

Institut für Parallele und Verteilte Systeme
Universitätsstraße 38
D-70569 Stuttgart

Diplomarbeit Nr. 3109

Korrespondenzbasierte Distanzbestimmung von 3D Scanpunkten

Aida Gusinjac

Studiengang: Informatik

Prüfer: PD. Dr. rer. nat. Michael Schanz

Betreuer: Dipl.-Inf. Alexander Kramer,
PD. Dr. rer. nat. Michael Schanz

begonnen am: 29.11.2010

beendet am: 31.05.2011

CR-Klassifikation: I.2.9, I.2.10, I.4.1, I.4.8

Inhaltsverzeichnis

1 Einleitung und Motivation	1
1.1 Motivation	1
1.2 Aufgabenstellung	2
1.3 Gliederung	3
2 Grundlagen und Begriffe	4
2.1 Lasercannen	4
2.2 Photogrammetrie	7
2.3 Triangulationen	11
2.4 Grundlagen dreidimensionaler Darstellungen	13
2.4.1 Koordinatensysteme	13
2.4.2 Homogene Koordinaten im 3D Raum	16
2.4.3 Koordinatentransformationen im 3D Raum	19
2.4.4 Vergleich zwischen Modell- und Weltkoordinaten	25
3 Epipolargeometrie	30
3.1 Achsparalleles Stereogeometrie	30
3.2 Die allgemeine Stereogeometrie	32
3.3 Beziehungen in Kamerakoordinaten	34
3.4 Bestimmung der Epipolarlinien	35
3.5 Beziehungen in Bildkoordinaten	37
3.6 Essential- und Fundamental-Matrix	38
4 Korrespondenzanalyse	39
4.1 Pixelbasierte Verfahren	40
4.2 Merkmalsbasierte Verfahrenen	44

5 Vorverarbeitung	46
5.1 Problembetrachtung	47
5.2 Methoden für Entfernungs berechnung	51
6 Block-Matching	54
6.1 Block-Matching Verfahren	54
6.2 Block-Matching Verfahren für Farbstereoanalyse	58
7 Zusammenfassung und Ausblick	60
Abbildungsverzeichnis	61
Literaturverzeichnis	62

1 Einleitung und Motivation

1.1 Motivation

Heutzutage spielen Robotik und Computer Vision, und deren Anwendung, eine wesentliche Rolle in vielen Teilen unseres Lebens. Ein wichtiger Teil von Robotik und Computer Vision ist die Bildverarbeitung beziehungsweise die 3D Umgebung. Die Erzeugung von präzisen 3D Modellen ist die Aufgabe von 3D Laserscannern. Mit einer Fotokamera können 2D Bilder erzeugt werden. Der FARO 3D Laserscanner erzeugt ein 3D Bild von einer komplexen Umgebung. Diese Umgebung besteht aus verschiedene geometrischen Objekten.

Der Laserscanner berechnet von seiner Position aus alle Objekte bis zu einer bestimmten Entfernung. Dazu wird die Umgebung zuerst als eine Menge von Punkten erfasst und zu Punktfolgen zusammengefasst. Letztendlich werden diese von einer Software interpretiert und als ein geometrisches Objekt dargestellt. Normalerweise kann der heutige 3D Laserscanner Objekte bis zu einer bestimmten Entfernung berechnen. Über Objekte die außerhalb dieser Entfernung liegen werden keine Information gespeichert.

Jetzt stellt sich die Frage, ob es möglich wäre, mit Hilfe von vorhandenen Daten, die Entfernung der weit entfernten Objekte zu berechnen. Um dieses Problem zu lösen werden mehrere Aufnahmen der gleichen Umgebung aus verschiedenen Positionen erstellt. Durch diese unterschiedlichen Betrachtungswinkel kann man mit Hilfe von Software die Entfernung zu weit entfernt liegenden Objekten berechnen(sehen Sie z.B.[1]).

1.2 Aufgabenstellung

In dieser Diplomarbeit sollen die Daten der Laserscanner, die aus unterschiedlichen Scanpositionen aufgenommen wurden, verifiziert oder ergänzt werden. Die zu untersuchenden Bereiche werden vorab durch den speziellen Marker markiert.

Betrachtet man nun zwei Scanpunktmenge A und B, die von unterschiedlichen Standorten aufgenommen wurden, so können folgende Szenarien entstehen:

- Ein Scanpunkt mit vorhandener Distanzinformation in A enthält nach der Projektion in B ebenfalls eine Distanz. Hier soll eine Gegenprüfung der beiden Messungen stattfinden.
- Ein Scanpunkt mit vorhandener Distanzinformation in A enthält nach der Projektion in B keine Distanz. In diesem Fall soll die Distanz in B berechnet werden.
- Ein Scanpunkt hat keine Distanzinformation in A und die Distanz konnte durch vorhergehende Schritte nicht bestimmt werden. Hier soll zuerst eine pixelweise Korrespondenz gefunden bzw. überprüft werden. Danach soll die Distanz in A und B berechnet werden. Im ersten Ansatz kann davon ausgegangen werden, dass die Marker exakt platziert wurden. Als nächstes soll die ungenaue Markerposition durch pixelweise Korrespondenzfindung kompensiert werden.

1.3 Gliederung

Die Arbeit ist in folgende Abschnitte gegliedert:

Kapitel 2 beschreibt die grundlegenden Konzepte, die in dieser Arbeit verwendet werden.

Kapitel 3 stellt ein geometrisches Modell vor, das die geometrischen Beziehungen zwischen verschiedenen Kamerabildern des gleichen Objekts aus verschiedene Winkelpositionen darstellt. Mit Hilfe dieser Beziehungen lässt sich die Abhängigkeit zwischen korrespondierenden Bildpunkten beschreiben.

Kapitel 4 beschreibt die Analyse zweier Ansichten eines Stereokamerasystems. Das Ziel besteht darin, korrespondierende Bildpunkte oder Bildmerkmale in zwei unterschiedlichen Bildern zu finden. Die Korrespondenzanalyse ist ein klassisches Aufgabefeld der Bildanalyse. Es gibt viele Verfahren für die Korrespondenzanalyse, wir werden aber nur die Pixel- und die Merkmalsbasierten Verfahren.

Kapitel 5 beschreibt die Problematik und Lösungswege für das vorgestellte Problem.

Kapitel 6 beschreibt ein Block Matching Verfahren.

Kapitel 7 fasst das Themengebiet nochmals zusammen.

2 Grundlagen und Begriffe

2.1 Lasercannes



Abbildung 2.1: FARO Laser Scanner Photon

Laser steht für 'light amplification by stimulated emission of radiation' und beschreibt den Vorgang der Lichtverstärkung durch stimuliert Emission. Dieses Ereignis führt zur Entstehung der Laserstrahlung. Anfang des 20. Jahrhunderts hat Albert Einstein die Idee der stimulierten Emission eingeführt, und 1960 hat Theodore Maiman den ersten funktionsfähigen Laser realisiert. Im Gegensatz zu anderen Lichtquellen, wie Glühlampen oder der Sonne, die Licht in alle Raumrichtungen mit relativ unbestimmten Frequenzen aussenden, emittiert der Laser einen gut gebündelten Strahl mit hoher Frequenzschärfe. Allgemein zeichnet sich Laserlicht durch folgende Eigenschaften aus (sehen Sie z.B. [2]):

- geringe spektrale Linienbreite
- starke Bündelung
- hohe Strahlintensität(-Energie)
- Eignung zur Erzeugung ultrakurzer Lichtimpulse

Laserscanning(sehen Sie z.B.[3])(auch Laserabtastung) bezeichnet das zeilen- oder rasterartige Überstreichen von Oberflächen oder Körpern mit einem Laserstrahl, um diese zu vermessen,zu bearbeiten, oder um ein Bild zu erzeugen. Geräte, die den Laserstrahl entsprechend ablenken, heißen Laserscanner. Ein Laserscanner besteht aus einem Scankopf und einer Treiber- und Ansteuerelektronik. Die Elektronik besteht aus einem leistungselektronischen Teil, der die Ströme für die Antriebe liefert, und aus einer eingebetteten Scannersoftware.

Bei Messanwendungen wird das Ergebnis des Scanvorgangs von Sensoren über den gleichen oder einem getrennten optischen Weg empfangen und von der Scannersoftware erfasst, die auch die anderen Komponenten anspricht und kontrolliert.

Beim 3D-Laserscanning wird die Oberflächengeometrie von Gegenständen mittels Pulslaufzeit, Phasendifferenz im Vergleich zu einer Referenz, oder durch Triangulation von Laserstrahlen digital erfasst. Dabei entsteht eine diskrete Menge von dreidimensionalen Abtastpunkten, die als Punktwolke bezeichnet wird. Die Koordinaten der gemessenen Punkte werden aus den Winkeln und der Entfernung in Bezug zum Ursprung (Gerätestandort) ermittelt.

Anhand der Punktwolke werden entweder Einzelmaßse wie z.B. Längen und Winkel bestimmt, oder es wird eine geschlossene Oberfläche aus Dreiecken konstruiert. Moderne Lasermesssysteme erreichen eine Punktgenauigkeit von bis zu 1 mm.

Im Bereich der Laserscanner, die nach dem Phasendifferenzverfahren (Phasenmessverfahren) arbeiten, wurden in den letzten Jahren enorme Fortschritte insbesondere in Hinblick auf die Abtastrate erzielt. So erreichen aktuelle Geräte Messgeschwindigkeiten von über 1 000 000 3D-Messpunkten (1MHz) pro Sekunde. Dabei wird, im Gegensatz zum Impulslaufzeitmessverfahren, ein kontinuierlicher Laserstrahl ausgesandt. Die Amplitude des ausgesandten Laserstrahls wird mit mehreren sinusförmigen Wellen unter-

schiedlicher Wellenlänge moduliert. Der entstehende zeitliche Abstand des empfangenen Signals gegenüber dem gesendeten Signal ist eine Folge der Entfernung zum Objekt. Bei gleichzeitiger Betrachtung der Phasenlage des gesendeten und des empfangenen Signals ergibt sich eine Phasendifferenz, die die Bestimmung des Objektabstandes erlaubt.

Der Einsatz des Laserscanning:

- Architekturvermessung mit Schwerpunkten in der Bauforschung und Denkmalpflege
- Bei möblierten Gebäude sind die Scanergebnisse de facto nicht auswertbar.
- andere Gebiete z.B. Rohrleitungsbau, Anlagenbau, Archäologie, Denkmalschutz, Tunnelbau, Forensik und Unfallforschung.

2.2 Photogrammetrie

Photogrammetrische Verfahren stützen sich auf photographisch erzeugte Abbildungen, Fotos oder Videobilder. Dadurch soll die Lage, Größe und Form von Objekten bestimmt werden. Die Messungen werden nicht direkt am Objekt, sondern indirekt aus deren Abbildern im Foto vorgenommen. Die Erfassung von geometrischen Objekten durch ein photogrammetrisches Verfahren besteht aus zwei Stufen. In der ersten Stufe ist das Foto vor Ort in der Art einer originären Erfassung herzustellen; d.h. das Objekt soll fotografiert werden. In der zweiten Stufe werden die interessante Objektmerkmale aus dem Foto im Sinne einer sekundären Erfassung bestimmt. Das Ableiten von Geometrieeinformati onen aus den Fotos setzt voraus, dass die Abbildungsvorschrift sehr genau bekannt ist. In der Regel werden daher Kameras benutzt, die annähernd eine ideale zentralperspektivische Abbildung gewährleisten. Sind die Abbildungseigenschaften einer Kamera über mehr oder minder lange Zeit konstant oder können diese auf einfache Weise, z. B. durch eine Resauplatte, aus den Fotos rekonstruiert werden, so spricht man von Messkammern bzw. Teilmesskammern. Aber auch aus Photos, die mit einer Amateurkamera mit unbekannten Abbildungseigenschaften aufgenommen wurden, sind geometrische Informationen bestimmbar, wenn zugleich die Abbildungseigenschaft mitbestimmt wird. Hierzu werden Referenzobjekte genutzt, deren abgebildete Form und Größe mit ihrer Sollgröße mathematisch in Verbindung gebracht werden .

Das Fotografieren ist ein physikalischer Vorgang, sodass Fotos die darauf abgebildeten Objekte objektiv dokumentieren. Ein Foto stellt ein Informationsträger sehr hoher Dichte dar und kann zu beliebiger Zeit und mehrmals unter verschiedenen Aspekten ausgewertet werden. Der Informationsträger Foto hält das vom Objekt zum Zeitpunkt des Fotografierens reflektierte Umgebungslicht fest. Es werden somit ausschließlich radiometrische Informationen in einem wohldefinierten Bildkoordinatensystem gespeichert. Der fotografische Messvorgang ist dem menschlichen Sehen nachempfunden, sodass die abgebildete Farb- und Helligkeitsinformation unmittelbar durch den Betrachter interpretiert werden können(sehen Sie z.B. [4]).

Die Photogrammetrie hat folgende spezielle Vorteile bei der geometrischen Erfassung:

- Das Fotografieren benötigt nur einen Bruchteil von Sekunden, sodass auch von beweglichen Plattformen aus fotografiert werden kann oder das Messobjekt sich bewegen darf.
- Die Fotos selbst sind bereits Endprodukte und eignen sich zur Dokumentation.
- Zur Aufnahme schwierig zu erreichender Abschnitte können Hubgeräte, Fesselballons, (Modell-) Hubschrauber etc. eingesetzt werden.
- Die zeitliche Trennung zwischen Aufnahme und Auswertung erlaubt die Vermessung von zerstörten oder bereits untergegangenen Gebäuden. Dieser Aspekte der Rekonstruktion von Gebäuden ist für die Denkmalpflege äußerst interessant.
- Moderne Digitalkameras und analytische Photogrammetrie-Auswerteprogramme erlauben eine preiswerte Erfassung und Auswertung.
- Das berührungslose Vermessen ist auch möglich wenn das Messobjekt: schwer zugänglich, sehr kompliziert strukturiert, sehr heiß, sehr weich, sehr empfindlich, sehr giftig oder radioaktiv ist.

Die größten Schwächen der photogrammetrischen Verfahren liegen darin begründet, dass

- für eine ausreichende Beleuchtung des Objektes zu sorgen ist, insbesondere im Gebäudeinneren,
- durch die Abbildung die Tiefeninformation der Objektgeometrie verloren geht und
- das Photo ein im Maßstab verkleinertes Zwischen-Modell auf dem Weg zum dreidimensionalen bzw. maßstabsgetreuen Modell darstellt.

Die erste Schwächen kann mittels zusätzlicher aktiver Beleuchtung durch Blitzgeräte überwunden werden. Die zweite führt dazu, dass zur Gewinnung von dreidimensionalen Informationen mindestens zwei Fotos erforderlich sind, auf denen jeweils das auszumesende Objekt abgebildet ist, und die von unterschiedlichen Aufnahmestandorten stammen müssen. Sind solche zwei sich Überdeckenden Fotos vorhanden und sind deren Aufnahmestandorte und Aufnahmerichtungen bekannt, so kann über ein modifiziertes dreidimensionales Vorwärtseinschneiden im Angesicht der Fotos dreidimensional gemessen werden. Durch dieses Verfahren (Beobachten, Analysieren und Dokumentieren) können

die Objekte virtuell nachempfunden werden. Die photogrammetrische Aufnahme ist ein relativ einfaches, preiswertes und hochgenaues Verfahren, mit dem ein detailreiches und schnelles Modell hergestellt werden kann. Der Modellmaßstab ist für Aufgaben der Bauaufnahme deutlich kleiner als der Maßstab des zu erstellenden Zielmodells. Während der, zum Teil recht aufwendigen, Auswertung ist damit eine Vergrößerung der originären Messinformation nötig. Im Vergleich zu anderen (direkten) Erfassungsverfahren wird das Endprodukt aus einem stark verkleinerten Zwischenmodell abgeleitet. Hierdurch ergeben sich theoretisch Genauigkeitsverluste. Um diese zu minimieren, muss der Bildmaßstab der Photos vergrößert und die Qualität der Messkammer, im Bezug zur Präzision der inneren Orientierung, erhöht werden. Die Detailerkennbarkeit und Genauigkeit der photogrammetrischen Auswertung ist wesentlich von der photographischen Auflösung und dem Bildmaßstab abhängig.

Die Photogrammetrie kann nicht zur Erfassung von Objekten eingesetzt werden bei denen die Oberfläche ohne Struktur ist, die sich im Abbild abzeichnen können, oder wenn sie durchsichtig sind. Andere Erfassungsmethoden wie z. B. das Laserscanning haben hier keine Probleme. Oftmals wird hervorgehoben, dass mittels Photogrammetrie flächenhaft vermessen werden kann. Dies gilt nur dann wenn die Objektoberfläche an jeder Stelle abbildbare Strukturen z.B. Farbänderungen aufweist.

Der größte Vorteil der Photogrammetrie ist der der berührungslosen Messung. So kann z.B. ein Museumsstück oder eine Skulptur, die nicht berührt werden sollte, anhand von markanten Punkten problemlos vermessen werden. Das zu vermessende Objekt kann zu jedem Zeitpunkt festgehalten werden. Der Zeitbedarf bleibt dabei sehr gering, da ein Bild innerhalb kürzester Zeit aufgenommen wird. Außerdem kann die Auswertung der Bilder örtlich und zeitlich getrennt von der Aufnahme erfolgen. Die Qualität und Genauigkeit der Aufnahmen ist sehr flexibel und kann je nach Bedarf variiert werden. Bei der digitalen Photogrammetrie ist die Speicherung sehr effektiv und geometrisch stabil. Detaillierte Informationen sind jederzeit extrahierbar und können schließlich vielfältig analysiert werden.

Es gibt jedoch auch einige Nachteile. Auf den Fotografien sind nur Oberflächendaten zu sehen. Eine notwendige Voraussetzung ist also, dass das Objekt abbildbar ist. Wenn dies der Fall ist, kann die Auswertung durch zufällige Verdeckungen, wie fallende Blätter, fliegende Vögel, oder auch Regen behindert werden. Bei der Luftbildphotogram-

metrie sind schon kleinste Wolken problematisch. Ein weiteres Problem besteht in fehlenden Kontrasten, bedingt durch Schnee oder fehlerhafte Beleuchtung. Je nach Wetterlage oder Beleuchtung im Raum ist indirekte Beleuchtung erforderlich. Außerdem sind für dreidimensionale Erfassungen mindestens zwei Aufnahmen von unterschiedlichen Standorten mit dem gleichen Bildinhalt erforderlich, sodass oft eine Flut von Bildmaterial entsteht(sehen Sie z.B. [5]).

2.3 Triangulationen

Das Grundprinzip der Triangulation ist in der Abbildung 2.2 vereinfacht für den zweidimensionalen Fall dargestellt. Von zwei verschiedenen Stationen an den Positionen S_1 (Laser) und S_2 (Kamera) wird der zu bestimmende Zielpunkt P (Objektpunkt) angepeilt. Dabei erhält man die beiden Winkel α und β mit der Genauigkeit $\Delta\alpha$ und $\Delta\beta$. Unter Kenntnis der Basislänge b kann man dann die Koordinaten von P relativ zum Koordinatenursprung bestimmen.

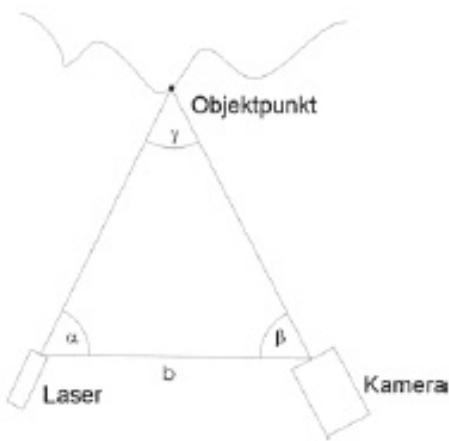


Abbildung 2.2: Triangulation

Vereinfacht könnte man auch sagen, dass von zwei Punkten auf einer Geraden, wobei der Abstand zwischen den beiden Punkten bekannt ist, Winkelmessungen zu beliebig anderen Punkten im Raum erfolgen, um deren Lage eindeutig zu bezeichnen.

Im dreidimensionalen Raum ist zu beachten, dass sich die zwei Sichtgeraden der beiden Basisstationen im Normalfall mathematisch nicht exakt schneiden, sondern windschief sind. Betrachtet man das lineare Gleichungssystem zur Bestimmung der Punktkoordinate P so ist dieses im zweidimensionalen Raum eindeutig lösbar, im dreidimensionalen Raum jedoch Überbestimmt und somit nicht mehr eindeutig lösbar. Das bedeutet, dass in den Gleichungen Zusatzinformationen enthalten sind, die man nutzen kann um beispielsweise den Messfehler von P aus dem Abstand der windschiefen Geraden zu schätzen.

Im Umkehrschluss ist es auch möglich die Anzahl der Beobachtungen zu reduzieren. Anstelle für beide Basisstationen jeweils zwei Richtungswinkel zu messen genügt es, dies nur bei einer Station zu tun. Bei der zweiten Station wird nur die Winkelkomponente in der Ebene die aus S_1 , S_2 und P aufgespannt wird bestimmt. Die graphische Lösung der Schnittgleichung enthält damit nicht mehr zwei Geraden sondern eine Ebene und eine Gerade und ist somit bei physikalisch sinnvollen Beobachtungen immer eindeutig lösbar. Eine technische Umsetzung dieses Prinzips findet man in der Streifenprojektion(sehen Sie z.B. [16]).

2.4 Grundlagen dreidimensionaler Darstellungen

In diesem Kapitel werden mathematische Grundlagen für 3D Umgebung beschrieben (sehen Sie z.B.[6]).

2.4.1 Koordinatensysteme

Zweidimensionales Polarkoordinatensystem

Ein zweidimensionales Polarkoordinatensystem besteht aus zwei Geraden, die orthogonal zueinander liegen (Abbildung 2.3). Ein Punkt in einem Polarkoordinatensystem wird durch ein zweier Tupel (r,θ) beschrieben, siehe Abbildung 2.4

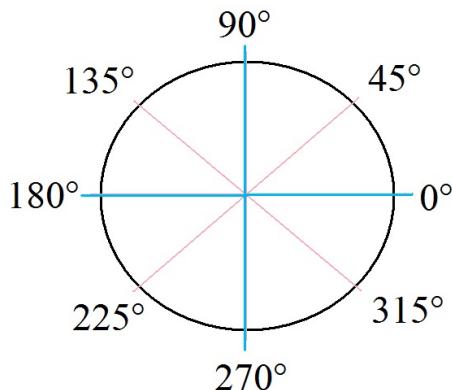


Abbildung 2.3: Polarkoordinatensystem

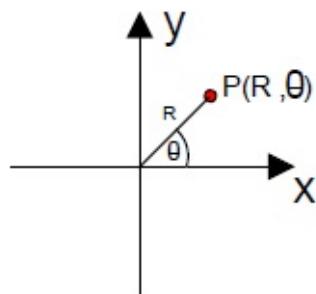


Abbildung 2.4: Punkt in einem Polarkoordinatensystem

Zweidimensionales kartesisches Koordinatensystem

Ein zweidimensionales kartesisches Koordinatensystem besteht aus zwei Achsen, die wie beim Polarkoordinatensystem orthogonal zueinander liegen. Ein Punkt in einem zweidimensionalen Koordinatensystem wird durch ein zweier Tupel (x, y) beschrieben, siehe Abbildung 2.5.

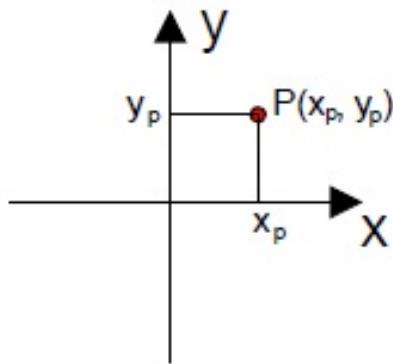


Abbildung 2.5: Punkt in einem zweidimensionalen kartesischen Koordinatensystem

Umrechnung von Polarkoordinaten in kartesischen Koordinaten (2D)

Um einen Punkt (r, θ) eines zweidimensionalen Polarkoordinatensystems in einen kartesischen Punkt (x, y) umzurechnen, werden folgende Formeln für die Berechnung des x und y Wertes genutzt:

$$x = r \cos(\theta) \quad (2.1)$$

$$y = r \sin(\theta) \quad (2.2)$$

Dreidimensionales kartesisches Koordinatensystem

Ein dreidimensionales Koordinatensystem besteht aus drei Achsen, die alle zueinander orthogonal liegen. Ein Punkt in einem dreidimensionalen Koordinatensystem wird durch ein dreier Tupel (x, y, z) beschrieben, siehe Abbildung 2.6.

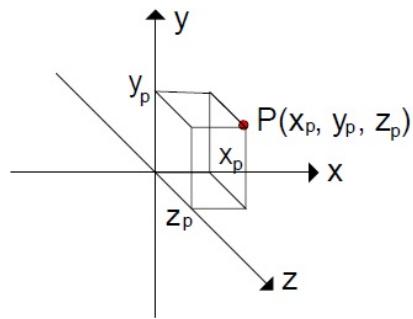


Abbildung 2.6: Punkt in einem dreidimensionalen Koordinatensystem

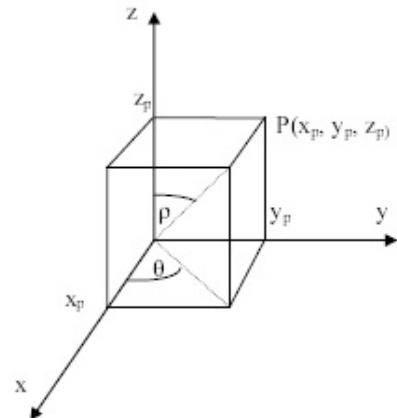
Umrechnung von Polarkoordinaten in kartesische Koordinaten (3D)

Abbildung 2.7: Punkt in einem dreidimensionalen Koordinatensystem

$$x = r \sin(\theta) \cos(\varphi) \quad (2.3)$$

$$y = r \sin(\theta) \sin(\varphi) \quad (2.4)$$

$$z = r \cos(\theta) \quad (2.5)$$

2.4.2 Homogene Koordinaten im 3D Raum

Objekte in der Ebene werden durch kartesische, Zylinder- oder Kugelkoordinaten dargestellt. Außerdem kann es auch mit **homogenen** Koordinaten beschreiben, diese haben einen großen Vorteil. Durch homogene Koordinaten kann man alle geometrischen Transformationen einheitlich durch Matrixmultiplikationen darstellen. Kartesische und homogene Koordinaten befinden sich in folgendem Zusammenhang:

- Ein Punkt $P = (x, y) R^2$ besitzt $[x, y, 1]$ als homogene Koordinaten.
- Ein Punkt $P = (x, y, z) R^3$ besitzt $[x, y, z, 1]$ als homogene Koordinaten.
- Ein zweidimensionaler Vektor wird als (3×3) -Matrix dargestellt.
- Ein dreidimensionaler Vektor wird als (4×4) -Matrix dargestellt.

Sowohl in zweidimensionalen Raum als auch in dreidimensionalen Raum kann man folgende Koordinatentransformationen einsetzen:

- Verschiebung von Objekten (Translation)
- Vergrößerung und Verkleinerung von Objekten (Skalierung)
- Drehung (bzw. Spiegelung an einem Punkt) von Objekten (Rotation)

Hiermit werden die Transformationen in dreidimensionalen Räumen beschrieben.

3D Translation um den Vektor (a,b,c)

$$P' = T \cdot P = \begin{bmatrix} x + a \\ y + b \\ z + c \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.6)$$

T = Translationmatrix

3D Skalierung mit den Faktoren S_x , S_y und S_z

$$P' = S \cdot P = \begin{bmatrix} S_x \cdot x \\ S_y \cdot y \\ S_z \cdot z \\ 1 \end{bmatrix} = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.7)$$

S = Skalierungsmatrix

3D Rotation um einen Winkel α

Wenn man in Richtung der positiven z-Achse blickt, bezeichnet man die Winkel im Gegenuhrzeigersinn als positiv.

Für eine Drehung um die z-Achse ergibt sich folgende Matrix:

$$P' = R_z \cdot P = \begin{bmatrix} x \cdot \cos(\alpha) + y \cdot \sin(\alpha) \\ y \cdot \cos(\alpha) - x \cdot \sin(\alpha) \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) & 0 & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.8)$$

R_z = Rotationmatrix

Entsprechend gilt für eine Drehung um die y-Achse

$$P' = R_y \cdot P = \begin{bmatrix} x \cdot \cos(\alpha) - z \cdot \sin(\alpha) \\ y \\ x \cdot \sin(\alpha) + z \cdot \cos(\alpha) \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.9)$$

R_y = Rotationmatrix

Eine Drehung um die x-Achse ergibt

$$P' = R_x \cdot P = \begin{bmatrix} x \\ y \cdot \cos(\alpha) + z \cdot \sin(\alpha) \\ -y \cdot \sin(\alpha) + z \cdot \cos(\alpha) \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) & 0 \\ 0 & -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.10)$$

R_x = Rotationmatrix

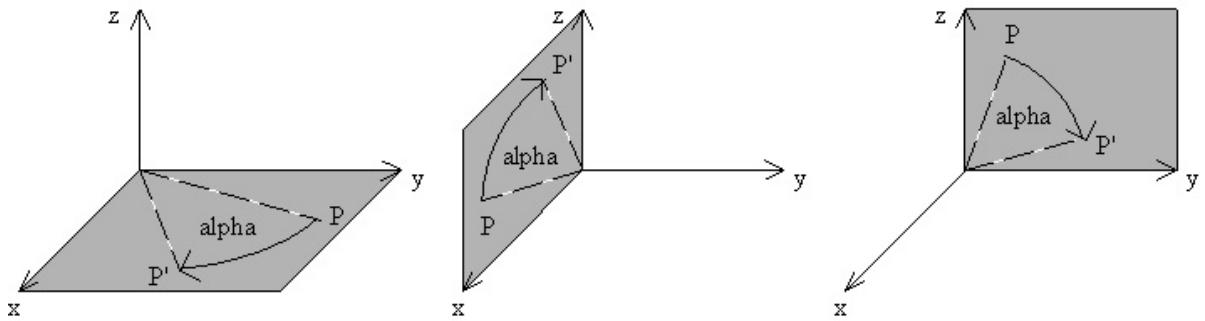


Abbildung 2.8: Drehung um die z-, y- und x-Achse im Gegenuhrzeigersinn

Spiegelungen

Spiegelungen sind Spezialfälle der Skalierung. Für eine Spiegelung am Nullpunkt des Koordinatensystems erhält man

$$P' = S_0 \cdot P = \begin{bmatrix} -x \\ -y \\ -z \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.11)$$

S_0 = Skalierungsmatrix am Nullpunkt

Für eine Spiegelung an der Achseebene $x = 0$ (also an der (y, z) -Ebene) erhält man

$$P' = S_x \cdot P = \begin{bmatrix} -x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.12)$$

S_x = Skalierungsmatrix an x-Achse

2.4.3 Koordinatentransformationen im 3D Raum

Transformation von Koordinatensystemen

Im vorhergehenden Abschnitt wurden Transformationen von Punkten bei festgehaltinem Koordinatensystem betrachtet. Oft ist es jedoch auch notwendig, die Koordinaten desselben Punktes in unterschiedlichen Koordinatensystemen zu ermitteln. Wir benötigen daher auch Transformationen von Koordinatensystemen.

Translation eines Koordinatensystems

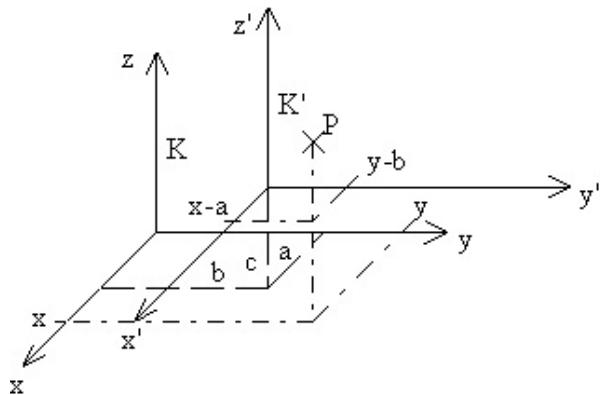


Abbildung 2.9: Translation eines Koordinatensystems

Zu einem gegebenen Punkt P mit den Koordinaten (x, y, z) in einem Koordinatensystem K sind die Koordinaten (x', y', z') desselben Punktes P bezüglich eines zweiten

Koordinatensystems K' gesucht. K' sei gegenüber K um den Vektor (a, b, c) verschoben.

$$P_{K'} = T \cdot P_K = \begin{bmatrix} x - a \\ y - b \\ z - c \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -a \\ 0 & 1 & 0 & -b \\ 0 & 0 & 1 & -c \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.13)$$

Eine Translation des Koordinatensystems um den Vektor (a, b, c) ist äquivalent zu einer Objekttranslation um den Vektor ($-a, -b, -c$).

Rotation des Koordinatensystems

Zu einem gegebenen Punkt P mit den Koordinaten (x, y, z) in einem Koordinatensystem K sind die Koordinaten (x', y', z') desselben Punktes P bezüglich eines zweiten Koordinatensystems K' gesucht, das gegenüber K um den Winkel α gedreht ist.

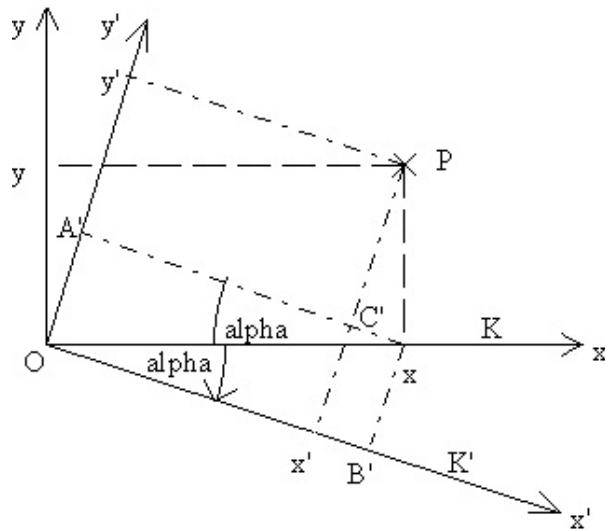


Abbildung 2.10: Rotation eines Koordinatensystems

Mit den Beziehungen für das

$$\text{Dreieck } OB'x : \quad \sin \alpha = \frac{\overline{B'x}}{x} = \frac{\overline{OA'}}{x} \text{ und } \cos \alpha = \frac{\overline{OB'}}{x} \quad (2.14)$$

$$\text{Dreieck } PC'x : \quad \sin \alpha = \frac{\overline{C'x}}{Px} = \frac{\overline{B'x'}}{y} \text{ und } \cos \alpha = \frac{\overline{PC'}}{Px} = \frac{\overline{A'y}}{y} \quad (2.15)$$

ergibt sich nach Abbildung 2.10 für eine Drehung um die z-Achse:

$$x' = \overline{OB'} - \overline{B'x'} = x \cos \alpha - y \sin \alpha \quad (2.16)$$

$$y' = \overline{OA'} - \overline{A'y'} = x \sin \alpha + y \cos \alpha \quad (2.17)$$

Dies lässt sich durch die folgende Rotation beschreiben:

$$P_{K'} = R \cdot P_K = \begin{bmatrix} x \cdot \cos(\alpha) - y \cdot \sin(\alpha) \\ x \cdot \sin(\alpha) + y \cdot \cos(\alpha) \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.18)$$

Daraus kann man schließen: Eine Rotation des Koordinatensystems um den Winkel α entspricht einer Rotation der Objekte um den Winkel $-\alpha$.

Drehung um eine beliebige Raumachse

Als typische Anwendung für Transformationen entwickeln wir nun die Rotationsmatrix R_a für die Drehung eines Punktes P um eine beliebig orientierte Achse im Raum um einen Winkel α . Die Drehachse ist dabei eine Gerade der Form $G = a + m \cdot b$, die durch die zwei Vektoren $a = (a_x, a_y, a_z)$ und $b = (b_x, b_y, b_z)$ und dem Skalar m bestimmt wird. Gesucht sind die Koordinaten eines Punktes $P = (x, y, z)$ nach einer Drehung um die Achse G um den Winkel α . Man geht dazu in folgenden Teilschritten vor:

1) Translation

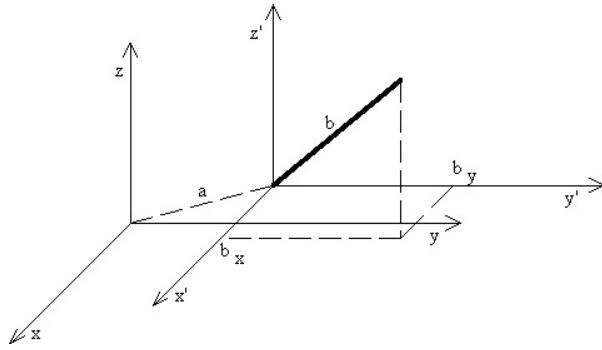


Abbildung 2.11: Translation in den Basispunkt a

Wir verschieben das Koordinatensystem (x, y, z) in den Basispunkt a der Drehachse durch eine Translation T_1 . Wir erhalten das Koordinatensystem (x', y', z') .

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & -a_x \\ 0 & 0 & 0 & -a_y \\ 0 & 0 & 0 & -a_z \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad (2.19)$$

2) Rotation um die z' -Achse

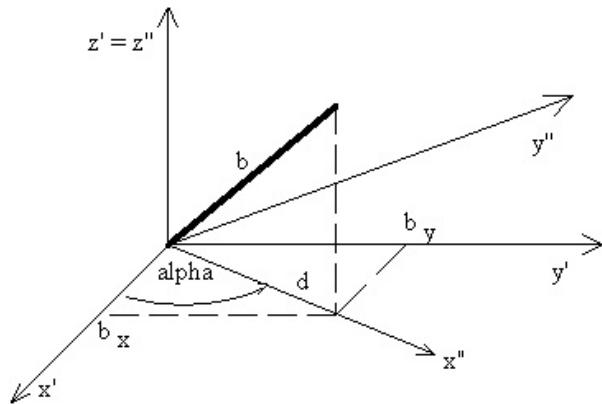


Abbildung 2.12: Rotation um z' -Achse

Nun drehen wir das Koordinatensystem (x', y', z') um seine z' -Achse, so dass der Vektor b in der (x'', z'') -Ebene liegt. Dann gilt:

$$\sin \alpha = \frac{b_y}{d}, \cos \alpha = \frac{b_x}{d}, d = \sqrt{b_x^2 + b_y^2} \quad (2.20)$$

Da das Koordinatensystem um $-\alpha$ gedreht wird, ergibt sich für die Transformation T_2 :

$$T_2 = \begin{bmatrix} \cos \alpha & \sin \alpha & 0 & 0 \\ -\sin \alpha & \cos \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \frac{1}{d} \begin{bmatrix} b_x & b_y & 0 & 0 \\ -b_y & b_x & 0 & 0 \\ 0 & 0 & d & 0 \\ 0 & 0 & 0 & d \end{bmatrix} \quad (2.21)$$

Rotation um die y'' -Achse

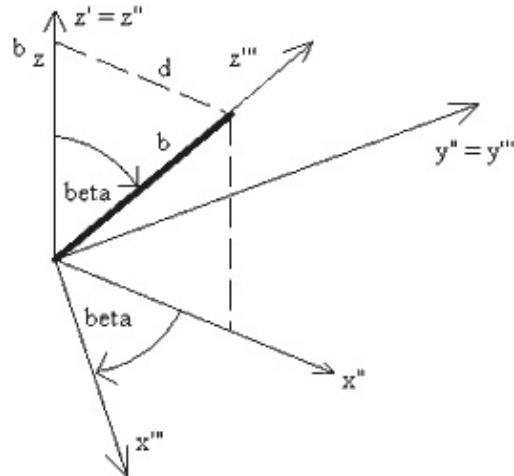


Abbildung 2.13: Rotation um y'' -Achse

Wir drehen das Koordinatensystem (x'' , y'' , z'') jetzt so um die y'' -Achse, dass die neue z -Achse in Richtung der gewünschten Drehachse zeigt, also mit dem Vektor b zusammenfällt. Die Rotationsmatrix für diese Drehung ist:

$$T_3 = \frac{1}{e} \begin{bmatrix} b_z & 0 & -d & 0 \\ 0 & e & 0 & 0 \\ d & 0 & b_z & 0 \\ 0 & 0 & 0 & e \end{bmatrix} \quad (2.22)$$

$$\text{mit } e = |b| = \sqrt{b_x^2 + b_y^2 + b_z^2}, \quad \sin \alpha = \frac{d}{e}, \quad \cos \alpha = \frac{b_z}{e}$$

Rotation um die z'' -Achse

In dem nun vorliegenden Koordinatensystem ist die gewünschte Drehung eine einfache Rotation des Objektpunktes um die neue z -Achse mit der Matrix:

$$T = \begin{bmatrix} \cos \alpha & \sin \alpha & 0 & 0 \\ -\sin \alpha & \cos \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.23)$$

Zusammenfassung zur Gesamttransformation

Die gesuchte Transformation T_a lässt sich durch die Verknüpfung der Transformationen

$$T_a = T \cdot T_3 \cdot T_2 \cdot T_1 \quad (2.24)$$

ermitteln. Damit der gedrehte Punkt wieder im Ausgangskoordinatensystem vorliegt, werden anschließend noch die inversen Transformationen T_1^{-1} , T_2^{-1} und T_3^{-1} darauf angewandt. Man erhält somit für die gesuchte Gesamttransformation:

$$T_a = T_1^{-1} \cdot T_2^{-1} \cdot T_3^{-1} \cdot T \cdot T_3 \cdot T_2 \cdot T_1 \quad (2.25)$$

2.4.4 Vergleich zwischen Modell- und Weltkoordinaten

Ein geometrisches Objekt lässt sich durch eine geschlossene Verbindung mehrerer Punkte über bestimmte Kurven beschreiben. Die Lage eines dieser Punkte P wird dabei durch seine Koordinaten (x_P, y_P, z_P) festgelegt.

Das hierfür benötigte Bezugssystem wird so gewählt, dass sich dieses spezielle Objekt möglichst einfach beschreiben lässt. Ein derartiges Koordinatensystem wird als Modellkoordinatensystem (x, y, z) oder auch MC (master coordinate system) bezeichnet. Bild 2.14 zeigt 8 Punkte, die in einem beispielhaften Modellkoordinatensystem definiert sind. Die Verbindung bestimmter Punkte durch Linien ergibt die Darstellung eines Quaders.

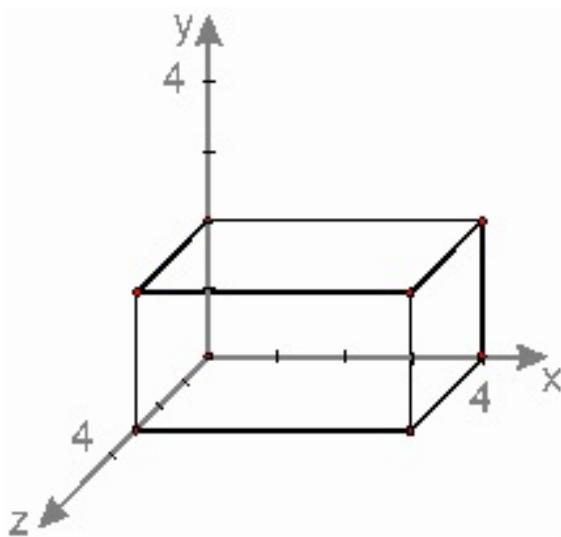


Abbildung 2.14: Darstellung eines Quaders im Modellkoordinatensystem (x, y, z)

Die einfache Beschreibung dieses Körpers ist dadurch gekennzeichnet, dass ein Teil der Punkte auf jeweils einer Koordinatenachse liegt. Beim Zusammenstellen des Objektes ist es leichter, die gewünschten Seitenlängen direkt auf die Achsen zu übertragen, als eine Umrechnung auf eine beliebig verschobene und verdrehte Position im Raum vorzunehmen.

Jetzt soll eine Szene, bestehend aus zwei Quaternen mit der gleichen Geometrie, aufgebaut werden. Eine Möglichkeit zu deren Realisierung ist die Definition eines zweiten Quaders bezüglich des Koordinatensystems des ersten Quaders. Damit sind zwei wesentliche Kon-

sequenzen verbunden: Der zweite Quader lässt sich nicht mehr einfach im Sinne weniger Punkte ausserhalb der Koordinatenachsen beschreiben. Außerdem muss der Verlauf der Verbindungslien (Geradengleichung: $y = m \cdot x + n$) zwischen den Punkten für diesen Quader neu bestimmt werden.

Eine andere Möglichkeit besteht nun darin, das bereits definierte MC einschliesslich des Quaders zu duplizieren und die Positionen dieser beiden Modellkoordinatensysteme in Bezug auf ein drittes Koordinatensystem zu beschreiben. Im Bereich der Computergrafik wird hierzu auf ein speziell definiertes unabhängiges Koordinatensystem zurückgegriffen, das als Weltkoordinatensystem bezeichnet wird. Die übliche Kurzschreibweise lautet WC (world coordinate system). Zur Unterscheidung von einem Modellkoordinatensystem werden dessen Achsen im folgenden mit Großbuchstaben (X, Y, Z) beschriftet.

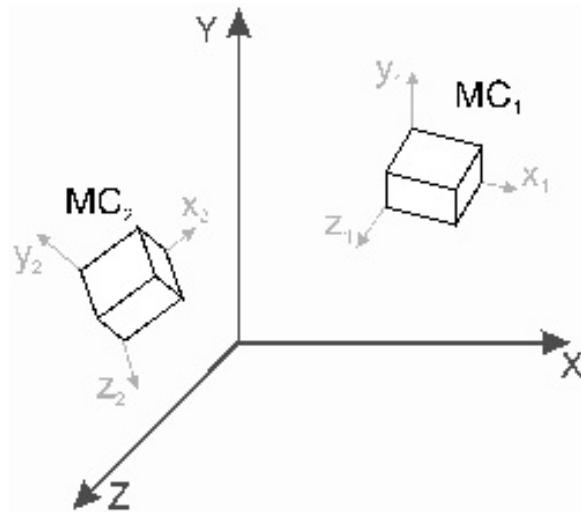


Abbildung 2.15: einfache, aus zwei gleichartigen Objekten aufgebaute Szene

In Bild 2.15 ist eine nach dem zweiten Verfahren - Bestimmung der relativen Positionen zusammengestellte Szene zu erkennen. Das aus Bild 2.14 bekannte Modellkoordinatensystem mit dem darin definierten Quader wurde hierzu dupliziert und die mit MC_1 und MC_2 bezeichneten Kopien an bestimmte Positionen im Raum verschoben.

Die Positionierung eines Objektes bezüglich des Weltkoordinatensystems bezeichnet man als Modelltransformation. Eine Modelltransformation besteht im allgemeinen aus mehreren, in einer bestimmten Reihenfolge ausgeführten Transformationsschritten.

Modelltransformation

Die Positionierung eines Objektes im Raum lässt sich durch eine Modelltransformation realisieren. Dabei wird ein deckungsgleich zum Weltkoordinatensystem liegendes MC durch die schrittweise Ausführung von Einzeltransformationen gedanklich in die gewünschte räumliche Lage gebracht. Soll jetzt ein Objekt in diesem MC definiert werden, bildet man dazu das Objekt zuerst im WC ab und multipliziert dann die Koordinaten aller Objektpunkte mit den einzelnen Transformationsmatrizen, die zur Positionierung des MC benötigt wurden. Als Ergebnis erhält man die Weltkoordinaten des Objektes, kann dieses also mit der korrekten Lage und Orientierung im Raum abbilden.

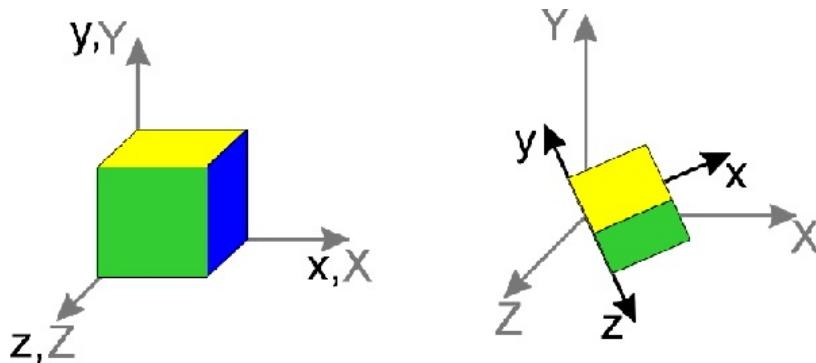


Abbildung 2.16: einfache, aus zwei gleichartigen Objekten aufgebaute Szene

Im allgemeinen besteht eine Modelltransformation aus einer aufeinanderfolgenden Ausführung von Rotation, Translation, Skalierung und Spiegelung. Für das folgende Beispiel soll aber eine mehrfache Rotation ausreichen. Bild 2.16 links zeigt dazu ein deckungsgleich zum WC(X, Y, Z) liegendes MC(x, y, z), in dem bereits ein würfelförmiges Objekt definiert ist. Das Objekt soll dabei nur der besseren Erkennbarkeit dienen, denn grundsätzlich ist die Definition eines Objektes vor abgeschlossener Positionierung des MC nicht notwendig. Im folgenden wird dieses MC von der in Bild 2.16 links erkennbaren Position in die Position gemäß Bild 2.16 rechts transformiert. Damit die Rotationen um die Achsen des Modellkoordinatensystems erfolgen können, werden die Richtungsvektoren der Achsen (x, y, z) durch die drei in Weltkoordinaten angegebenen Punkte $P_x(1, 0, 0)$, $P_y(0, 1, 0)$ und $P_z(0, 0, 1)$ beschrieben. Der erste Schritt der Modelltransformation soll nun eine Rotation um die z-Achse, beschrieben durch $P_z(0, 0, 1)$, und dem Winkel sein.

$$R_M = \begin{pmatrix} x^2(1 - \cos \varphi) + \cos \varphi & xy(1 - \cos \varphi) - z \sin \varphi & xz(1 - \cos \varphi) + y \sin \varphi & 0 \\ xy(1 + \cos \varphi) + \sin \varphi & y^2(1 - \cos \varphi) + \cos \varphi & yz(1 - \cos \varphi) - x \sin \varphi & 0 \\ xz(1 - \cos \varphi) + \sin \varphi & yz(1 - \cos \varphi) + x \sin \varphi & z^2(1 - \cos \varphi) + \cos \varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Dafür lässt sich gemäß R_M eine Rotationsmatrix R_{M1} aufstellen

$$R_{M1} = \begin{pmatrix} 0.87 & -0.5 & 0 & 0 \\ 0.5 & 0.87 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Mit der Transformationsmatrix R_{M1} lassen sich jetzt die Koordinaten der Punkte P_x , P_y und P_z berechnen. P_z bleibt für diesen Transformationsschritt unverändert, so dass nur die Punkte P_x auf (0.87, 0.5, 0) und P_y auf (-0.5, 0.87, 0) transformiert werden. Die Zahlenangaben ermöglichen ein einfaches Skizzieren des gedrehten Koordinatensystems (Bild 2.17 links).

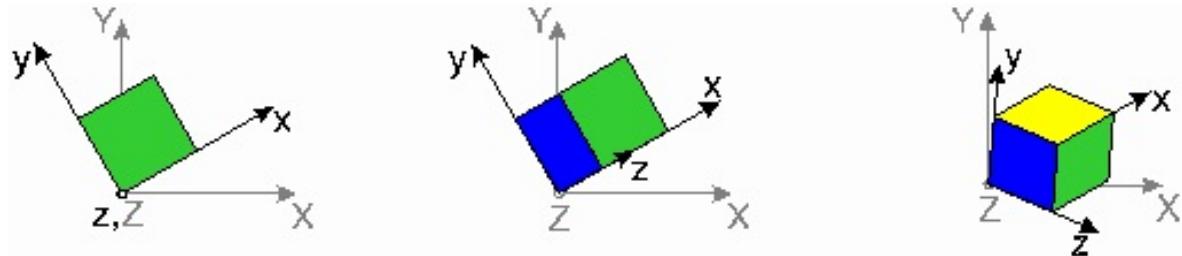


Abbildung 2.17: Rotation mit R_{M2} , R_{M2} und R_{M3}

Der nächste Schritt soll aus einer Drehung um die y-Achse mit dem Winkel $\varphi_y = 30^\circ$ bestehen. Dabei ist zu beachten, dass die Koordinaten des transformierten Punktes P_y einzusetzen sind. Die zweite Transformationsmatrix ergibt sich deshalb zu

$$R_{M2} = \begin{pmatrix} 0.90 & -0.06 & 0.43 & 0 \\ -0.06 & 0.97 & 0.25 & 0 \\ -0.43 & -0.25 & 0.87 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Durch Multiplikation der Koordinaten mit R_{M2} berechnen sich die neuen Positionen der Punkte zu $P_x(0.75, 0.43, -0.5)$ und $P_z(0.43, 0.25, 0.87)$; P_y bleibt unverändert. Die abschließende Rotation um die x-Achse mit $\varphi_x = 30^\circ$ ergibt die Rotationsmatrix R_{M3}

$$R_{M3} = \begin{pmatrix} 0.94 & 0.29 & 0.17 & 0 \\ -0.21 & 0.89 & -0.40 & 0 \\ -0.27 & 0.35 & 0.90 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

und die endgültigen Punktkoordinaten $P_x(0.75, 0.43, -0.5)$, $P_y(-0.22, 0.88, 0.43)$ und $P_z(0.63, -0.22, 0.75)$.

Nach Durchführung dieses Transformationsschrittes befindet sich das MC an der erwarteten Position (Abbildung 2.17 rechts), die Modelltransformation ist damit abgeschlossen. Es sind jetzt alle notwendigen Matrizen bekannt, um die Position eines Objektes in Weltkoordinaten anzugeben.

Für das vorangegangene Beispiel berechnet man die Weltkoordinaten eines beliebigen Punktes $Q(x, y, z, 1)$ nach der Gleichung

$$(Q_{WC})^T = R_{M3} \cdot R_{M2} \cdot R_{M1} \cdot (Q_{MC})^T \quad (2.26)$$

Die Modelltransformation eines beliebig komplexen Objektes lässt sich realisieren, indem man mit Gleichung (2.26) alle Objektkoordinaten bezüglich des MC in die entsprechenden Weltkoordinaten umrechnet.

3 Epipolargeometrie

Hiermit wird eine geometrische Beziehung zwischen mindestens zwei unterschiedlichen Kameras beziehungsweise mindestens zwei Laserscannern beobachtet. Die Anordnung von zwei Kameras (Laserscannern) wird als Stereosystem bezeichnet und lässt sich in Abhängigkeit von ihren räumlichen Anordnungen in zwei grundsätzliche Klassen unterteilen. Davon ist die eine das achsparallele Stereosystem, die andere die konvergente Anordnung. Im ersten Fall liegen beide Kameras auf einer optischen Achsen und sind parallel ausgerichtet. Im zweiten Fall können die Kameras beliebig positioniert werden. In dieser Arbeit wird der zweite Fall beschrieben(sehen Sie z.B. [7]).

3.1 Achsparalleles Stereogeometrie

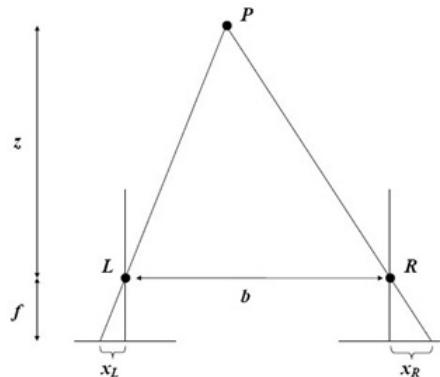


Abbildung 3.1: Achsparallele Stereogeometrie

Das achsparallele Stereosystem ist durch zwei Kameras ausgeprägt, die nur horizontal verschoben sind. Abbildung 3.1 zeigt zwei parallel ausgerichtete Kameras mit den Brenn-

punkten L bzw. R. Ihre Bildebenen liegen in einer gemeinsamen Ebene eingebettet. Die Variable b bezeichnet den horizontalen Abstand der Kameras und f die Brennweite. Punkt P stellt das beobachtete Objekt dar. Die Entfernung von P zur Basislinie ist z. Als horizontale Bildkoordinate von P ergibt sich für die linke Kamera der Wert x_L und für die rechte Kamera der Wert x_R .

$$d := x_L - x_R \quad (3.1)$$

Den Abstand d bezeichnet man als Disparität d.h. die Verschiebung zwischen korrespondierenden Bildpunkten. Durch Verwendung des Strahlensatzes ergibt sich für die Entfernung z die Beziehung

$$z = \frac{f * b}{d} \quad (3.2)$$

Hat man z berechnet, lassen sich auch die x- und y-Koordinate von P im dreidimensionalen Raum berechnen. In einem Koordinatensystem, dessen Ursprung im Brennpunkt der linken Kamera liegt, gilt:

$$x = \frac{x_L}{f} z, \quad y = \frac{y_L}{f} z \quad (3.3)$$

Die Entfernung verhält sich umgekehrt proportional zur Disparität. Das bedeutet, dass eine Disparität von 0 für unendliche Entfernung steht. Bei abnehmender Entfernung nimmt die Disparität zu. Bei bekannten Kameraparametern (f, b) ist die Disparität damit ein eindeutiges Maß für die Objektentfernung.

3.2 Die allgemeine Stereogeometrie

Die Epipolargeometrie bearbeitet einen Teilbereich der projektiven Geometrie. Sie wird auch als allgemeine Stereogeometrie bezeichnet. Die Epipolargeometrie deckt im Wesentlichen alle Projektionen und Operationen im zwei- bzw. dreidimensionalen euklidischen Raum ab. Sie beschreibt vollständig die geometrischen Informationen korrespondierender Punkte zwischen zwei perspektivischen Bildern zueinander.

Es ist nicht möglich, die Position eines Objektes aus den Pixeln eines Bildes zu bestimmen. Um räumliche Informationen aus digitalen Bildern zu bekommen, werden deshalb mindestens zwei Ansichten, manchmal auch mehrere Ansichten benötigt. Nach der Aufnahme eines Bildpaars werden korrespondierende Punkte gesucht, d.h. Punkte bei denen im linken und rechten Bild der gleiche Weltpunkt abgebildet wird. Es müsste also für jeden Bildpunkt im linken Bild das komplette rechte Bild nach dem korrespondierenden Punkt durchsucht werden.

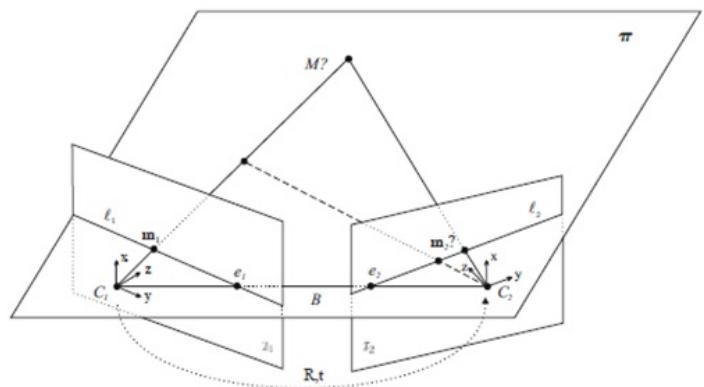


Abbildung 3.2: Epipolargeometrie

Sei m_1 ein Bildpunkt einer perspektivischen Abbildung des 3D-Raumpunktes M auf eine Bildebene I_1 . Können darauf aufbauend Einschränkungen der Positionen des korrespondierenden Bildpunktes m_2 in einer weiteren Bildebene I_2 gemacht werden? Es soll gezeigt werden, dass Einschränkungen aus der Kalibrierung, und bekannter Lage zweier Kameras zueinander formuliert werden können. Die Suche über den ganzen Bildraum kann auf eine Suche über eine Linie reduziert werden. Durch eine Transformation der Bilder kann zudem erreicht werden, dass diese Linien mit den Zeilen der transformier-

ten Bilder zusammenfallen und folglich eine horizontale Suche ermöglicht wird. Diese Transformation wird Rektifikation genannt.

Seien nun zwei Kameras mit ihren optischen Zentren C_1, C_2 auf den gleichen Raumpunkt M ausgerichtet wie in Abbildung 3.2 dargestellt. Die Verbindungsline zwischen den beiden optischen Zentren C_1, C_2 wird Basislinie B genannt. Aufgrund der leicht zueinander gedrehten Bildebenen I_1 und I_2 schneidet die Basislinie beide Bildebenen. Die Schnittpunkte der Basisline mit den Bildebenen werden Epipole genannt und mit e_1 und e_2 kennzeichnet. Anders formuliert stellen die Epipole die Projektion der optischen Zentren in die jeweils andere Bildebene dar. Ihre Position in den Bildebenen hängt ausschließlich von der Anordnung der Kameras zueinander ab. Sie können, müssen sich aber nicht, in den jeweils aufgenommenen Bildern befinden. Die beiden optischen Zentren C_1, C_2 spannen gemeinsam mit dem 3D-Raumpunkt M eine Ebene auf, die als Epipolarebene bezeichnet wird. Es ist offensichtlich, dass auch die Punkte m_1 und m_2 auf dieser Ebene π liegen, da die optischen Strahlen von M zu C_1 und M zu C_2 Geraden auf der Epipolarebene sind.

Der 3D-Raumpunkt M liegt auf der Geraden, die durch das Projektionszentrum C und seinem Bildpunkt m definiert wird. Diese Gerade, in das andere Bild projiziert, kennzeichnet die Epipolarlinie l_i und entspricht genau der Schnittgeraden der Bildebenen mit der Epipolarebene π . Da die Bildpunkte m_1, m_2 komplanar in der Epipolarebene π sind, müssen Sie sich folglich auf den Epipolarlinien l_1, l_2 befinden. Alle Abbildungen $m_{1\pi}, m_{2\pi}$ der zu π komplanaren Raumpunkte M_π müssen auf den Epipolarlinien $l_{1\pi}, l_{2\pi}$ liegen. Rotiert man die Epipolarebene um die Basislinie, so kann die Gesamtheit aller 3D Punkte im Raum erfasst werden. Aus jeder neuen Epipolarebene resultieren jeweils neue Schnittgeraden mit den Bildebenen und somit neue Epipolarlinien. Das gemeinsame an diesen Epipolarlinien ist, dass sie sich in den jeweiligen Epipolen ihrer Bildebene schneiden. Die Gesamtheit der Epipolarlinien in einer Bildebene wird Epipolarlinienbüschel (engl. pencil of epipolar lines) genannt. Als wesentliche Eigenschaft erhält man somit folgende wichtige Beziehung. Wird ein 3D Punkt M in einem Bild an einer bestimmten Bildposition abgebildet, so beschränkt sich die Suche des korrespondierenden Punktes m_2 auf die Epipolarlinie l_2 , statt auf das gesamte Bild(sehen Sie z.B. [8]).

3.3 Beziehungen in Kamerakoordinaten

Bei Stereogeometrie muss ein Kamerasystem existieren, das mindestens aus zwei Kameras besteht, wobei die Kameras zueinander gedreht sind. Diese allgemeine Positions- und Orientierungsänderung wird durch Koordinatentransformation angegeben (sehen Sie z.B. [7]).

$$M_{C_2} = RM_{C_1} + t \quad (3.4)$$

wobei R eine orthogonale Drehmatrix und t der dreidimensionale Verschiebungsvektor ist.

In Kamerakoordinaten stehen Systeme über eine Transformation in Beziehung, wobei der Ursprung in das Koordinatensystem der Kamera 1 gelegt wird.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + t, \text{ mit } R = \begin{bmatrix} x_1^T \\ y_1^T \\ z_1^T \end{bmatrix} \quad (3.5)$$

Aus dem Lochkameramodell ergeben sich die Abbildungen eines 3D Punktes M in Sensorskoordinaten:

$$\widetilde{m}'_1 = \frac{M_1}{z_1}, \widetilde{m}'_2 = \frac{M_2}{z_2} \quad (3.6)$$

aus (3.5) und (3.6) erfolgt

$$\widetilde{m}'_2 = \frac{1}{z_2}(z_1 R \widetilde{m}'_1 + t) \quad (3.7)$$

Das Kreuzprodukt der Gleichung mit dem Vektor t und weiterhin das innere Produkt (Skalarprodukt) mit \widetilde{m}'_2 sowie die Eliminierung der Strukturparameter gibt schließlich die folgende Gleichung.

$$t \times \widetilde{m}'_2 = \frac{z_1}{z_2} t \times R \widetilde{m}'_1 \quad (3.8)$$

$$\widetilde{m}'_2(t \times \widetilde{m}'_2) = \widetilde{m}'_2 t \times (R \widetilde{m}'_1) = 0 \quad (3.9)$$

$$\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}_x = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}, \text{ mit } [t]_x = -[t]^T_x \quad (3.10)$$

Mit dieser Schreibweise erhält man schließlich die zentrale Gleichung für die mathematische Beziehung zwischen den Abbildungen eines 3D Punktes in zwei Kameras, die Epipolargleichung

$$\widetilde{m'_2}^T E \widetilde{m'_2} = 0, \text{ mit } E = [t]_x R \quad (3.11)$$

3.4 Bestimmung der Epipolarlinien

Eine Linie in einer Ebene kann mit dem dreidimensionalen Vektor v auf folgende Weise beschrieben werden:

$$ax + by + cz = 0, \quad \text{mit } v = [a, b, c]^T \quad (3.12)$$

Für Punkte auf diese Linie müssen die Ausführungen zur Projektive Ebene gelten:

$$p_1^T v = 0 \quad \text{und} \quad p_2^T v = 0, \quad \text{mit } p_i = [x_i, y_i, 1]^T \quad (3.13)$$

Eine Linie kann durch das Kreuzprodukt zweier Punkte definiert werden.

$$l = p_1 \times p_2 \quad (3.14)$$

Jeder 3D Punkt der zugehörigen Bildpunkte kann mit dem Kameramodell in homogenen Koordinaten und auch mit entsprechendem Skalierungsfaktor λ darstellt werden.

$$\lambda_1 \widetilde{m'_1} = M_1 \quad \text{und} \quad \lambda_2 \widetilde{m'_2} = M_2, \quad \text{mit } \lambda_i \in (0, \infty) \quad (3.15)$$

Die beiden Kameras können über die Euklidische Transformation verknüpft werden.

$$M_2 = RM_1 + t = \lambda_1 R \widetilde{m'_1} + t \quad (3.16)$$

$$\lambda_2 \widetilde{m'_2} = \lambda_1 R \widetilde{m'_1} + t \quad (3.17)$$

Nun kann die Epipolarlinie in Kamera 2 über zwei ausgezeichnete Punkte berechnet werden. Dies ist der Epipol in Ansicht2, der die Projektion des optischen Zentrums von Kamera1 in der Bildebene 2 darstellt. Er entspricht damit einer Skalierung mit $\lambda_1 = 0$ und man erhält den Epipol dann aus Gl. 3.17, der bis auf einen Faktor dem Vektor t entspricht:

$$\widetilde{e'_2} = t \quad (3.18)$$

Die Abbildung eines Punktes in Ansicht 2 ist der zweite Punkt und entspricht einer Skalierung von $\widetilde{m'_1}$ mit $\lambda_1 = \infty$. Hier kann man den Vektor t vernachlässigen und man erhält die Projektion des Punktes $\widetilde{m'_{2\infty}}$

$$\widetilde{m'_{2\infty}} = R\widetilde{m'_1} \quad (3.19)$$

Aus beiden Abbildungen kann die Epipolarlinien in Ansicht 2 berechnen werden.

$$l'_2 = \widetilde{e'_2} \times \widetilde{m'_{2\infty}} = t \times \widetilde{Rm'_1} = E\widetilde{m'_1} \quad (3.20)$$

Die korrespondierende Epipolarlinie in Ansicht 1 lautet:

$$l'_1 = E^T \widetilde{m'_2} \quad (3.21)$$

3.5 Beziehungen in Bildkoordinaten

Die intrinsische Transformation, die durch Matrix A hergestellt wird, kann man durch den Zusammenhang zwischen normierten Kamerakoordinaten und Pixelkoordinaten beschreiben. Sie lauten für die beiden Kameras(sehen Sie z.B. [8]):

$$\widetilde{m}_1 = \widetilde{A}_1 \widetilde{m}'_1 \quad \text{und} \quad \widetilde{m}_2 = \widetilde{A}_2 \widetilde{m}'_2 \quad (3.22)$$

Falls man nun diese Beziehung in die Epipolargleichung einsetzt, bekommt man folgende Gleichung:

$$\widetilde{m}_2^T \widetilde{A}_2^{-T} E \widetilde{A}_1^{-T} \widetilde{m}_1 = \widetilde{m}_2^T F \widetilde{m}_1 = 0, \quad \text{und} \quad F = \widetilde{A}_2^{-T} E \widetilde{A}_1^{-1} \quad (3.23)$$

Die 3x3 Matrix F wird Fundamental-Matrix genannt. Die Fundamental-Matrix beschreibt die Epipolargeometrie vollständig in Pixelkoordinaten, weil sie sowohl die intrinsischen Parameter der beiden Kameras als auch die extrinsischen Parameter der euklidischen Transformation enthält. Entsprechend für die Epipolarlinien gilt folgende Gleichung:

$$l_2 = F \widetilde{m}_1 \quad \text{und} \quad l_1 = F^T \widetilde{m}_2 \quad (3.24)$$

Die Epipole in Bildkoordinaten ergeben sich unter Einbeziehung der intrinsischen Matrizen wie folgt:

$$\widetilde{e}_2 = \widetilde{A}_2 t \quad \text{und} \quad \widetilde{e}_1 = \widetilde{A}_1 R^T t \quad (3.25)$$

Die Definition der Fundamental-Matrix kann umformuliert werden. Unter Verwendung einer Beziehung zwischen einer nicht-singulären Matrix und einem Vektor stellt man sie als antisymmetrische Matrix dar:

$$F = \widetilde{A}_2^{-T} [t]_x R \widetilde{A}_1^{-1} = [\widetilde{A}_2 t]_x \widetilde{A}_2 R \widetilde{A}_1^{-1}, \quad \text{mit} \quad [\widetilde{A}_2 t]_x \widetilde{A}_2 = \widetilde{A}_2^{-T} = \widetilde{A}_2^{-T} [\widetilde{A}_2 t]_x \quad (3.26)$$

Wenn diese Definition auf Ansicht 2 angewendet wird, bekommt man folgende Gleichung:

$$F = [\widetilde{e}_2]_x \widetilde{A}_2 R \widetilde{A}_1^{-1} \quad (3.27)$$

Analog für die Fundamental-Matrix und die Epipole in Bildkoordinaten erhält man:

$$F \widetilde{e}_1 = 0 \quad \text{und} \quad F^T \widetilde{e}_2 = 0 \quad (3.28)$$

3.6 Essential- und Fundamental-Matrix

Die geometrischen Beziehungen zwischen korrespondierenden Punkten in zwei Ansichten eines Stereokamerasytems werden sowohl mit Essential- als auch mit der Fundamental-Matrix beschrieben.

Die Essential-Matrix beschreibt die Beziehung der Punkte in Sensorkoordinaten. Dagegen liefert die Fundamental-Matrix die Beziehung für korrespondierende Bildpunkte in Bildkoordinaten. Für die Essential-Matrix ist es wichtig, dass die intrinsischen Parameter bekannt sind, um von den Bildkoordinaten in Pixel, in Sensorkoordinaten transformieren zu können. In diesem Fall spricht man vom kalibrierten Fall, demgegenüber spricht man bei der Fundamental-Matrix vom unkalibrierten Fall(sehen Sie z.B. [7]).

Bei der Essential-Matrix sind drei für Rotation, und zwei für Translation unbekannte Parameter gegeben. Die Essential-Matrix enthält zwei Bedingungen:

1. Da die Determinante der antisymmetrischen Matrix Null ist, verschwindet auch die Determinante der Essential-Matrix, denn nach der Determinantenregel gilt:

$$\det(E) = \det([x]_x) \det(R) = 0 \quad (3.29)$$

2. Aufgrund von $\det(E) = 0$ existieren nur zwei linear unabhängige Zeilen- oder Spaltenvektoren. Damit ist der Rang der Matrix $R_g(E) = 2$. Die beiden von Null verschiedenen Eigenwerte der Essential-Matrix sind gleich.

Die Fundamental-Matrix enthält zwei Eigenschaften:

1. da die Determinante der Essential-Matrix verschwindet, gilt $\det(F) = 0$;
2. deswegen gilt für die Fundamental-Matrix $R_g(F) = 2$;

4 Korrespondenzanalyse

Sei ein Element der einen Ansicht gegeben, dann wird das dazu korrespondierende Element der anderen Ansicht gesucht. Hierbei müssen zwei Entscheidungen getroffen werden:

1. Welches Bildelement soll verglichen werden?
2. Welches Ähnlichkeitsmaß kann dafür verwendet werden?

Das Korrespondenzproblem besteht darin herauszufinden, welcher Punkt eines Bildes zu welchem Punkt eines anderen Bildes korrespondiert, unter der Annahme, dass es sich um Abbilder desselben Raumpunktes handelt. Die Korrespondenzanalyse ist ein klassisches Aufgabenfeld der Bildanalyse.

Die Algorithmen der Korrespondenzanalyse können grob in zwei Klassen unterteilt werden. Es handelt sich dabei um die pixel- und die merkmalsbasierten Verfahren.

Bei den pixelbasierten Verfahren werden Regionen fester Größe um den Pixel herum miteinander verglichen und deren Korrelation über ein Ähnlichkeitsmaß bestimmt. Das korrespondierende Bildelement befindet sich an der Stelle, an der ein Maximum der Ähnlichkeitsfunktion herrscht. Die pixelbasierten Verfahren betrachten die Menge aller Pixel. Wegen ihres Bezugs auf einen Bildblock werden sie auch als Block-Matching bezeichnet.

Die merkmalsbasierten Verfahren beschränken den Suchraum auf eine Menge von Merkmalen im Bild. Merkmale können unter anderem Ecken- oder Liniensegmente sein. Im Gegensatz zu der Korrelationsmessung nutzen diese Verfahren die Abstände zwischen den Merkmalen als Ähnlichkeitskriterium. Als Vorverarbeitungsschritt ist somit eine Merkmalsextraktion notwendig.

Zusammenfassend ist also die Zielsetzung der Korrespondenzanalyse eine robuste, fehlerfreie und eindeutige Zuordnung zwischen Bildmerkmalen zweier unterschiedlicher Bilder derselben Szene. Ein Scanpunkt hat keine Distanzinformation in A und die Distanz konnte durch vorhergehende Schritte nicht bestimmt werden. Hier soll zuerst eine pixelweise Korrespondenz gefunden werden. Danach soll die Distanz in A und B berechnet werden. Im ersten Ansatz kann davon ausgegangen werden, dass die Marker exakt platziert wurden. Als nächstes soll die ungenaue Markerposition durch pixelweise Korrespondenzfindung ausglichen werden (sehen Sie z.B [9]).

4.1 Pixelbasierte Verfahren

Eine einfachste Korrespondenzanalyse kann auf mindestens zwei Bilder anwenden. Dabei vergleicht man die Intensitätswerte, und bei Farbbildern zusätzlich die Farbwerte, des einen Bildes, mit den jeweiligen Werten des anderen Bildes. Grundlegend aussagekräftiger jedoch ist die Bildstruktur in der Umgebung dieses Bildpunktes. Deshalb wird bei der Korrespondenzanalyse nicht der einzelne Bildpunkt, sondern ein Fenster um diesen Bildpunkt herum, ein sogenannter Block, betrachtet. Diese Methode wird als Block-Matching bezeichnet.

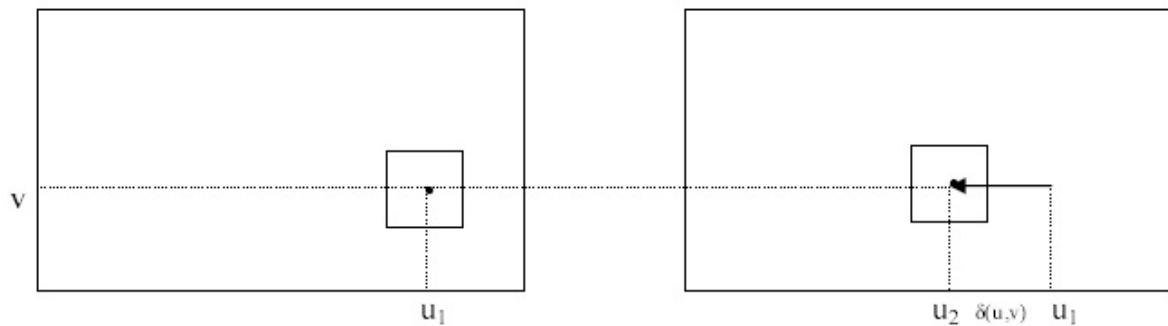


Abbildung 4.1: Achsparallele Stereogeometrie

Die Vorgehensweise für das Block-Matching ist dabei folgende: für jede Position (u_1, v) in der ersten Ansicht wird ein Referenzblock der Größe (m, n) um den Aufpunkt gewählt und mit entsprechenden Musterblöcken in der zweiten Ansicht an der verschobenen Po-

sition (u_2, v) verglichen. Damit ist die Ähnlichkeit zwischen zwei Blöcken gleicher Größe zu bestimmen. Damit wir noch einen definierten Ähnlichkeitsmaß verwenden können, werden dann die Bildpunktposition ermittelt, die für den gewählten Referenzblock in der ersten Ansicht den ähnlichsten Musterblock in der zweiten Ansicht angibt. Solche Verschiebung bezeichnet man dann als Disparität $\delta(u, v)$. Bei einer achsparallelen Stereogeometrie enthält die Disparität nur eine Komponente, im allgemeinen Fall hat die Disparität aber eine horizontale und vertikale Komponente. Wurden für alle Bildpunkte in einer Ansicht die Disparitäten berechnet, so werden diese als Disparitätskarte oder Disparitätsfeld bezeichnet. In der folgenden Abbildung 4.1 ist ein Referenzblock und ein, um die Disparität $\delta(u, v)$ verschobener, Musterblock dargestellt.

Die pixelbasierte Verfahren sind leichter zu implementieren. Wie gesagt, basiert dieses Verfahren auf einer Korrespondenzanalyse von Fenstern, bzw. Blöcken, welche sich um den Aufpunkt befinden. Solche Methoden sind translatorische Schätzverfahren. Es werden nur geradlinige Bewegungen der einzelnen betrachteten Blöcke berücksichtigt. Bewegungs- und Helligkeitsänderungen zweier Bilder (Rotation, Zoom, starke Helligkeitsänderung) führen zu schwerwiegenden Problemen. Es gibt zwar Schätzverfahren, die genau auf solche Probleme abgestimmt sind, jedoch sind diese schwierig zu implementieren (sehen Sie z.B. [10]).

Parametrische Ähnlichkeitsmaße für das Block-Matching

Zur Berechnung der Ähnlichkeiten von zwei Bildblöcken kann die allgemeine Normdefinition herangezogen werden, wobei der Index p die entsprechende Norm bezeichnet.

$$l_p = \left[\left(\sum_m \sum_n |f_1(u + m, v + n) - f_2(u + \Delta u + m, v + \Delta v + n)| \right)^p \right]^{1/p} \quad (4.1)$$

Mittlere absolute Fehler

Setzt man $p = 1$ so ergibt sich die l_1 -Norm, die auch als mittlerer absoluter Fehler (engl. *sum of absolute differences (SAD)*) bezeichnet wird. Mit dieser Norm berechnet man die absolute Differenz zwischen zwei Blöcken, wobei die Ähnlichkeit dort am größten

ist, und folglich die Differenz minimal wird.

$$\delta(u, v) = \underset{\delta(u, v)}{\operatorname{argmin}} \frac{1}{|\Lambda|} \sum_m \sum_n |f_1(u + m, v + n) - f_2(u + \Delta u + m, v + \Delta v + n)| \quad (4.2)$$

Mittlere quadratischer Fehler

Sehr häufig wird der mittlere quadratische Fehler (engl. *sum of squared differences* (SSD)), der sich entsprechend der Normdefinition für $p = 2$ ergibt, als Abstandsmaß gewählt. Die Vergrößerung der Norm führt zu einem höheren Exponenten in der Normdefinition aus Gl.(4.1), der wiederum zu einer starken Gewichtung von größeren Fehlern führt.

$$\delta_{opt}(u, v) = \underset{\delta(u, v)}{\operatorname{argmin}} \frac{1}{|\Lambda|} \sum_m \sum_n [f_1(u + m, v + n) - f_2(u + \Delta u + m, v + \Delta v + n)]^2 \quad (4.3)$$

Mit dem mittleren quadratischen Fehler wird die Intensitäts der Bilddifferenz minimiert. Multipliziert man Gl. (4.3) aus, so ergibt sich folgender Term:

$$\begin{aligned} \delta_{opt}(u, v) = & \underset{\delta(u, v)}{\operatorname{argmin}} \frac{1}{|\Lambda|} \sum_m \sum_n [f_1(u + m, v + n)]^2 + \\ & \sum_m \sum_n [f_2(u + \delta(u, v) + m, v + n)]^2 - \\ & 2 \sum_m \sum_n [f_1(u + m, v + n) f_2(u + \delta(u, v) + m, v + n)] \end{aligned} \quad (4.4)$$

Die beiden ersten Summanden stellen jeweils die Intensitäts der Referenz und des Musterblockes dar und sind konstant. Der dritte Term beschreibt die Korrelation, als die Ähnlichkeit zwischen dem Referenz- und Musterblock. Somit wird der quadratische Fehler minimal, wo die Ähnlichkeit am größten ist. Die ausführliche Schreibweise des quadratischen Fehlers zeigt jedoch, dass der Fehler auch von der Intensität der beiden Blöcke abhängt.

Normierte Kreuzkorrelation

Um eine Abhängigkeit von der Intensität des Muster- und Referenzblockes zu vermeiden, wird die normierte Kreuzkorrelation (engl. normalized crosscorrelation, NCC) verwendet. Ein Vergleich der relativen Unterschiede zwischen den Bildblöcken wird ermittelt, da der Nenner auf die Intensität der beiden Blöcke normiert wird. Das Maximum wird von der normierten Kreuzkorrelation geliefert, bei der die Ähnlichkeit am größten ist. Auch die NCC weist ähnlich der SSD ein sensibles Verhalten gegenüber Ausreißern auf. Die Intensitätsbilder der ersten und zweiten Ansicht der Szene sind mit $f_i(u, v)$, i gekennzeichnet.

$$\delta_{opt}(u, v) = \operatorname{argmin}_{\delta(u,v)} \frac{\sum_m \sum_n f_1(u + m, v + n) \cdot f_2(u + \delta(u, v) + m, v + n)}{\sqrt{\sum_m \sum_n (f_1(u + m, v + n))^2 \cdot \sum_m \sum_n (f_2(u + \delta(u, v) + m, v + n))^2}} \quad (4.5)$$

Neben den parametrischen Bewertungsfunktionen gibt es auch noch nicht-parametrische Ähnlichkeitsmaße. Bei der Rank-Transformation beispielsweise wird die Anzahl jener Bildpunkte berechnet, die einen geringeren Intensitätswert aufweisen, als der Pixel des Aufpunktes. Es wird ein neues Bild berechnet, bei dem diese Werte an der Pixelposition abgetragen werden. Die neuen Bilder werden schließlich mit einem parametrischen Ähnlichkeitsmaß (wie im Abschnitt erläutert) verglichen. Dieses Verfahren ist invariant gegenüber Rotation, Reflektion und monotoner Grauwerttransformation. Bei der Census-Transformation wird jedem Fenster eine Bitkette zugewiesen. Sie beschreibt die Relation der Intensitäten der Bildpunkte im Messfenster bezüglich der Intensität des Aufpunktes. Die Länge der Bitkette entspricht daher der Anzahl der verglichenen Pixel im Messfenster. Ähnlich den parametrischen Ähnlichkeitsmaße wird die Ähnlichkeit aus der Summe von Hamming-Distanzen berechnet.

4.2 Merkmalsbasierte Verfahrenen

Bei den merkmalbasierten Verfahren werden Bildmerkmale bezüglich ihrer Korrespondenz untersucht. Die merkmalbasierten Verfahren nutzen unterschiedliche Methoden, je nachdem ob sie sich auf eine Korrespondenzanalyse von Punktmerkmalen oder Liniensegmenten stützen.

Korrespondenzanalyse von Punktmerkmalen

Als Punktmerkmale eines Bildes sind Ecken geeignet, die man auf beiden Bildern finden kann. Punktmerkmale kann man durch Standardverfahren, wie Moravec-Operator oder Harris-Ecken-Detektor extrahieren. Wenn die interessanten Punkten für beide Bilder vorliegen, so wird die Auswahl für alle Punkte auf dem linken Bild, die nur einen entsprechenden Punkt auf der Epipolarlinie im rechten Bild haben, durch die Epipolarbedingung

$$\tilde{m}_2^T F \tilde{m}_1 = \tilde{m}_2^T \tilde{l}_2 = 0 \quad (4.6)$$

definiert. Der Moravec-Operator analysiert die mittlere Änderung der Bildintensitäten um einen Bildpunkt. Sei $f(u,v)$ das Intensitätsbild, dann ist das Ergebnis des Moravec-Detektors wie folgt:

$$MO(u,v) = \frac{1}{8} \sum_{k=-1}^1 \sum_{l=-1}^1 |f(u+k, v+l) - f(u, v)| \quad (4.7)$$

Der Moravec-Operator exportiert den größten Wert, bei dem die Änderung in mehreren Richtungen besonders groß ist, z.B. bei Ecken. Ein weiterer Eckendetektor ist der Harris-Ecken Detektor. Bei diesem wird zuerst der Gradient in horizontaler und vertikaler Richtung berechnet. Die diskrete Approximation des horizontalen und vertikalen Gradienten ist

$$\frac{df}{du} = f(u-1, v) - f(u+1, v) \quad (4.8)$$

$$\frac{df}{dv} = f(u, v-1) - f(u, v+1) \quad (4.9)$$

Die Detektion kann man unempfindlich gegenüber Rauschstörungen machen, falls die Quadrate der örtlichen Ableitungen einer Tiefpass-Filterung unterzogen werden. Durch

eine Faltung (\otimes) mit der Gewichtsfunktion g ergibt sich das gelieferte Bild, das folgende Ergebnisse hat:

$$\left(\frac{\partial f}{\partial u}\right)^2 = g \otimes \left(\frac{df}{du} \cdot \frac{df}{du}\right) \quad (4.10)$$

$$\left(\frac{\partial f}{\partial v}\right)^2 = g \otimes \left(\frac{df}{dv} \cdot \frac{df}{dv}\right) \quad (4.11)$$

$$\left(\frac{\partial f}{\partial uv}\right)^2 = g \otimes \left(\frac{df}{du} \cdot \frac{df}{dv}\right) \quad (4.12)$$

Damit kann dann folgende Matrix aufgestellt werden:

$$M = \begin{bmatrix} \left(\frac{\partial f}{\partial u}\right)^2 & \left(\frac{\partial f}{\partial uv}\right) \\ \left(\frac{\partial f}{\partial v}\right) & \left(\frac{\partial f}{\partial v}\right)^2 \end{bmatrix} \quad (4.13)$$

Aus dieser Matrix kann man folgendes Kriterium für die Eckenerkennung ableiten. Die Determinante lautet:

$$\det M = \left(\frac{\partial f}{\partial u}\right)^2 \left(\frac{\partial f}{\partial v}\right)^2 - \left(\frac{\partial f}{\partial uv}\right)^2 \quad (4.14)$$

Liegt nun eine große Änderung in horizontaler und vertikaler Richtung vor, so wird die Determinante einen von Null verschiedenen Wert annehmen. Um nun zwischen Kanten und Ecken zu unterscheiden wird die Spur der Matrix, d.h. die Summe der Hauptdiagonalelemente, herangezogen.

$$\text{trace}(M) = \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2 \quad (4.15)$$

Ist die Spur der Matrix groß, so liegt eine Intensitätsänderung in horizontaler und vertikaler Richtung vor. Bei Kanten hingegen verläuft die Intensitätsänderung vorzugsweise nur in einer Richtung. Damit lässt sich folgendes Auswahlkriterium für einen ausgewiesenen Eckpunkt definieren, wobei k ein Gewichtungsfaktor ist (sehen Sie z.B. [7]):

$$K = \det M - k(\text{trace} M)^2 \quad (4.16)$$

5 Vorverarbeitung

SCENE ist eine Software für den FARO Laser Scanner. SCENE verarbeitet und verwaltet gescannte Daten besonders effizient und einfach durch die neuartige automatische Objekterkennung, Scan Registrierung und Platzierung. SCENE hat die Möglichkeit Scans in Farbe zu übernehmen. In SCENE existieren massenhaft Funktionen und Möglichkeiten für die Datenbearbeitung in einer 3D Umgebung (sehen Sie z.B [11]).

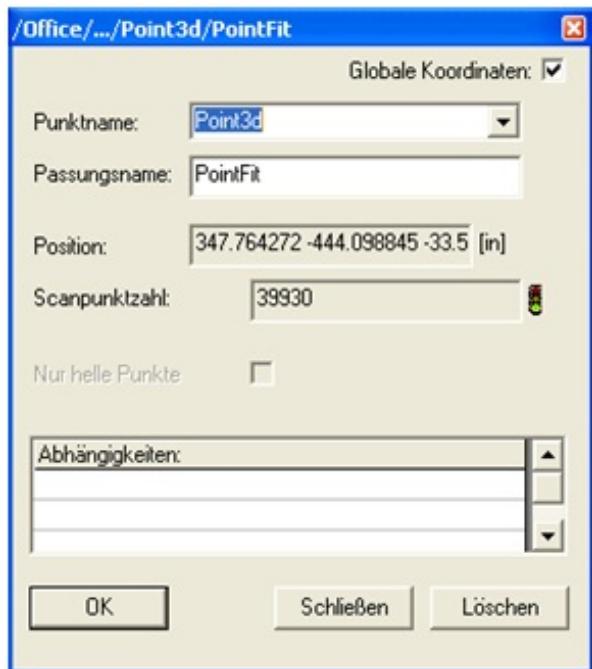


Abbildung 5.1: Passungsobjekt für einen Punkt

Eine von zahlreichen Möglichkeiten der Software ist die Erkennung von CAD Objekten, wie beispielsweise ein Punkt, eine Ebene, ein Wand, eine Kugel, ein Zylinder, eine Kante, eine Region und ein Objekt-Marker. Derartige geometrische Objekte sind in den

Scann vorhanden, und es ist interessant, die Parameter der Objekte aus den Scanpunkten zu gewinnen. Hierbei markiert man einen Bereich der Scanpunkte und trägt dann SCENE auf, daraus die Parameter eines ausgewählten geometrischen Objekts zu bestimmen, indem man im Kontextmenü der Markierung den Befehl zur Anpassung des entsprechenden Objekts auswählt .

Der Punkt ist das einfachste geometrische Objekt. Er ist gekennzeichnet durch seinen Namen und seiner Position im Raum. Die Ebene gibt es in SCENE in zwei Varianten. Eine ist als idealisierte Ebene, die unendlich groß ist und keinen Rand besitzt, definiert. Die andere Variante der Ebene hat eine Randbegrenzung. In beiden Fällen werden Position und Lage durch die Position eines Punkts der Ebene sowie eine so genannte Normale beschrieben. Das Wandobjekt ist dem Ebenenobjekt ähnlich. Wandobjekte können aus Wänden, Decken oder Böden erstellt werden. Aber im Gegensatz zur Ebene, wird das Objekt Wand dazu verwendet, um Scann zueinander auszurichten, die von den gegenüberliegenden Seiten des Wandobjektes aufgenommen wurden. Die Kugel ist ein Objekt, die durch ihre Position und Größe, d.h. dem Radius dargestellt wird. Der Zylinder wird durch seinen Außendurchmesser und seine Länge definiert. Die Kantenerkennung basiert auf einer kombinierten Betrachtung von Unterschieden in den Helligkeits- und Entfernungswerten. Die Region wird dazu benutzt eine Markierung zu speichern. Zum Beispiel kann man einen ganz bestimmten Bereich an Scanpunkten im Arbeitsbereich speichern, um ihn später näher zu untersuchen(sehen Sie z.B [11]).

5.1 Problembetrachtung

In dieser Arbeit wird der Objekt-Marker interessant. Mit dem Objekt-Marker kann man schnell automatische Markierungen festlegen und daraus unterschiedliche Objekttypen im Scan erzeugen. Vordefinierte Typen sind Sphere (Kugel), CircularFlatTarget (Kreiszielmarke), Plane (Ebene), CheckerboardTarget (Schachbrettzielmarke) und Slab (Wand). Ein weiterer Objekt-Marker wäre der DistanzCheck gewesen, der die Entfernung eines angeklickten Objektes beziehungsweise Punktes berechnet. Die Idee besteht darin, dass man zwei Scans beobachtet, die die gleiche Umgebung aus zwei unterschiedlichen Positionen darstellen. Der DistanzCheck wird auf ein Objekt des Scann, beispielsweise ein Punkt, angewendet. SCENE soll entweder automatisch oder manuell korrespondierende

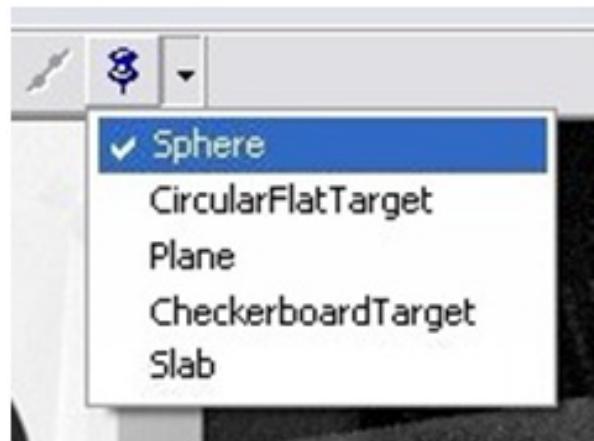


Abbildung 5.2: Objekt-Marker

Punkte auf dem zweiten Scann finden. Um dieses Problem zu lösen, gibt es mehrere verschiedene Analyse- und Bearbeitungsmöglichkeiten.

Wie schon bekannt ist, hat ein Laserscanner eine Umgebung, in der er die Daten erfassen kann, d.h. er kann bis zu einem bestimmten Radius die Daten scannen, erfassen und letztendlich speichern. So erfasste Daten sind Millionen von Punkten, von denen jeder Punkt durch XYZ Koordinaten dargestellt und gespeichert wird. Der Laserscanner hat jedoch die Möglichkeit, die Daten zu scannen, erfassen und zu speichern, die außerhalb seiner Radiusumgebung liegen. Solche Daten werden auch von SCENE übernommen, wobei sie wie ein Bild gespeichert werden. Die Punkte außerhalb der Radiusumgebung sind schon von SCENE gespeichert, man kann aber von diesen keine Information ablesen, d.h. man kann keine XYZ Koordinaten der Punkte bestimmen, und somit auch keine Entfernung. Es existieren aber auch die Punkte, die in der Radiusumgebung des Lasescanners liegen, und trotzdem keine Entfernung haben. Dies kann passieren wenn der Laserstrahl an bestimmten Punkten nicht zurückgeworfen wird.

Man unterscheidet drei Fälle:

- Die korrespondierende Punkte in beiden Ansichten besitzen eine Distanzinformation. Hier kann man einen Punkt aus der ersten Ansicht durch einen Klick wählen. Hier bekommt man automatisch von SCENE die Daten über den korrespondierenden Punkt aus der zweiten Ansicht, d.h. man bekommt nicht nur die Position des Punktes im Weltkoordinatensystem, sondern auch die Entfernung des Punk-

tes zur Position des zweiten Laserscanners. Man kann den Punkt aus der zweiten Ansicht natürlich auch manuell bestimmen. Dazu nutzt man die Methode der zwei windschiefen Geraden. Um diese Methode einzusetzen muss man zwei Richtungsvektoren bestimmen. Man kann durch die Informationen aus der ersten Ansicht einen entsprechenden Richtungsvektor bestimmen. Danach muss man den korrespondierenden Richtungsvektor in der zweiten Ansicht finden. Hierbei nutzt man die Rotation, die Translation und die Skalierung. Nun ist die Voraussetzung für die Methode der zwei windschiefen Geraden erfüllt.

- Der Punkt in der ersten Ansicht besitzt eine Distanzinformation, der korrespondierende Punkt in der zweiten Ansicht jedoch nicht. Auch in diesem Fall gibt es eine Möglichkeit den korrespondierenden Punkt in der zweiten Ansicht automatisch zu finden. Dazu nutzt SCENE schon vorhandene Algorithmen. Das manuelle Finden des korrespondierenden Punktes funktioniert wie im ersten Fall.
- Die korrespondierende Punkte in beiden Ansichten besitzen keine Distanzinformation. Auch in diesem Fall gibt es eine automatische und eine manuelle Methode, wobei die manuelle Methode wie im ersten und zweiten Fall funktioniert. Bei der automatischen Methode exportiert man jeweils ein Bild aus beiden Ansichten und wendet auf diese Bilder das Block-Matching Verfahren an. Mit Hilfe von SCENE kann man die Bilder in verschiedene Formate exportieren. Wenn man ein Objekt in einem Bild betrachtet bzw. man darauf klickt, kann man korrespondierende Pixel im zweitem Bild mit Hilfe des Block-Matchig Algorithmus finden.



Abbildung 5.3: Das Schloss: erste Ansicht

In dieser Arbeit behandle ich den dritten Fall. Wie betrachten ein Gebäude aus zwei verschiedene Scannpositionen. In der Abbildung 5.3 wird ein Objekt des Gebäudes, das keine Distanzinformation hat, markiert. Um diese Distanzinformation zu bestimmen, wird ein Richtungsvektor des Objektes definiert. Diese Richtungsvektor ist in der Abbildung 5.4 als eine Gerade dargestellt. Nun wird das Block Matching Verfahren entlang der Länge dieser Gerade ausgeführt, um das korrespondierende Objekt zu finden. Auf dieses Block Matching Verfahren werde ich im nächsten Kapitel näher eingehen.



Abbildung 5.4: Das Schloss: zweite Ansicht

5.2 Methoden für Entfernungs berechnung

Distanz zwischen zwei Punkten in Raum

Im kartesischen Koordinatensystem berechnet man den Abstand (euklidischer Abstand) zweier Punkte mit Hilfe des Satzes von Pythagoras:

$$d(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5.1)$$

wobei $X = (x_1, x_2, \dots, x_n) \in R^n$ und $Y = (y_1, y_2, \dots, y_n) \in R^n$

$$d_{XY} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} \text{ im } R_3 \quad (5.2)$$

Distanz zwischen zwei windschiefen Geraden im Raum

Im dreidimensionalen Raum existieren zwei Geraden die sowohl keinen Schnittpunkt besitzen und auch nicht Parallel sind. Diese nennt man windschiefe Geraden. Hier wird diese Methode beschrieben. Die l_1 und l_2 sind zwei Geraden die folgendermaßen definiert sind:

$$l_1 : Q_1 = pos_1 + s * dir_1 \quad (5.3)$$

$$l_2 : Q_2 = pos_2 + t * dir_2 \quad (5.4)$$

wobei dir_1 , dir_2 die Richtung der Vektoren der zwei Geraden und pos_1 , pos_2 die Ursprungspunkten der Vektoren der zwei Geraden sind.

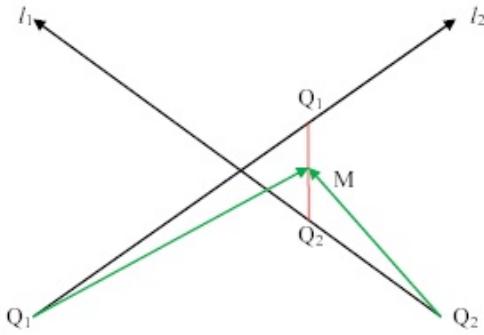


Abbildung 5.5: Abstand windschiefer Geraden

$$l_1 : \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} pos_{1x} \\ pos_{1y} \\ pos_{1z} \end{bmatrix} + s * \begin{bmatrix} dir_{1x} \\ dir_{1y} \\ dir_{1z} \end{bmatrix} \quad (5.5)$$

$$l_2 : \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} pos_{2x} \\ pos_{2y} \\ pos_{2z} \end{bmatrix} + t * \begin{bmatrix} dir_{2x} \\ dir_{2y} \\ dir_{2z} \end{bmatrix} \quad (5.6)$$

$$F(s, t) = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 \quad (5.7)$$

$$F(s, t) = f^2 + 2 \cdot f \cdot t \cdot dir_2 - 2 \cdot f \cdot s \cdot dir_1 + dir_2^2 t^2 - 2 \cdot t \cdot s \cdot dir_1 \cdot dir_2 + dir_1^2 s^2 \quad (5.8)$$

wobei $f = (pos1 - pos2)$ und die erste Ableitung aus Formel (5.8) folgendermaßen lautet:

$$F_s(s, t) = -2 \cdot f \cdot dir_1 - 2 \cdot t \cdot dir_1 \cdot dir_2 + 2 \cdot s \cdot dir_1^2 \quad (5.9)$$

$$F_t(s, t) = -2 \cdot f \cdot dir_2 - 2 \cdot s \cdot dir_1 \cdot dir_2 + 2 \cdot t \cdot dir_1^2 \quad (5.10)$$

Damit wir Koeffizienten s, t finden können, müssen wir die Formeln (5.9) und (5.10) mit Null verglichen.

$$F_s(s, t) = 0 \quad (5.11)$$

$$F_t(s, t) = 0 \quad (5.12)$$

Durch Lösen dieses Gleichungssystems bekommt man die gewünschten Koeffizienten (s, t). Die minimale Distanz zwischen den zwei windschiefen Geraden ist:

$$D = \sqrt{F(s, t)} \quad \text{oder} \quad D = \frac{|\overrightarrow{\text{pos}_1 \text{pos}_2} \cdot \vec{n}|}{\|\vec{n}\|} \quad (5.13)$$

wobei $\vec{n} = \overrightarrow{\text{dir}_1} \times \overrightarrow{\text{dir}_2}$

6 Block-Matching

Korrespondierende Pixel besitzen einen ähnlichen Intensitätswert, aber in der Regel gibt es eine große Anzahl von Pixeln, die identische Intensitätswerte in einem Bild haben. Um dieses Problem zu vermeiden, wird die Unterteilung in Blöcke empfohlen. Wenn die Blöcke zugeordnet wurden, kann man die Block Matching Methode auch auf die Pixel in diesem Block anwenden. Die Wahl der Fenstergröße entscheidet über die Qualität solcher Algorithmen

6.1 Block-Matching Verfahren

Hauptaufgabe dieser Verfahren wird der Ähnlichkeitsvergleich von Grauverteilungen zwischen zwei gleichen großen Pixelmatrizen ($n \times m$). Alle Pixel der Matrix sollen gleichen Disparitätswert haben. Deswegen muss für jeden Block nur ein Disparitätswert berechnet werden (sehen Sie z.B. [12]).

1. Im ersten Verarbeitungsschritt wird ein Bild des Bildpaars (zum Beispiel das linke) in eine konstante Anzahl von gleich großen Blöcken aufgeteilt. Die Suche nach einem korrespondierendem Block im rechten Bild wird nur für die festgelegten Blöcke des linken Bildes durchgeführt.
2. Das Maß für die Ähnlichkeit zwischen den Intensitätswerten der Pixel in den entsprechenden Blöcken wird mit Hilfe von der „mean square error“ Formel berechnet.

$$e(x, y, \Delta) = \frac{1}{mn} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} |E_R(x + i, y + j) - E_L(x + i + \Delta, y + j)|^2 \quad (6.1)$$

wobei die Intensitätsfunktionen des linken bzw. rechten Bildes mit E_L bzw. E_R bezeichnet werden. Das Ähnlichkeitsmaß ist für einen Offset Δ , der die Differenz $(x_L - x_R)$ zwischen den Spaltenpositionen im rechten und im linken Bild angibt, und eine Blockgröße von $(n \times m)$ Pixeln durch $e(x, y, \Delta)$ definiert. (x, y) bezeichnet hier jeweils die linke obere Ecke eines Blockes im linken Bild. Die Disparität D zwischen den Blöcken ist definiert durch den Abstand zwischen den Positionen der Blöcke, die die minimale Abweichung aufweisen.

3. Der Suchbereich in horizontaler Richtung im rechten Bild wird zusätzlich durch ein Disparitätslimit d_{max} beschränkt. d_{max} ergibt sich direkt aus der definierten Einsatzumgebung des Verfahrens. Je dichter Objekte der beobachteten Szene am Kamerasystem liegen, desto größer ist die Disparität der zum Objekt gehörenden Pixel. Durch die Festlegung einer Mindestdistanz lässt sich d_{max} also leicht ermitteln.
4. Innerhalb des Suchbereiches (zwischen 0 und d_{max}) wird der $(n \times m)$ -große Block punktweise verschoben. Der Verschiebungswert Δ für den die e-Funktion ihr Minimum annimmt, bestimmt den so genannten Blockdisparitätswert D . Ein Disparitätswert ist ein Vektor, welcher die Änderung der Lage eines Blockes (später eines Bildpunktes) zwischen den Bildern des Bildpaars beschreibt. Er ist nur dann eindeutig bestimmt, wenn die e-Funktion im Suchbereich ein eindeutiges Minimum besitzt. In den Fällen, in denen kein eindeutiges Minimum existiert, wird ein zusätzliches Entscheidungskriterium verwendet.
5. Unter der Annahme, dass sich die Disparitätswerte benachbarter Blöcke nur geringfügig unterscheiden, werden alle Disparitätswerte, für die die e-Funktion ein Minimum annimmt, mit dem Wert des benachbarten Blocks verglichen. Ausgewählt wird die Disparität mit dem geringsten Unterschied zu der Disparität des Nachbarblockes.

Das Ergebnis der Anwendung des Block-Matching Verfahrens ist eine Disparitätsmatrix, in der jeweils Blöcke fester Größe einen identischen Wert besitzen. Dieses Ergebnis lässt sich unter Verwendung eines Pixelselektionsverfahrens weiter verfeinern, sodass für jedes Pixel ein Disparitätswert bestimmt werden kann. Das Verfahren setzt sich aus drei Verarbeitungsschritten zusammen:

1. Anwendung des Medianoperators auf die Disparitätswerte der Blöcke
2. Pixelselektion
3. Anwendung des Medianoperators auf die Disparitätswerte, die für jedes einzelne Pixel bestimmt wurden.

Der Medianoperator bestimmt in einer Menge von Werten denjenigen Wert, der bei deiner sortierten Reihenfolge der Werte die mittlere Position einnehmen würde. Als erster Schritt wird zunächst der Medianoperator auf die Blockdisparitäten innerhalb einer 3x3-Blockumgebung angewendet, um einzelne Ausreißer in den Disparitätswerten zu eliminieren. Anschließend wird bei der Pixelselektion die Disparität für jeden Pixel (x', y') eines Blockes unter Verwendung der Disparitätswerte dieses und der benachbarten Blöcke bestimmt (sehen Sie z.B. [8]). Für die Bestimmung der Disparität eines einzelnen

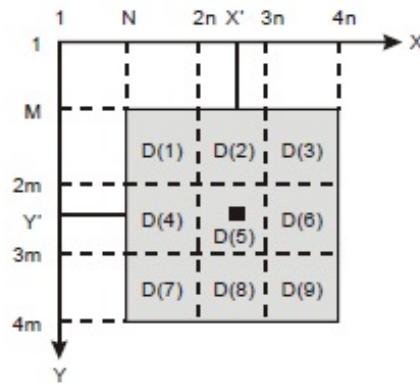


Abbildung 6.1: Pixelselektion des Disparitätswertes an der Position (x', y') unter Verwendung der Blockdisparitäten $D(k)$

Pixels an der Position (x', y') werden die Differenzen $D(k)$ zwischen dem Intensitätswert des linken Bildes an der Position (x', y') und den Intensitätswerten des rechten Bildes an den Positionen $(x' + D(k), y')$ für alle Disparitäten $D(k)$ mit $(1 < k < 9)$ aus der (3x3)-Blockumgebung gebildet (vgl. Abbildung 6.1):

$$D(k) = |E_R(x', y') - E_L(x' + D(k), y')| \quad \text{mit } k = 1, \dots, 9 \quad (6.2)$$

Der Disparitätswert $\text{DISP}(x', y')$ ist definiert durch den Wert $D(k)$, für den die Betragsdifferenz $D(k)$ ihr Minimum annimmt. Bei der Anwendung der Pixelselektion auf jeden

Bildpunkt ergibt sich eine Disparitätsmatrix in Originalbildgröße. Abschließend wird der Medianoperator auf die mittels Pixelselektion berechneten Disparitätswerte angewendet.

Hiermit wird der Block Matching Algorihmus mit anschließender Pixelselektion beschrieben.

```

begin
    unterteile linkes Bild in Blöcke der Größe n x m;
    for (jeden Block im linken Bild) do begin {BM}
        initialisiere min und D;
        setze (x, y) gemäß linker oberer Ecke des aktuellen Blocks;
        for  $\Delta := 1$  to  $d_{max}$  do begin
            berechne  $e(x, y, \Delta)$ ;
            if  $e(x, y, \Delta) < min$  then
                 $min := e(x, y, \Delta); D := \Delta$ 
            else
                if  $e(x, y, \Delta) = min$  then
                     $D :=$  Disparität mit geringstem Unterschied
                    zum Disparitätswert des Nachbarblockes
                end if
            end if;
            speichere Blockdisparitätswert
        end for
    end for;
    filtere Blockdisparitätswerte mit Medianoperator;
    for (jedes Pixel im linken Bild) do begin Pixelselektion
        for  $k := 1$  to 9 do begin berechne  $D(k)$ ;
        bestimme Wert D für den  $D(k)$  minimal ist;
         $DISPARITÄT(\text{Pixelposition}) := D$ 
    end for
    end for;
    filtere Matrix DISPARITÄT mit Medianoperator
end

```

6.2 Block-Matching Verfahren für Farbstereoanalyse

1. Wahl eines geeigneten Farbraums bzw. Koordinatensystems, hier wird der RGB-Raum vorgestellt.
2. Differenzen zwischen zwei Pixeln im Raum beschreibbar machen. RGB ist ein euklidischer Raum \Rightarrow Abstandsmaße für die Farbdifferenz:

$$D_1(F_1F_2) = \sqrt{(r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2} \quad (6.3)$$

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B} \quad (6.4)$$

wo $F_1 = (r_1, g_1, b_1)$, $F_2 = (r_2, g_2, b_2)$ r,g,b sind durch Intensität normierte Farbwertanteile

3. weitere Ablauf wie bei Block-Matching Verfahren, nur daß mit Farbabstand D_1 die $e_{Farbe}(x, y, \Delta)$ berechnet wird, also

statt $|E_R(x + i, y + j) - E_L(x + i + \Delta, y + j)|^2$ wird

$$\begin{aligned} & [((r_R(x + i, y + j) - r_L(x + i + \Delta, y + j))^2 + \\ & ((g_R(x + i, y + j) - g_L(x + i + \Delta, y + j))^2 + \\ & ((b_R(x + i, y + j) - b_L(x + i + \Delta, y + j))^2]_2^l \quad \text{eingesetzt.} \end{aligned}$$

Wenn statt D_1, D_3 benutzt wird, welches eine Approximation von D_1 ist, mit:

$$\begin{aligned} D_3(F_1, F_2) = & |r_1 - r_2|^2 + |g_1 - g_2|^2 + |b_1 - b_2|^2, \text{ dann statt} \\ & |E_R(x + i, y + j) - E_L(x + i + \Delta, y + j)|^2 \quad \text{wird} \end{aligned}$$

$$\begin{aligned} & [| (r_R(x + i, y + j) - r_L(x + i + \Delta, y + j))|^2 + \\ & |(g_R(x + i, y + j) - g_L(x + i + \Delta, y + j))|^2 + \\ & |(b_R(x + i, y + j) - b_L(x + i + \Delta, y + j))|^2] \quad \text{eingesetzt.} \end{aligned}$$

4. Wenn auch hier Pixelbetrachtung durchgeführt werden soll, dann geht man wie beim Pixelselektionsverfahren vor, aber hier muß zusätzlich das Farbabstandsmaß berücksichtigt werden.

RGB ist ein euklidischer Raum => Abstandsmaße für die Farbdifferenz : r,g,b sind durch Intensität normierte Farbwertanteile $F_1 = (r_1, g_1, b_1)$ $F_2 = (r_2, g_2, b_2)$ Der Farbabstand D_1 repräsentiert den Winkel zwischen Farbvektoren F_1 und F_2

7 Zusammenfassung und Ausblick

In dieser Arbeit werden einen 3D Lasersysteme beobachtet und die Anwendung von Laserscanner vorgestellt. Außerdem werden die Algorithmen und die Verfahren beschrieben, die die Distanz zwischen korrespondierenden Objekten in einer 3D Umgebung bestimmen. Um diese Aufgabe zu lösen, soll man die Methoden von Model- und Weltkoordinatensystemen im 3D Umgebung anwenden. Außerdem wird eine geometrische Beziehung zwischen zwei Laserscanner Positionen beobachtet.

Diese Beziehungen werden durch die Epipolargeometrie ausgeprägt. Mit ihrer Hilfe lässt sich die Abhängigkeit zwischen korrespondierenden Bildpunkten beschreiben, also den Punkten, die ein einzelner Objektpunkt in den beiden Bildern erzeugt. Die Korrespondenzanalyse Algorithmen bestimmen die Ähnlichkeit zwischen den Intensitätsvergleich der Pixel in dem entsprechenden Blocke des Bildes. Es gibt jedoch in der Regel eine große Anzahl von Pixel, die identische Werte in einem Bild haben. Um diese Problem zu vermeiden, muss man die Block-Matching Verfahren verwendet werden. Mit diesen Algorithmen kann man die Objekte, die außer Radiusumgebung des Laserscanner liegen, beobachten.

Dieses System kann erweitert und optimiert werden. Möglichkeiten für die Erweiterungen des 3D Lasersystems sind: Entwicklung anderer Scan-Methoden, um komplexe Objekte detektieren und modellieren zu können, andere Verarbeitungsalgorithmen zur Analyse der 3D Punktfolke.

Abbildungsverzeichnis

2.1	FARO Laser Scanner Photon	4
2.2	Triangulation	11
2.3	Polarkoordinatensystem	13
2.4	Punkt in einem Polarkoordinatensystem	13
2.5	Punkt in einem zweidimensionalen kartesischen Koordinatensystem	14
2.6	Punkt in einem dreidimensionalen Koordinatensystem	15
2.7	Punkt in einem dreidimensionalen Koordinatensystem	15
2.8	Drehung um die z-, y- und x-Achse im Gegenuhrzeigersinn	18
2.9	Translation eines Koordinatensystems	19
2.10	Rotation eines Koordinatensystems	20
2.11	Translation in den Basispunkt a	22
2.12	Rotation um z'-Achse	22
2.13	Rotation um y"-Achse	23
2.14	Darstellung eines Quaders im Modellkoordinatensystem (x, y, z)	25
2.15	einfache, aus zwei gleichartigen Objekten aufgebaute Szene	26
2.16	einfache, aus zwei gleichartigen Objekten aufgebaute Szene	27
2.17	Rotation mit R_{M2} , R_{M2} und R_{M3}	28
3.1	Achsparallele Stereogeometrie	30
3.2	Epipolargemetrie	32
4.1	Achsparallele Stereogeometrie	40
5.1	Passungsobjekt für einen Punkt	46
5.2	Objekt-Marker	48
5.3	Das Schloss: erste Ansicht	50
5.4	Das Schloss: zweite Ansicht	50

5.5	Abstand windschiefer Geraden	52
6.1	Pixelselektion des Disparitätswertes an der Position (x', y') unter Verwendung der Blockdisparitäten $D(k)$	56

Literaturverzeichnis

- [1] FARO: FARO Europe GmbH Co. KG; <http://www.faroeurope.com>, 2010.
- [2] Eichler, J., Eichler, H.-J.; Laser, Bauformen, Strahlführung, Anwendungen; Springer, 5. aktualisierte Ausgabe, 2003.
- [3] Teutsch,C.; Model-based Analysis and Evaluation of Point Sets from Optical 3D Laser Scanners, volume 1 of Magdeburger Schriften zur Visualisierung; Shaker Verlag, 2007.
- [4] Luhmann, T., Müller, C.; Photogrammetrie- Laserscanning- Optische 3D-Messtechnik; Wichmann Herbert, 2007
- [5] Rodehorst, V.: Photogrammetrische 3D-Rekonstruktion im Nahbereich durch Auto-Kalibrierung mit projektiver Geometrie; Dissertation, Berlin 2004
- [6] <http://www.vvl.htwg-konstanz.de/3Dbildver/Labor/3/032E.html>, (Stand 24.05.2011)
- [7] O. Schreer: Stereoanalyse und Bildsynthese; Springer-Verlag Berlin Heidelberg; 1. Auflage; 2005.
- [8] Faugeras, Olivier: Three-dimensional computer vision: a geometric viewpoint. Cambridge, MA, USA : MIT Press, 1993.
- [9] Klette, R., Koschan A., Schlüns, K.: „Computer Vision – Räumliche Information aus digitalen Bildern“; Vieweg Technik Verlag; 1996.
- [10] Hartley, R., Zisserman,A.: Multiple View Geometry in Computer Vision;

- Cambridge University Press; 2004.
- [11] FARO: E743 FARO Scene 4.6 Manual Benutzerhandbuch, 2009
- [12] Zhang, C.; "A Survey on Stereo Vision for Mobile Robots"; Dept. Of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh; USA
- [10] Bogicevic, I.; Korrespondenzfindung zwischen Targetmengen im 3D-Raum, Diplomarbeit, 2008
- [11] Bertele, S.; Extraktion charakteristischer Positionen und Objekte in hochauflösten 3D-Scandaten, Diplomarbeit, 2009
- [12] Röhle, N.; Merkmalsbasierte Registrierung von hochauflösten 3D-Punktwolken auf projizierten Daten, Diplomarbeit, 2010
- [13] Teuch, C.; Model-based Analysis and Evaluation of Point Set from Optical 3D Laser Scanners; Shaker Verlag, 2007
- [14] Abraham,S.;Kamera-Kalibrierung und metrische Auswertung monokularer Bildfolgen; Universität Bonn; Dissertation; 2000.
- [15] Faugeras, O., Mourrain, D.; On the geometry and algebra of the point and line correspondences between N images; International Conference on Computer Vision; 1995
- [16] <http://de.wikipedia.org/wiki/Streifenprojektion> (Stand 24.05.2011)

Erklärung

Hiermit versichere ich, diese Arbeit
selbständig verfasst und nur die
angegebenen Quellen benutzt zu haben.

(Aida Gusinjac)