflektor (links) und Sensor (rechts)) (Liebske 2005, S. 113).

Einen weiteren Ansatz für eine flexible Navigation und Kollisionsvermeidung birgt der Einsatz von Kameras, welche an jedem autonomen Gerät montiert werden. Die Kollisionskontrolle erfolgt lediglich über den Abgleich der korrespondierenden Kamerabilder. Die Distanzberechnung zu potenziellen Kollisionsobjekten basiert hierbei auf dem Einsatz von zwei Kameras. Eine praxistaugliche Lösung muss gewährleisten, dass mögliche Kollisionsobjekte auf den von ihnen aufgenommenen Bildern erfasst werden und so die Entfernung zum jeweiligen Objekt berechnet werden kann. Das „scale-invariant feature transform“ Verfahren (SIFT-Verfahren) das von David Lowe (Lowe 2004) entwickelt wurde, bietet hierfür eine Umsetzungsmöglichkeit. Die grundlegende Idee des Verfahrens ist die Wiedererkennung von markanten Punkten (SIFT-Punkten) in korrespondierenden Bildern, über die der Abstand zu Kollisionsobjekten bestimmt werden kann. Seine Anwendbarkeit hat das SIFT-Verfahren bereits im praktischen Einsatz, wie beispielsweise der Erstellung von Panoramabildern, unter Be-

weis gestellt¹. Hierbei werden fortlaufend Fotos von dem gewünschten Zielobjekt aufgenommen. Die Bilder müssen über Bildüberschneidungen verfügen und werden später anhand von gemeinsamen SIFT-Punkten zusammengefügt (vergleiche hierzu Abbildung 1.2).



Abbildung 1.2.: Panoramabildherzeugung - Vorstufe: Teilbilder mit SIFT-Punkten (links) und übereinander gelegte Teilbilder anhand gefundener SIFT-Punktepaare (rechts) (Gremse 2005).

Nachdem die Bilder erfolgreich zusammengefügt worden sind, erhält man ein Panoramabild wie in Abbildung 1.3.



Abbildung 1.3.: Panoramabildherzeugung - Endstufe: Fertiges Panoramabild (Gremse 2005)

1.2. Zielsetzung

Ziel dieser Arbeit ist es, mittels einer Stereokamera und unter Verwendung des SIFT-Verfahrens, ein räumliches Tiefenbild einer beliebigen Umgebung zu erzeugen. Im Zuge dieser Entwicklung soll die Genauigkeit und die Zuverlässigkeit des Verfahrens untersucht und ein möglicher Einsatz in praktischen Anwendungen, speziell im Bereich Kollisionserkennung und -vermeidung, geprüft werden.

¹Weitere Einblicke in das Thema SIFT bieten unter anderem folgende Arbeiten: Im Bereich der Panoramabildherzeugung (Gremse 2005) und (Mildner 2008). Im Bereich visuelle Navigation und Positionsbestimmung (Schatz 2006) und (Hering 2008).

Die zentrale Rolle spielt hierbei die Berechnung von Raumpunkten aus korrespondierenden SIFT-Punkten sowie die Genauigkeitsanalyse des SIFT-Verfahrens und die Anwendbarkeit in verschiedenen Szenarien.

1.3. Gliederung der Arbeit

Die Arbeit ist in folgende Abschnitte unterteilt:

- **Kapitel 2 - Grundlagen:**

Erläuterung der grundlegenden Funktionsweise des SIFT-Algorithmus, Grundlagen aus den Bereichenameratechnik und Koordinatensysteme, sowie der Herleitung des vollständig projektiven Kameramodells für die Berechnung der Raumpunkte.

- **Kapitel 3 - Anforderungsanalyse:**

Ermittlung der generellen Anforderungen für die Realisierung eines Tiefenbildes und an das hierfür benötigte Verfahren zur Ermittlung von korrespondierenden Bildpunkten sowie der Anforderungen an ein Programm zur Realisierung dieser Anforderungen.

- **Kapitel 4 - Konzeption:**

Erstellung eines Konzeptentwurfs für die praktische Umsetzung der in der Anforderungsanalyse aufgestellten Anforderungen. Festlegung des Verfahrens zur Ermittlung von korrespondierenden Bildpunkten sowie der benötigten Hard- und Software.

- **Kapitel 5 - Versuchsdurchführung:**

Darstellung der Entwicklungsumgebung und Entwicklung eines C++ Programms für die Berechnung der Raumpunkte. Erläuterung der einmaligen und sich wiederholenden Arbeitsvorgänge im realen Einsatz.

- **Kapitel 6 - Analyse:**

Genauigkeitsanalyse der durch den SIFT-Algorithmus gewonnenen Raumpunkte und Analyse der Anwendbarkeit in der Praxis in Hinblick auf die gefundene Menge an Merkmalspunkten. Des Weiteren werden die Faktoren untersucht und beschrieben, die einen Einfluss auf die Punktequalität haben.

- **Kapitel 7 - Diskussion und Ausblick:**

Am Ende findet eine Zusammenfassung über die gewonnenen Erkenntnisse statt und es wird ein Ausblick auf mögliche Entwicklungen und Verbesserungen gegeben.

2. Grundlagen

In diesem Kapitel wird ein Einblick in das Thema Tiefenbilder gegeben. Es werden die benötigten Grundlagen aus den Bereichen der Bildverarbeitung undameratechnik erläutert und die zentrale Formel zur Berechnung von Raumpunkten aus zwei oder mehreren Bildpunkten hergeleitet. Des Weiteren wird das für diese Bachelorarbeit wichtige „scale-invariant feature transform“ Verfahren (SIFT-Verfahren) beschrieben.

2.1. Tiefenbild

In diesem Abschnitt soll ein Einblick in das Thema Tiefenbilder gegeben und einige Verfahren angeschnitten werden, die für eine Tiefenbilderstellung in Frage kommen. In dieser Arbeit wird jedoch nur das Stereoverfahren zur Berechnung der Tiefeninformationen verwendet.

Unter einem Tiefenbild versteht man ein Bild, das dem Betrachter Tiefeninformationen über den dargestellten Inhalt bietet. Die Gestalt der betrachteten dreidimensionalen Objekte kann mit Hilfe eines Tiefenbildes explizit dargestellt werden (Jiang und Bunke 2007, S.3 ff.). Die Entfernungen werden hierbei durch verschiedene Grauabstufungen repräsentiert.

Bei der Erstellung von Tiefenbildern unterscheidet man zwischen passiven und aktiven Verfahren (Jiang u. a. 2007).

2.1.1. Passive Verfahren

Bei den passiven Verfahren werden die Objekte mit Hilfe einer Kamera aufgenommen und die erhaltenen Bilder dann ausgewertet. In diesem Abschnitt werden zwei passive Verfahren vorgestellt.

Stereoverfahren:

Das Stereoverfahren basiert auf der Verwendung von Stereobildern. Die Grundlage für die Berechnung der Entfernungen von Objekten bildet hierbei das Triangulationsprinzip¹. Dieses setzt voraus, dass korrespondierende Bildpunkte in den Bildpaaren gefunden werden. Mit Hilfe der korrespondierenden Bildpunkte und den inneren und äußeren Kameraparametern können dann die zugehörigen Entfernungen berechnet werden. Erläuterungen zu den inneren und äußeren Kameraparametern sowie zu der Berechnung der Entfernung folgen im Verlauf des Grundlagenkapitels in den Abschnitten 2.4 (Kamera) und 2.5 (projektives Kameramodell). Der theoretische Aufbau des Verfahrens ist in Abbildung 2.1 zu sehen. Eine mögliche Form des Tiefenbilds ist in Abbildung 2.2 dargestellt.

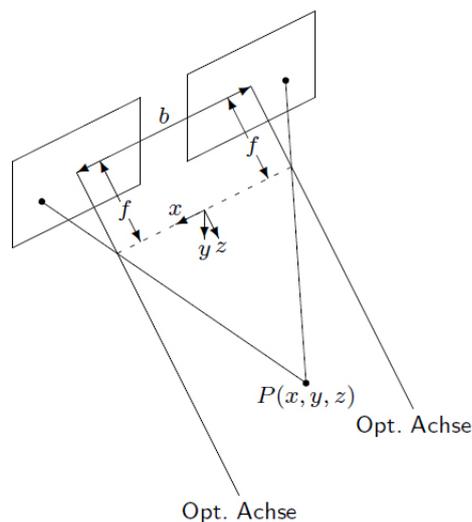


Abbildung 2.1.: Theoretischer Aufbau des Stereoverfahrens (Jiang u. a. 2007, S.3). Die verwendeten Kameras sind parallel zueinander angeordnet. Die Entfernung eines Punktes $P(x, y, z)$ kann ermittelt werden, wenn dieser in beiden Kameraaufnahmen identifizierbar ist.

Form aus Schattierung:

Dieses passive Verfahren basiert auf der Darstellung der reflektierten Lichtintensität betrachteter Objekte. Die reflektierte Lichtintensität wird hierbei von einer Kamera aufgenommen. Das Ziel bei diesem Verfahren ist die Berechnung der Oberflächenorientierung. Der Versuchsaufbau ist in Abbildung 2.3 zu sehen.

¹Triangulationsverfahren: Das Verfahren ermöglicht die optische Abstandsmessung eines Punktes mit Hilfe von trigonometrischen Funktionen. Hierbei werden von zwei, auf einer Gerade liegenden, Punkten, mit bekanntem Abstand, Winkelmessungen zu einem beliebigen Punkt im Raum durchgeführt, um dessen Lage zu bestimmen. Zwischen den Punkten entsteht ein gedachtes Dreieck.



Abbildung 2.2.: Tiefenbild aus Stereobildern (Fusiello u. a. 2000). Beim Tiefenbild (rechts) repräsentieren die unterschiedlichen Graustufen die Entfernung der Objekte zur Kamera.

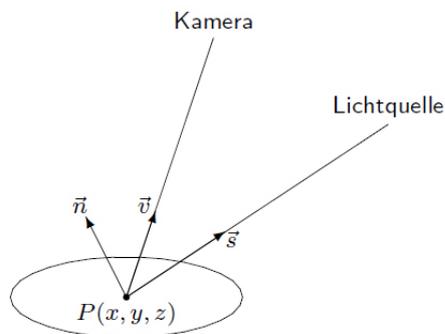


Abbildung 2.3.: Das Bild zeigt den Versuchsaufbau für die Bestimmung von Bildinformationen aus Schattierungen (Jiang u. a. 2007, S.9). Die von der Lichtquelle ausgehenden Lichtstrahlen werden durch Objekte reflektiert.

2.1.2. Aktive Verfahren

Im Gegensatz zu den passiven Verfahren, wird bei den aktiven Verfahren die Kamera durch eine aktive Energiequelle ersetzt. Im Folgenden werden drei mögliche Verfahren erläutert.

Laufzeitverfahren:

Beim Laufzeitverfahren wird ein Signal von einem Emitter ausgestrahlt und über einen sogenannten Ablenker abgelenkt. Die abgelenkten Strahlen treffen auf ein Objekt und werden von diesem reflektiert. Diese wiederum abgelenkten Strahlen werden dann von einem Empfänger aufgefangen. Der Empfänger ist an einen Detektor angeschlossen, an den die Daten übermittelt werden. Durch den Vergleich mit einem Referenzsignal des Emitters ermittelt dieser die Laufzeit und die entsprechende Objektentfernung. Der schematische Versuchsaufbau eines solchen Systems ist in Abbildung 2.4 dargestellt.

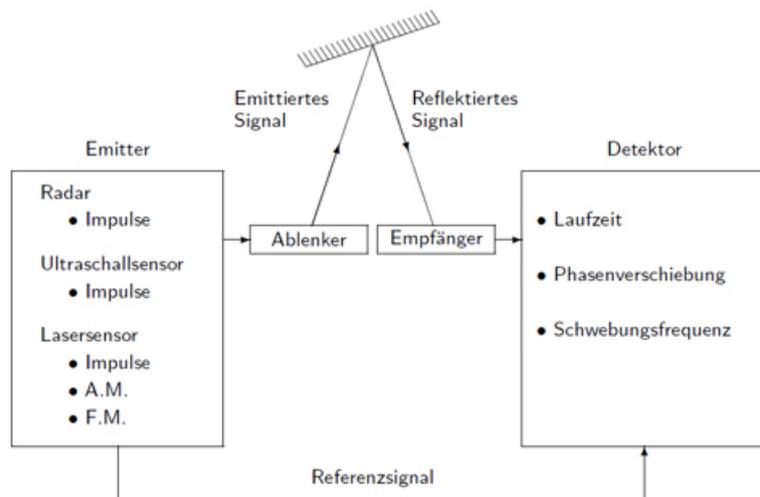


Abbildung 2.4.: Schematischer Aufbau des Laufzeitverfahrens (Jiang u. a. 2007, S.12).

Triangulationsverfahren:

Bei den Triangulationsverfahren unterscheidet man zwischen der Projektion von Lichtstrahlen und von Lichtebenen. Die beiden Verfahren sollen hier kurz dargestellt werden.

Projektion von Lichtstrahlen:

Die Realisierung dieses Verfahrens basiert auf dem Einsatz eines Senders und eines Empfängers. Hierbei sendet der Sender ein Signal aus, welches von einem Hindernis reflektiert und von dem Empfänger aufgenommen wird. Durch diese Anordnung ist das benötigte Dreieck zur Bestimmung des Objektabstands durch Triangulation vollständig gegeben. Eine einfache Implementierung des Verfahrens besteht aus einem Lichtprojektor als Sender und

einer Kamera als Empfänger. Der Lichtprojektor projiziert einen Lichtstrahl auf die betrachtete Szene und der entstehende Lichtfleck wird von der Kamera aufgezeichnet. Nachdem die Position des Lichtflecks in der Bildebene der Kamera gemessen worden ist, liegt das benötigte Dreieck für die Positionsbestimmung des entstandenen Lichtflecks vor. Um eine komplette Szene erfassen zu können, müssen die Lichtstrahlen in horizontaler und vertikaler Richtung abgelenkt werden. Abbildung 2.5 zeigt eine mögliche Versuchsanordnung.

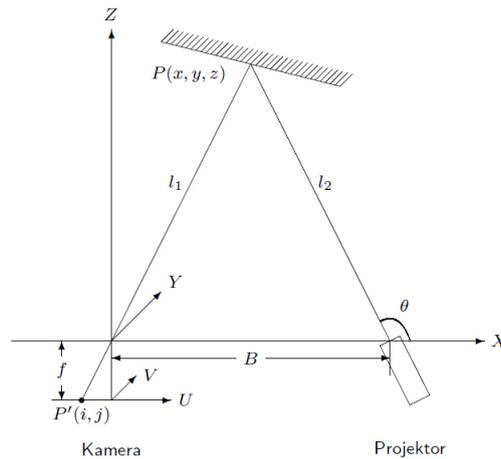


Abbildung 2.5.: Eine mögliche Triangulation mit einem Sender (Lichtstrahl) und einem Empfänger (Kamera) (Jiang u. a. 2007, S.13). Der vom Sender ausgehende Lichtstrahl wird vom Objekt reflektiert und von der Kamera erfasst.

Projektion von Lichtebenen:

Im Gegensatz zur Projektion von Lichtstrahlen kommt bei der Projektion von Lichtebenen ein Laser als Sender zum Einsatz. Dieser projiziert eine Lichtebene auf ein Objekt oder in den Raum, welche als Profillinie im Kamerabild aufgezeichnet wird. Für jeden Punkt auf dieser Profillinie wird dann die entsprechende Entfernung bestimmt. Um eine komplette Szene erfassen zu können, müssen die Lichtstrahlen in horizontaler Richtung abgelenkt werden. Abbildung 2.6 zeigt einen Versuchsaufbau des Verfahrens.

2.2. Scale-Invariant Feature Transform (SIFT)

Scale-invariant feature transform wurde von David G. Lowe entwickelt (Lowe 2004) und beschreibt ein Verfahren zum auffinden von markanten unveränderlichen Formen und Punkten in Bildern. Es eignet sich besonders zum Vergleich von verschiedenen Bildern eines Objektes oder einer Szene. Das Verfahren ist unempfindlich gegenüber Skalierungen und Rotationen des Bildes und relativ stabil gegenüber affinen Verzerrungen und Verformungen, z.B.

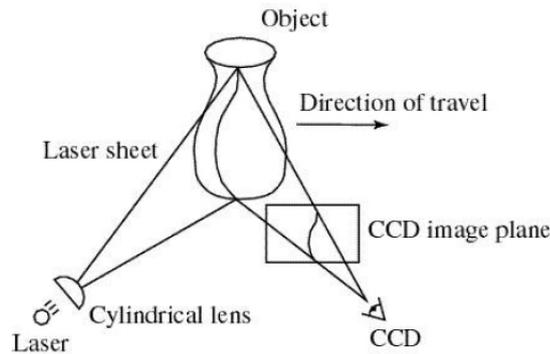


Abbildung 2.6.: Triangulation mit Lichtebenen (Jiang u. a. 2007, S.14). Die vom Laser ausgehende Lichtebene wird vom Objekt reflektiert und als Profillinie in der Kamera wahrgenommen.

durch verschiedene Aufnahmestandpunkte, und gegenüber unterschiedlichen Lichtverhältnissen. Diese Eigenschaften sind besonders wichtig, wenn das eingesetzte Stereokamerasystem eine große Basisbreite aufweist, die einzelnen Kameras also nicht direkt nebeneinander angeordnet sind. Hierbei können eine Schrägstellung des Fahrzeugs oder unterschiedliche Lichtverhältnisse an den einzelnen Kameras zu Unterschieden zwischen den Aufnahmen führen. Durch die Eigenschaften des SIFT-Verfahrens haben diese Unterschiede beim erkennen von Merkmalspunkten keine Auswirkung beziehungsweise werden diese minimiert. Die gefundenen einzelnen Merkmalspunkte werden durch folgende Parameter beschrieben:

- Die (x, y) Bildkoordinaten des Merkmalspunktes: Ein Bildpunkt kommt nur als potenzieller Merkmalspunkt in Frage, wenn er ein Extremum in seiner Nachbarschaft der ihn umgebenden Bildpunkte darstellt. Dieses wird im Laufe des Verfahrens über verschiedene Stufen einer Gaußpyramide ermittelt.
- Die Lage des Merkmalspunktes im Raum: Die Lage wird durch die Länge (Skalierung) und Richtung (Orientierung) seines Gradienten² beschrieben.
- Einen 128-stelligen Merkmalsvektor: Der Merkmalsvektor beschreibt den Merkmalspunkt über dessen Umgebung. Hierfür werden mehrere Orientierungs-Histogramme erstellt. Diese spiegeln die Orientierung der Gradienten in der Umgebung des Merkmalspunktes wieder. Die Daten werden in einem 4x4 großen Array mit jeweils 8 Elementen abgespeichert.

²Ein Gradient ist ein Vektor, der in die Richtung des größten Funktionswertes zeigt und dessen Länge ein Maß für die Steigung (Anstieg des Funktionswertes) ist. Für jeden Bildpunkt zeigt der Gradient, wie groß die stärkste Kontraständerung ist und in welcher Richtung diese liegt (Lange 2008, S. 23).

Durch die relativ gute Stabilität der Merkmalsvektoren über verschiedene Bilder einer aufgenommenen Szene hinweg, ist es möglich, gleiche Punkte in den verschiedenen Aufnahmen einander zuzuordnen. Dies geschieht im Normalfall über den Vergleich der Euklid'schen Distanz der Merkmalsvektoren.

Die Berechnung der markanten Raumpunkte erfolgt in vier Schritten:

1. **Extrema werden im Skalenraum erkannt:** Das Programm sucht nach in Frage kommenden Punkten über alle Bildbereiche und Skalierungen hinweg. Hierbei wird eine Funktion verwendet, die einen Vergleich zwischen mehreren Gauss geglätteten Bildern (difference of Gaussian) vornimmt. Es wird sichergestellt, dass die gefundenen Punkte gegenüber Skalierung und Orientierung invariant sind.
2. **Die Merkmalspunkte werden lokalisiert:** Die genaue Position und Skalierung der Kandidaten wird bestimmt. Die Auswahl findet anhand von Stabilitätsmerkmalen statt, welche vor dem Suchlauf festgelegt werden können.
3. **Die Orientierung wird bestimmt:** Jedem Schlüsselpunkt werden ein oder mehrere Orientierungen zugewiesen, die auf den lokalen Gradienten-Richtungen basieren. Die Punkte die gegenüber Transformationen bezüglich der Orientierung, Skalierung und Position invariant sind können im Weiteren verwendet werden.
4. **Der Merkmalsvektor wird erzeugt:** Die Merkmalsvektoren werden über einen vorgegebenen Radius des Merkmalspunktes ermittelt und stellen damit die Eindeutigkeit gegenüber Verzerrungen des Bildes und Änderungen der Lichtverhältnisse sicher.

Das SIFT-Verfahren wird in dieser Arbeit in Zusammenhang mit der Software Impresario (Anhang B) verwendet. Das hierfür benötigte Impresario-Makro „SIFTSearchFor“ stammt von Emmanuel Lange (Lange 2008), welches auf der ursprünglich von Andrea Vedaldi (Vedaldi 2006) implementierten Version des SIFT-Verfahrens „SIFT++“ beruht. Das SIFT-Verfahren spielt in dieser Arbeit eine zentrale Rolle, da die Bildpunkte der gefundenen SIFT-Punkte für die Distanzberechnung des Tiefenbildes verwendet werden.

2.3. Koordinatensysteme

Die Position eines Punktes kann anhand seiner Koordinaten in einem beliebigen Bild oder Raum ermittelt werden. Die Koordinaten erhält man, indem jeder Raumrichtung eine Koordinatenachse zugeordnet wird. Wenn man die Lage von Punkten in einem 3D-Raum beschreiben will, so gibt es in diesem Fall zwei sinnvolle Koordinatensysteme, die dies ermöglichen. Das Weltkoordinatensystem und das Projektionszentrumskordinatensystem (Kamerakordinatensystem). Für die Darstellung von 2D-Punkten bietet sich das Bildkoordinatensystem

an. Der Zusammenhang der verschiedenen Koordinatensysteme ist in Abbildung 2.7 zu sehen. Für das Thema Koordinatensysteme vergleiche auch (Jähne 1997, S. 190 ff.).

Folgende Notationen werden in den nächsten Abschnitten verwendet:

- \mathbf{x} : Vektoren / 2D-Bildpunkte
- x : Koordinate eines 2D-Bildpunktes
- \mathbf{X} : Matrizen / 3D-Punkt in Welt- oder Projektionszentrumskordinatensystem (Kamerakordinatensystem)
- X : Koordinate eines 3D-Punktes

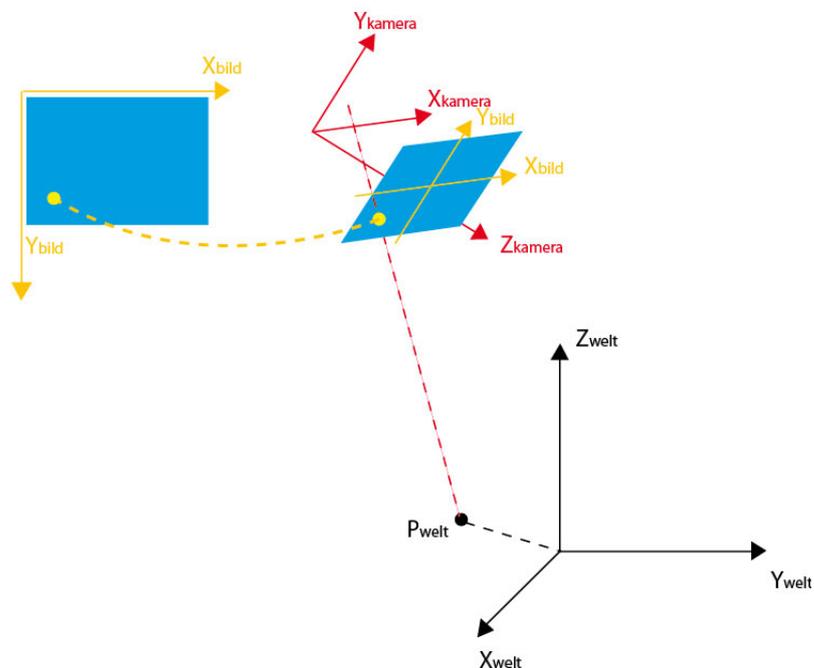


Abbildung 2.7.: Transformation von Weltkoordinaten in Kamerakordinaten und von Kamerakordinaten in Pixelkoordinaten (Hofmann 2008/09)

2.3.1. Weltkoordinatensystem

Bei dem Weltkoordinatensystem handelt es sich um die Sicht auf eine bestimmte Szene mit einem frei zu wählenden Ursprung. Die Lage des Punktes P_w im Raum lässt sich folgender-

maßen darstellen:

$$\mathbf{P}_W = [X_W, Y_W, Z_W].$$

Jede Kamera kann auf ein Weltkoordinatensystem (Weltkoordinatensystem der Kamera) kalibriert werden. Ergebnisse späterer Berechnungen beziehen sich dann auf den Ursprung des Weltkoordinatensystems der Kamera.

2.3.2. Projektionszentrumskoordinatensystem

Das Projektionszentrumskoordinatensystem, manchmal auch Kamerakoordinatensystem genannt, ist mit der sich bewegenden Kamera verknüpft. Hierbei fällt die Z-Achse mit der optischen Achse des Kamerasystems zusammen. Die Kurzschreibweise in Projektionszentrumskoordinaten lautet

$$\mathbf{P}_C = [X_C, Y_C, Z_C].$$

2.3.3. Bildhauptpunktkoordinatensystem

Das Bildhauptpunktkoordinatensystem beschreibt ein zweidimensionales Koordinatensystem mit dem Ursprung in Zentrum. Hierbei beschreibt x die horizontale und y die vertikale Lage eines Punktes. Diese Anordnung entspricht der Abbildung von 2D-Koordinaten in der Kamera.

2.3.4. Bildpunktkoordinatensystem

Das Bildpunktkoordinatensystem beschreibt ein zweidimensionales Koordinatensystem mit dem Ursprung in der oberen linken Ecke. Hierbei beschreibt x die horizontale und y die vertikale Lage eines Punktes. Diese Anordnung entspricht der in dieser Arbeit verwendeten Darstellung von Bildpunkten für die späteren Berechnungen und ist auch Grundlage gängiger Bildverarbeitungsprogramme.

2.4. Kamera

In diesem Abschnitt werden einige Grundlagen zu Kameras beschrieben. Der Schwerpunkt liegt hierbei auf bestimmten Parametern und Gegebenheiten von Kameras, welche für die späteren Berechnungen der 3D-Koordinaten der Raumpunkte benötigt werden, beziehungsweise Einfluss auf die Ergebnisse haben können.

2.4.1. Kameraparameter

Bildet man mit einer Kamera einen 3D-Raum auf ein 2D-Bild ab, so hängt das Bild von den Parametern der Kamera ab. Je nach Kameramodell unterscheiden sich die zu bestimmenden Kameraparameter. Man unterscheidet zwischen inneren und äußeren Parametern.

Innere Parameter

Innere oder intrinsische Parameter, sind Parameter der inneren Orientierung und sind unabhängig von der Lage und Ausrichtung der Kamera im Weltkoordinatensystem. Sie definieren den Zusammenhang zwischen dem 3D-Projektionszentrumskoordinatensystem (Kamerakoordinatensystem (metrisch)) und dem 2D-Bildkoordinatensystem (pixel). Es gibt folgende innere Kameraparameter:

- Auftreffpunkt der optischen Achse³ der Kamera auf der Bildebene
- Chipgröße beziehungsweise Pixelanzahl (z.B. 640x480)
- Kamerakonstante (Bildweite f)
- Verzerrungskoeffizienten der radialen Linsenverzeichnung (k_1, k_2, k_4)
- Verzerrungskonstanten der tangentialen Linsenverzeichnung (p_1, p_2)

Äußere Parameter

Äußere Parameter oder extrinsische Parameter, sind Angaben zur äußeren Orientierung der Kamera, wie Position und Ausrichtung, in einem Weltkoordinatensystem. Sie dienen der Wiederherstellung des Zusammenhangs zwischen dem 3D-Weltkoordinatensystem und dem 3D-Projektionszentrumskoordinatensystem (Kamerakoordinatensystem). Es gibt folgende äußere Kameraparameter:

- Translationsvektor oder Raumpunkt der Kamera (X, Y, Z)
- Kamerawinkel oder die Ausrichtung der Kamera (ω, ϕ, κ)

³Optische Achse: Eine in Längsrichtung durch die Mitte eines Objektivs verlaufende gedachte Linie, welche senkrecht auf der Bildebene auftrifft.

2.4.2. Linsenverzeichnung

Im Gegensatz zum Modell der Lochkamera, auf welches im Abschnitt 2.5.1 eingegangen wird, verfügen reale Kameras über ein Objektiv. Hierdurch kommt es zu verschiedenen optischen Effekten, die auch als Linsenverzeichnung bezeichnet werden. Diese Verzeichnung ist bei Weitwinkelobjektiven besonders groß. Bei manchen Kameras, z.B. bei Kameras mit einem so genannten Fischaugenobjektiv, ist dieser Effekt sogar erwünscht, um einen möglichst großen Sichtbereich aufnehmen zu können. Um von einem unverzerrten Bild ausgehen zu können und um Abbildungsfehler zu vermeiden, muss diese Verzerrung bei der späteren Berechnung beachtet und ausgeglichen werden.

In Abbildung 2.8 sieht man die resultierenden Fehler, die aufgrund der Linsenverzeichnung entstehen können.

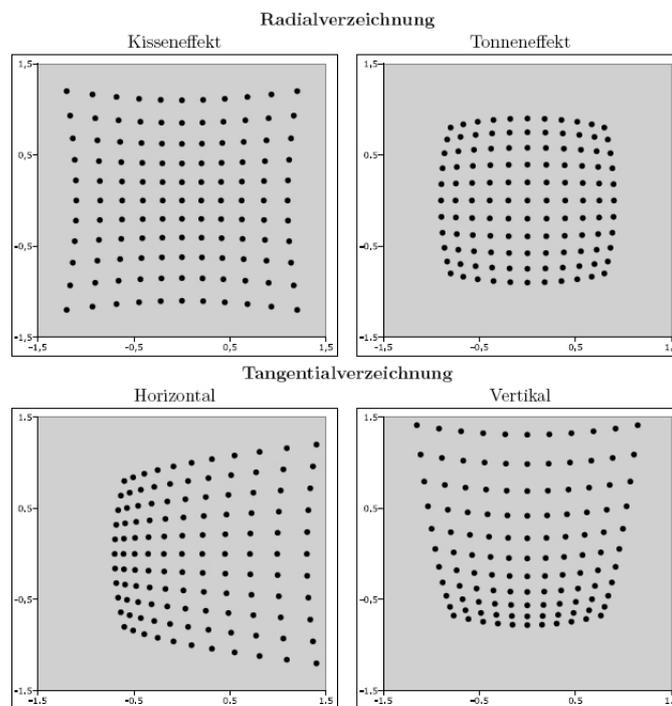


Abbildung 2.8.: Die Abbildung zeigt die Effekte der radialen und tangentialen Linsenverzeichnung (Eschenburg 2006, S. 25)

Radiale Linsenverzeichnung

Die radiale Linsenverzeichnung resultiert aus der Umlenkung beziehungsweise Ablenkung der Lichtstrahlen im Objektiv. Die verschiedenen Punkte werden ausgehend von der opti-

schen Achse kreisförmig verzerrt. Bei der radialen Verzerrung unterscheidet man zwei Formen der Verzerrung:

- tonnenförmige Verzerrung (barrel distortion): Der Punkt ist in einem Abstand vom optischen Zentrum abgebildet, der kleiner ist als bei einer idealen Projektion.
- kissenförmige Verzerrung (pin-cushion distortion): Der Punkt ist in einem Abstand vom optischen Zentrum abgebildet, der größer ist als bei einer idealen Projektion.

Mathematisch lässt sich die radiale Verzerrung so darstellen (Hartley und Zissermann 2003, S. 191):

$$\mathbf{p}_u = \mathbf{p}_h + L(r)(\mathbf{p}_v - \mathbf{p}_h), \text{ wobei } L(r) = 1 + k_1 r + k_2 r^2 + k_4 r^4, \text{ } r \geq 0 \text{ und } L(0) = 1.$$

- \mathbf{p}_u = der unverzerrte Bildpunkt
- \mathbf{p}_v = der verzerrte Bildpunkt
- \mathbf{p}_h = Bildhauptpunkt
- k_1, k_2, k_4 = Verzerrungskoeffizienten
- $L(r)$ = Verzerrungsfunktion
- r = radialer Abstand zur optischen Achse

Tangentiale Verzerrung

Eine tangentielle Verzerrung tritt durch eine Verkippung des Kamerachips auf. Dies geschieht, wenn der Chip in der Kamera nicht eben montiert ist. Die Folge ist, dass die einfallenden Lichtstrahlen nicht senkrecht auf dem Chip auftreffen.

Der Einfluss der tangentialen Verzerrung ist im Vergleich zur radialen Verzerrung meist geringer. Bei hohen Genauigkeitsanforderungen muss diese Verzerrung jedoch mitbestimmt werden. Im Gegensatz zur radialen lässt sich die tangentielle Verzerrung nicht durch eine einfache mathematische Funktion des Bildradius darstellen. Die Verzerrungskonstanten q_1 und q_2 werden direkt mit den Koordinaten des betrachteten Bildpunktes verknüpft.

Mathematisch lässt sich die tangentielle Verzerrung auf folgende Weise darstellen (Mircheva 2004, S. 10):

$$x_v = x_u + x_u(2q_1 x_u y_u + q_2(r^2 + 2x_u^2))$$

$$y_v = y_u + y_u(2q_2 x_u y_u + q_1(r^2 + 2y_u^2)).$$

2.4.3. Kamerakalibrierung

Unter der Kamerakalibrierung versteht man die Ermittlung der inneren und der äußeren Parameter einer Kamera. Auf diese Weise erhält man die Abbildung der Kamera. Dieser Vorgang ist sehr wichtig und sollte sorgfältig durchgeführt werden, da die Parameter für die späteren Rechnungen vonnöten sind und bei schlecht gemessenen Parametern falsche Werte berechnet werden. Der Vorgang der Kalibrierung wird im Kapitel 5 (Versuchsdurchführung) beschrieben.

2.5. Herleitung des vollständig projektiven Kameramodells

In diesem Abschnitt wird nun das vollständige projektive Kameramodell, sowie die daraus abgeleitete Formel für die Raumpunktvermessung mit Stereobildern, hergeleitet. Die Herleitung orientiert sich hierbei an den Vorlesungsfolien von Herrn Professor Meisel (vergleiche Meisel 2008, S. 19 ff.). Die nötigen Arbeitsschritte lassen sich in folgende Abschnitte gliedern:

- Abbildung eines Raumpunktes auf Bildkoordinaten
- Bildhauptpunktverschiebung
- Umwandlung des Weltkoordinatensystems in das Projektionszentrumskordinatensystem
- Vollständiges projektives Kameramodell
- Raumpunktvermessung mit Stereobildern

2.5.1. Abbildung eines Raumpunktes auf Bildkoordinaten

Interessiert man sich für die 3D-Koordinaten eines Punktes in einem aufgenommenen Foto, so kann man sich die Beziehung zwischen dem Foto und der realen Welt mit Hilfe des Lochkameramodells veranschaulichen, welches in Abbildung 2.9 dargestellt ist.

Nur der von einem Punkt $\mathbf{P} = (X, Y, Z)$ des Objektes ausgehende Lichtstrahl, der durch die Lochblende fällt, trifft die Bildebene im Punkt $\mathbf{p} = (x, y)$ und erzeugt dort ein unverzerrtes Abbild des 3D-Punktes. Hierbei geht jedoch die Information über den Abstand, also die dritte Dimension, verloren. Die mathematischen Zusammenhänge lassen sich mit Hilfe der Abbildung 2.10 nachvollziehen.

Nach dem Strahlensatz gilt:

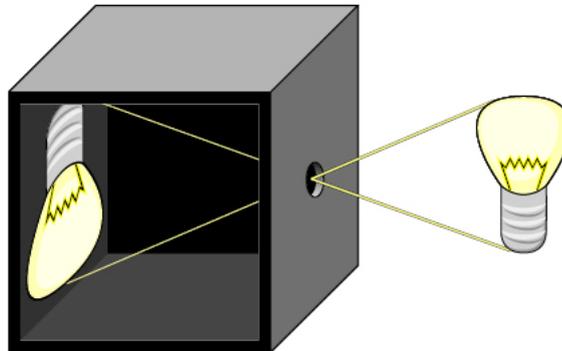


Abbildung 2.9.: Aufbau der Lochkamera (Eschenburg 2006, S. 22). Im Gegensatz zu einer richtigen Kamera mit Objektiv, tritt das Licht durch eine einfache Öffnung in die Kamera ein.

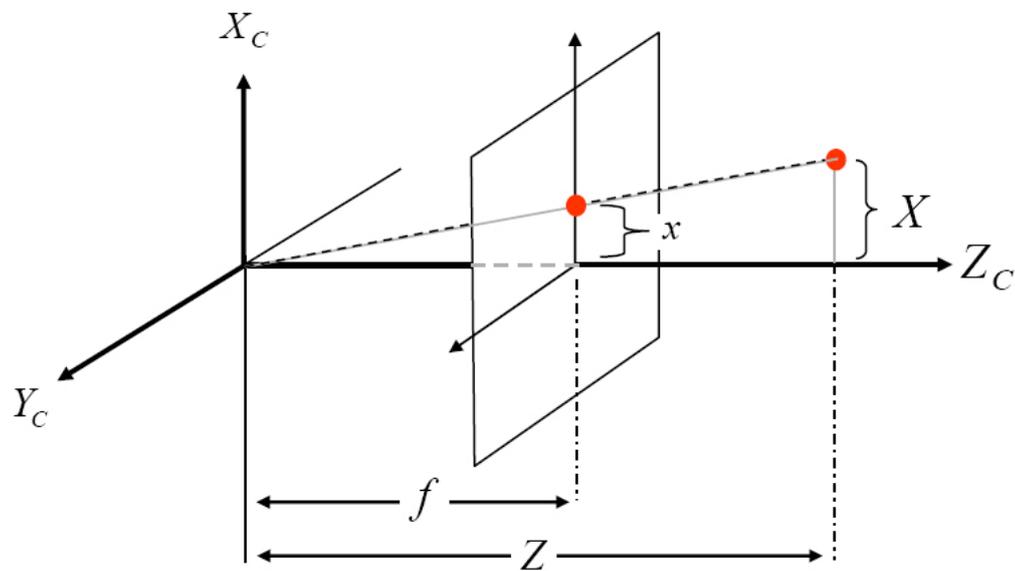


Abbildung 2.10.: Mathematisches Modell der Lochkamera (Meisel 2008, S. 19). Es zeigt die Beziehung einer Abbildung der 3D-Koordinaten eines Raumpunktes in Projektionszentrumskordinaten auf seine zugehörigen 2D-Bildhauptpunktkoordinaten.

$$\frac{X}{Z} = \frac{x}{f} \Leftrightarrow x = f \frac{X}{Z}.$$

Durch die Gesamtheit der auftreffenden Lichtstrahlen wird ein Bild erzeugt. Aus dem gegebenen Modell kann man folgende Beziehung zwischen den 3D-Koordinaten des abzubildenden Punktes und den 2D-Bildhauptpunktkoordinaten ableiten (für y kann man eine zur obigen Gleichung analoge Relation herleiten):

$$\mathbf{p} = (x, y)^T = \left(f \frac{X}{Z}, f \frac{Y}{Z} \right)^T.$$

In homogenen Koordinaten (siehe Anhang A) erhält man:

$$\begin{pmatrix} x_{bh} \\ y_{bh} \\ 1 \end{pmatrix} \sim \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{pmatrix}.$$

- $\mathbf{p}(x_{bh}, y_{bh})$ = Bildpunkt in Bildhauptpunktkoordinaten
- $\mathbf{P}(X_C, Y_C, Z_C)$ = Raumpunkt in Projektionszentrumskoordinaten
- f = Bildweite

Hier wird bereits die Darstellung in homogenen Koordinaten verwendet, da die späteren Rechnungen ebenfalls in homogenen Koordinaten durchgeführt werden. Bei dem Modell der Lochkamera handelt es um ein idealisiertes Kameramodell. Durch das Fehlen einer Linse entfällt hier die bereits erwähnte Problematik der Linsenverzeichnung. Bei den späteren Berechnungen muss diese jedoch ausgeglichen werden.

2.5.2. Bildhauptpunktverschiebung

Bei den nun vorhandenen Bildkoordinaten geht man von der Annahme aus, dass der Ursprung des Bildkoordinatensystems gleich dem Durchstoßpunkt der optischen Achse der Kamera (Bildhauptpunkt) ist, die Koordinaten also in Bildhauptpunktkoordinatenform vorliegen. Da der Ursprung des Bildkoordinatensystems jedoch aus bautechnischen Gründen meist nicht mit der optischen Achse zusammenfällt, muss diese Ungenauigkeit korrigiert werden. Hinzu kommt, dass die Koordinaten eines Bildpunktes meistens in den Koordinaten des Bildverarbeitungssystems angegeben werden, also in Bildpunktkoordinaten. Im Zuge der Bildhauptpunktverschiebung erfolgt auch die Verschiebung der Koordinaten in das Bildpunktkoordinatensystem. Der Ursprung des Bildpunktkoordinatensystems befindet sich bei

den gängigen Systemen oben links und damit um (p_x, p_y) versetzt zum Bildhauptpunkt (vergleiche Abbildung 2.11). Für die Koordinaten des Bildverarbeitungssystems gilt daher:

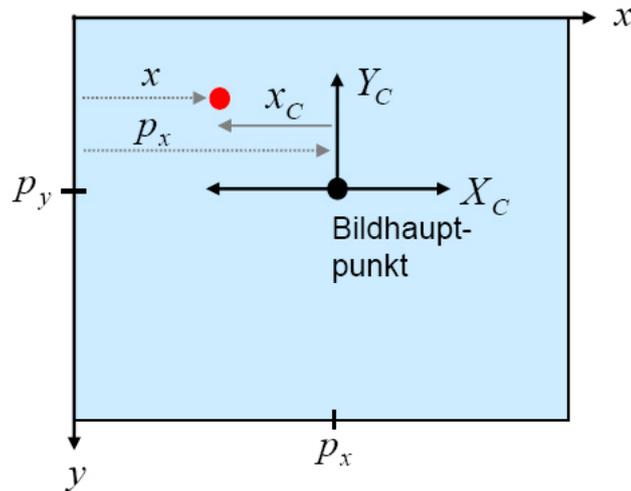


Abbildung 2.11.: Bildhauptpunktverschiebung (Meisel 2008, S. 21): Da der Bildhauptpunkt (Auftrittspunkt der optischen Achse auf der Bildebene) und der Mittelpunkt des Bildhauptpunktkoordinatensystems meistens nicht übereinstimmen, müssen die Koordinaten entsprechend dieser Abweichung korrigiert werden. Zusätzlich werden die Koordinaten in das Bildpunktkoordinatensystem (Ursprung oben links) verschoben.

$$x = +x_{bh} + p_x = \left(+f \frac{X_C}{Z_C} + p_x\right)$$

$$y = -y_{bh} + p_y = \left(-f \frac{Y_C}{Z_C} + p_y\right).$$

oder in homogenen Koordinaten:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -f & 0 & p_x & 0 \\ 0 & +f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{pmatrix}.$$

- $\mathbf{p}(x, y)$ = Bildpunkt in Bildpunktkoordinaten
- $\mathbf{P}(X_C, Y_C, Z_C)$ = Raumpunkt in Projektionszentrumskoordinaten

Der Vorteil dieser Darstellung ist, dass alle Bildpunkte positive Vorzeichen haben bzw. im positiven Wertebereich liegen.

2.5.3. Weltkoordinatensystem zu Projektionszentrumskordinatensystem

Da die meisten Raumpunktkoordinaten bezüglich eines Weltkoordinatensystems gegeben sind, bedarf es für die Rechnung einer Umwandlung der Koordinaten in Projektionszentrumskordinaten, welche durch eine Translation \mathbf{t} und eine Rotation \mathbf{R} beschrieben werden kann. Mit der Translation wird die Verschiebung des Ursprungs des Weltkoordinatensystems in den Ursprung des Projektionszentrumskordinatensystems erreicht. Die Anpassung der Achsen des Weltkoordinatensystems an das Projektionszentrumskordinatensystem erfolgt durch eine Rotation der selbigen. Es lässt sich also folgende Beziehung herstellen:

$$\mathbf{P}_C = \mathbf{R}(\mathbf{P}_W - \mathbf{t}).$$

Rotation

Die Rotation beschreibt eine Verdrehung der Koordinatenachsen zur Anpassung an ein anderes Koordinatensystem. Aus der Rotation der drei Achsen lässt sich folgende Matrix ableiten:

$$\mathbf{R} = \begin{bmatrix} \cos\beta \cos\gamma & \cos\alpha \sin\gamma + \sin\alpha \sin\beta \cos\gamma & \sin\alpha \sin\gamma - \cos\alpha \sin\beta \cos\gamma \\ -\cos\beta \sin\gamma & \cos\alpha \cos\gamma - \sin\alpha \sin\beta \sin\gamma & \sin\alpha \cos\gamma + \cos\alpha \sin\beta \sin\gamma \\ \sin\beta & -\sin\alpha \cos\beta & \cos\alpha \cos\beta \end{bmatrix}.$$

- α, β, γ beschreiben die Rotation der Achsen des Koordinatensystems.

Translation

Die Translation beschreibt die Verschiebung eines Punktes in einen anderen Punkt um einen Vektor \mathbf{t} .

$$\mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

2.5.4. Vollständig projektives Kameramodell

Nachdem die Zusammenhänge zwischen den einzelnen Koordinatensystemen und Punktdarstellungen geklärt sind, können die entwickelten physikalischen Abhängigkeiten in einer Rechnung zusammengefasst werden. Die Rechnung enthält nicht den Ausgleich der Linsenverzeichnung. Dieser muss bei der späteren Rechnung, durch die Anpassung der Bildpunktkoordinaten, gesondert vorgenommen werden.

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -f & 0 & p_x & 0 \\ 0 & +f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{R} & -\mathbf{Rt} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

- $\mathbf{p}(x, y)$ = Bildpunkt in Bildpunktkoordinaten
- $\mathbf{P}(X_W, Y_W, Z_W)$ = Raumpunkt in Weltkoordinaten

Mit Hilfe dieser Formel lässt sich ein 3D-Raumpunkt auf einen 2D-Bildpunkt, in Bildpunktkoordinaten, abbilden. Durch ausmultiplizieren der Matrizen lässt sich das physikalische Modell in Form der so genannten Projektionsmatrix (H-Matrix) darstellen:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \sim \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \end{pmatrix} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}.$$

- h_{ij} = Variable für den Inhalt der oben beschriebenen und ausmultiplizierten Matrizen.

Da die Koordinaten noch in homogenen Koordinaten vorliegen, müssen diese wieder in die euklidische Form gebracht werden.

$$x = (h_{11}X_W + h_{12}Y_W + h_{13}Z_W + h_{14}) / (h_{31}X_W + h_{32}Y_W + h_{33}Z_W + h_{34}) \quad (1)$$

$$y = (h_{21}X_W + h_{22}Y_W + h_{23}Z_W + h_{24}) / (h_{31}X_W + h_{32}Y_W + h_{33}Z_W + h_{34}) \quad (2)$$

2.5.5. Raumpunktvermessung mit Stereobildpaaren

Ziel dieser Bachelorarbeit ist die Berechnung von Raumpunkten aus Stereobildpaaren. Um die benötigten Raumpunkte, in Weltkoordinaten, aus den gegebenen Bildpunktkoordinaten berechnen zu können, wird die hergeleitete Formel entsprechend nach X_W , Y_W und Z_W umgestellt.

Umformung von (1) und (2) nach X, Y und Z:

$$xh_{31}X_W + xh_{32}Y_W + xh_{33}Z_W + xh_{34} = h_{11}X_W + h_{12}Y_W + h_{13}Z_W + h_{14}$$

$$\Leftrightarrow (xh_{31} - h_{11})X_W + (xh_{32} - h_{12})Y_W + (xh_{33} - h_{13})Z_W = h_{14} - xh_{34}$$

$$yh_{31}X_W + yh_{32}Y_W + yh_{33}Z_W + yh_{34} = h_{21}X_W + h_{22}Y_W + h_{23}Z_W + h_{24}$$

$$\Leftrightarrow (yh_{31} - h_{21})X_W + (yh_{32} - h_{22})Y_W + (yh_{33} - h_{23})Z_W = h_{24} - yh_{34}.$$

Aus den Gleichungen ist zu erkennen, dass es drei Unbekannte X, Y, Z gibt, aber nur zwei Gleichungen. Daraus folgt, dass mindestens zwei Bilder beziehungsweise Bildpunkte benötigt werden, um einen Raumpunkt zu berechnen.

Dieses Vorgehen entspricht der Idee des räumlichen Vorwärtsschnitts (vergleiche hierzu Abbildung 2.12). Dieses beschreibt, bei gegebener innerer und äußerer Orientierung der Kameras, die Berechnung eines Punktes $\mathbf{P}(X, Y, Z)$ aus den in korrespondierenden Bildern gemessenen vier Bildkoordinaten (x_1, y_1, x_2, y_2) (vergleiche Suthau 2002). Die benötigten Informationen über die innere und äußere Orientierung der Kameras sind durch die später beschriebene Kamerakalibrierung bekannt.

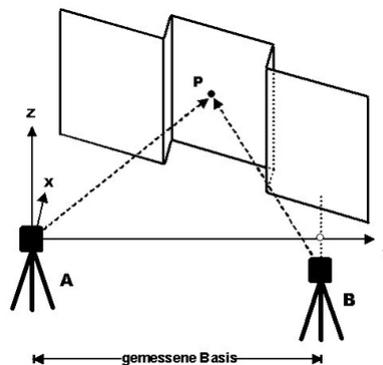


Abbildung 2.12.: Grundprinzip des räumlichen Vorwärtsschnitts (GIK 2006). Durch die Aufnahme eines Punktes von zwei verschiedenen Standpunkten und den vorhandenen inneren und äußeren Kameraparametern kann dessen Entfernung berechnet werden.

In Matrixschreibweise lässt sich die Formel für zwei oder mehr Quellbilder wie folgt aufstellen:

$$\begin{bmatrix} ({}^1x^1h_{31} - {}^1h_{11}) & ({}^1x^1h_{32} - {}^1h_{12}) & ({}^1x^1h_{33} - {}^1h_{13}) \\ \vdots & \vdots & \vdots \\ ({}^nx^n h_{31} - {}^nh_{11}) & ({}^nx^n h_{32} - {}^nh_{12}) & ({}^nx^n h_{33} - {}^nh_{13}) \\ ({}^1y^1h_{31} - {}^1h_{21}) & ({}^1y^1h_{32} - {}^1h_{22}) & ({}^1y^1h_{33} - {}^1h_{23}) \\ \vdots & \vdots & \vdots \\ ({}^ny^n h_{31} - {}^nh_{21}) & ({}^ny^n h_{32} - {}^nh_{22}) & ({}^ny^n h_{33} - {}^nh_{23}) \end{bmatrix} \begin{pmatrix} X_W \\ Y_W \\ Z_W \end{pmatrix} = \begin{bmatrix} {}^1h_{14} - {}^1x^1h_{34} \\ \vdots \\ {}^nh_{14} - {}^nx^n h_{34} \\ {}^1h_{24} - {}^1y^1h_{34} \\ \vdots \\ {}^nh_{24} - {}^ny^n h_{34} \end{bmatrix}$$

- 1x bis nx = Koeffizienten für ein zusammengehöriges Paar x,y Gleichungen. Für jeden zur Berechnung verwendeten Bildpunkt kommt ein Paar x,y Gleichungen hinzu.
- x,y = Koordinaten des verwendeten Bildpunktes
- h_{ij} = Kameraparameter ij der Kamera K
- $\mathbf{P}(X_W, Y_W, Z_W)$ = Raumpunkt in Weltkoordinaten

Diese Darstellung entspricht der zentralen mathematischen Formel, mit welcher die Berechnungen in dieser Bachelorarbeit durchgeführt werden. Die berechneten Koordinaten liegen in Weltkoordinaten vor. Da in diesem Fall jedoch die Entfernung der Punkte zur Kamera entscheidend ist, müssen diese noch in Projektionszentrumskoordinaten umgewandelt werden. Diese Anpassung erfolgt durch eine erneute Rotation und Translation des Lösungsvektors (berechneter Raumpunkt in Weltkoordinaten).

3. Anforderungsanalyse

In diesem Kapitel werden die Anforderungen für eine Realisierung des im Grundlagenkapitel beschriebenen Tiefenbildes (Abschnitt 2.1) ermittelt. Ein besonderes Interesse gilt hier einem möglichen Einsatz im Bereich des Kollisionsschutzes, speziell im Bereich eines Hochregallagers oder einer vergleichbaren Situation. Die Analyse gliedert sich in die generellen Anforderungen an ein Tiefenbild, das Kamerasystem zur Erstellung von Bildpaaren und das Verfahren zur Bestimmung von korrespondierenden Bildpunkten sowie die hieraus abgeleiteten Anforderungen an ein Programm, das diese Anforderungen umsetzt.

Da die Realisierung über das passive Stereoverfahren erfolgen soll, wird dieses bereits im Zuge der Analyse berücksichtigt.

3.1. Anforderungen an ein Tiefenbild und das Verfahren zur Bestimmung von korrespondierenden Bildpunkten

Vor der praktischen Umsetzung werden die Anforderungen bestimmt, die ein Tiefenbild für den Einsatz im Bereich des Kollisionsschutzes erfüllen muss. Da die Anforderungen an das für die Berechnung der Tiefeninformationen verwendete Stereoverfahren eng mit den Anforderungen des Tiefenbildes zusammenhängen, werden diese hier ebenfalls erläutert. Das Stereoverfahren basiert auf dem Vorhandensein von korrespondierenden Bildpunkten in Stereobildern. Es wird also ein Verfahren gesucht, das korrespondierende Bildpunkte findet. Aus diesen Anforderungen und den Anforderungen an ein Tiefenbild, ergeben sich die im Folgenden beschriebenen Bedingungen für einen sicheren Kollisionsschutz.

3.1.1. Große Menge an Tiefeninformationen

Damit die Umgebung im Bereich vor der Kamera möglichst genau dargestellt werden kann, wird eine große Menge an Tiefeninformationen benötigt. Ist die Menge der Informationen zu gering, werden potenzielle Kollisionsobjekte möglicherweise nicht erfasst und ein sicherer

Kollisionsschutz ist nicht gegeben. Für das Programm zur Bestimmung der korrespondierenden Bildpunkte heißt das, dass dieses möglichst viele korrespondierende Bildpunkte finden muss.

3.1.2. Homogene Verteilung der Tiefeninformationen

Noch wichtiger als die Menge an Tiefeninformationen ist die Verteilung dieser Informationen über das gesamte Bild. Liegen Tiefeninformationen nur für Teilbereiche des Bildes vor, so ist eine verlässliche Darstellung der Umgebung nicht möglich. Potenzielle Kollisionsobjekte werden möglicherweise nicht erkannt. Es besteht eine Kollisionsgefahr in allen, nicht mit Tiefeninformationen versehenen, Bereichen. Für einen sicheren Kollisionsschutz ist somit eine homogene Verteilung der Tiefeninformationen notwendig. Diese Anforderung gilt besonders im Bereich der Bildmitte, also der unmittelbaren Fahrtrichtung des Fahrzeugs.

3.1.3. Distanzgenauigkeit

Die dargestellten Tiefeninformationen müssen besonders für Objekte im Nahbereich eine hohe Genauigkeit gewährleisten. Hier kommt es darauf an, das Objekt zu umgehen oder rechtzeitig vor dem Objekt zu stoppen. Bei größeren Distanzen ist die Genauigkeit nicht so wichtig. Hier reicht es, die Objekte generell zu erfassen.

3.1.4. Verlässlichkeit des Verfahrens und Verfügbarkeit der Tiefeninformationen

Ein mögliches Verfahren muss sicherstellen, dass es verlässliche korrespondierende Bildpunkte in verschiedenen Anwendungsfällen liefert. Ein Verfahren muss immer verlässliche Tiefeninformationen liefern, um für den Kollisionsschutz verwendet zu werden. Einen hohen Stellenwert hat hierbei die Verlässlichkeit in der Umgebung von Menschen.

Für eine mögliche Verwendung der Tiefeninformationen in weiteren Arbeitsschritten, ist eine Ablage der zusammengehörigen Bilddaten und Tiefeninformationen wünschenswert.

3.1.5. Unabhängigkeit des Verfahrens gegenüber äußeren Einflüssen

Je nachdem welches Stereokamerasystem zur Aufnahme der Bildpaare verwendet wird, können äußere Faktoren, wie unterschiedliche Lichtverhältnisse oder Skalierungen der Bilder, Einfluss auf die einzelnen Bildaufnahmen haben. Ein mögliches Verfahren muss sicherstellen, dass diese Einflüsse keine beziehungsweise nur möglichst geringe Auswirkungen auf die Ermittlung der korrespondierenden Bildpunkte haben.

3.2. Anforderungen an das Kamerasystem

Nach der Bestimmung der Anforderungen an ein Tiefenbild und dem daran gekoppelten Verfahren zur Bestimmung von korrespondierenden Bildpunkten, muss nun ein Kamerasystem gefunden werden, das geeignete Bildpaare liefert.

3.2.1. Gleichheit der verwendeten Kameras und der Kameraparameter

Damit die Aufnahmen der Stereobildpaare möglichst gleich sind, ist es erforderlich, dass die im Kamerasystem verwendeten Kameras vom selben Typ sind. Hiermit wird sichergestellt, dass es keine Unterschiede im Bereich der Qualität der Bilder und der Art der Darstellung gibt. Möglicherweise wirkt sich die Verwendung unterschiedlicher Kameras auch auf die Qualität und Menge der korrespondierenden Bildpunkte aus.

Durch die Gleichheit der Kameras wird ebenso sichergestellt, dass auch die Kameraparameter gleich sind. Diese werden für die spätere Berechnung der Raumpunkte mit dem Stereoverfahren benötigt. Eine Ausnahme bilden hier Parameter, die in jedem Kamerasystem einmalig sind. Diese müssen separat für beide Kameras ermittelt werden.

3.2.2. Stabilität und Unveränderlichkeit des Kamerasystems

Wie bereits erwähnt, werden für die Berechnung der Raumpunkte verschiedenen Kameraparameter benötigt. Damit sich diese durch einen manipulierten Versuchsaufbau nicht verändern, muss gewährleistet sein, dass das verwendete Kamerasystem entsprechend stabil beziehungsweise unveränderlich ist.

3.3. Anforderungen an das Programm zur Berechnung der Raumpunkte und Erstellung des Tiefenbildes

Aus den bisher ermittelten Anforderungen leiten sich die Anforderungen für ein Programm zur Berechnung der Raumpunkte und zur Erstellung des Tiefenbildes ab.

3.3.1. Bedienbarkeit

Eine wünschenswerte Fähigkeit des Programms ist die automatische Aufnahme und Auswertung der Bilder und die daraus folgenden Erstellung eines Tiefenbildes. In einem ersten Schritt steht jedoch die Entwicklung eines manuellen Prototyps im Vordergrund.

3.3.2. Geschwindigkeit

Um das System in möglichst vielen Bereichen einsetzen zu können, spielt die Geschwindigkeit eine Rolle mit der das Programm die Bilder bearbeitet. Soll eine Auswertung der Bildinformationen in Echtzeit erfolgen, ist dies eine kritische Komponente.

4. Konzeption

In diesem Kapitel wird aus den ermittelten Anforderungen (Kapitel 3) ein Konzeptentwurf für die praktische Umsetzung erstellt. Hierfür wird das zentrale Verfahren zur Auffindung der korrespondierenden Bildpunkte ermittelt und die Hard- und Softwarekomponenten ausgewählt. Des Weiteren werden verschiedene Testszenarien entwickelt, welche überprüfen sollen, ob das Verfahren die Anforderungen aus der Anforderungsanalyse erfüllt.

4.1. Bestimmung des Verfahren zur Auffindung von korrespondierenden Bildpunkten

Da in dieser Arbeit das Stereoverfahren zur Berechnung der Raumpunkte verwendet werden soll, wird ein Verfahren benötigt, das korrespondierende Bildpunkte in Stereobildern finden kann. Durch die Anforderungen an diese Bachelorarbeit ist dieses Verfahren bereits vorgegeben. Hier kommt das „scale-invariant feature transform“ Verfahren (SIFT-Verfahren) zum Einsatz, welches im Grundlagenkapitel (siehe 2.2) beschrieben ist.

Durch die gegebenen Anforderungen, muss dieses Verfahren bestimmte Eigenschaften aufweisen. Diese gliedern sich in die Bereiche Menge und Verteilung der gefundenen korrespondierenden Bildpunkte in verschiedenen Anwendungsfällen und die Unabhängigkeit des Verfahrens gegenüber äußeren Einflüssen. Die nötigen Schritte zur Sicherstellung dieser Eigenschaften werden im Folgenden analysiert.

4.1.1. Analyse der Punktemenge und der Punkteverteilung in verschiedenen Anwendungsfällen

Um die Verlässlichkeit des Verfahrens im Bereich der Punktemenge und deren Verteilung sicherzustellen, werden verschiedene Testszenarien aufgestellt. Diese werden im Folgenden erläutert.

Situation im Hochregallager

Damit das Szenario der Situation in einem Hochregallager mit langen schmalen Gängen nahe kommt, wird hierfür ein Flur im Universitätsgebäude gewählt. Um eine größere Aussagekraft des Szenarios herbeizuführen, wird der Flur einmal im leeren Zustand und einmal in einem Zustand mit künstlichen Hindernissen analysiert.

Beliebige Situationen

Das Verfahren wird zusätzlich in Situationen getestet, um eine allgemeingültige Aussage über die Punkteanzahl und -verteilung treffen zu können. Als Szenarien dienen hier ein Labor und ein Arbeitsraum. Diese Orte werden in der Zukunft vielleicht ein Anwendungsgebiet darstellen. Gleichzeitig liefern sie in ihrer Beschaffenheit eine gute Grundlage zur Bestimmung von korrespondierenden Bildpunkten im Sinne des SIFT-Verfahrens und damit die entsprechenden Voraussetzungen für eine erfolgreiche Analyse.

Menschen

Wie bereits in der Einleitung erwähnt, spielt die Erkennung von Menschen beim Kollisionsschutz eine wichtige Rolle. Zur Überprüfung des Verfahrens in diesem Bereich, wird ein Mensch im Flur-Szenario (gestellte Situation im Gang eines Hochregallagers) als potenzielles Hindernis eingesetzt. Zusätzlich findet eine gesonderte Untersuchung an einer Person statt.

4.1.2. Unabhängig gegenüber äußeren Einflüssen

Das SIFT-Verfahren stellt bereits sicher, dass äußere Faktoren, wie unterschiedliche Lichtverhältnisse, und Unterschiede in den Bildaufnahmen, wie Bildverkippen oder unterschiedliche Skalierungen, keine bis minimale Einflüsse auf die Zuordnung korrespondierender Bildpunkte haben. Bildverkippen können beispielsweise durch eine Schrägstellung des Fahrzeugs auftreten.

4.2. Bestimmung der zu verwendenden Software

An der HAW wurde bereits eine Bachelorarbeit zum Thema SIFT angefertigt. Im Zuge dieser Arbeit wurde ein Impresario-Modul (Anhang B) entwickelt. Um eine erneute Implementierung

zu vermeiden, wird dieses Modul wieder verwendet. Sowohl Impresario als auch das Modul basieren auf C++, daher fällt die Wahl der Programmiersprache ebenfalls auf C++. Da die vorhandenen Sourcen in einer kompatiblen Version vorliegen, wird Microsoft Visual Studio 2003 als Entwicklungsumgebung verwendet.

4.3. Bestimmung des Stereokamerasystems

Für die Bachelorarbeit steht die im Labor der HAW vorhandene Bumblebee-Stereokamera zur Verfügung. Auf die Bumblebee-Stereokamera wird im Kapitel Versuchsdurchführung (Kameraanbindung 5.1.2) genauer eingegangen. Die Bauart der Stereokamera stellt sicher, dass die Anforderungen bezüglich der Kameras und der Kameraparameter erfüllt werden.

4.4. Bestimmung der Kalibrierungssoftware

Zur genauen Bestimmung der Kameraparameter eignet sich die vorhandene Software PhotoModeler. Mit dieser Software lassen sich die internen und externen Kameraparameter bestimmen, die für die Berechnung der Raumpunkte benötigt werden. Die Software wird ebenfalls im Kapitel Versuchsdurchführung (Kamerakalibrierung 5.1.3) genauer beschrieben.

4.5. Genauigkeitsanalyse der berechneten Raumpunkte

Die Genauigkeit der berechneten Raumpunkte nach dem Stereoverfahren spielt für die Einsetzbarkeit des Verfahrens eine wichtige Rolle. Es ist zu überprüfen, ob die Berechnung mit den durch das SIFT-Verfahren gegebenen korrespondierenden Bildpunkten und den mit Hilfe des PhotoModelers ermittelten Kameraparametern diesen Anforderungen genügt. Die Genauigkeit wird nur in einem Versuch untersucht. Es ist anzunehmen, dass sich diese bei gleich bleibender Qualität der korrespondierenden Bildpunkte nicht ändert. Aus den Anforderungen an die Genauigkeit des Verfahrens in verschiedenen Distanzen wird folgendes Testszenario entwickelt.

Szenario der Genauigkeitsanalyse

Zu Bestimmung der Genauigkeit des Verfahrens wird ein Hindernis in verschiedenen Entfernungen zur Kamera aufgestellt. Für diese Entfernungen werden jeweils Stereobilder aufge-

nommen und korrespondierende Bildpunkte ermittelt. Hieraus erfolgt dann die Berechnung der Entfernung und eine Auswertung der Ergebnisse.

5. Versuchsdurchführung

In diesem Kapitel wird der im vorherigen Kapitel entwickelte Konzeptentwurf für die Erstellung eines C++-Programms zur Berechnung der Raumpunkte praktisch umgesetzt. Die Versuchsdurchführung umfasst sowohl die Beschreibung der Entwicklungsumgebung als auch die nötigen Arbeitsschritte zur Bestimmung der SIFT-Punkte und die daraus resultierende Berechnung der Raumpunkte. Hierbei lässt sich zwischen solchen Arbeiten unterscheiden, welche einmalig zu Beginn des Projektes durchgeführt werden müssen und solchen, welche bei der Raumpunktbestimmung für jedes neue Stereobildpaar anfallen.

Einmalige Arbeitsschritte:

- Entwicklung des Stereomessmoduls
- Anbindung der Kamera
- Kamerakalibrierung
- Erstellung des 3D-Modells der Kamerakalibrierung

Wiederkehrende Arbeitsschritte:

- Merkmalsextraktion
- Ergebnisfilterung und Raumpunktvermessung
- Erstellung des Tiefenbildes

5.1. Entwicklungsumgebung des Projekts

Dieser Abschnitt beschreibt die Entwicklungsumgebung in welcher die Bachelorarbeit durchgeführt wird, sowie die Ermittlung der benötigten Parameter für die Raumpunktberechnung.

5.1.1. Entwicklung des Stereomessmoduls

Das zu entwickelnde Messmodul wird mit Microsoft Visual Studio 2003 unter Verwendung der LTI-Lib Bibliothek (Anhang C) entwickelt. Visual Studio kommt in dieser Version zum Einsatz, da die verwendeten LTI-Komponenten in der für 2003 kompatiblen Form vorliegen. Die Wahl der Programmiersprache fällt auf C++, da auch die LTI-Lib auf C++ basiert. Der zentrale Teil der Anwendung besteht aus der Realisierung der im Grundlagenkapitel entwickelten Formel zur Berechnung eines Raumpunktes aus zwei korrespondierenden Bildpunkten (siehe Kapitel 2.5.4). Die berechneten Punkte sind Linsenverzeichniskorrigiert und liegen in Projektionszentrumskoordinaten vor. Hierdurch ist gewährleistet, dass sich die Entfernung der potenziellen Hindernisse auf die Kameraposition bezieht.

Der Ablauf des Programms ist in Abbildung 5.1 zu sehen.

5.1.2. Kameraanbindung

Dieser Abschnitt beschreibt die Anbindung der Kamera und die Software zum Aufnehmen der Stereobildpaare.

Bumblebee-Stereokamera

Bei der Bumblebee (siehe Foto 5.2) handelt es sich um eine Stereo-Farbkamera der kanadischen Firma Point Grey Research Inc. (PointGrey). Das Kamerasystem enthält zwei vorkalibrierte CCD-Kameras mit einer maximalen Auflösung von 640x480 Pixel. Bei dem verwendeten CCD-Chip handelt es sich um den Sony ICX424AQ. Die Kamera verfügt über eine Firewire-Schnittstelle, über welche sie an den PC angeschlossen wird. Weitere Informationen zur Kamera können dem Datenblatt (siehe CD-ROM) entnommen werden. Zur Ansteuerung der Kameras wird die Software Digiclops verwendet.

Digiclops

Digiclops ist eine im Lieferumfang der Bumblebee enthaltene Software zum Ansteuern der einzelnen Kameras und zum Aufnehmen von Fotos. Über die Software können verschiedenen Parameter der Kamera beeinflusst werden. Zum Beispiel:

- Helligkeit
- Kontrast
- Weißabgleich

Ablauf des C++ Programms zur Berechnung von Raumpunkten aus SIFT-Punktepaaren



Funktion zur Berechnung der Raumpunkte

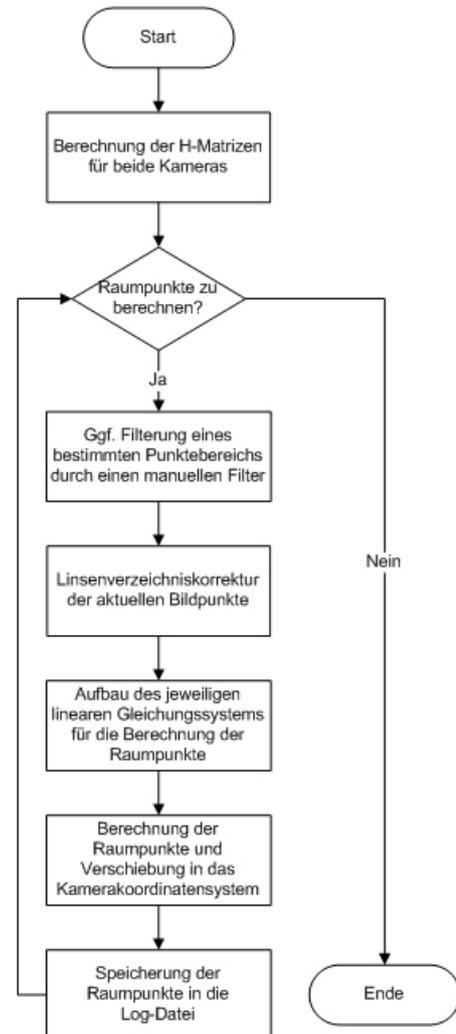


Abbildung 5.1.: Das Ablaufdiagramm auf der linken Seite beschreibt den kompletten Ablauf des C++ Programms während der Raumpunktberechnung für ein Stereobildpaar. Auf der rechten Seite ist die Funktion zur Raumpunktberechnung dargestellt. Die für die H-Matrix benötigten Matrizen und Vektoren sind im Grundlagenkapitel beschrieben.



Abbildung 5.2.: Bumblebee-Stereokamera der kanadischen Firma Point Grey Research Inc. (PointGrey)

- Lichtempfindlichkeit

Der Haupteinsatzbereich der Software liegt in diesem Fall in der Aufnahme der Stereobildpaare für die Kamerakalibrierung und für die späteren Raumpunktberechnungen.

5.1.3. Kamerakalibrierung

Wie bereits im Grundlagenkapitel erwähnt, werden für die Rückberechnung eines Raumpunktes aus mehreren Bildpunkten die inneren und äußeren Kameraparameter benötigt. Dieser Abschnitt beschreibt die Kamerakalibrierung und die Erschließung der benötigten Parameter mittels der Software PhotoModeler.

PhotoModeler

PhotoModeler (EOS) ist eine Software zum erstellen und vermessen von genauen und hochwertigen 3D-Modellen mit einer Kamera. So bietet PhotoModeler eine Funktion zur Kalibrierung der verwendeten Kamera und liefert über diese die inneren Parameter der Kamera. Die externen Kameraparameter können nach der Erstellung eines 3D-Modells des Kalibrierungsprojektes ausgelesen werden. In diesem Fall wird für die Bestimmung der äußeren Parameter statt der Kalibrierungsvorlage ein eigenes Modell verwendet, um später eine bessere Kontrollmöglichkeit für die Raumpunktberechnung zu erhalten. So können die berechneten Punkte mit den gemessenen Punkten des PhotoModelers verglichen und am realen Modell überprüft werden. Die Kameraparameter werden für die spätere Berechnung der Raumpunkte benötigt.

Innere Kameraparameter

Für die Kalibrierung der inneren Kameraparameter liefert PhotoModeler bereits die in Abbildung 5.3 abgebildete Kalibrierungsvorlage mit.

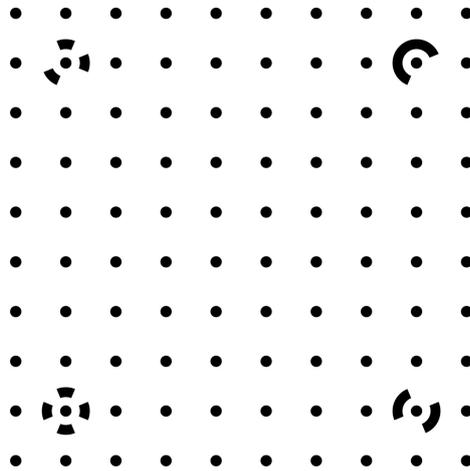


Abbildung 5.3.: Mitgeliefertes Kalibrierungsmuster für die Kamerakalibrierung (EOS). Die Vorlage wird aus verschiedenen Positionen aufgenommen und ermöglicht dem PhotoModeler die Ermittlung der inneren und äußeren Kameraparameter.

Diese Vorlage muss mit der Kamera aus bestimmten Positionen und Perspektiven aufgenommen werden und in ein so genanntes Kalibrierungsprojekt eingefügt werden. Entspricht die Qualität der Aufnahmen den Anforderungen des PhotoModelers, so erfolgt die Kalibrierung automatisch. Die Kalibrierung ergab die in Abbildung 5.4 dargestellten inneren Parameter.

Äußere Kameraparameter

Die Kalibrierung der äußeren Kameraparameter entspricht einer Kalibrierung der Kamera auf ein (Kamera-) Weltkoordinatensystem. Anhand dieses Kameramodells kann auch die Korrektheit des später programmierten C++-Programms überprüft werden. Alle später berechneten Raumpunkte beziehen sich auf den Ursprung dieses kalibrierten Weltkoordinatensystems. Es kann also nachgemessen werden, ob die Ergebnisse der Rechnung mit den Angaben des PhotoModelers übereinstimmen.

Um das entsprechende Kalibrierungsprojekt erstellen zu können, wird eine Szene in Form eines Legomodells aufgebaut, welches in Abbildung 5.5 zu sehen ist. Dieses Modell wird dann ähnlich wie bei der Kalibrierung der inneren Parameter aus verschiedenen Positionen (Winkel und Entfernung) aufgenommen und in das Projekt importiert.

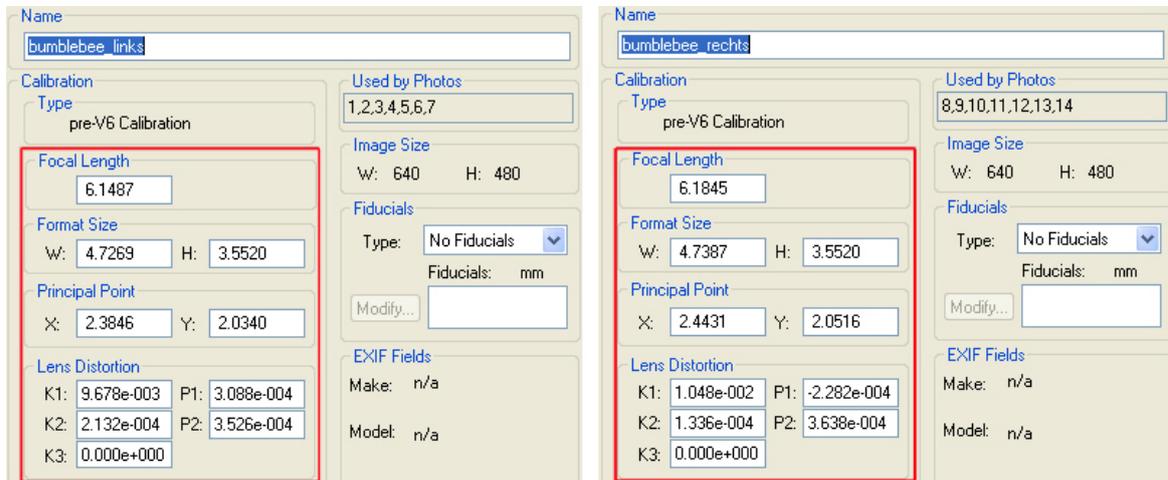


Abbildung 5.4.: Darstellung der inneren Kameraparameter nach der Kalibrierung durch den PhotoModeler (linke und rechte Kamera).

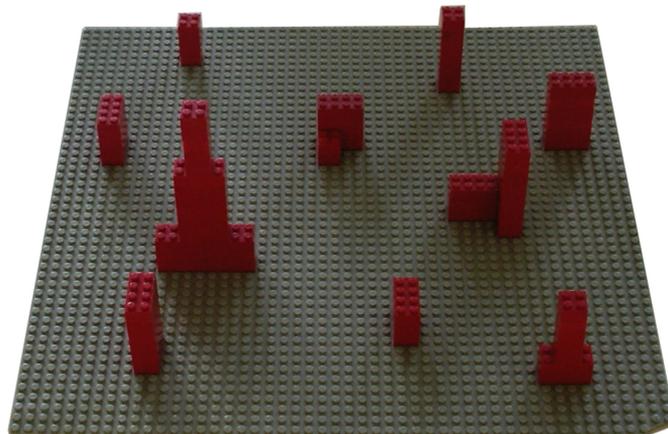


Abbildung 5.5.: Kalibrierungsmodell aus Lego: Für die Bestimmung der äußeren Kameraparameter wird die Kamera auf ein Weltkoordinatensystem kalibriert.

Als nächstes wird ein beliebiges Bild ausgewählt und ein Punkt in diesem als Ursprungspunkt definiert. Des Weiteren werden noch einige Fixpunkte benötigt, deren Entfernungen zum Ursprung vorher manuell gemessen werden und einige Referenzpunkte, für die keine Entfernungsangaben gemacht werden müssen. Um die Kalibrierung abschließen zu können, werden nun die markierten Punkte in den Kalibrierungsbildern einander zugeordnet, soweit sie in diesen sichtbar sind (vergleiche Abbildung 5.6).

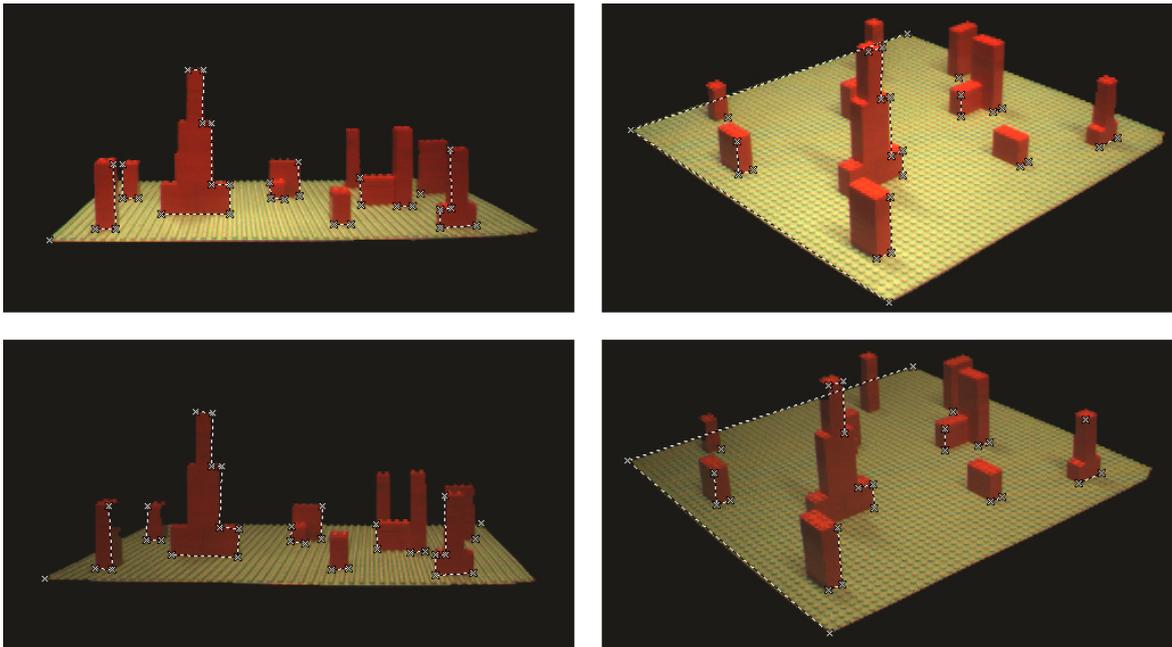


Abbildung 5.6.: Kalibrierungsprojekt im PhotoModeler: Kalibrierungsbilder mit Referenzpunkten zur Bestimmung der äußeren Kameraparameter (linke Kamera (oben) und rechte Kamera (unten)).

Nach diesem Schritt ist die Kalibrierung abgeschlossen und man erhält die folgenden in Abbildung 5.7 dargestellten externen Parameter.

5.1.4. 3D-Modell der Kalibrierungsumgebung

Nach erfolgreich abgeschlossener Kamerakalibrierung und unter Verwendung einer entsprechenden Anzahl an Referenzpunkten, lässt sich im PhotoModeler aus den Kalibrierungsbildern ein 3D-Modell der Kalibrierungsumgebung erstellen. In diesem können unter anderem sämtliche Positionen der Kameras angezeigt werden, welche bei der Bildaufnahme der Kalibrierungsbilder zum Einsatz gekommen sind. Abbildung 5.8 zeigt, wie ein solches Modell aussehen kann.

Camera station position	0.155, -1.123, 0.136	Camera station position	0.275, -1.120, 0.134
X	0.155049	X	0.275319
Y	-1.122836	Y	-1.120181
Z	0.135920	Z	0.134476
Camera station angles (degrees)	87.966, 1.101, -0.582	Camera station angles (degrees)	87.444, 1.834, -0.248
Omega (deg.)	87.965772	Omega (deg.)	87.443613
Phi (deg.)	1.101287	Phi (deg.)	1.833609
Kappa (deg.)	-0.582157	Kappa (deg.)	-0.248200

Abbildung 5.7.: Darstellung der äußeren Kameraparameter nach der Kalibrierung durch den PhotoModeler (linke und rechte Kamera).

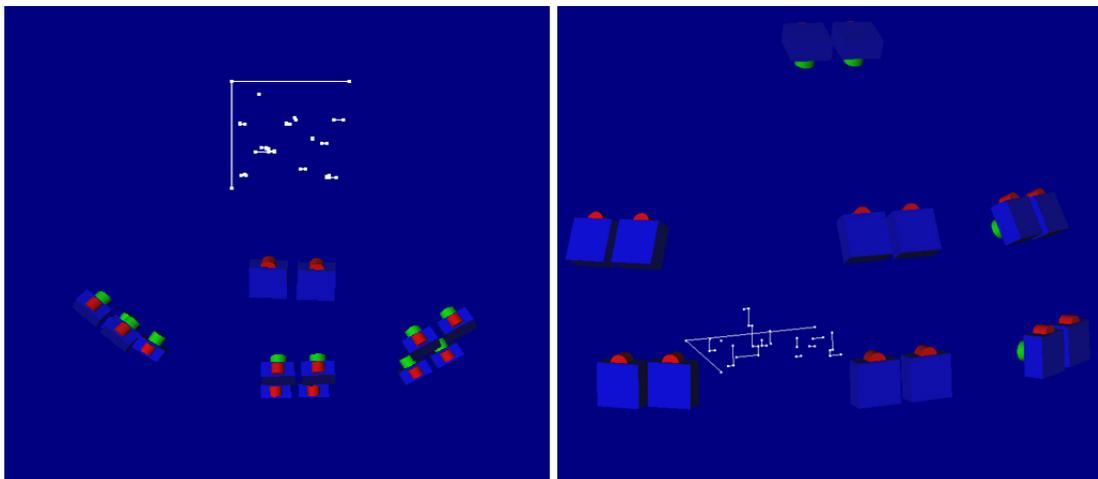


Abbildung 5.8.: 3D-Modell des kalibrierten Kamera-Weltkoordinatensystems im PhotoModeler (Ansicht von oben und vorne links). Zusätzlich sind die Kamerastandpunkte sichtbar, von denen die verwendeten Kalibrierungsbilder aufgenommen worden sind.

5.2. Merkmalsextraktion

Um die SIFT-Punkte zu erhalten, sind mehrere Arbeitsschritte notwendig. Diese lassen sich grob in die Bildaufnahme, Bildkonvertierung und die eigentliche Punktbestimmung unterteilen. Die einzelnen Schritte von der Bildaufnahme bis hin zur Bestimmung der SIFT-Punkte für die Raumpunktberechnung sind in Abbildung 5.9 dargestellt.

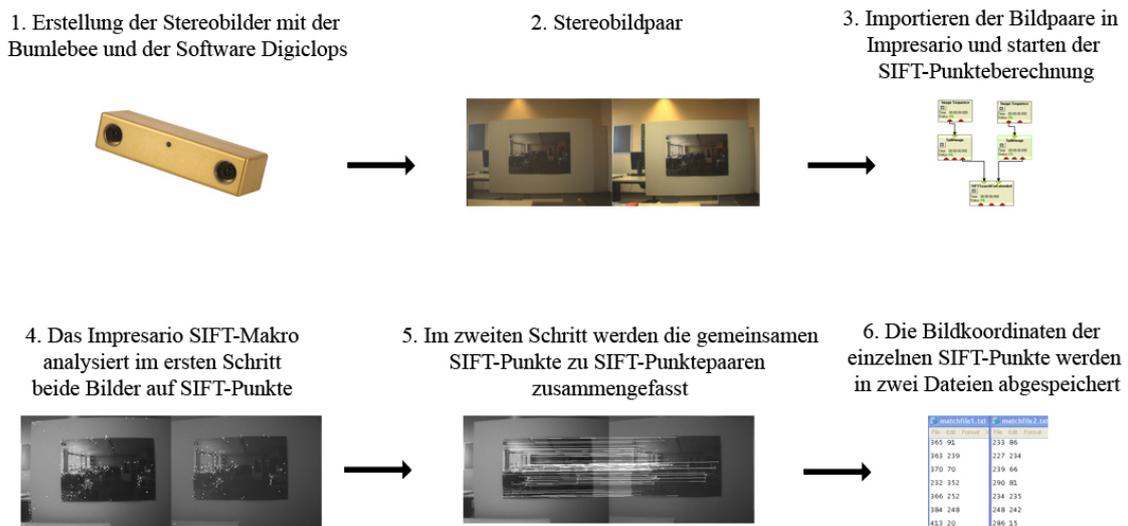


Abbildung 5.9.: Die Grafik zeigt die einzelnen Schritte von der Bildaufnahme, über die SIFT-Punktbestimmung in Impresario, bis hin zur Sicherung der ermittelten Punkte in zwei Dateien.

5.2.1. Bildaufnahme

Die Bildaufnahme erfolgt mit der im Lieferumfang der Bumblebee enthaltenen Software Digiclops. Nach der Aufnahme liegen die Bilder als Portable Pixmap (PPM) vor. Da Impresario diesen Datentyp jedoch nicht unterstützt, werden die Bilder in 24 Bit Bitmaps (BMP) konvertiert. Die Aufnahme der Bilder erfolgt immer paarweise, je ein Foto der linken und ein Foto der rechten Kamera.

5.2.2. Bestimmung der korrespondierenden Bildpunktpaare

Für die Bestimmung der SIFT-Punkte kommt die Software Impresario zum Einsatz. Anhand dieser Software wird ein Projekt erstellt, welches die Quellbilder einliest, in Grauwertbilder konvertiert und anschließend den SIFT-Algorithmus auf diese anwendet. Abbildung 5.10 zeigt das erstellte Impresario-Projekt.

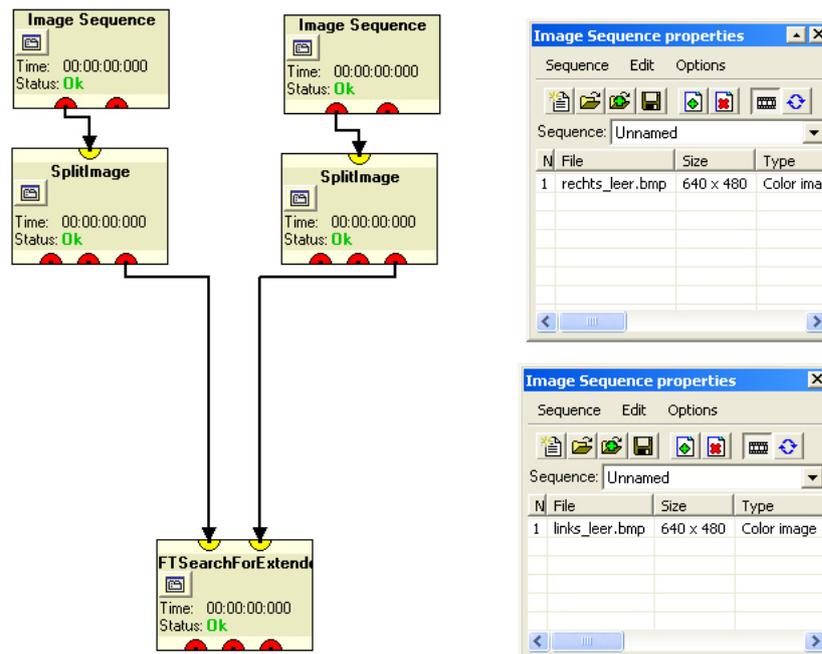


Abbildung 5.10.: Das Bild zeigt das in dieser Bachelorarbeit verwendete Impresario-Projekt. Über das Makro „Image Sequence“ werden die korrespondierenden Bildpaare geladen und an das Makro „SplitImage“ übergeben. Hier werden die Farbbilder für die weitere Verarbeitung in Grauwertbilder konvertiert. Im Makro „FTSearchForExtended“ werden die SIFT-Punkte der einzelnen Bilder und daraus die zusammengehörigen SIFT-Punktpaare bestimmt.

In diesem Zusammenhang sucht der SIFT-Algorithmus nach markanten SIFT-Punkten in den beiden in Abbildung 5.11 dargestellten Quellbildern. Da die Grauwertbilder in den verschiedenen Szenarien teilweise sehr dunkel sind, werden diese für eine bessere Darstellung im weiteren Verlauf dieser Bachelorarbeit aufgehellt gezeigt.

Im zweiten Schritt werden die SIFT-Punkte miteinander verglichen und die passenden Partner gesucht. Die gefundenen SIFT-Punktpaare werden wie in Abbildung 5.12 markiert.

Die gefundenen SIFT-Punktpaare werden in zwei Dateien, je eine für das linke und rechte Quellbild, abgespeichert. Die Punkte zu denen kein Partner gefunden werden kann, werden



Abbildung 5.11.: Die Fotos zeigen die Aufnahme einer Szene aus der später beschriebenen Genauigkeitsanalyse des SIFT-Verfahren. Die weißen Punkte markieren die gefundenen SIFT-Punkte im Bild der linken und rechten Kamera nach der Verarbeitung durch das Impresario-Projekt.

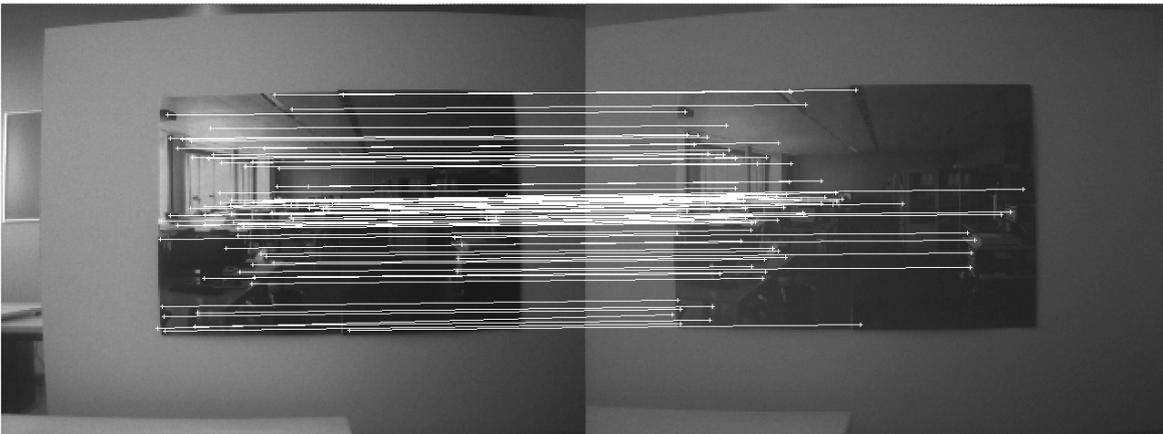
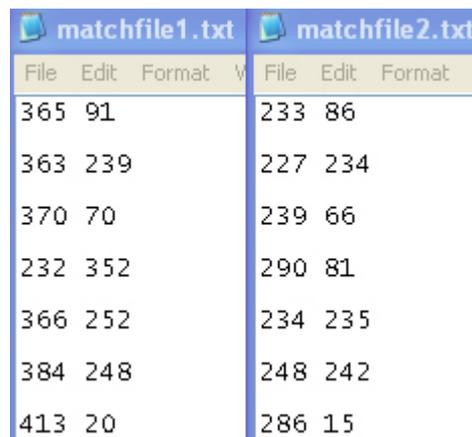


Abbildung 5.12.: Gefundene SIFT-Punktpaare: Es werden nur die Punkte verbunden, zu denen ein Partner im anderen Bild gefunden wird.

nicht abgespeichert. Die Punkte liegen in ihren x- und y-Koordinaten vor und sind in den Dateien so sortiert, dass immer die zusammengehörigen SIFT-Punkte an derselben Stelle stehen. Hierfür wurde das Impresario-Modul, das von Emanuell Lange entwickelt wurde, diesen Bedürfnissen angepasst. Ein Auszug der SIFT-Punktelisten ist in Abbildung 5.13 zu sehen.



File	Edit	Format	File	Edit	Format
365	91		233	86	
363	239		227	234	
370	70		239	66	
232	352		290	81	
366	252		234	235	
384	248		248	242	
413	20		286	15	

Abbildung 5.13.: Liste von SIFT-Punktepaaren: Die Dateien enthalten jeweils die Koordinaten (x, y) der gefundenen SIFT-Punkte, die zu einem SIFT-Punktepaar zusammengefasst worden sind. Die linke Datei enthält die Koordinaten aus der linken Aufnahme und die rechte Datei entsprechend die zugehörigen Koordinaten aus der rechten Aufnahme.

Nach diesen Schritten liegen nun die für die Berechnung der Raumpunkte benötigten und in Frage kommenden Punktepaare vor.

5.3. Ergebnisfilterung und Raumpunktvermessung

Dieser Abschnitt beschreibt, wie aus den SIFT-Punkten die zugehörigen Raumpunkte berechnet werden. Der prinzipielle Ablauf ist in Abbildung 5.14 beschrieben.

Als erstes müssen die ermittelten SIFT-Punkte eingelesen werden. In diesem Zusammenhang erfolgt bereits eine Filterung bezüglich ihrer Verwendbarkeit. Der Ablauf wird im Folgenden beschrieben.

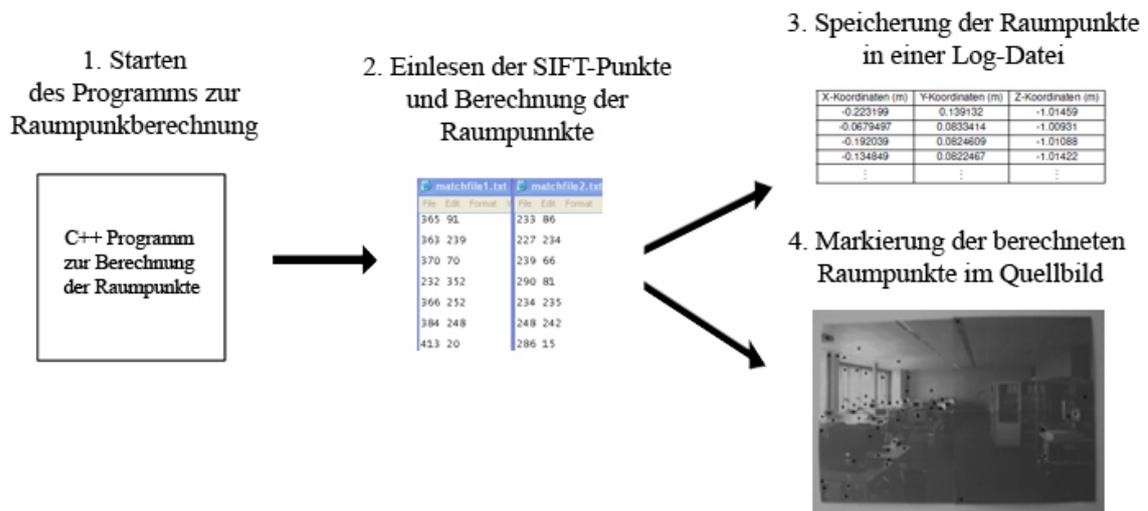


Abbildung 5.14.: Die Grafik zeigt schematisch den Ablauf des C++ Programms zur Berechnung der Raumpunkte aus den vorhandenen SIFT-Punktelisten.

Filterung doppelter Einträge

Es kommt vor, dass die Koordinaten eines Ausgangspunktes, wie in Abbildung 5.15 zu sehen, gleich zwei oder mehreren Zielpunkten zugeordnet werden. Hierdurch ist die Eindeutigkeit der Punkte nicht mehr gegeben, so dass diese Einträge nicht gespeichert werden. Um sicher zu stellen, dass alle doppelten Punkte gefiltert werden, muss die Filtermethode zwei mal aufgerufen werden, wobei einmal die Daten der linken SIFT-Punktliste und einmal die Daten der rechten SIFT-Punktliste ausschlaggebend sind.

Filterung bei horizontaler Abweichung

Ein weiteres Kriterium für die Verwendbarkeit eines SIFT-Punktpaares liegt in ihrer Abweichung der y-Koordinaten zueinander. Es kommt vor, dass die y-Koordinaten bei einem SIFT-Punktpaar, wie in Abbildung 5.16, eine große Differenz aufweisen. Theoretisch darf das nicht passieren, da die beiden Kameras im Falle der Bumblebee-Stereokamera nahezu auf einer Ebene montiert sind. Bei diesen Punkten handelt es sich folglich um falsch zugeordnete Punktpaare, welche gefiltert werden. Eine geringe Abweichung hingegen ist erlaubt, da die Kameras nicht genau auf einer Ebene liegen. Die maximal gewünschte y-Abweichung kann manuell über einen Parameter des C++ Programms eingestellt werden.

Wenn das verwendete Stereokamerasystem aus zwei räumlich voneinander getrennten Kameras besteht, z.B. je eine Kamera auf einer Seite des Fahrzeugs, so kann die Situation

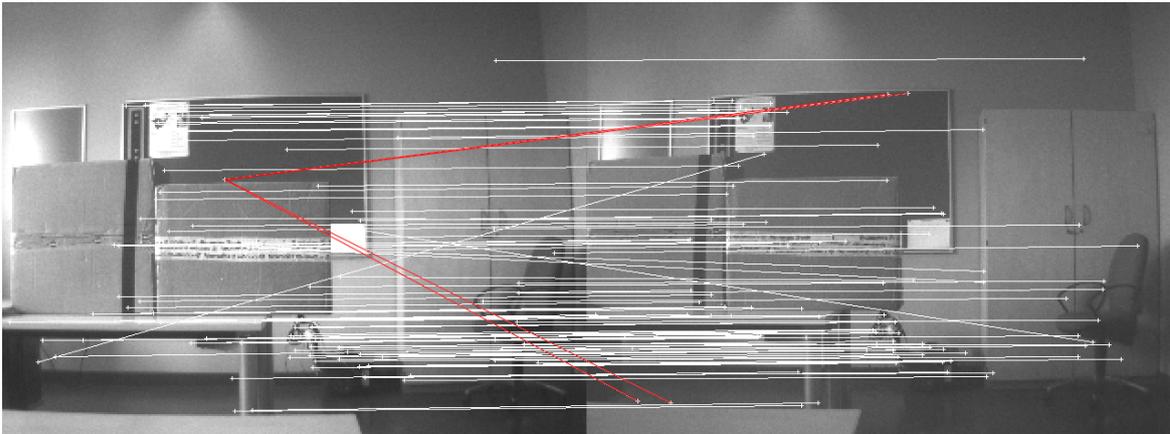


Abbildung 5.15.: Zuordnung eines Quellpunktes zu mehreren Zielpunkten. Die Berechnung eines Raumpunktes ist hier nicht eindeutig möglich und somit werden diese fehlerhaften Punkte gefiltert.

eintreten, dass sich die Kameras zum Aufnahmezeitpunkt nicht auf einer Ebene befinden (Schrägstellung des Fahrzeugs). Im Falle eines solchen Systems muss diese Methode gegebenenfalls angepasst werden.

Mit den verbleibenden SIFT-Punktpaaren und den vorhandenen inneren und äußeren Kameraparametern kann nun die eigentliche Raumpunktberechnung erfolgen. Hierfür werden die im Grundlagenkapitel beschriebenen mathematischen Grundlagen in einem C++ Programm implementiert.

5.3.1. Raumpunkte

Aus der Berechnungsfunktion erhält man alle Raumpunkte, deren Ausgangs-Bildpaare nicht gefiltert werden. Diese werden wie die Quellpunkte in einer Textdatei mit X, Y und Z-Koordinate abgespeichert. Die Anordnung der Daten ähnelt der Darstellung in Tabelle 5.1. Falsch berechnete Punkte, deren Z-Koordinate im positiven Bereich liegt, also räumlich hinter der Kamera, werden nicht gespeichert. Die nun verbleibenden Punkte sind relativ verlässlich. Eine Analyse zur Genauigkeit der berechneten Raumpunkte und zur Menge der markanten Punkte im praktischen Einsatz, findet im folgenden Kapitel statt.

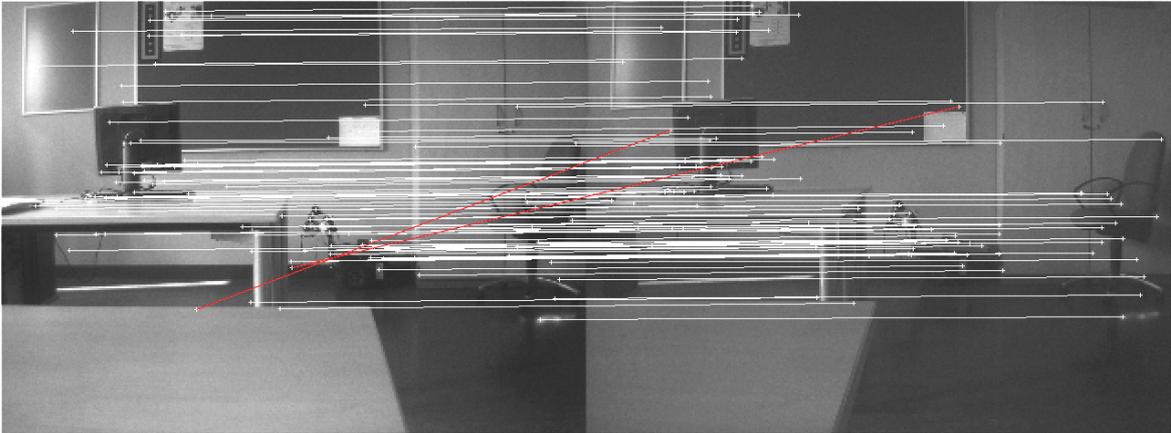


Abbildung 5.16.: Die Abbildung zeigt SIFT-Punktepaare mit einer deutlichen horizontalen Abweichung zwischen den Punkten im linken und rechten Bild. Diese ist auf Grund der Bauweise der Stereokamera nur im Bereich von einigen Pixel möglich. Bei zu großen Abweichungen zwischen den y-Koordinaten ist somit von einer falschen Zuordnung auszugehen.

X-Koordinaten (m)	Y-Koordinaten (m)	Z-Koordinaten (m)
-0.223199	0.139132	-1.01459
-0.0679497	0.0833414	-1.00931
-0.192039	0.0824609	-1.01088
-0.134849	0.0822467	-1.01422
⋮	⋮	⋮

Tabelle 5.1.: Berechnete Raumpunkte: Die Tabelle zeigt die Koordinaten der berechneten Raumpunkte in einer Log-Datei.

5.4. Erstellung des Tiefenbildes

An dieser Stelle sollen die berechneten Raumpunkte in Form eines Tiefenbildes dargestellt werden. Dies ist jedoch nicht in einer der Realität entsprechenden Form möglich, so dass darauf verzichtet werden muss. Die Ursache hierfür liegt sowohl an der geringen Dichte der gefundenen SIFT-Punkte als auch an deren Verteilung im Bild. Diese sind nicht homogen über die Bildfläche verteilt, sondern konzentrieren sich teilweise stark auf einzelne Bildbereiche, so dass für große Teile des Bildes keine Raumpunkte und damit auch keine Tiefeninformationen über den Raum, vorliegen. Selbst in den Bereichen mit Tiefeninformationen kann diese nicht aussagekräftig dargestellt werden. Es liegen keine Informationen über die direkte Umgebung der Raumpunkte vor und ein Rückschluss auf die Formen und Objekte, auf denen die Punkte gefunden werden, ist in der jetzigen Form nicht möglich, zumal nicht jedes Objekt einen SIFT-Punkt aufweist. Die Realität kann also ganz anders aussehen, als es ein mögliches Tiefenbild darstellen würde.

Nach dem jetzigen Stand des Projektes werden die ermittelten Raumpunkte als Kreuze in einer Kopie des Ausgangsbildes der rechten Kamera dargestellt. Abbildung 5.17 zeigt die Markierung der Punkte anhand von Beispielbildern.



Abbildung 5.17.: Ergebnisbild mit markierten Raumpunkten. Jedes Kreuz markiert einen Bildpunkt, zu dem ein Raumpunkt berechnet werden konnte und somit eine Tiefeninformation vorliegt. Für eine bessere Betrachtung sind die Kreuze rot eingefärbt.

6. Analyse

Im vorangegangenen Kapitel wurde ein Programm zur Bestimmung der Raumpunkte entwickelt. Mit Hilfe dieses Programms findet in diesem Kapitel eine Analyse bezüglich der Genauigkeit und Verwendbarkeit der berechneten Raumpunkte, die mit dem SIFT-Verfahren gefunden werden, in Hinsicht auf eine mögliche Anwendung im Bereich der Abstandsmessung und Kollisionsvermeidung statt. Außerdem wird untersucht, wie die allgemeine Ausbeute der SIFT-Punkte unter verschiedenen Bedingungen ausfällt und welche Faktoren diese beeinflussen.

6.1. Genauigkeitsanalyse der Raumpunkte

In diesem Abschnitt wird untersucht, wie verlässlich die berechneten Raumpunkte sind beziehungsweise wie groß die Abweichungen zu den tatsächlichen Raumpunkten sind. Im speziellen wird die Abweichung im Bereich der Distanz betrachtet. Diese ist für eine Kollisionsvermeidung von entscheidender Wichtigkeit.

Versuchsaufbau

Für diesen Versuch wird ein Foto einer realen Situation, Labor, aufgenommen und auf die Größe von vier DIN A4 Blättern vergrößert und mittels einer Holzplatte auf einem Karton befestigt. Das so erstellte Messobjekt ist in Abbildung 6.1 zu sehen. Dieses Szenario wird gewählt, da es genügend unterschiedliche Formen und Strukturen aufweist und der SIFT-Algorithmus somit ausreichend viele markante Punkte finden kann. Die Montage des Bildes auf der Holzplatte stellt sicher, dass das Bild plan aufgeklebt ist und durch den Karton auch vertikal steht. Hiermit soll ein Auftreten von Messfehlern durch einen schlechten Versuchsaufbau minimiert werden.

Im Folgenden wird der Karton in Entfernungen von 37 cm bis 503 cm (maximale Tiefe des Raumes vor der Kamera) vor die Kamera gestellt und werden paarweise Fotos der linken und rechten Kamera aufgenommen. Die Versuchsreihe aus der Perspektive der linken Kamera ist in Abbildung 6.2 dargestellt.



Abbildung 6.1.: Versuchsaufbau der Genauigkeitsanalyse: Holzplatte mit Testbild. Hierfür wurde das Foto eines Labors auf vier DIN A4 Blätter vergrößert und auf einer Holzplatte montiert. Dieser Versuchsaufbau bietet durch die unterschiedlichen Formen und Strukturen eine gute Ausgangsbasis für die Anwendung des SIFT-Algorithmus.



Abbildung 6.2.: Fotoserie der Genauigkeitsanalyse aus Sicht der linken Kamera mit einer Messobjektentfernung von 37 cm (oben links) bis 503 cm (unten rechts).

Bildabstand (in cm)	Anzahl der Messpunkte	berechneter Mittelwert (in cm)	mittlere Abweichung (in cm)	mittlere relative Abweichung (in %)	Standard- abweichung (in cm)
37	4	38,26	+1,26	+3,40	0,22
64,3	54	65,39	+1,09	+1,69	0,58
100	63	101,48	+1,48	+1,48	1,47
150	30	150,15	+0,15	+0,10	2,80
200	23	198,53	-1,46	-0,74	4,62
300	14	292,08	-7,91	-2,63	5,03
403	12	393,45	-9,54	-2,37	8,70
503	9	487,04	-15,95	-3,17	13,71

Tabelle 6.1.: Die Tabelle zeigt die Auswertung der durch die Fotoserie zu Genauigkeitsanalyse gewonnen Messergebnisse.

Jedes Bildpaar wird daraufhin auf SIFT-Punktepaare untersucht und die entsprechenden Raumpunkte werden berechnet, so dass am Ende mehrere Listen mit den berechneten Raumpunkten vorliegen.

Analyse der Messwerte

Im folgenden wird aus den gewonnen Werten das arithmetische Mittel, die mittlere und mittlere relative Abweichung, so wie die Standardabweichung bestimmt.

Für das arithmetische Mittel (Mittelwert) gilt:

$$\bar{x}_{arithm} = \frac{1}{n} \cdot \sum_{i=1}^n x_i.$$

Hieraus ergeben sich für die beschriebenen Testabstände die in Tabelle 6.1 dargestellten Messwerte (in cm).

Grafisch lässt sich die Entwicklung der mittleren absoluten Abweichung mit steigender Distanz wie in Abbildung 6.3 darstellen. Es ist zu erkennen, dass die mittlere Abweichung mit größer werdender Entfernung zum Testobjekt zunimmt.

Wie man der vorherigen Tabelle entnehmen kann, sind die berechneten Raumpunkte im Bereich unterhalb von zwei Metern für viele Anwendungen ausreichend genau. Die Ungenauigkeit nimmt jedoch mit größer werdender Entfernung und in unmittelbarer Kameranähe zu. Dies liegt unter anderem daran, dass der Überschneidungswinkel zwischen den beiden

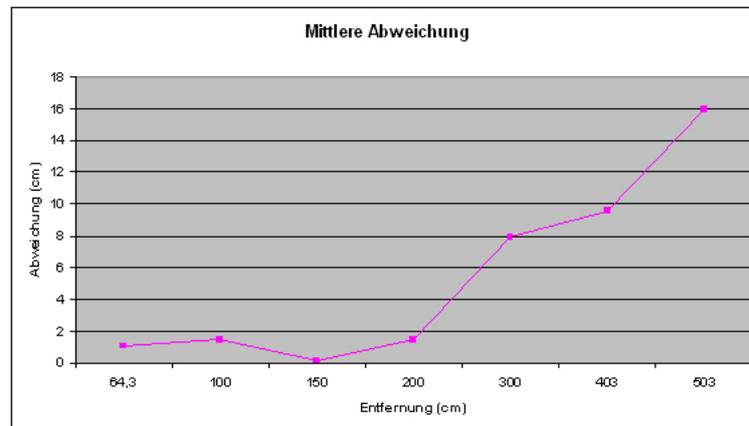


Abbildung 6.3.: Die Grafik zeigt die berechnete mittlere Abweichung aus der Genauigkeitsanalyse. Die Quadrate markieren die Messwerte. Die Linie stellt keine Messwerte dar, sondern soll lediglich eine bessere Augenführung ermöglichen. Aufgrund der sehr kleinen Standardabweichungen der Messwerte werde diese nicht in der Grafik angezeigt.

Kameraaufnahmen und dem zu berechnenden Punkt immer spitzer beziehungsweise flacher wird. Ungenauigkeiten zwischen den Bildpunkten der einzelnen Quellbilder wirken sich bei der Berechnung der Raumpunkte stärker aus. Es ist zu untersuchen, ob dieser Effekt durch eine höhere Bildauflösung gelindert werden kann. Für einen solchen Test ist jedoch ein anderes Kamerasystem erforderlich.

Die mittlere relative Abweichung ist wie in Abbildung 6.4 dargestellt. Für eine bessere Ansicht werden die Beträge der Werte verwendet. Es ist zu erkennen, dass die relative Abweichung bei einer Entfernung von 150 cm ein Minimum aufweist, die Genauigkeit hier also am größten ist. Diese Tatsache unterstützt die Aussage aus dem vorherigen Absatz, dass die Genauigkeit der berechneten Raumpunkte im direkten Nahbereich und bei größerer Entfernung zur Kamera abnimmt.

6.2. Analyse zur SIFT-Punkte-Verteilung

In diesem Abschnitt soll überprüft werden, ob in verschiedenen Situationen genügend SIFT-Punktepaare ermittelt werden können, um daraus eine aussagekräftige Kollisionskontrolle abzuleiten.

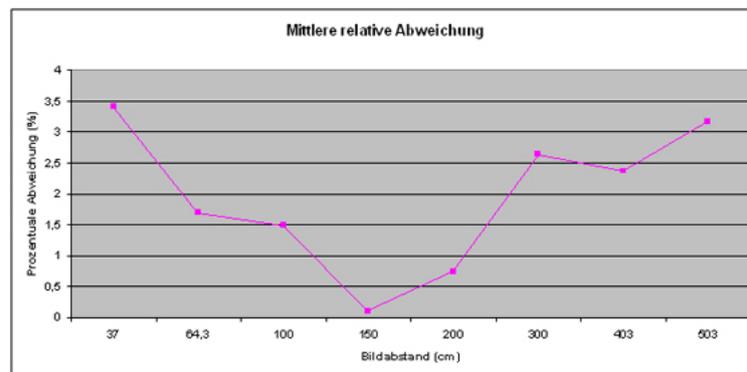


Abbildung 6.4.: Die Grafik zeigt die berechnete mittlere relative Abweichung aus der Genauigkeitsanalyse. Die Quadrate markieren die Messwerte. Die Linie stellt keine Messwerte dar, sondern soll lediglich eine bessere Augenführung ermöglichen. Für eine bessere Ansicht werden die Beträge der Werte verwendet. Es ist zu erkennen, dass die mittlere relative Abweichung bei einer Entfernung von 150 cm ein Minimum aufweist, die Genauigkeit in diesem Bereich also am größten ist.

6.2.1. Flur

Dieses Szenario unterteilt sich in die Untersuchung des Flurs im leeren Zustand und einer Untersuchung mit Hindernissen (Kollisionsobjekte).

Leerer Flur:

In diesem Szenario dient der in Abbildung 6.5 dargestellte Flur als Testobjekt. Die Tapete ist einheitlich gestrichen und verfügt über kein besonderes Muster. An der Wand hängen lediglich einige Pinnwände. Diese Tatsachen sind wichtig, da es hierdurch zu einer geringeren Ausbeute an SIFT-Punkten kommt, als wenn die Wand über ein auffälliges und kontrastreiches Muster verfügen würde. In beiden Fällen wird die Kamera in verschiedenen Entfernungen zur Stirnwand aufgestellt.

Nach der Anwendung des SIFT-Algorithmus auf das Versuchsbild in Abbildung 6.6 ist zu sehen, dass sich der Großteil der gefundenen SIFT-Punktpaare im Bereich der Seitenwände befindet. Diese bieten anscheinend die für den Algorithmus nötige Struktur beziehungsweise eine entsprechend großen Differenz an Grauwerten. An der Stirnwand sind hingegen nur sehr wenige bis fast keine Punktpaare zu finden. Die wenigen vorhandenen Punkte werden jedoch beim Einlesen durch das C++ Programm bereits gefiltert, da es sich hierbei oftmals um doppelt belegte Ausgangspunkte handelt.

Wird die Entfernung zur Stirnwand größer, so verschwinden selbst die wenigen vorher an



Abbildung 6.5.: Szenario - Leerer Flur: Die roten Linien zeigen den Standpunkt der Kamera in verschiedenen Entfernungen zur Stirnwand des Flurs. Hiermit soll die Einordnung der verschiedenen Versuchsbilder erleichtert werden.



Abbildung 6.6.: Leerer Flur mit SIFT-Punktepaaren (Distanz 5m): Es ist zu erkennen, dass sehr wenige SIFT-Punktepaare gefunden worden sind. Besonders im Bildzentrum sind kaum Punkte vorhanden. Dies ist vor allem auf die gleichmäßige Struktur der Stirnwand und die ähnlichen Grautöne in diesem Bereich zurückzuführen, welche wenig Ansatzpunkte für den SIFT-Algorithmus bieten.

der Stirnwand vorhandenen SIFT-Punktpaare, wie in Abbildung 6.7 zu sehen ist. Die SIFT-Punkte sind ausschließlich im Wand- und Deckenbereich zu finden.



Abbildung 6.7.: Leerer Flur mit SIFT-Punktpaaren (Distanz 20m): Die gefundenen SIFT-Punktpaare befinden sich hauptsächlich im Bereich der Decke oder der rechten Seitenwand. Dies liegt vor allem an dem höheren Kontrast durch die Deckenbeleuchtung und die unterschiedlichen Strukturen der Pinnwände.

Die Tabellen, die im Folgenden zu den einzelnen Versuchen gezeigt werden, geben die Anzahl der ermittelten SIFT-Punkte in den verschiedenen Entfernungen zu den Messobjekten wieder. Darüber hinaus ist zu erkennen, wie sich die Anzahl der Punkte mit den verschiedenen Filterungsstufen während der Raumpunkteberechnung verändert. In der letzten Spalte wird dann die Anzahl der tatsächlich berechneten Raumpunkte aufgelistet. Ob sich diese Punkte in der Praxis verwenden lassen, wird hier nicht mehr gesondert untersucht und hängt von der jeweiligen Punkteposition ab.

In den Tabellen 6.2 bis 6.5 wird die SIFT-Punkteverteilung der verschiedenen Testszenarien dargestellt. Stimmen in der jeweiligen Tabelle die Anzahl der Punkte nach dem y -Filter, Filterung der horizontalen Abweichung, und die Anzahl der berechneten Raumpunkte überein, dann wurden keine Punkte falsch berechnet (beispielsweise Punkte hinter der Kamera) oder durch den manuellen Filter herausgefiltert.

Tabelle 6.2 zeigt die Punktverteilung über die verschiedenen Distanzen im Flur-Szenario. Mit zunehmender Entfernung zur Stirnwand nimmt auch die Anzahl der gefundenen SIFT-Punktpaare zu. Gleichzeitig werden aber auch mehr Punkte herausgefiltert. Dies liegt häufig an doppelt zugeordneten Quellpunkten.

Flur mit Hindernissen:

Um dem SIFT-Algorithmus eine bessere Vorlage zur Markierung von SIFT-Punkten zu geben und das Szenario realistischer zu gestalten, werden einige Hindernisse zwischen die

Bildabstand (in m)	Anz. Punkte nach einlesen	Anz. Punkte nach Doppelfilter	Anz. Punkte nach y-Filter	Anz. Raumpunkte
5	10	8	8	8
10	9	7	7	7
15	19	9	8	8
20	18	14	13	13

Tabelle 6.2.: Die Tabelle zeigt die Punkteverteilung im leeren Flur über die verschiedenen Stufen der Raumpunkteberechnung. Die Anzahl der eingelesenen Punkte entspricht der Anzahl der gefundenen SIFT-Punktepaare.

Kamera und die Stirnwand gebracht (siehe Abbildung 6.8). Hierbei handelt es sich um einen Menschen und den Karton aus der Genauigkeitsmessung. Zusätzlich wird ein Bild an der Stirnwand des Flures angebracht. Hierdurch soll diese etwas mehr Kontrast erhalten und dient bei einer möglichen Distanzberechnung mit ihren SIFT-Punkten zur Bestimmung der Entfernung zur Kamera. Diese ist im Sinne des Versuches somit leicht zu kontrollieren.



Abbildung 6.8.: Szenario - Flur mit Hindernissen: Um die Situation realer zu gestalten und mehr SIFT-Punkte zu erhalten, werden verschiedene Hindernisse im Flur aufgestellt.

Die Messergebnisse der SIFT-Punktemenge im veränderten Flur Szenario sind in Tabelle 6.3 zu sehen. Mit zunehmender Entfernung zur Stirnwand nimmt auch die Anzahl der gefundenen SIFT-Punktepaare zu. Dies liegt vor allem daran, dass ab einer bestimmten Entfernung mehr Hindernisse erfasst werden. Die Filterung wirkt sich bei diesem Test nicht so stark aus,

Bildabstand (in m)	Anz. Punkte nach einlesen	Anz. Punkte nach Doppelfilter	Anz. Punkte nach y-Filter	Anz. Raumpunkte
2	5	1	1	1
3	9	3	3	3
4	12	8	8	8
5	5	5	4	4
6	10	3	3	3
7	10	4	4	4
8	16	10	10	10
9	20	14	13	13
10	13	11	10	10
15	16	16	14	14
20	21	17	15	15

Tabelle 6.3.: Die Tabelle zeigt die Punkteverteilung im Flur mit Hindernissen über die verschiedenen Stufen der Raumpunktberechnung.

da die meisten Punkte im Bereich der Hindernisse gefunden werden. Hier gibt es weniger mehrfach zugeordnete Bildpunkte.

6.2.2. Labor

In diesem Fall wird ein Labor mit Menschen als Testobjekt gewählt. Dieses Szenario, siehe Abbildung 6.9, ist von seiner räumlichen Aufteilung vielleicht nicht unbedingt mit einer Anwendung im Bereich Kollisionsschutz vergleichbar, bietet aber durch seine Beschaffenheit eine große Vielfalt an Formen und Strukturen, welche durchaus in einem realen Anwendungsszenario vorkommen kann. Der Schwerpunkt der Untersuchung liegt hier ebenfalls im Bereich der Erkennungsmenge an SIFT-Punktepaaren.

Die Untersuchung ergibt, ähnlich wie beim Flur, dass die Menge der SIFT-Punktepaare mit größer werdender Entfernung zur Stirnwand zunehmen (vergleiche hierzu Abbildungen 6.10 und 6.11). Dies liegt vor allem daran, dass die Wand dem SIFT-Algorithmus keine gute Grundlage für das Auffinden von markanten Punkten bietet.

Tabelle 6.4 zeigt die Auswertung der gefundenen SIFT-Punkte im Labor. Wie in den beiden Flur-Szenarien nimmt auch hier die Anzahl der gefundenen SIFT-Punktepaare mit größer werdender Entfernung zur Stirnwand zu. Durch falsche Punktezuordnungen auf Grund von ähnlichen Strukturen steigt aber auch die Anzahl der gefilterten Punkte im Laufe der Raumpunktberechnung. Die Entwicklung der Raumpunkte mit steigenden Entfernungen ist nicht

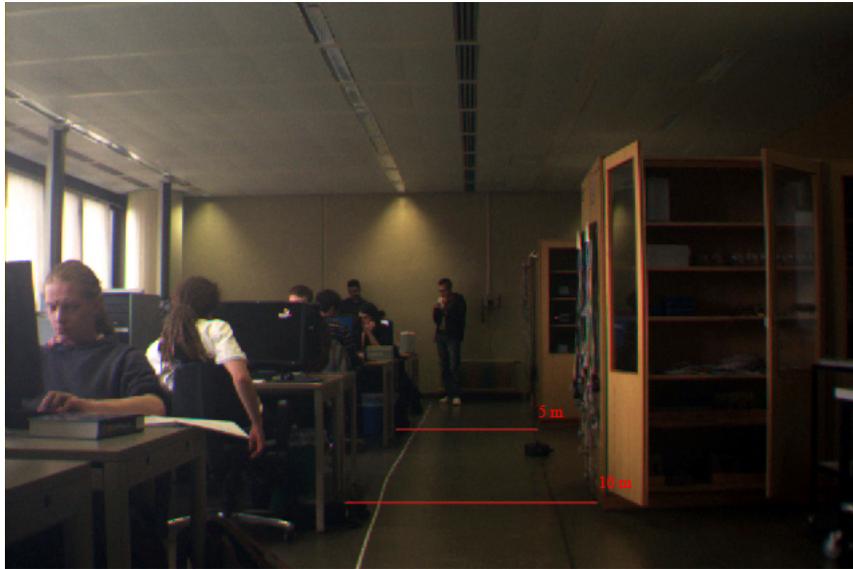


Abbildung 6.9.: Szenario - Labor: Hierbei soll die Wirksamkeit des SIFT-Algorithmus in einer Situation untersucht werden, die auf Grund ihrer Vielfalt an Formen und Strukturen einem realen Anwendungsszenario nahe kommt.

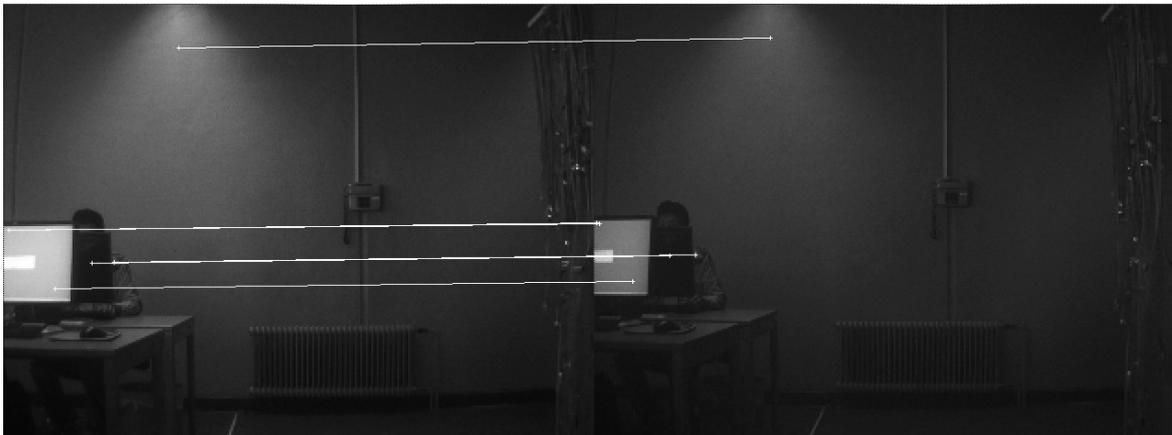


Abbildung 6.10.: Labor mit SIFT-Punktpaaren (Distanz 5m): Es werden so gut wie keine SIFT-Punktpaare gefunden. Hier gibt es, wie beim leeren Flur, zu wenige struktur- und kontrastreiche Flächen.

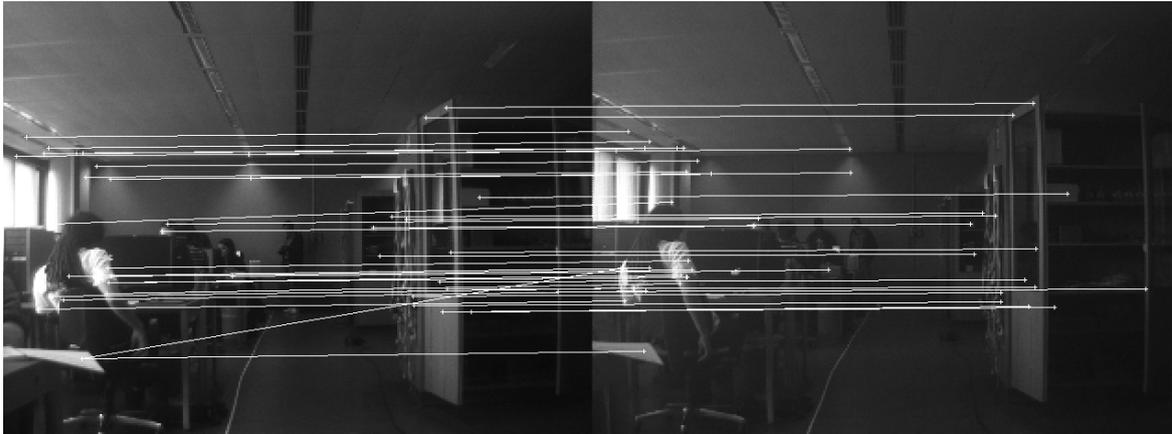


Abbildung 6.11.: Labor mit SIFT-Punktpaaren (Distanz 15m): Die Menge der gefundenen SIFT-Punktpaare ist erheblich größer als im Bereich von 5m, da nun der volle Umfang des Labors mit seiner Einrichtung erfasst werden kann. Im direkten Bildzentrum werden jedoch kaum markante Punkte erkannt, da sich dort die Rückwand des Labors befindet.

Bildabstand (in m)	Anz. Punkte nach einlesen	Anz. Punkte nach Doppelfilter	Anz. Punkte nach y-Filter	Anz. Raumpunkte
5	8	2	2	2
10	16	10	9	9
15	47	32	26	26

Tabelle 6.4.: Die Tabelle zeigt die Punkteverteilung im Labor über die verschiedenen Stufen der Raumpunkteberechnung.

unproblematisch für den Einsatz des Verfahrens im Bereich des Kollisionsschutzes, da dort der Schwerpunkt auf kürzeren bis mittleren Distanzen liegt.

6.2.3. Arbeitsraum

Diese Testumgebung zeigt zum einen den Weg der Kamera in einen Arbeitsraum (Abbildung 6.12) und zum anderen die Aufnahme des gleichen Raums aus der Innenansicht (Abbildung 6.13). Die Menge der potenziellen Punktpaare ist in den Abbildungen ebenfalls zu erkennen. Mit diesem Szenario soll eine weitere Situation in den Test aufgenommen werden, wie sie durchaus in der Praxis bei der Bewegung eines autonomen Arbeitsgeräts vorkommen kann. Das Ziel bleibt hier auch die Auswertung der potenziellen Menge an SIFT-Punktpaaren.

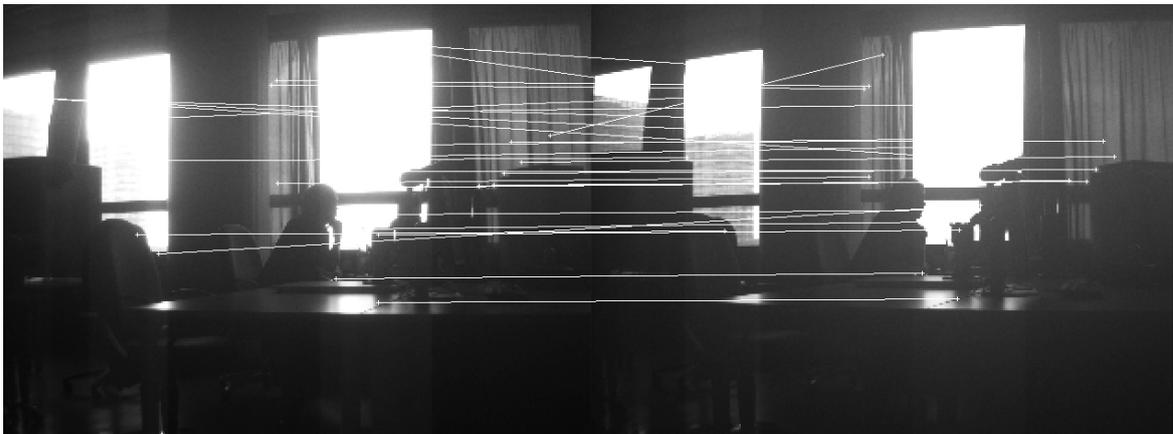


Abbildung 6.12.: Szenario - Reinfahrt in einen Arbeitsraum: Diese Aufnahme simuliert die Fahrt eines Fahrzeuges mit montierter Stereokamera in einen Arbeitsraum. Bei dieser Aufnahme steht die Anzahl gefundener SIFT-Punktepaare in einer alltäglichen Umgebung im Vordergrund. Die Anzahl der SIFT-Punktepaare wird durch die dunkle Aufnahme und dem daraus resultierenden geringen Kontrast beeinträchtigt.



Abbildung 6.13.: Szenario - Arbeitsraum aus statischer Position: Diese Aufnahme simuliert den Blick eines Fahrzeuges in einem Arbeitsraum. Bei dieser Aufnahme steht ebenfalls die Anzahl gefundener SIFT-Punktepaare in einer alltäglichen Umgebung im Vordergrund. Es ist zu erkennen, dass eine große Menge an SIFT-Punktepaaren gefunden wird. Dies wird vor allem durch die Vielzahl der vorhandenen Objekte und andere Lichtverhältnisse begünstigt.

Kamera- position	Anz. Punkte nach einlesen	Anz. Punkte nach Doppelfilter	Anz. Punkte nach y-Filter	Anz. Raumpunkte
Tür	30	23	19	19
im Raum	148	100	93	93

Tabelle 6.5.: Auswertung der Punkteverteilung im Arbeitsraum: Die Tabelle stellt die Anzahl der SIFT-Punktepaare im Laufe der Raumpunktberechnung dar.

Hierbei fällt auf, dass im Vergleich zum Labor und vor allem zum Flur eine Menge an SIFT-Punktepaaren gefunden wird. Die Berechnung und Kontrolle der möglichen Raumpunkte wird hier nicht erneut vorgenommen. Man sieht aber, dass die hohe Anzahl der Punktepaare eine große Menge an potenziellen Raumpunkten und damit genügend Material für eine mögliche Anwendung zur Kollisionsvermeidung (vergleiche Tabelle 6.5) bietet. Die Verringerung der Punktemenge im Bereich der Innenaufnahme ist auf die Filterung doppelter und falsch zugeordneter Ausgangspunkte zurückzuführen.

6.2.4. Auswertung der SIFT-Punktemenge

Die Messwerte der Versuche „Flur“, „Flur mit Hindernissen“ und „Labor“, sind in Abbildung 6.14 zusammengefasst dargestellt. Es ist zu erkennen, dass sich die Verwertbarkeit der gefundenen SIFT-Punkte in diesen Versuchen je nach Entfernung zwischen 60% und 70% einpendelt. Für den Nahbereich, welche im Szenario „Flur mit Hindernissen“ untersucht wurde, liegt die Verwertbarkeit teilweise deutlich unter diesen Werten. Der Grund liegt vermutlich an der ungünstigen Flurumgebung.

Für die beiden Aufnahmen des Arbeitsraums ergeben sich folgenden Werte:

- Arbeitsraum Reinfahrt: 63%
- Arbeitsraum Statisch: 62%

Im Mittel über alle Versuche beträgt der Anteil der berechneten Raumpunkte etwa 60% bis 70% der ursprünglich gefundenen SIFT-Punkte.

6.2.5. Analyse des Verfahrens am Menschen

Wie in den beschriebenen Testszenarien erläutert, ist die Ausbeute der SIFT-Punktepaare durchaus unterschiedlich. Diese hängt besonders von der Vielfalt der Strukturen ab, wie z.B. Personen, Tapeten, Gerätschaften etc. Einfache Räume ohne diese Vielfalt liefern wenige

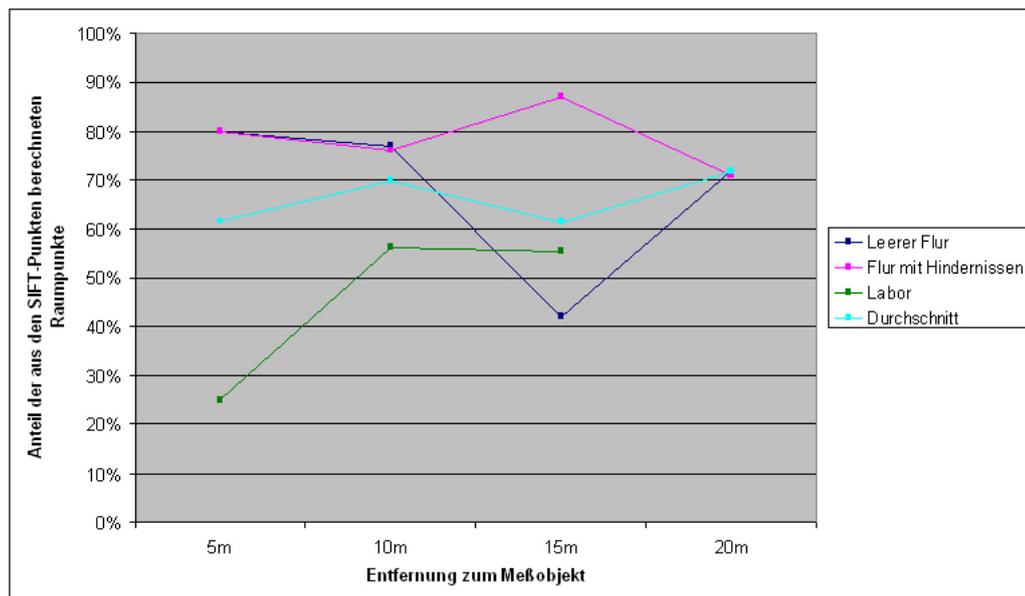


Abbildung 6.14.: Die Grafik zeigt den Anteil der SIFT-Punkte die zur Berechnung eines Raumpunktes führen. Die Quadrate markieren die Messwerte. Die Linien stellen keine Messwerte dar, sondern sollen lediglich eine bessere Augenführung ermöglichen. Auf Grund der Raumgröße des Labors liegen hier die Messwerte nur bis zu einer Entfernung von 15m vor. Die Abweichung im Bereich des leeren Flures ist auf eine ungünstige Aufnahmesituation zurückzuführen, bei der es viele doppelte Punktezuordnungen gegeben hat.

Punktepaare und diese können durch ihre Positionierung auch häufig nicht für den Kollisionsschutz verwendet werden. Bei einer ausreichend vorhandenen Struktur müssen die Punkte gegebenenfalls auf Verlässlichkeit geprüft werden.

Ein nicht zu vernachlässigender Punkt, gerade im Bereich Mensch-Maschine-Interaktion, ist die Tatsache, dass an Menschen kaum SIFT-Punkte gefunden werden (siehe Abbildung 6.15). Die Anzahl könnte jedoch möglicherweise durch eine andere Kleidungswahl erhöht werden. Hierzu wurde eine Versuchsreihe durchgeführt. Abbildung 6.16 zeigt die Testbilder dieser Versuchsreihe. Die Bilder zeigen eine Person, die einmal mit einem alltäglichen Pullover und einmal mit einer Signaljacke bekleidet ist. Die Signaljacke liefert eine deutlich höhere Anzahl an SIFT-Punkten. Die Anzahl der gefundenen SIFT-Punkte wird zusätzlich mit zunehmender Entfernung geringer. So lange im Nahbereich aber einige verlässliche Punkte an Menschen gefunden werden, kann dies für eine Kollisionsvermeidung ausreichen. Kritisch wird es jedoch, wenn selbst diese wenigen Punkte nicht mehr gefunden werden.

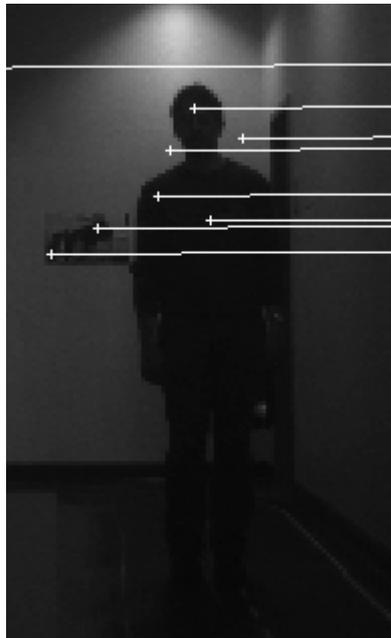


Abbildung 6.15.: SIFT-Punkte an einer Person: Die Abbildung verdeutlicht, wie wenig SIFT-Punkte an einem Menschen gefunden werden. Es liegt die Vermutung nahe, dass die Kleidung nicht genug Struktur aufweist, beziehungsweise diese im Verlauf der Grauwertbilderstellung verloren geht.

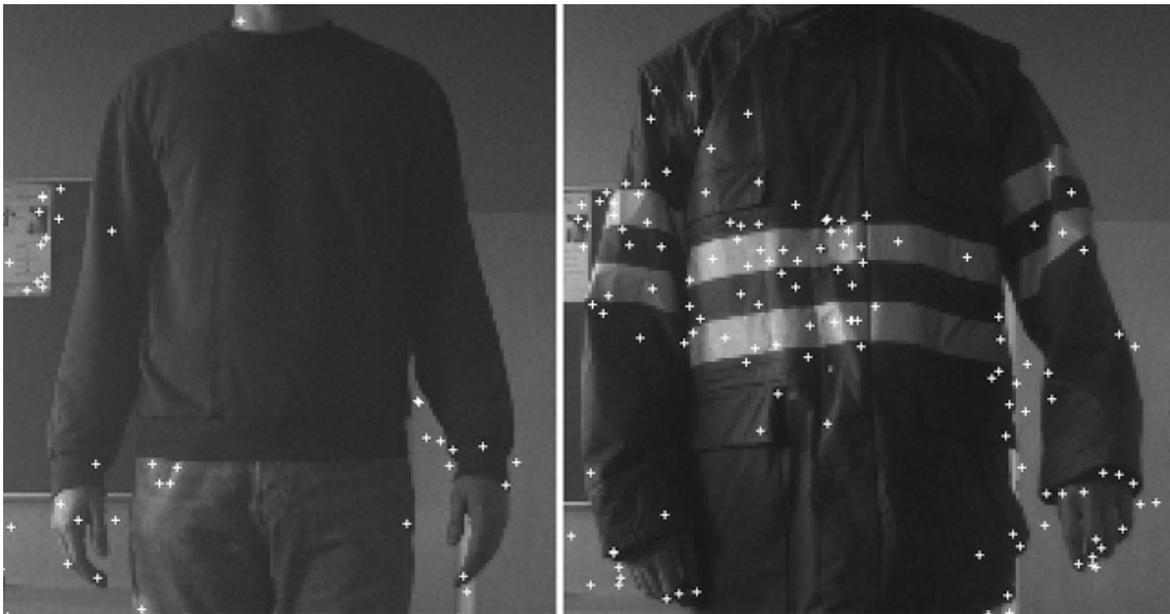


Abbildung 6.16.: SIFT-Punkte an einer Person (links: alltäglicher Pullover, rechts: Signaljacke): Es ist zu sehen, dass die Art der Bekleidung einen Einfluss auf die Anzahl der gefundenen SIFT-Punktepaare hat.

6.3. Qualitätseinflüsse

Im Bereich der Zuverlässigkeit und Masse der gewonnen Raumpunkte spielen mehrere Faktoren eine Rolle. Bei den Faktoren „Downsampling“ und „Bildauflösung“ wird die Auswirkung von unterschiedlichen Auflösungen der Quellbilder in verschiedenen Stufen der SIFT-Punktermittlung untersucht. Der Abschnitt „Helligkeit und Kontrast“ beschäftigt sich mit der Manipulation der Quellbilder und die Auswirkungen einer alternativen Kameraanordnung für die Aufnahme der Stereobilder wird unter dem Punkt „Basisbreite“ untersucht. Zum Schluss findet im Abschnitt „Auffälligkeiten der SIFT-Punkteverteilung“ eine Betrachtung der Verteilung der gefundenen SIFT-Punkte statt.

6.3.1. Downsampling

Unter Downsampling versteht man die Verringerung der Bildpunkte einer Grafik. Das Bild wird also verkleinert. Hiermit wird ein schnelleres Auffinden von SIFT-Punkten erreicht. In Impresario wird hierfür das bereits erwähnte Plugin von Emanuell Lange verwendet. Das Modul zum Downsampling wird zwischen die Konvertierung des Originalbilds in ein Grauwert-

bild und den SIFT-Algorithmus integriert. Abbildung 6.17 zeigt den Aufbau des Impresario-Projekts mit und ohne Downsampling.

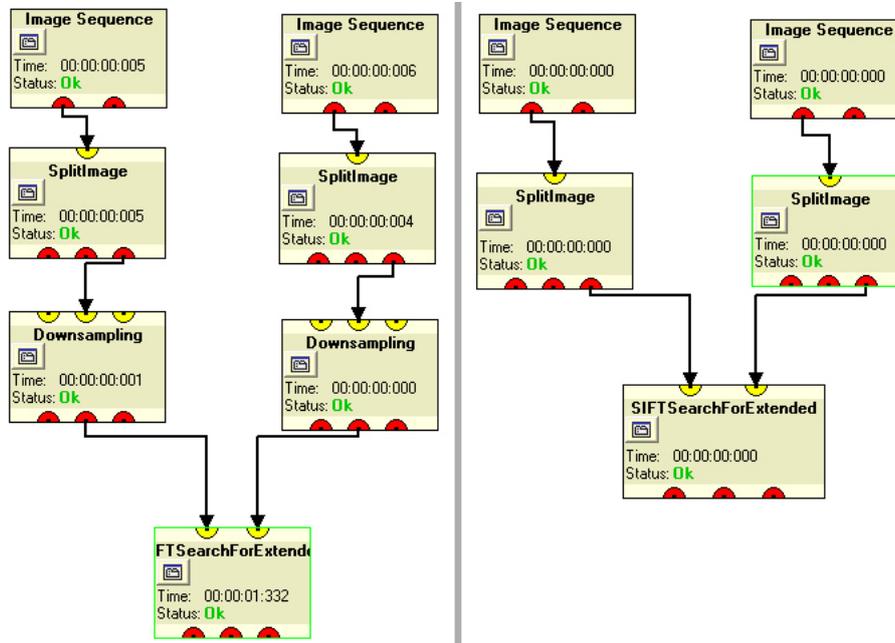


Abbildung 6.17.: Die abgebildeten Impresario-Projekte zeigen einmal das bereits erwähnte und verwendete Projekt ohne Downsampling (rechte Seite) und einmal das gleiche Projekt mit der Erweiterung der Downsampling-Funktion. Hierbei werden die in Graubilder umgewandelten Quellbilder (Split-Image) vor der Weiterleitung an das Makro zur SIFT-Punktebestimmung (FTSearchForExtended) an das Makro „Downsampling“ übergeben. Hier wird die Bildgröße halbiert, bevor die Suche nach den SIFT-Punkten erfolgt.

Folgen des Downsamplings:

Es werden weniger SIFT-Punkte gefunden, da das Bild kleiner wird und ggf. markante Punkte vor der Analyse verloren gehen (siehe Abbildung 6.18).

Durch die geringere Menge an SIFT-Punkten werden auch weniger SIFT-Punktpaare gefunden. Abbildung 6.19 und 6.20 zeigen die gefundenen SIFT-Punktpaare in verschiedene Szenarien im Original und unter Verwendung von Downsampling.

Hiermit werden potenzielle SIFT-Punkte bereits vor der Berechnung der Raumpunkte und dem damit verbundenen Filtervorgang ausgeschlossen. Ein Nachteil, wenn auf das Downsampling verzichtet wird, ist, dass die Ermittlung der SIFT-Punktpaare länger dauert. Die Dauer hängt von der Auflösung der verwendeten Kamera ab. In diesem Fall liegt der Unterschied bei der Ermittlung der Raumpunkte mit und ohne Downsampling-Option im Sekun-



Abbildung 6.18.: SIFT-Punkte mit Downsampling (links) und ohne Downsampling (rechts): Es ist zu erkennen, dass im linken Bild deutlich weniger SIFT-Punkte gefunden werden.

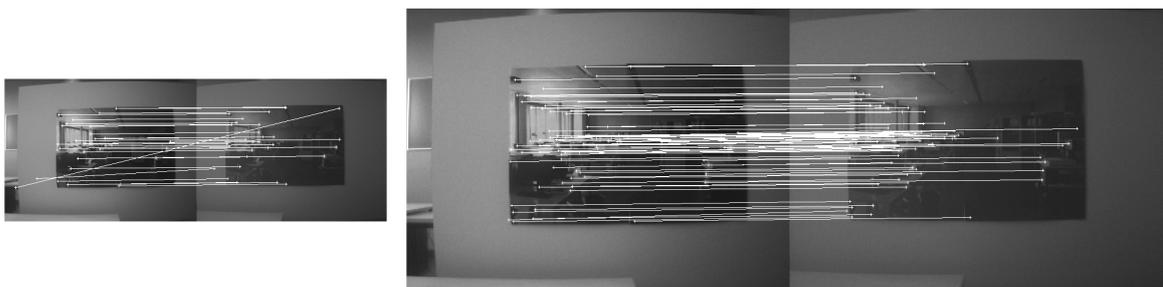


Abbildung 6.19.: SIFT-Punktpaare mit Downsampling (links) und ohne Downsampling (rechts): Es ist zu erkennen, dass sich die geringere Anzahl an gefundenen SIFT-Punkten auch in der Menge der gefundenen SIFT-Punktpaare niederschlägt. Die Auswirkung ist noch größer, wenn zusätzlich falsch zugeordnete Punkte gefiltert werden.

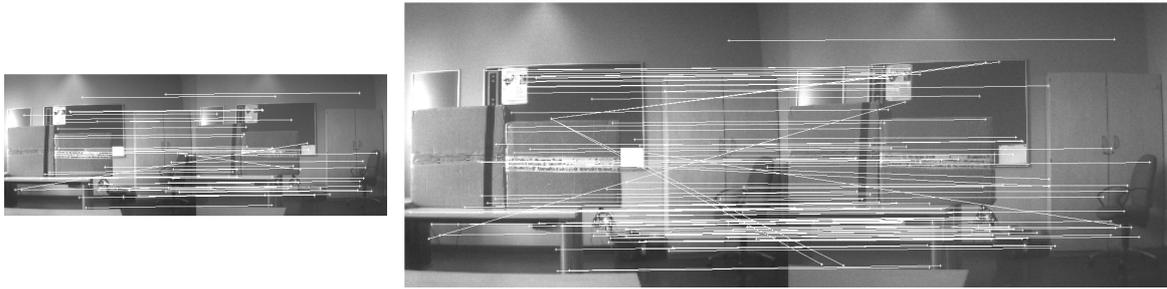


Abbildung 6.20.: SIFT-Punktpaare mit Downsampling (links) und ohne Downsampling (rechts) in einem Arbeitsraum: Dieses Szenario bestätigt die Annahme, dass durch das Downsampling eine geringere Menge an SIFT-Punktpaaren gefunden wird.

denbereich. Schnelle Reaktionszeiten spielen unter anderem in Eingebetteten-Systemen eine wichtige Rolle. Des Weiteren ist es aufgrund der begrenzten Kapazität dieser Systeme häufig notwendig, den Stromverbrauch durch eine Verringerung der benötigten Rechenleistung zu senken.

Die Auswirkung des Downsamplings auf die Anzahl der gefundenen SIFT-Punkte könnte bei einer hohen Qualität der Quellbilder eine weniger starke Auswirkung haben als bei niedriger Qualität wie in diesem Fall. Eine Kombination von höher aufgelösten Quellbildern und Downsampling muss jedoch erst getestet werden.

6.3.2. Bildauflösung

Werden die Bilder von einer Kamera mit geringer Auflösung aufgenommen, so werden mögliche SIFT-Punkte von Anfang an ausgeschlossen, da Strukturen und Details im Bild nicht so deutlich abgebildet werden wie bei einer hohen Auflösung. Dies führt meist dazu, dass weniger SIFT-Punkte gefunden werden. Des Weiteren besteht die Gefahr, dass potenzielle, bei einer geringen Auflösung ermittelte SIFT-Punkte im Zuge des SIFT-Algorithmus verloren gehen, da das Bild durch die geringere Qualität weniger Unterschiede im Grauwertbereich aufweist. Diese Unterschiede werden für die Unterscheidung der einzelnen SIFT-Punkte benötigt. Durch die starke Ähnlichkeit der Merkmalsvektoren sind die Punkte nicht für den weiteren Gebrauch geeignet.

Ein Bild mit geringer Auflösung ähnelt einem Bild mit doppelter Auflösung und darauf angewandter Downsampling-Funktion. Der Unterschied ist, dass das Bild mit der geringen Auflösung schon beim Erstellen ungenauere Bildinformationen liefert. Ist die Auflösung der Kamera also bereits niedrig und wird zusätzlich noch die Funktion des Downsamplings verwendet, so wird die Ausbeute der SIFT-Punktpaare immer geringer.

Im Zuge dieser Bachelorarbeit wird außer der Bumblebee kein weiteres Stereokamerasystem verwendet, so dass die Auswirkungen einer höheren Auflösung nicht getestet werden kann.

6.3.3. Helligkeit und Kontrast

Je ähnlicher die verwendeten Kameras des Stereokamerasystems und damit die Qualität der aufgenommenen Bilder sind, desto besser gelingt die Zuordnung der zusammengehörigen SIFT-Punkte. Bei der in dieser Arbeit verwendeten Kamera gibt es merkliche Unterschiede zwischen den einzelnen Kameraaufnahmen. Dies wird besonders im Bereich der Bildhelligkeit deutlich. Woraus dieser Unterschied resultiert konnte nicht festgestellt werden. Eine Folge dieser Tatsache ist, dass einige der SIFT-Punkte die in einem Bild gefunden werden, auf Grund der schlechten Qualität im anderen Bild nicht mehr erkannt werden. Die tatsächliche Anzahl der SIFT-Punkte bleibt hinter der möglichen Anzahl zurück.

Versuche zur Bildverbesserung im Bereich der Ausgangsbilder, speziell mit den Faktoren Helligkeit und Kontrast, ergaben bei der ermittelten SIFT-Punkte menge keinen nennenswerten Vorteil gegenüber den unbearbeiteten Bildern. Eine Ausnahme bilden hier extrem dunkle Quellbildaufnahmen. Hier konnten nach einer Aufhellung wenigstens einige SIFT-Punkte ermittelt werden.

In einem Versuch konnte die Menge der gefundenen SIFT-Punkte aber durch einen automatischen Weißabgleich¹ der Quellbilder erheblich gesteigert werden, siehe hierzu Abbildung 6.21.

Die Abbildungen 6.22 und 6.23 zeigen die Menge der SIFT-Punktepaare nach einem automatischen Weißabgleich im Flur-Szenario und die SIFT-Punkte an einem Menschen.

Ob sich der Weißabgleich im Allgemeinen anwenden lässt und die Genauigkeit der Raumpunkte mit der gemessenen übereinstimmt beziehungsweise diese sogar übertrifft, ist zu überprüfen. Eventuell ist die Menge der SIFT-Punkte unter anderen Aufnahmebedingungen oder mit einem anderen Stereokamerasystem von vornherein größer.

6.3.4. Basisbreite des Stereokamerasystems

Unter der Basisbreite versteht man den Abstand der einzelnen Kameras des Stereokamerasystems zueinander. Um ein optimales Messergebnis zu erzielen, darf der Winkel zwischen

¹Das Licht lässt sich in verschiedene Farbtemperaturen einteilen. So haben verschiedene Lichtsituationen am Aufnahmeort auch unterschiedliche Einflüsse bei der Aufnahme von Fotos. Der Weißabgleich bietet eine Möglichkeit die Farbtemperaturen den vorherrschenden Lichtverhältnissen anzupassen.



Abbildung 6.21.: Die Abbildung zeigt einmal das Quellbild vor dem Weißabgleich (links) und danach (rechts). Die Menge der gefundenen SIFT-Punkte konnte mit diesem Verfahren erheblich gesteigert werden.



Abbildung 6.22.: Die Abbildung zeigt die Menge der SIFT-Punktpaare im Flur mit Hindernissen vor und nach einem Weißabgleich (links das Originalbild, rechts nach einem Weißabgleich).



Abbildung 6.23.: Durch den Weißabgleich werden auch an einem Menschen mehr SIFT-Punkte gefunden. Das Bild zeigt die Menge der gefundenen SIFT-Punkte vor (links) und nach (rechts) dem Abgleich.

den Stereobildern nicht zu klein, spitzer Winkel bei großer Entfernung, und nicht zu groß, flacher Winkel bei nahen Distanzen, werden, da sich Abweichungen der Bildkoordinaten in diesen Bereichen besonders stark bei der Berechnung der Raumpunkte auswirken (vergleiche hierzu Abbildung 6.24 und Abbildung 6.25). Wenn ein Objekt zu nah ist, als das es von beiden Kameras erfasst werden kann, dann ist die Basisbreite zu groß und muss verringert werden. Da meist eine feste Basisbreite vorliegt, ist somit auch der minimale Abstand gegeben, in dem Objekte in beiden Kameras erkannt werden können. Liegt eine geringe Basisbreite der Kameras vor, können Objekte auch in sehr naher Entfernung wahrgenommen werden. Dafür wird das Schätzen von Entfernungen mit zunehmender Distanz immer ungenauer. Wählt man eine größere Basisbreite, so können nahe gelegene Objekte erst ab einer bestimmten Entfernung erfasst werden. Dafür lassen sich Objekte in größeren Entfernungen besser erfassen und Ungenauigkeiten treten erst später ein. Für verschiedene Anwendungen gibt es also verschiedene optimale Basisbreiten, je nachdem in welchen Bereichen die Entfernung genau gemessen werden soll.

6.3.5. Auffälligkeiten der SIFT-Punkteverteilung

Die verschiedenen Versuche haben gezeigt, dass die Verteilung der Merkmalspunkte nicht gleichmäßig über das ganze Bild erfolgt. Teilweise wird in ganzen Teilbereichen des Bildes kein einziger markanter Punkt ermittelt. Dies ist besonders kritisch, wenn das Verfahren für

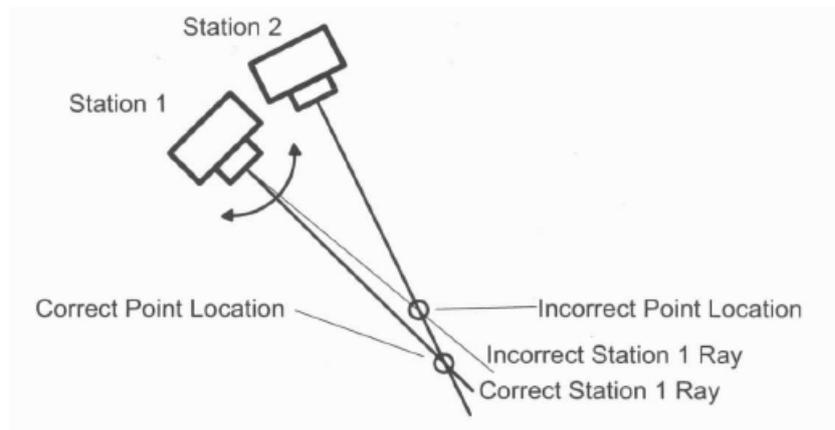


Abbildung 6.24.: Die Abbildung zeigt die Anordnung der einzelnen Kameras eines Stereokamerasystems in einem ungünstigen Winkel zueinander. Wenn Kamera 1 (Station 1) die Zuordnung des korrekten Punktes falsch vornimmt, wirkt sich dieses gleich in einer großen Abweichung zwischen wirklicher und berechneter Entfernung aus (Systems 2007).

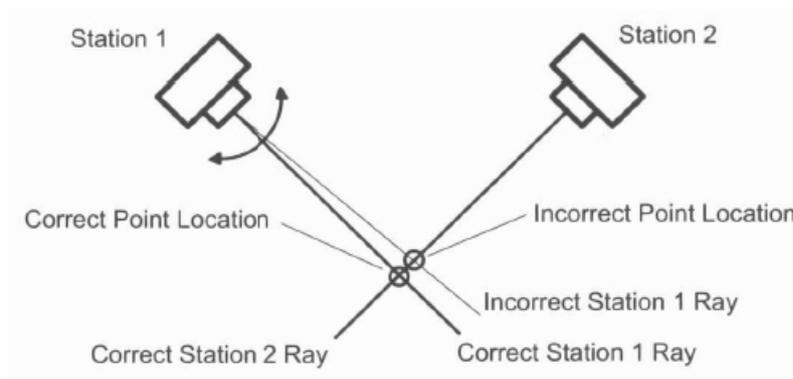


Abbildung 6.25.: Die Abbildung zeigt die Anordnung der einzelnen Kameras eines Stereokamerasystems in einem guten Winkel zueinander. Wenn Kamera 1 (Station 1) die Zuordnung des korrekten Punktes falsch vornimmt, wirkt sich dieses nicht so stark auf die Abweichung zwischen wirklicher und berechneter Entfernung aus (Systems 2007).

den Kollisionsschutz verwendet werden soll. Es kann passieren, dass im direkten Bereich vor der Kamera und damit vor dem Fahrzeug keine markanten Punkte ermittelt werden. Ein potenzielles Hindernis wird also möglicherweise nicht erkannt. Es ist zu überprüfen, ob die räumliche Verteilung der Punkte beeinflusst werden kann.

7. Diskussion und Ausblick

Abweichend zum Titel dieser Arbeit, bildet das beschriebene Tiefenbild nicht mehr den zentralen Aspekt dieser Bachelorarbeit. Es hat sich gezeigt, dass ein auf SIFT-Punkten basiertes Tiefenbild in seiner momentan möglichen Form kein verlässliches Abbild der Umgebung liefern kann. Dies liegt vor allem an der Verteilung der gefundenen Punkte, welche nicht homogen über das Bild verteilt sind und, je nach Raumumgebung, teilweise ganze Raumbereiche freilassen sowie an der geringen Punktedichte im Allgemeinen. Somit liegen für große Teile des Bildes keine Tiefeninformationen vor. Selbst in den Bereichen mit Tiefeninformationen kann diese nicht aussagekräftig dargestellt werden, da es keine Informationen über die direkte Umgebung der Punkte gibt. Es ist nicht sichergestellt, dass jedes Objekt im Bild einen SIFT-Punkt aufweist und damit alle potenziellen Hindernisse erkannt werden. Die Realität kann also ganz anders aussehen, als es durch ein Tiefenbild dargestellt werden könnte. Das Tiefenbild sieht in seiner momentanen Form so aus, dass die Punkte im Bild markiert werden, für die ein Raumpunkt und damit eine Tiefeninformation ermittelt worden ist. Zusätzlich gibt es eine Log-Datei in welcher die Raumpunkte abgespeichert werden. In Folge dieser Erkenntnis und in Anlehnung an die für ein Tiefenbild geknüpften Genauigkeitsuntersuchungen wurde die Aufgabe dahingehend erweitert, eine Aussage über die Anwendbarkeit des SIFT-Algorithmus in realen Umgebungen zu treffen. Das Augenmerk liegt hierbei auf der Anzahl der gefundenen SIFT-Punkte, welche für eine praktische Anwendung von entscheidender Wichtigkeit sind.

Die Genauigkeitsanalyse der durch die SIFT-Punkte gewonnenen Raumpunkte hat ergeben, dass die Entfernung zur Kamera relativ genau berechnet werden kann. Auch wenn diese mit zunehmender Entfernung ungenauer wird, so ist das Verfahren im entscheidenden Nahbereich sehr genau. Hiermit könnte es beispielsweise für den Einsatz im Bereich der Kollisionskontrolle geeignet sein.

Entscheidend für einen realen Einsatz ist zusätzlich die Verfügbarkeit einer genügenden Anzahl von SIFT-Punkten zur Ermittlung von Hindernissen. Die Analyse in diesem Bereich ergab, dass in durchschnittlichen Umgebungen eine entsprechende Menge an Punkten gefunden wird. In manchen Umgebungen werden allerdings nur wenige SIFT-Punkte gefunden. Ob ein Arbeitsumfeld für den praktischen Einsatz geeignet ist, sollte daher anhand von Probeaufnahmen der jeweiligen Umgebung überprüft werden.

Da die SIFT-Punkte teilweise weit über das Bild verstreut sind, muss für einen reellen Einsatz gegebenenfalls der Einzugsbereich der zu verwendenden Punkte entsprechend eingegrenzt werden, um nur die für die Kollisionsvermeidung relevanten Punkte zu verwenden. Für die Praxis heißt dies, dass bei einer großen Entfernung zu einem Kollisionsobjekt nur die Punkte in einem kleinen Bildausschnitt in der Bildmitte analysiert werden. Bei einer Annäherung der Kamera an das Kollisionsobjekt muss dieser Bildausschnitt entsprechend vergrößert werden. Der zu analysierende Bildbereich kann über einen manuellen Filter eingegrenzt werden.

Die Quantität der gefundenen SIFT-Punkte hängt vor allem von der im Einsatzgebiet vorherrschenden Umgebungsstruktur (Wand, Regale etc.) ab. Hier kommt es in ungünstigen Fällen dazu, dass Punkte nur im Randbereich des Bildes oder im Deckenbereich gefunden werden. Dies wären für einen Kollisionsschutz unwichtige beziehungsweise eher vernachlässigbare Informationen. Für den Einsatz im Bereich Mensch und Maschine ist zu sagen, dass Menschen nach Erfahrungen mit dem SIFT-Algorithmus generell sehr wenige und teilweise auch nicht verwertbare Punkte aufweisen. Dieser Nachteil kann jedoch durch eine entsprechend kontrastreiche Kleidung verringert werden. Im Zweifelsfall würde aber auch ein einzelner richtig bestimmter Punkt zur Kollisionsvermeidung ausreichen. Kritisch wird es, wenn selbst die wenigen vorhandenen Punkte wegfallen. In geringen Entfernungen wurden generell weniger Punkte gefunden. Diese Tatsache ist nicht zu vernachlässigen, da die Gefahr von Kollisionen gerade im Nahbereich liegt.

Einen Ansatz für eine generelle Verbesserung des gesamten Verfahrens könnte ein Stereokamerasystem mit variabler Basisbreite darstellen. Die Kameras des Systems wären hierbei nicht mehr in einem gemeinsamen Gehäuse untergebracht. Durch die variable Basisbreite könnte vor allem die Reichweite des Systems und die Genauigkeit bei größeren Distanzen gesteigert werden. Es ist jedoch zu prüfen, in wie weit die Genauigkeit bei größeren Entfernungen für ein Kollisionsschutzsystem von Bedeutung ist. Eventuell würde dieser Ansatz über den Einsatz in einem Hochregallager hinausgehen.

Wahrscheinlich kann die Menge der SIFT-Punkte positiv durch den Einsatz von hochauflösenden Kameras beeinflusst werden. Handelt es sich bei dem Einsatzgebiet um einen Anwendungsbereich, in der Zeit ein kritischer Faktor ist, dann benötigt der SIFT-Algorithmus aufgrund der Bildqualität, bei einer generell besseren Punkteausbeute, eine deutlich längere Laufzeit. Es wäre auch zu prüfen, wie sich der Einsatz von zwei einzelnen am Fahrzeug befestigten Kameras (größere Basisbreite) in Verbindung mit einer höheren Kameraauflösung auf die Punktequalität auswirken würde. Allerdings ist hier zu beachten, dass der minimal berechenbare Abstand zum Fahrzeug abnimmt. Eine Alternative für die Reduzierung der minimalen messbaren Entfernung könnte im Einsatz von weitwinkligen Objektiven liegen. Diese Thesen sollten aber noch genauer überprüft werden.

Da ein Großteil der Vorarbeiten für die Berechnung der Raumpunkte bisher manuell durchgeführt wird und sich die Berechnung momentan auf einzelne Bildpaare beschränkt, liegt

hier ein weiterer Ansatzpunkt für eine Optimierung des Verfahrens. Es wird eine Schnittstelle zwischen Impresario und den einzelnen Kameras der Bumblebee benötigt, um einen Zugriff auf die Videostreams zu bekommen. Anschließend kann man sich überlegen, ob man den kompletten Stream für die Berechnung verwenden möchte oder ob in einem definierten Zeitintervall Bildpaare aufgenommen werden. Hiermit könnten dann auch automatisch die SIFT-Punktlisten zu den eingehenden Bildern erstellt werden. Eventuell muss das C++ Programm und die Übertragung der Daten über Dateien überarbeitet werden. Dies ist besonders wichtig, wenn man mit Videostreams arbeiten will, da dort ständig auf die Dateien zugegriffen werden müsste.

Abschließend soll noch einmal ein Blick auf ein mögliches Szenario geworfen werden, welches die Erstellung eines aussagekräftigen Tiefenbildes ermöglichen könnte. Der Ansatz basiert hier auf einer Kombination des SIFT-Algorithmus mit einem extra Kantendetektor. Der SIFT-Algorithmus erkennt beispielsweise einen Punkt an einem Schrank. Als nächstes wird dessen Entfernung zur Kamera berechnet. Anschließend erkennt der Kantendetektor den Umriss des Schranks. Alle Punkte innerhalb des Rahmens könnten auf die gleiche Entfernung gesetzt werden. Ergänzend könnte auch noch ein Eckendetektor eingesetzt werden, der die richtige Zuordnung der Kanten zueinander übernimmt beziehungsweise mögliche Lücken zwischen diesen schließt. Dies ist jedoch nur ein theoretischer Ansatz und müsste sowohl in Hinsicht auf die benötigten Kanten- und Eckenfilter, als auch auf das daraus entwickelte Tiefenbild überprüft werden. Ein weiteres Problem tritt im Bereich von schrägen Flächen auf. Hier müsste ebenfalls geprüft werden, wie diese erkannt und die Entfernungen berechnet und markiert werden könnten. Ein solches Verfahren hätte vermutlich einen deutlichen Einfluss auf die Berechnungsdauer. Vermutlich würde es jedoch auch zu einer erheblich besseren Abdeckung der Umgebung mit Entfernungsinformationen führen.

Literaturverzeichnis

- [Davies 2005] DAVIES, E. R.: *Machine Vision*. 3. Auflage. Amsterdam : Elsevier, 2005. – ISBN 978-3527405848
- [Duif 2006] DUIF, Thomas: *Entwicklung von Bildverarbeitungsalgorithmen mit Impresario*. 1 2006. – URL http://www.rz.rwth-aachen.de/global/show_document.asp?id=aaaaaaaaaavhxq. – Abruf: 2009-07-08
- [EOS] EOS: *Photomodeler*. – URL <http://www.photomodeler.com/index.htm>. – Abruf: 2009-06-30
- [Erhardt 2008] ERHARDT, Angelika: *Einführung in die Digitale Bildverarbeitung*. Wiesbaden : Vieweg+Teubner, 2008. – ISBN 978-3519004783
- [Eschenburg 2006] ESCHENBURG, Jonas: *Optisches Kameratracking anhand natürlicher Merkmale*, Universität Augsburg, Diplomarbeit, 2006. – URL http://www.informatik.uni-augsburg.de/lehrstuehle/mml/publications/reports/2006_13_esc/tr2006-13.pdf. – Abruf: 2009-06-30
- [Fusiello u. a. 2000] FUSIELLO, A. ; ROBERTO, V. ; TRUCCO, E.: Symmetric stereo with multiple windowing. In: *International Journal of Pattern Recognition and Artificial Intelligence* 14 (2000), S. 1053–1066
- [GIK 2006] GIK, Geodätisches I.: *Vermessungssysteme für Aufgaben der Industrie*. 3 2006. – URL <http://www.gik.uni-karlsruhe.de/361.html?&P=>. – Abruf: 2009-06-30
- [Gremse 2005] GREMSE, Felix: *Skaleninvariante Merkmalstransformation - SIFT Merkmale*. 8 2005. – URL http://phobos.imib.rwth-aachen.de/lehmann/seminare/bv_2005-08.pdf. – Abruf: 2009-06-30
- [Hartley und Zissermann 2003] HARTLEY, Richard ; ZISSERMANN, Andrew: *Multiple View Geometry in computer vision*. 2. Auflage. Cambridge University Press, 2003. – ISBN 0-521-54051-8

- [Hering 2008] HERING, Nils: *Bildverstehen mittels selbstähnlichen SIFT-Features*, Uni Koblenz, Diplomarbeit, 2008. – URL <http://www.uni-koblenz.de/~nilshering/sift.pdf>. – Abruf: 2009-07-08
- [Hofmann 2008/09] HOFMANN, Tim: *Geometrische Kamerakalibrierung*. 2008/09. – URL <http://www.mi.fh-wiesbaden.de/~schwan/Projects/CG/CarreraCV/doku/intrinsisch/intrinsisch.htm>. – Abruf: 2009-06-30
- [Hornberg 2006] HORNBERG, Alexander (Hrsg.): *Handbook of Machine Vision*. Weinheim : Wiley-VCH, 2006. – ISBN 978-3527405848
- [Jähne 1997] JÄHNE, Bernd: *Digitale Bildverarbeitung*. 4. Auflage. Berlin : Springer-Verlag, 1997. – ISBN 3-540-61379-x
- [Jiang und Bunke 2007] JIANG, Xiaoyi ; BUNKE, Horst: *Dreidimensionales Computersehen: Gewinnung und Analyse von Tiefenbildern*. Berlin : Springer, 2007. – ISBN 3-540-60797-8
- [Jiang u. a. 2007] JIANG, Xiaoyi ; ROTHUS, Kai ; WACHENFELD, Steffen: *Computer Vision*. 2007. – URL <http://cvpr.uni-muenster.de/teaching/ss07/computerVisionSS07/script/CV11-2.pdf>. – Abruf: 2009-07-29
- [Lange 2008] LANGE, Emmanuel: *Verfolgung markanter Raumpunkte in Videobildsequenzen anhand skalierungs- und rotationsinvarianter Merkmale*, HAW Hamburg, Diplomarbeit, 2008. – URL http://opus.haw-hamburg.de/volltexte/2008/611/pdf/bachelorarbeit_emmanuel_lange.pdf. – Abruf: 2009-07-08
- [Liebske 2005] LIEBSKE, Roland: *Statischer und dynamischer Kollisionsschutz mit Hilfe von Laser-Distanz-Sensoren*. 2005. – URL http://www.idm-sensoren.de/NR/rdonlyres/4B4BB273-141D-4C5B-AF70-668078D7CC0A/0/IDM869_04_07_TD_Kollisionsschutz_AKL_PLE_DE.pdf. – Abruf: 2009-06-30
- [Lowe 2004] LOWE, David G.: Distinctive image features from scale-invariant keypoints. In: *International Journal of Computer Vision* 60 (2004), Nr. 2, S. 91–110
- [LTI 2005] LTI: *LTI-Lib und Impresario*. 11 2005. – URL <http://ltilib.sourceforge.net/doc/homepage/index.shtml>. – Abruf: 2009-06-30
- [Meisel 2008] MEISEL, Andreas: *Anwendungen 3D-Bildverarbeitung*. 3 2008. – URL http://www.informatik.haw-hamburg.de/uploads/media/AW_3DBV_V02.pdf. – Abruf: 2009-07-08
- [Mildner 2008] MILDNER, Philip: *Suche nach korrespondierenden Pixeln in zwei Bildern*. 2008. – URL http://www.informatik.uni-mannheim.de/pi4.data/content/courses/2008-fss/mm_seminar/2008_Seminar_Korrespondenzen.pdf. – Abruf: 2009-06-30

- [Mircheva 2004] MIRCHEVA, Aneliya: *Kamerakalibration*. 2004. – URL http://www.ipr.ira.uka.de/~megi/SEMINAR/SS_04/SEMINAR/S-AneliaMircheva-Kamerakalib.pdf. – Abruf: 2009-07-08
- [Nef 2002] NEF, Marco: *Entwicklung eines Systems zur Simulation eines multi-Kamera-basierten Gesichtsscanners*, ETH Zürich, Departement Informatik, Diplomarbeit, 2002. – URL <http://www.shima.ch/papers/GDV2002/>. – Abruf: 2009-07-11
- [PointGrey] POINTGREY: *Bumblebee*. – URL <http://www.ptgrey.com/index.asp>. – Abruf: 2009-06-30
- [Schatz 2006] SCHATZ, Alexej: *Visuelle Navigation mit Scale Invariant Feature Transform*, Universität Bielefeld, Diplomarbeit, 2006. – URL http://www.ti.uni-bielefeld.de/downloads/publications/diplomatheses/dal5_alexej_schatz_visuelle_nav_sift.pdf. – Abruf: 2009-06-30
- [Suthau 2002] SUTHAU, Tim: *Photogrammetrische Vermessung von superelastischen Drähten mit CCD-Kameras*, TU Berlin, Diplomarbeit, 2002. – URL <http://www.fpk.tu-berlin.de/~tim/diplom/diplom.phtml>. – Abruf: 2009-06-30
- [Systems 2007] SYSTEMS, EOS: *PhotoModeler 6 User Guide*. 1. 210 - 1847 West Broadway Vancouver, British Columbia Canada V6J 1Y6: EOS Systems Inc. (Veranst.), Januar 2007
- [Vedaldi 2006] VEDALDI, Andrea: *A lightweight C++ implementation of SIFT*. 2006. – URL <http://www.vlfeat.org/~vedaldi/code/siftpp.html>. – Abruf: 2009-06-30

A. Homogene Koordinaten

In der linearen Algebra erfolgt die Darstellung von dreidimensionalen Daten durch dreidimensionale Vektoren und Transformationsmatrizen. Im Gegensatz hierzu verwendet die Computergrafik meist die Darstellung in homogenen Koordinaten. Hierbei werden die Vektoren und Matrizen um eine so genannte homogene Dimension erweitert. Im Normalfall geschieht dies durch das Hinzufügen einer 1 in der vierten Spalte. Homogene Koordinaten werden also durch einen Spaltenvektor mit vier Komponenten dargestellt:

$$\mathbf{x} = [t_x, t_y, t_z, t]^T.$$

Um die ursprünglichen 3D-Koordinaten zu erhalten, werden die ersten drei Komponenten durch die vierte Komponente dividiert (vergleiche Jähne 1997, S. 196 ff).

Mit der homogenen Darstellung lassen sich verschiedene Transformationen in einer einzigen 4x4 Matrix abbilden. Sie kommt häufig bei der Umrechnung von Koordinatensystemen zum Einsatz.

Die drei wichtigsten Transformationsmatrizen beschreiben Translationen, Skalierungen und Rotationen (Nef 2002, S.25 ff.). Soll ein homogener Vektor $\mathbf{v} = (x, y, z, 1)^T$ um $(t_x, t_y, t_z)^T$ verschoben werden, so kann man ihn mit der Translationsmatrix

$$\mathbf{T}(t_x, t_y, t_z) = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

multiplizieren. Der resultierende Vektor erhält die Form

$$\mathbf{v}' = (x + t_x, y + t_y, z + t_z, 1)^T,$$

was genau dem gewünschten Resultat entspricht. Ähnlich sieht die Skalierungsmatrix aus, welche die Komponenten des homogenen Vektors \mathbf{v} um die reellen Faktoren $(s_x, s_y, s_z)^T$ skaliert:

$$\mathbf{S}(s_x, s_y, s_z) = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Für die Darstellung von Rotationen um einen Winkel Θ existiert für jede Koordinatenachse eine gesonderte Variante einer homogenen Rotationsmatrix, welche jeweils einen Vektor um den gegebenen Winkel um die gefragte Achse rotiert.

$$\mathbf{R}_x(\Theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \Theta & -\sin \Theta & 0 \\ 0 & \sin \Theta & \cos \Theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ für die x-Achse}$$

$$\mathbf{R}_y(\Theta) = \begin{bmatrix} \cos \Theta & 0 & \sin \Theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \Theta & 0 & \cos \Theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ für die y-Achse}$$

$$\mathbf{R}_z(\Theta) = \begin{bmatrix} \cos \Theta & -\sin \Theta & 0 & 0 \\ \sin \Theta & \cos \Theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ für die z-Achse}$$

B. Impresario

Impresario (Image Processing and Pattern Recognition Engineering System) ist ein grafisches Werkzeug für das schnelle Erstellen und Testen von einfachen Bildverarbeitungsprogrammen. Der Aufbau auf Templates ermöglicht das einfache Wiederverwenden von einmal geschriebenem Programmcode und Funktionen. Die Erstellung von Projekten erfolgt per Drag and Drop der vorhandenen Templates auf die Arbeitsfläche. Es wurde wie die LTI-Lib (vergleiche Anhang C) an der RWTH Aachen (LTI 2005) entwickelt und basiert in Teilen auf der LTI-Lib. Das Programm steht nicht mehr zum Download bereit. Eine Arbeit die an der RWTH Aachen im Zuge eines Seminars durchgeführt wurde (Duif 2006), beschreibt, wenn auch nicht als Schwerpunkt, einige grundlegende Bestandteile von Impresario, inklusive dem Erstellen von Makros. Hiermit kann einem eventuell der Einstieg in die Software erleichtert werden.

C. LTI-Library

Die LTI-Lib (LTI = Lehrstuhl für Technische Informatik) wurde am Lehrstuhl für Technische Informatik der Rheinisch-Westfälischen Technischen Hochschule Aachen (RWTH Aachen) entwickelt (LTI 2005) und umfasst eine Reihe von Algorithmen und Datenstrukturen wie sie häufig im Bereich der Bildverarbeitung zum Einsatz kommen. Sie ist das Ergebnis jahrelanger Forschung im Bereich der Computervisualisierung wie z.B. Objekterkennung, Robotersicht und Gestenerkennung. Die Programmierung erfolgte komplett in C++, was ein effizientes Arbeiten unter einem objektorientierten Ansatz ermöglicht. Die weitere Entwicklung der Bibliothek wurde mittlerweile eingestellt.

D. Inhalt der CD-ROM

Dieser Arbeit liegt eine CD-ROM mit folgender Verzeichnisstruktur bei:

- **[Ausarbeitung]** beinhaltet diese Arbeit im PDF-Format.
- **[Literatur]** beinhaltet die in dieser Arbeit verwendete Literatur, soweit diese in elektronischer Form vorliegt.
- **[Quellcode]** beinhaltet den Quellcode für das Impresario-Modul zur Erstellung der SIFT-Punktlisten sowie für das C++ Programm zur Berechnung der Raumpunkte.

Versicherung über Selbstständigkeit

Hiermit versichere ich, dass ich die vorliegende Arbeit im Sinne der Prüfungsordnung nach §24(5) ohne fremde Hilfe selbstständig verfasst und nur die angegebenen Hilfsmittel benutzt habe.

Hamburg, 23. August 2009

Ort, Datum

Unterschrift