
Stereokalibrierung und Szenenrekonstruktion aus heterogenen Bildquellen in Bezug auf Kameras gleicher und unterschiedlicher Auflösungen

Erarbeitet von Studenten und Studentinnen
im Rahmen der Abschlussarbeit
Masterarbeit der Fakultät

Studiengang Semester Max Mustermann 222222

Betreut von: Prof. Dr. Mustermann

[Disclaim here](#)



Fakultät XYZ der Hochschule Furtwangen
Sommersemester - Wintersemester 2015

Inhaltsverzeichnis

1 Einleitung	5
2 Kameramodelle	9
2.0.1 Lochkameramodell	10
2.0.2 Kameraparameter	11
3 Basis Transformationen	15
3.1 Koordinatentransformation durch Basistransformationen	15
3.2 Aufbau der Koordinatensysteme	18
4 Homographien	22
4.1 Abbildungsunterschiede bei verschobenen Rotationsachsen	24
4.2 Berechnung der Homographiematrix aus Bildebenenkoordinaten	28
5 Epipolar Geometrie	31
5.1 Geometrische Erläuterung der Fundamentalmatrix und der Essentiellen Matrix	33
6 Minimalbeispiel 3D-Stereokalibrierung und Szenenrekonstruktion bei Kameras gleicher Auflösung	36
6.1 Vorgehen: Projektion eines Quaders in zwei verschiedenen transformierte Kameras	37
6.2 Berechnung der Projektionsmatrizen	37
6.3 Transformation der Objektpunkte von Weltkoordinaten in Kamerakoordinaten	38
6.4 Umrechnung von Bildebenenkoordinaten in Sensorkoordinaten	40
6.5 Fundamentalmatrix und der eight-Point-Algorithmus	40
6.6 Epipole und Epipolargeraden	41
6.6.1 Geometrische Konstruktion der Epipole und der Epipolarlinien	42
6.7 Essentiellen Matrix	43
6.8 Exterenen Kameraparameter mit essentieller Matrix	44
6.9 Szenenrekonstruktion durch Triangulation	46
6.10 Rektifizierung	49
6.10.1 Projektive Transformation	53
6.10.2 Ähnlichkeitstransformation	58
6.10.3 Scherungstransformation	60
7 Auswirkungen von unterschiedlichen Kameraauflösungen	64
7.1 Abbildungsunterschiede	64
7.2 Auswirkungen auf die Epipolargeometrie	66
7.3 Minimalbeispiel mit unterschiedlichen Kameraauflösungen	67
8 3D-Stereokalibrierung und Szenenrekonstruktion mit reellen Daten und Kameras gleicher Auflösung	71
8.1 Arbeitsprozess	71
8.2 Normalized-eight-Point-Algorithm	73
8.2.1 Singularity-Constraint der Fundamentalmatrix	75
8.2.2 Singularity- Constraint der essentiellen Matrix	77
8.3 Szenenrekonstruktion	77
8.3.1 Minimieren der Kostenfunktion durch Sampson-approximation	78

9 3D-Stereokalibrierung und Szenenrekonstruktion mit reellen Daten und Kameras unterschiedlicher Auflösung	84
9.1 Ergebnisse einer Stereoanalyse mit Kameras unterschiedlicher Auflösung	84
10 Punktesortierung in Schachbrettmustern	88
10.1 Algorithmus zur Punktesortierung in verzeichneten Schachbrettbildern	88
10.1.1 Vorläufiges Klassendiagramm	89
10.1.2 Beispiele	93
11 Appendix	96
11.1 Homographien bei Drehungen um ein Projektionszentrum	96
11.2 Realbeispiel der Homographieberechnung bei Drehung um einen Drehpunkt	103
12 Fazit - Conclusion	107
13 Nächste Schritte - next steps	108
14 Protocol - 10.11.2015	109
15 Abkürzungsverzeichnis - List of Abbreviations	110

Over the last decades computer vision scientist have taken a new approach to vision. They build different computational models of what shoud be computed, what can really be computed, and how these computations can be realized by computer programs, and they use computers to test their models are correct. The result is a better understandung of vision from a different point of view, and at the same time some working artificial vision systems are built that can be used in idustry, medicine, etc. The knowledge obtained on neirophysiology and psychophysics have given hints to and influenced computer vision scientists, helping find solutions to the design of specific algorithms and implementation of vision systems. On the other Hand coputational vision has also given nerophysologists and psychophysicsist a mathematical framework for modeling vision processes.[1]

1 Einleitung

Die Computer Vision gehört zum Fachbereich der Computer Science und beinhaltet die Entwicklung von künstlichen Intelligenzen, allen voran dem Versuch Computern ein visuelles Verständnis ihrer Umgebung zu geben. Computer Vision behandelt den umgekehrten Weg des Bildes aus der Realität in den Rechner[2]. Die Fähigkeit die gesehenen Bilder zu verarbeiten und zu interpretieren erlaubt es dem Menschen ein Verständnis über die ihn umgebenden Welt zu erlangen. Computer, welche die Fähigkeit besitzen die Welt um sich herum wie das menschliche Auge zu betrachtet und aus diesen entstehenden Bildern Informationen zu extrahieren, wären dem zu folge auch in der Lage Entscheidungen auf Grund von visuellen Eindrücken zu fällen und entsprechend zu handeln. Ein derartiges Verhalten lässt sich in drei Schritte unterteilen. Die Bildaufnahme und Speicherung, die Bildverarbeitung und die Bildanalyse[2]. Als erstes wird die Umgebung in digitale, für der Computer verständliche, Bilder umgewandelt werden. Dies erfolgt mit Hilfe von Kameras, Sensoren oder Lasern. Die Bildverarbeitung und die Bildanalysen lassen sich unterteilen in die Detektion elementarer Merkmale wie Kanten, gerade Linien und Eckpunkte, Segmentierung, die Analyse elementarer Formen und die Identifikation von Objekten[7]. Für die Identifikation von Objekten, werden die zuvor berechneten speziellen Punkte aus den verschiedenen Bildern analysiert, um die 3D Strukturen der Objekte zu erhalten. Beispiele für auf Computer Vision basierende Applikationen sind im Bereich des Autonomen Fahrens, Motion-Capturing-Anlagen, Bewegungserkennungen oder Service Robotern zu finden. Mit dem Entwickeln von Algorithmen für Computer Vision Applikationen, sieht man sich mit immer wieder mit komplizierten Aufgaben und Herausforderungen konfrontiert. Bei der Aufnahme von Bildern, kann es immer wieder zu unvorhersehbaren Bildfehlern wie beispielsweise Rauschen oder Verzerrungen durch die Kameralinse kommen, was auch nicht oft zum Verlust von Referenzdaten führt. In den Ergebnisse, des in dieser Arbeit verfassten Algorithmus für die Rekonstruktion einer stereoskopisch aufgenommenen Szene, sind Auswirkungen eben solcher Fehler zu erkennen. Im Kapitel 3D-Stereokalibrierung und Szenenrekonstruktion mit reellen Daten und Kameras gleicher Auflösung wird aufgeführt, wie mit solchen Fehlern umgegangen werden kann. Computer Vision kommt immer dann zum Einsatz, sobald Informationen aus Bildern gewonnen werden sollen. Dabei kann man unterscheiden unter der Informationsgewinnung aus einzel Bildern oder bewegt Bildern, genauso wie single-view, two-view oder multiple-view Aufnahmen. Single-View beschreibt die Bildverarbeitung anhand der Aufnahmen einer Kamera, two-view befasst sich mit Stereoskopischen Aufnahmen und involviert zwei Kameras und multiple-view ist die Zusammenfassung von mehr als zwei Kameras, wie sie beispielsweise im Motion-Capturing verwendet werden[3]. Zur Analyse verschiedener Bilder ist es essentiell die verschiedenen Kameraparameter, wie Position und Auflösung, zu kennen. Diese können zuvor bestimmt werden, oder aus den Bildern mit Hilfe spezieller Modelle abgeleitet werden. Die Position und Rotation einer Kamera im Raum werden als die extrinsischen Kameraparameter bezeichnet und können durch eine 3×4 -Matrix beschrieben werden. Die Parameter wie die Auflösung oder Brennweiten, werden als die intrinsischen Kameraparameter bezeichnet und können in einer 3×3 -Matrix, die sogenannte Kameramatrix, ausgedrückt werden werden. Sind die Parameter nicht bekannt, so können sie durch die aus Bilddaten gewonnenen Informationen geschätzt werden. Das Verfahren wird als Kamerakalibrierung bezeichnet[3, 4, 5, 1]. Die extrinsischen und intrinsischen Matrizen bilden zusammen die sogenannte Projektionsmatrix P , welche die komplette Transformation eines Dreidimensionalen Punktes M im Raum in einen Zweidimensionalen Bildpunkt m beschreibt mit $m = PM$. Im Kapitel Kameramodelle wird das hier verwendete Kameramodell der Lockkamera und die extrinsischen und intrinsischen Kameraparameter und ihre jeweiligen Matrizen noch genauer vorgestellt. Sind die Kameraparameter bekannt, kann eine Szenenrekonstruktion erfolgen. Die Szenenrekonstruktion beschreibt den Vorgang, in welchem Anhand von Bildpunkten und den bekannten Kameraparameter die Dreidimensionale Szene der Bilder rekonstruiert wird. Das Ziel dieser Masterarbeit ist es zum einen, mit Hilfe einer Stereoaufnahme einer Szene, eine Kamerakalibrierung und anschließend eine Szenenrekonstruktion durchzuführen. Die eigens imple-

mentierte Kamerakalibrierung befasst sich ausschließlich mit der Schätzung der extrinsischen Kamera-parameter, für die intrinsischen wird auf ein bereits vorhandenes externes Programm zurückgegriffen. Danach erfolgt die Szenenrekonstruktion. Der eigens entwickelte Ansatz wird in zwei Schritten in dieser Arbeit vorgestellt. Im erste Schritt wird ein sogenanntes Minimalbeispiel aufgebaut. Das Minimalbeispiel beinhaltet eine synthetisch im Programm aufgebaute 3D-Szene. Die 3D-Szene wird dann in zwei voneinander unterschiedlich positionierten simulierten Kameras projiziert, so dass pro Kamera ein 2D-Bild des 2D-Objektes existiert. Anhand dieser 2D-Bilddaten soll dann eine Kalibrierung der extrinsischen Kameraparameter und anschließend eine Rekonstruktion der 3D-Szene erfolgen. Vorteil an diesem Minimalbeispiel ist, dass das Ergebnis der Rekonstruktion leicht abgeglichen werden kann und somit die Funktionalität des implementierten Algorithmus an einem reinen nicht verfälschten Fall getestet werden kann. Des Weiteren können Fehler welche im späteren Beispiel mit realen Bilddaten auftreten, in diesem synthetischen Aufbau rekonstruiert und somit mögliche Lösungen für diese gefunden werden. Abbildung 6.2 zeigt noch einmal Schematisch den Arbeitsprozess des Minimalbeispiels. Die einzelnen Schritte in dieser Abbildung werden im Kapitel Minimalbeispiel 3D-Stereokalibrierung und Szenenrekonstruktion bei Kameras gleicher Auflösung genauer erläutert.

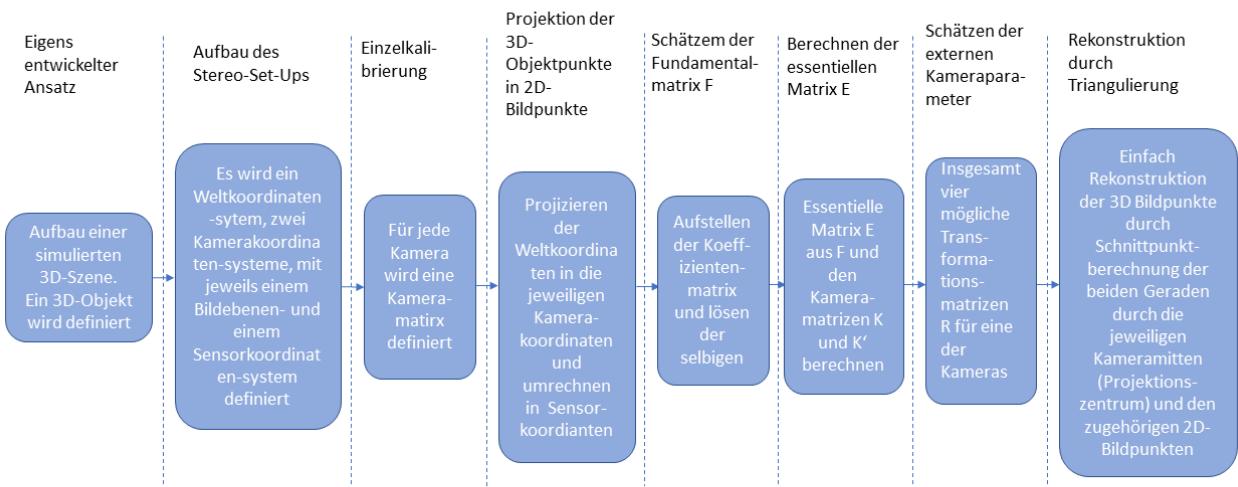


Abbildung 1.1: Arbeitsprozesse der Stereoanalyse bei Verwendung von eigens erstellten synthetischen Bilddaten

Das Minimalbeispiel soll dabei einmal mit Kameras gleicher Auflösung und einmal mit Kameras unterschiedlicher Auflösung durchgerechnet werden. Eine Frage, die ebenfalls im Zuge der Arbeit beantwortet werden soll, ist welche Auswirkungen zwei Kameras mit unterschiedlichen Auflösungen auf den Prozess der Kamerakalibrierung und der Szenenrekonstruktion haben. Die Stereokalibrierungs Applikation, welche von *MatLab* bereitgestellt wird, ist nicht in der Lage eine Stereokalibrierung und Szenenrekonstruktion von zwei Bildern unterschiedlicher Auflösung zu machen. Im Zuge dessen, wurde die Vorgehensweise der Applikation genauer betrachtet und den Auslöser dieses Problems gesucht. Der Arbeitsprozess, welcher in *Matlab* für die Kamerakalibrierung und der anschließenden Szenenrekonstruktion verfolgt wurde ist in Abbildung 1.2 zu sehen. Der Grund warum die Stereoanalyse bei unterschiedlichen Kameraauflösungen in *MatLab* nicht funktioniert wurde im Schritt der Rektifizierung ausgemacht. Im Kapitel Rektifizierung wird nochmal explizit auf das Problem eingegangen und ein möglicher anderer Ansatz aufgezeigt, welcher das Problem beheben könnte.

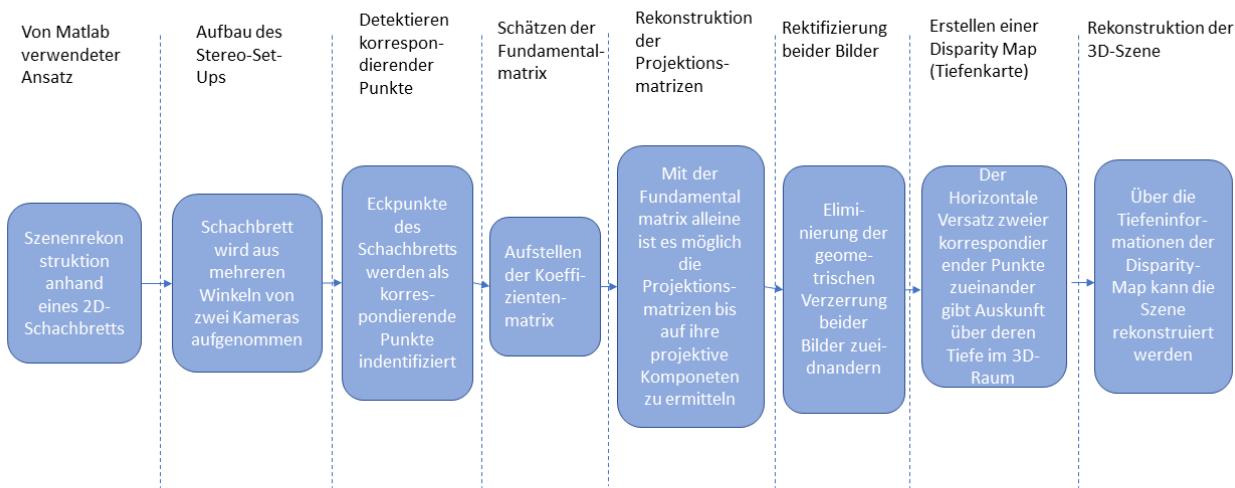


Abbildung 1.2: Arbeitsprozesse der Stereoanalyse bei Verwendung von realen Bilddaten in einem unkalibrierten Fall

Im zweiten großen Teil der Arbeit, soll der im Minimalbeispiel aufgebaute Algorithmus in einer realen Umgebung getestet werden. Es wird also ein Stereosystem mit zwei Kameras im Labor aufgebaut. Die intrinsischen Kameraparameter werden zuvor mit Hilfe eines externen Programms ermittelt. Danach nehmen die Kameras gleichzeitig einen realen Szenenaufbau aus zwei unterschiedlichen Positionen und Winkeln auf. Ab hier setzt der Algorithmus ein, welcher auch für das Minimalbeispiel genutzt wurde. Zuerst wird eine Kalibrierung der extrinsischen Kameraparameter durchgeführt und anschließend die Szene rekonstruiert. Im Kapitel 3D-Stereokalibrierung und Szenenrekonstruktion mit reellen Daten und Kameras gleicher Auflösung wird des Weiteren noch drauf eingegangen, wie mit den genannten Bildfehlern, welche in Realaufnahmen entstehen können umgegangen wird, um die Auswirkungen dieser Fehler auf die Resultate so gering wie möglich zu halten. In Abbildung 1.3 ist der Arbeitsprozess für das Realbeispiel zu sehen.

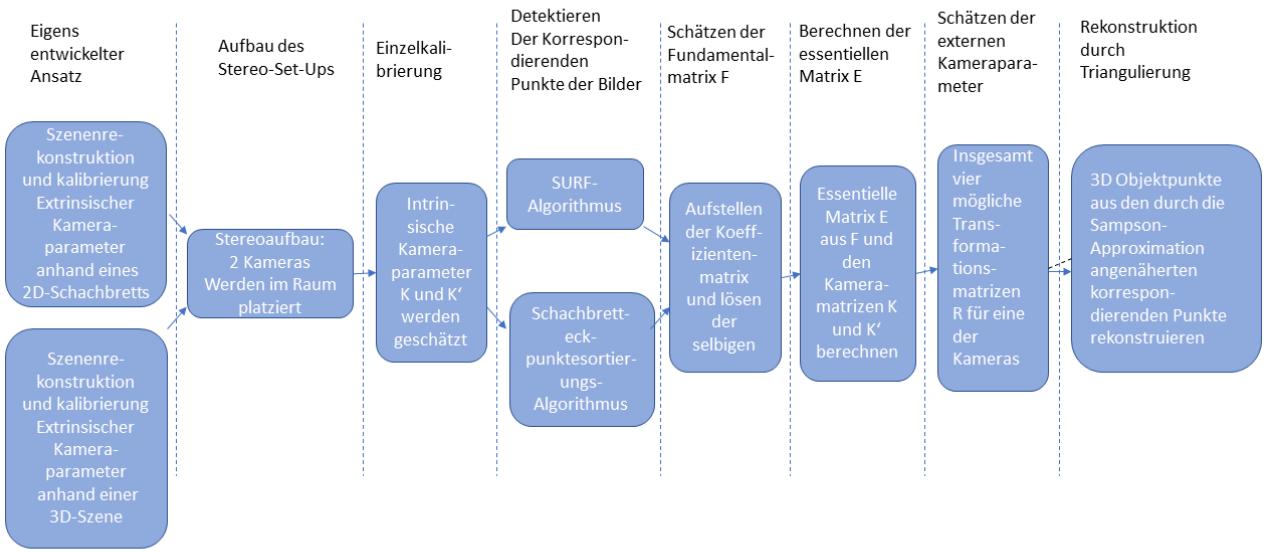


Abbildung 1.3: Arbeitsprozesse der Stereoanalyse bei Verwendung von realen Bilddaten in einem kalibrierten Fall

Auch im Realbeispiel wird geprüft, ob unterschiedlicher Kameraauflösungen sich auf die Szeneriekonstruktion auswirken und mit den Ergebnissen aus dem Minimalbeispiel verglichen.

Die Kapitel 1 bis 5 umfassen die theoretischen Mathematischen Hintergründe, welche für das Verständnis der implementierten Kalibrierungs- und Rekonstruktionsalgorithmen vorhanden sein müssen. Kapitel 6 beinhaltet die einzelnen Schritte des in Abbildung 1.1 aufgezeigten Kalibrierungs- und Rekonstruktionsalgorithms des Minimalbeispiels. Mit diesem Kapitel werden die zuvor beschriebenen mathematischen Werkzeuge zusammengefasst und in Bezug der Stereoskopischen Szeneriekonstruktion angewandt. Der Grund der Implementierung dieses Minimalbeispiels war unter anderem der, dass somit Fehler, welche im Realbeispiel aufgetreten sind rekonstruiert und eine mögliche Lösung dieser ermittelt werden konnte. Kapitel 7 geht auf die Problematik der unterschiedlichen Kameraauflösungen im Bezug auf die im Minimalbeispiel konstruierten Szene ein und führt auch noch auf, wie genau Auflösungen zustande kommen. Kapitel 8 und 9 transferieren dann den im Minimalbeispiel implementierten Algorithmus auf ein reales Beispiel mit einer echten Stereoskopischen Aufnahme. Die einzelnen Schritte aus Abbildung 1.3 werden hier ausführlichst besprochen. Das letzte Kapitel, Kapitel 10, bezieht sich auf einen Algorithmus welcher im Zuge des Realbeispiels entstanden ist. In Abbildung 1.3 ist im Ansatz der Punkt Kalibrierung anhand eines 2D-Schachbretts aufgeführt. Der in Kapitel 10 beschriebene Algorithmus, kann die zuvor detektierten Punkte eines Schachbretts in eine Gitterstruktur sortieren, selbst wenn es zu starken Bildfehlern wie Verzeichnungen kommt und das Schachbrett auf Grund seiner Lage auch noch projektiv verzerrt ist. Als Gitterstruktur sortiert bedeutet, dass die Punkte später wissen in welcher Zeile beziehungsweise Spalte sie sich im Gitternetz befinden und welches ihre Nachbarn sind.

2 Kameramodelle

Die folgenden Kapitel 2 bis 5 sollen die theoretischen Grundlagen für die bevorstehende Stereobildanalyse bilden. Im jetzigen Kapitel wird das Kameramodell vorgestellt. Des Weiteren wird die Notwendigkeit der klaren Differenzierung zwischen den einzelnen im Modell verwendeten Koordinatensysteme aufgezeigt. Sie werden dazu benötigt um zu zeigen wie später die Zusammenhänge im Weltkoordinatensystem gerechnet werden. Im Kapitel Basis Transformationen wird dann aufgezeigt, wie der Wechsel zwischen den Koordinatensystemen funktioniert und wie Punkte bezüglich unterschiedlich definierter Koordinatensysteme dargestellt werden. Dies bildet den Übergang zu den darauf folgenden Kameratransformationen durch die jeweiligen Kameramatrix und Transformationsmatrix der Kameras, welche die Grundbausteine der gesamten Projektionsmatrix bilden. Im darauf folgenden Kapitel Homographien, wird zum einen gezeigt wie sich Bildpunkte, welche sich auf einer Ebene befinden, ohne bekannte Projektionsmatrizen ineinander überführen lassen. Homographien werden in dieser Arbeit eingeführt, da das Prinzip hinter deren Funktion vor allem im Kapitel Rektifizierung bekannt sein sollte, um die Vorgehensweise dessen zu verstehen. Des weiteren bildet vor allem der geometrische Aufbau von Homographien eine gute Überleitung zu den für die Stereobildanalyse wichtigen Epipolar-geometrie. Die Epipolargeometrie ist das ausschlaggebende Element, wenn es um die Kalibrierung der extrinsischen Kameraparameter und der darauf folgenden Szenenrekonstruktion geht. Sie beschreibt eine intrinsische projektive Geometrie zwischen zwei Bildern und ist insbesondere für die Korrespondenzanalyse zwischen den Punkten der Bilder ausschlaggebend.

(Übergang??) Um visuelle Daten aufzunehmen bedarf es optischen Geräten wie Kamera, Laser oder Sensoren. Die Geräte bilden die Dreidimensionale Welt auf eine Zweidimensionale Bildebene ab. Diese Abbildung lässt sich mithilfe von Kameramodellen beschreiben. Ein häufig benutztes Kameramodell ist das Lochkameramodell. Das Modell beruht ausschließlich auf der geometrischen Optik und vernachlässigt physikalische Effekte wie Beugung oder die Auswirkung der Linse[8]. In diesem Modell, werden die Punkte im 3D-Raum durch eine sogenannte Zentralprojektion auf die Bildebene projiziert und anders herum[7, 3]. Bei einer Zentralprojektion wird ein Strahl von einem Punkt im 3D-Raum durch das Projektionszentrum der Kamera gezogen. Das Projektionszentrum der Kamera beschreibt den Punkt in der Kamera an welchem sich alle Strahlen treffen. Ein Strahl der durch einen Punkt im 3D-Raum und durch das Projektionszentrum einer Kamera geht, schneidet eine Ebene, welche im weiteren Verlauf als Bildebene bezeichnet wird. Der Schnittpunkt des Strahls mit der Bildebene wird Bildpunkt genannt [7, 3].

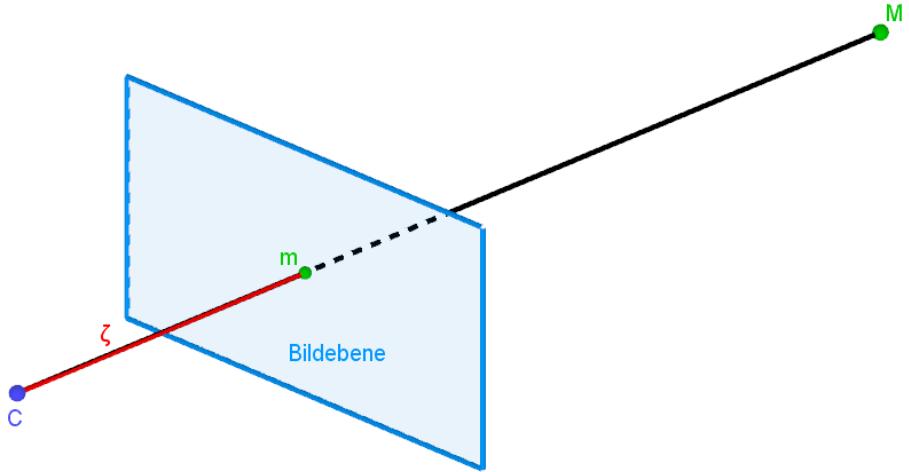


Abbildung 2.1: C ist die Kamera beziehungsweise das Projektionszentrum der Kamera. In rot ist die Strecke ζ markiert. M ist der Dreidimensionale Objektpunkt in Weltkoordinaten und m sein projizierter Punkt auf der 2D-Bildebene

2.0.1 Lochkameramodell

Das Kameramodell, auf welchem diese Arbeit aufbaut, ist das sogenannten Lochkameramodell. Es ist eines der am simpelsten aufgebauten Kameramodelle und wird standardmäßig zur Beschreibung optischer Kameras verwendet. Das Lochkameramodell beinhaltet ein Projektionszentrum C , welches auch gleichzeitig das Kamerazentrum bildet. Die Kamera an sich wird im Lochkameramodell als Punkt angegeben und ist meist gleichbedeutend mit dem Punkt C . C besitzt ein eigenes dreidimensionales Koordinatensystem, welches ab jetzt als Kamerakoordinatensystem bezeichnet wird. Das Kamerakoordinatensystem und alle weiteren Koordinatensysteme, die noch eingeführt werden, dienen dazu die Position eines Punktes innerhalb ihres geometrischen Raumes eindeutig zu bezeichnen. Das Kamerakoordinatensystem hat seinen Ursprung im Projektionszentrum C . Die Ausrichtung der Achsen des Koordinatensystems ist frei definierbar. Im weiteren Verlauf der Arbeit wird das Kamerakoordinatensystem als kartesisches und rechtsdrehendes Koordinatensystem definiert. In Abbildung 2.2 zeigt die X-Achse des Kamerakoordinatensystems in das Bild rein, die Y-Achse ist vertikal nach oben ausgerichtet und die Z-Achse zeigt in Richtung der Bildebene. Die Z-Achse wird hier auch als Hauptachse bezeichnet. Der Punkt in welchem sich die Hauptachse mit der Bildebene schneidet, wird Hauptpunkt HP genannt. Der Abstand des Hauptpunktes zum Projektionszentrum wird im folgenden mit ζ bezeichnet und bildet einen elementaren Baustein der intrinsischen Kameraparameter. Der Hauptpunkt HP bildet den Ursprung eines neuen zweidimensionalen Koordinatensystems, dem Bildebene-Koordinatensystem [3]. Als vorläufig letztes Koordinatensystem wird noch ein dreidimensionales Weltkoordinatensystem eingeführt, dessen Position und Orientierung frei wählbar sind. Ein Punkt im 3D-Raum wird ab jetzt als ein Punkt in Bezug auf ein Weltkoordinatensystem bezeichnet. Des Weiteren wird der gesamte stereoskopische Aufbau mit zwei Lochkameramodellen und den 3D-Objektpunkten im Raum immer im Bezug auf ein Weltkoordinatensystem beschrieben werden. Dieses Weltkoordinatensystem wird diesem Aufbau wiederum meist deckungsgleich mit einem der im Stereoaufbau benutzten Kamerakoordinatensystemen gesetzt.

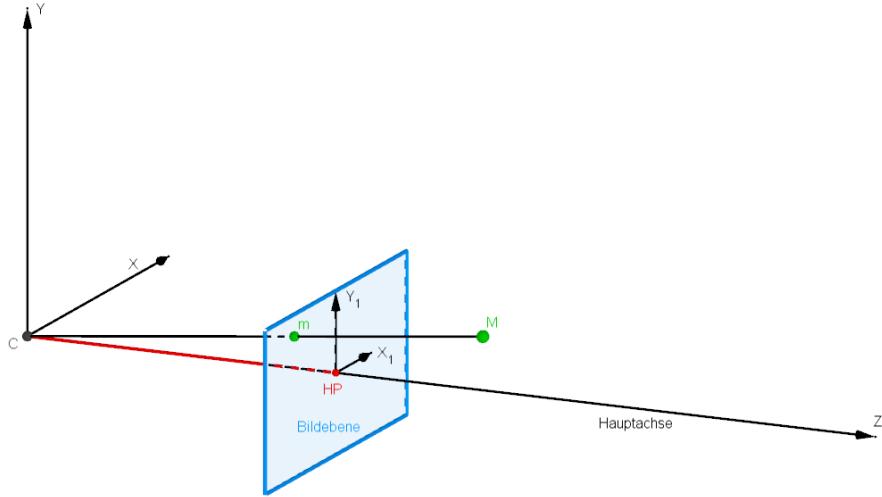


Abbildung 2.2: Lochkameramodell: C ist das Kamerazentrum

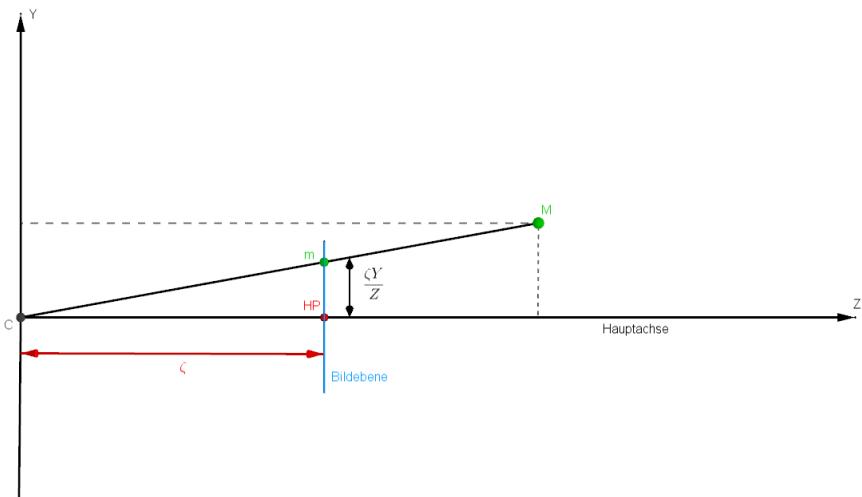


Abbildung 2.3: Lochkameramodell [9]

2.0.2 Kameraparameter

Was bedeutet nun Szenenrekonstruktion und Kamerakalibrierung eines Stereoaufbaus im Zusammenhang mit dem gerade vorgestellten Kameramodell. In Abbildung 2.4 ist ein Stereoaufbau aus zwei Lochkameramodellen dargestellt. Jede Kamera besitzt ein eigenes Kamerakoordinatensystem und eine eigene Bildebene, welche wiederum ein eigenes Bildnebenkoordinatensystem besitzt. Des Weiteren ist ein Objektpunkt M in Bezug auf ein Weltkoordinatensystem zu sehen. Die zwei Strahlen, die jeweils von M ausgehen, beschreiben die Projektion des 3D-Objektpunktes M auf zwei somit korrespondierende Bildpunkte m und m' auf den jeweiligen Bildebenen. Diese Projektion funktioniert auch in die andere Richtung. Sind zwei zueinander korrespondierende Bildpunkte auf den beiden Bildebenen gegeben, so kann jeweils ein Strahl durch die jeweiligen Projektionszentren C und C' und deren zugehörigen Bildebenenpunkte m und m' geschickt werden, welche sich dann in einem Punkt M im Raum treffen. Diese umgekehrte Projektion wird allgemein als Triangulation bezeichnet[3]. Hier kommen jetzt die Projektionsmatrizen P und P' ins Spiel, welche die Projektion M auf m und m' beschreiben mit $m = PM$ und $m' = P'M$ und genauso auch anders herum gelten für die Rückprojektion von m und m' auf M [7, 3].

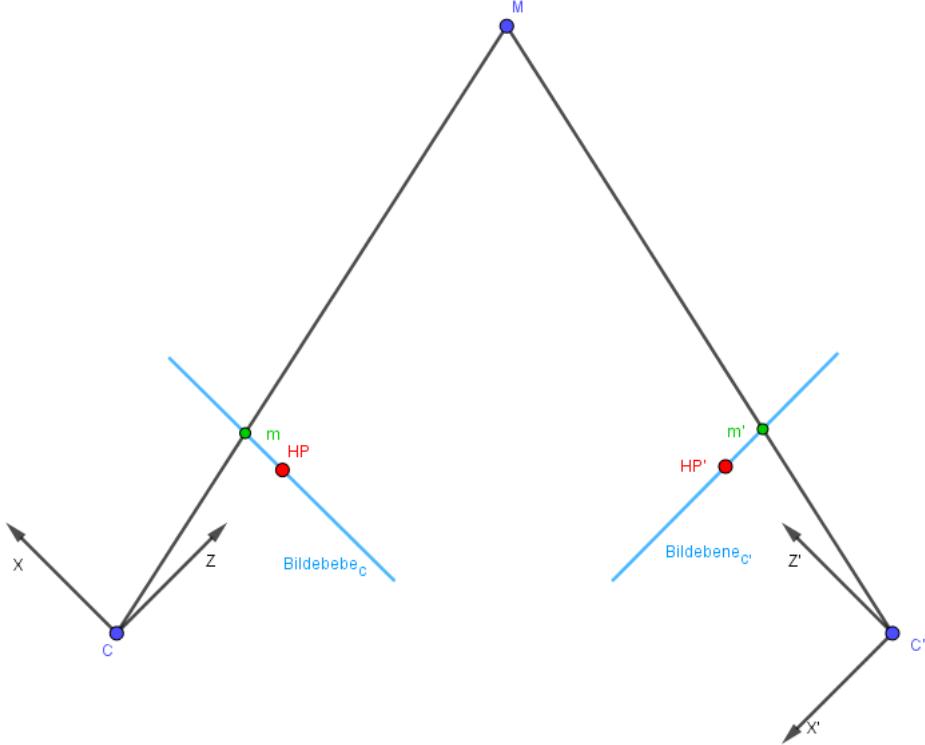


Abbildung 2.4: Stereoskopischer Aufbau mit Lochkameramodellen. Der Punkt M im Raum kann mit den jeweiligen Projektionsmatrizen P und P' der Kameras C und C' in die Bildpunkte m und m' projiziert werden. $m = PM$ und $m' = P'M$. Eine Rückprojektion der Punkte m und m' zum Punkt M ist mit bekannten Projektionsmatrizen genauso möglich.

Die Projektionsmatrizen P und P' bilden sich aus den extrinsischen und intrinsischen Kameraparametern. Bei den extrinsischen Kameraparametern, handelt es sich um die jeweilige Orientierung einer Kamera bezüglich eines Weltkoordinatensystems, was in einer 3×4 -Transformationsmatrix R beschrieben ist. R kann sowohl eine Rotation als auch eine Translation beinhalten. Die intrinsischen Kameraparameter werden in der sogenannten Kameramatrix K ausgedrückt. In der Literatur findet man diese wie folgt definiert[3].

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.1)$$

Im folgenden werden die Elemente der Kameramatrix K im erläutert. Wird nochmal von dem in Abbildung 2.3 und 2.2 beschriebenen simplen Kameramodell ausgegangen, so lässt sich ein dreidimensionaler Punkt M mit $M = [X \ Y \ Z]^T$ durch aufstellen einer Matrix in einen homogenen 2D-Bildebenepunkt m mit $m = [\zeta X \ \zeta Y]^T$ projizieren.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \mapsto \begin{pmatrix} \zeta X \\ \zeta Y \end{pmatrix} = \begin{bmatrix} \zeta & 0 & 0 & 0 \\ 0 & \zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightsquigarrow \begin{pmatrix} \zeta \frac{X}{Z} \\ \zeta \frac{Y}{Z} \end{pmatrix} \quad (2.2)$$

Abbildung 2.5 und die Gleichungen 2.3 bis 2.5 sollen Aufschluss darüber geben, wie der Abstand ζ geometrisch zustande kommt.

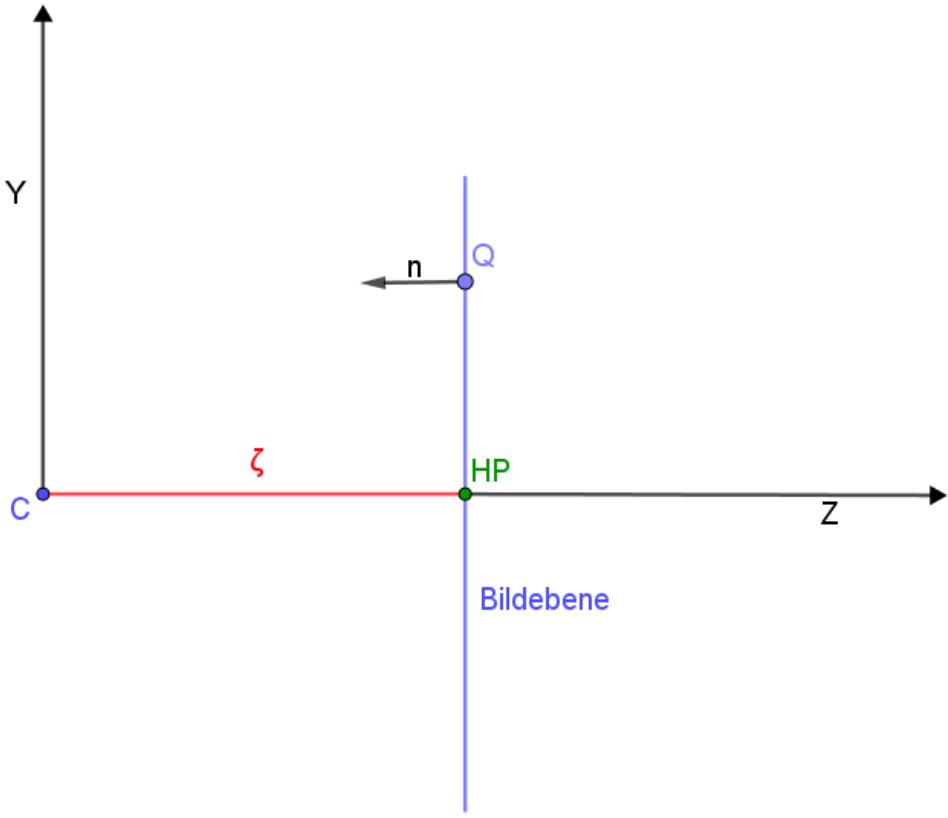


Abbildung 2.5: In blau ist die Bildebene dargestellt auf ihr befinden sich die Punkte Q und HP . HP steht in dieser Abbildung für den Hauptpunkt. Z liegt nicht auf der Ebene, Das Projektionszentrum C liegt hinter der Bildebene und somit auch hinter dem Sensor. n ist die Normale der Bildebene

Wie in Abbildung 2.5 ersichtlich kann ζ , durch folgende Gleichung ausgedrückt werden.

$$\vec{p} + |\vec{QZ} \cdot \vec{n}| \vec{n} = \vec{Z} \quad (2.3)$$

Q ist hierbei ein beliebigen Punkt auf der Bildebene und \vec{n} der normalenvektor der Bildebene.

$$\vec{QZ} \cdot \hat{n} = \zeta \quad (2.4)$$

$$\vec{p} = \vec{Z}\zeta\hat{n} \quad (2.5)$$

Bestimmt man nun, dass die Hauptachse $= \hat{n}$ ist, kann daraus geschlossen werden, dass $\vec{p} = \vec{Z} - \zeta\hat{n}$. Die allgemeine Form der Kameramatrix bezogen auf das Lochkameramodell wird wie in Gleichung 2.1 dargestellt beschrieben. Anstelle von f steht in der Kameramatrix aus Gleichung 2.1 α_x und α_y . Die Werte welche α_x und α_y annehmen können, hängen von der geometrischen Form der Pixel des in der Kamera verbauten Sensors ab[3, 10]. α_x und α_y sind genau dann einheitliche Werte, wenn die Form der verbauten Pixel, ebenfalls einheitlich quadratisch ist. α_x und α_y sind also genau dann ungleich, wenn die auf dem Sensor verbauten Pixel beispielsweise rechteckig oder die geometrische Form eines Parallelogramms besitzen[3]. Im Stereoaufbau dieser Arbeit, wurden Kameras mit CMOS-Chips mit quadratischen Pixeln verwendet, weshalb α_x und α_y einheitlich sein werden. α_x und α_y entstehen aus der Multiplikation der Werte für ζ_x und ζ_y mit einem jeweiligen Skalierungsfaktor m_x und m_y . In der Literatur wird ζ meist als f für *focal length* bezeichnet und die Bildebene auch *Focal plane* genannt [3]. Der Begriff *focal length* ist in diesem Zusammenhang jedoch etwas missverständlich. ζ sowie f geben in der Kameramatrix wie sie in Gleichung 2.2 genau genommen zwei verschiedene Abstände an. Die Matrix müsste genau genommen die Form

$$\begin{bmatrix} \zeta_x & 0 & 0 & 0 \\ 0 & \zeta_y & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

haben. Bis jetzt war immer nur von der Bildebene die rede. Später wird jedoch auf der Bildebene noch ein weiteres zweidimensionales Koordinatensystem definiert und zwar das sogenannten Sensorkoordinatensystem. Dieses Koordinatensystem beschreibt den Aufbau des in der Kamera montierten Sensorchips. Dieser besteht meist aus einer Viereckigen Platte auf welcher sich die Pixelbausteine befinden. Diese Pixel können sowohl quadratisch, rechteckige oder auch abstrakte Formen wie ein Parallelogramm haben. Das Sensorkoordinatensystem Skaliert und unterteilt seine Achsen nach den Kantenlängen der jeweiligen Pixel. Die geometrische Form dieser Pixel sind dafür verantwortlich, dass es nicht immer nur ein einheitlichen ζ geben kann. Nur wenn die Pixel quadratisch sind haben ζ_x und ζ_y die selben Werte, sind sie beispielsweise rechteckig ändern sich dementsprechend die Werte. ζ_x und ζ_y beschreiben also genaugenommen den Abstand des Kamerazentrums zu den jeweiligen Pixelkanten.

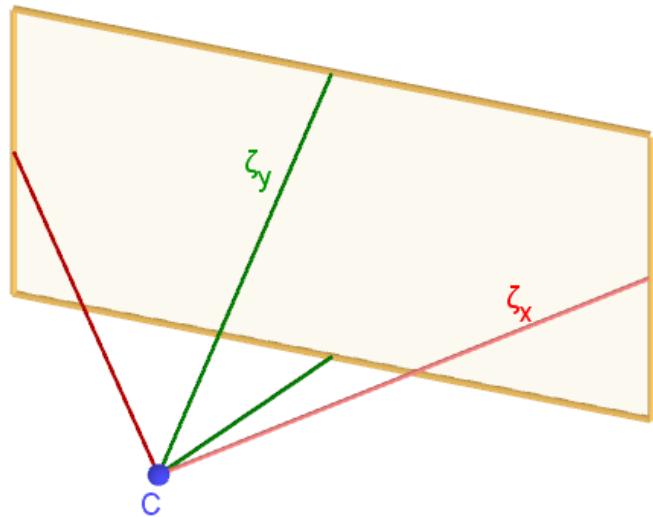


Abbildung 2.6: ζ_x und ζ_y beschreiben den Abstand des Kamerazentrums zu den jeweiligen Pixelkanten
Sind die Pixel nicht quadratisch, wird auf die Werte f_x und f_y jeweils eine Skalierung m_x und m_y drauf multipliziert so dass $\alpha_x = f_x \cdot m_x$ und $\alpha_y = f_y \cdot m_y$ entspricht[3]. Die Matrixeinträge x_0 und y_0 , bilden einen Translationsvektor. Sie beinhalten die Position des Hauptpunkts auf der Bildebene. x_0 und y_0 sind definiert als $x_0 = p_x \cdot m_x$ und $y_0 = p_y \cdot m_y$. Der Matrixeintrag s wird dem sogenannten *skew-Faktor* zugeordnet, welcher nur dann zum Einsatz kommt, sollte die optische Achse nicht orthogonal auf den Bildsensor auftreffen. Sprich wenn der Chip geneigt in der Kamera montiert wurde[3]. Die komplette Projektionsmatrix $P = KR$ [3] besteht aus der Matrixmultiplikation der hergeleiteten Transformationsmatrix R welche die externen Kameraparameter repräsentiert und der Kameramatrix K , welche die internen Kameraparameter repräsentiert[3, 1].

3 Basis Transformationen

(Neue Einleitung schreiben)

In diesem Kapitel wird der Fokus auf die benötigten Koordinatensysteme für die Stereobildanalyse gelegt. Als aller erstes werden anhand praxisnaher Beispiele die Notwendigkeit der Basis Transformation und damit einhergehend der Spezialfall einer Transformation von einem Koordinatensystem wie zum Beispiel einem dreidimensionalen Weltkoordinatensystem (O, δ) mit $\delta = (\hat{d}_1, \hat{d}_2, \hat{d}_3, O)$ in ein dreidimensionales Kamerakoordinatensystem (C, β) mit $\beta = (\hat{b}_1, \hat{b}_2, \hat{b}_3, C)$ aufgezeigt. Die gewählte Notation zur Beschreibung der Koordinatensysteme setzt sich am Beispiel des Kamerakoordinatensystems aus der Bezeichnung für den Ursprung C und den entsprechenden Koordinatenachsen $(\hat{b}_1, \hat{b}_2, \hat{b}_3)$ mit wiederum $\hat{b}_1 = \begin{pmatrix} b_{11} \\ b_{12} \\ b_{13} \end{pmatrix}$, $\hat{b}_2 = \begin{pmatrix} b_{21} \\ b_{22} \\ b_{23} \end{pmatrix}$ und $\hat{b}_3 = \begin{pmatrix} b_{31} \\ b_{32} \\ b_{33} \end{pmatrix}$, welche mit β zusammengefasst werden. In

Abbildung 3.1 sind die beiden kartesischen Koordinatensysteme (O, δ) und (C, β) schematisch dargestellt.

3.1 Koordinatentransformation durch Basistransformationen

Anhand eines Beispiels wird im folgenden eine Transformation eines Punktes bezüglich des Weltkoordinatensystems in das Kamerakoordinatensystem veranschaulicht. Es soll in einer symbolischen Schreibweise Punkte aus einem Weltkoordinatensystem (O, δ) mit $\delta = (\hat{d}_1, \hat{d}_2, \hat{d}_3)$ in ein Kamerakoordinatensystem (C, β) mit $\beta = (\hat{b}_1, \hat{b}_2, \hat{b}_3)$ überführt werden, dessen Ursprung eine Verschiebung und Rotation zum Weltkoordinatensystem aufweist. Beispielhaft wird dies in Abbildung 3.1 aufgezeigt.

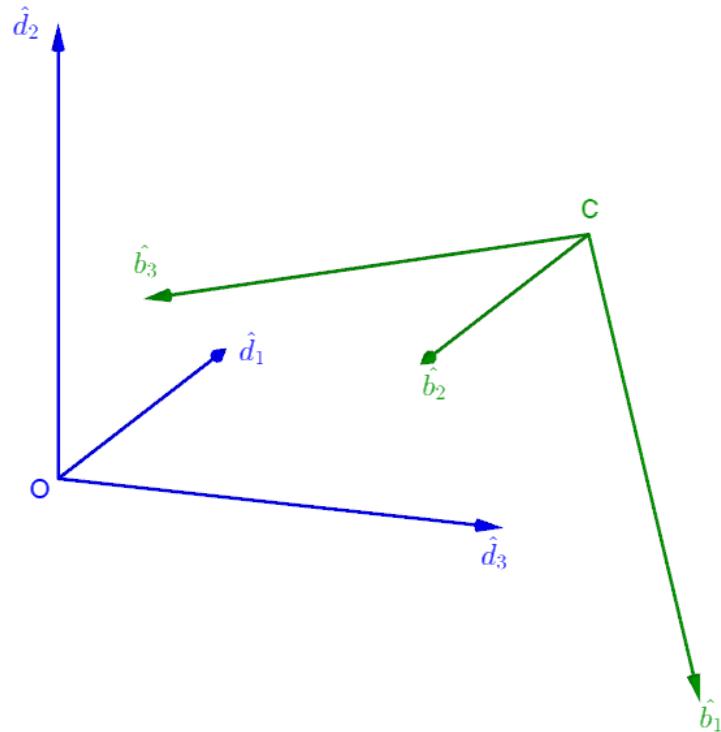


Abbildung 3.1: Weltkoordinatensystem (O, δ) mit $\delta = (\hat{d}_1, \hat{d}_2, \hat{d}_3)$ und Kamerakoordinatensystem (C, β) mit $\beta = (\hat{b}_1, \hat{b}_2, \hat{b}_3)$

Zunächst wird eine sogenannte Koordinatisierung von Punkten im Weltkoordinatensystem vorgenommen. Ein Punkt P_δ bezüglich des Weltkoordinatensystems wird dann wie folgt beschrieben.

$$P_\delta = O + p_{1\delta}\hat{d}_1 + p_{2\delta}\hat{d}_2 + p_{3\delta}\hat{d}_3 \quad (3.1)$$

$$\rightsquigarrow P_\delta = (p_{1\delta}, p_{2\delta}, p_{3\delta})^T = \begin{pmatrix} p_{1\delta} \\ p_{2\delta} \\ p_{3\delta} \end{pmatrix} \quad (3.2)$$

Um das Koordinatentupel in einem projektiven Raum darstellen zu können, wird es projektiv erweitert. Die Entstehenden projektiv erweiterten Koordinaten werden als homogene Koordinaten bezeichnet. Ein Punkt P_δ bezüglich des Weltkoordinatensystems wird um eine vierte homogene Komponente erweitert. Diese kann einen Wert zwischen 0 und 1 annehmen. Nimmt die Komponente den Wert 0 an, so handelt es sich um einen Fernpunkt, sprich der Punkt liegt im unendlichen.

$$P_\delta = \begin{bmatrix} p_{1\delta} \\ p_{2\delta} \\ p_{3\delta} \\ 1 \end{bmatrix} = \left\{ k \begin{bmatrix} p_{1\delta} \\ p_{2\delta} \\ p_{3\delta} \\ 1 \end{bmatrix} \in \mathbb{R}^4 \mid k \in \mathbb{R} \right\} \quad (3.3)$$

$$\begin{bmatrix} \lambda p_{1\delta} \\ \lambda p_{2\delta} \\ \lambda p_{3\delta} \\ 1 \end{bmatrix} = \begin{bmatrix} p_{1\delta} \\ p_{2\delta} \\ p_{3\delta} \\ 1 \end{bmatrix} \text{ für } \lambda \neq 0 \quad (3.4)$$

Ein Punkt P_β bezüglich des Kamerakoordinatensystem wird im Vergleich wie folgt beschrieben.

$$P_\beta = C + p_{1\beta}\hat{b}_1 + p_{2\beta}\hat{b}_2 + p_{3\beta}\hat{b}_3 \quad (3.5)$$

$$P_\beta = \begin{pmatrix} p_{1\beta} \\ p_{2\beta} \\ p_{3\beta} \end{pmatrix} \quad (3.6)$$

Zwischen den beiden Koordinatensystemen (O, δ) und (C, β) werden die folgenden Beziehungsgleichungen aufgestellt.

$$C_\beta = O_\delta + C_{\beta,1}\hat{d}_1 + C_{\beta,2}\hat{d}_2 + C_{\beta,3}\hat{d}_3 \quad (3.7)$$

$$\hat{b}_1 = b_{11}\hat{d}_1 + b_{12}\hat{d}_2 + b_{13}\hat{d}_3 \quad (3.8)$$

$$\hat{b}_2 = b_{21}\hat{d}_1 + b_{22}\hat{d}_2 + b_{23}\hat{d}_3 \quad (3.9)$$

$$\hat{b}_3 = b_{31}\hat{d}_1 + b_{32}\hat{d}_2 + b_{33}\hat{d}_3 \quad (3.10)$$

Diese Beziehungsgleichungen werden in Gleichung 3.1 eingesetzt.

$$\begin{aligned} P_\delta &= O + (C_{\beta,1} + p_{1\beta}b_{11} + p_{2\beta}b_{21} + p_{3\beta}b_{31}) \cdot \hat{d}_1 \\ &\quad + (C_{\beta,2} + p_{1\beta}b_{12} + p_{2\beta}b_{22} + p_{3\beta}b_{32}) \cdot \hat{d}_2 \\ &\quad + (C_{\beta,3} + p_{1\beta}b_{13} + p_{2\beta}b_{23} + p_{3\beta}b_{33}) \cdot \hat{d}_3 \end{aligned} \quad (3.11)$$

Aus Gleichung 3.11 wird ein Gleichungssystem in der Form von Gleichung 3.12 aufgestellt und gelöst.

$$\begin{aligned} p_{1\delta} &= C_{\beta,1} + (C_{\beta,1} + p_{1\beta}b_{11} + p_{2\beta}b_{21} + p_{3\beta}b_{31}) \\ \rightsquigarrow p_{1\delta} - C_{\beta,1} &= (C_{\beta,1} + p_{1\beta}b_{11} + p_{2\beta}b_{21} + p_{3\beta}b_{31}) \end{aligned} \quad (3.12)$$

Wenn P_β gegeben ist, erhält man auf diese Weise direkt P_δ . Wenn jedoch P_δ gegeben ist, so muss das in Gleichung 3.13 aufgestellte lineare Gleichungssystem gelöst werden.

$$\begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} \begin{pmatrix} p_{1\beta} \\ p_{2\beta} \\ p_{3\beta} \end{pmatrix} = \begin{pmatrix} p_{1\delta} - C_{\beta,1} \\ p_{2\delta} - C_{\beta,2} \\ p_{3\delta} - C_{\beta,3} \end{pmatrix} \quad (3.13)$$

Wenn (C, β) ein kartesisches Koordinatensystem ist, so ist die entstehende koeffizientenmatrix D_β orthogonal und es gilt $D_\beta^{-1} = D_\beta^T$.

$$D_\beta^T = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \quad (3.14)$$

$$\rightsquigarrow \begin{pmatrix} p_{1\beta} \\ p_{2\beta} \\ p_{3\beta} \end{pmatrix} = D_\beta^T \begin{pmatrix} p_{1\delta} - C_{\beta,1} \\ p_{2\delta} - C_{\beta,2} \\ p_{3\delta} - C_{\beta,3} \end{pmatrix} \quad (3.15)$$

Handelt es sich um kein kartesisches Koordinatensystem, so muss lediglich die Inverse D_β^{-1} anstatt der Transponierten D_β^T gebildet und wie gehabt verfahren werden. Im folgenden wird noch einmal kompakt und in einer symbolischen Schreibweise die Transformation von Welt- in Kamerakoordinaten festgehalten. Einmal wird wie bereits angefangen mit Spaltenvektoren gearbeitet, das selbe Verfahren wird dann noch einmal mit Zeilenvektoren dargestellt. Beide Ansätze funktionieren nach dem selben Prinzip, der einzige Unterschied ist die Darstellung der Matrizen. Es ist jedoch gerade wenn man mit Programmen wie Beispielsweise *Matlab* arbeitet wichtig zu wissen welche Darstellung benutzt wird und worin sie sich unterscheiden. Auf diese weise passieren keine Fehler beim Interpretieren der späteren Resultate. *MatLab* arbeitet beispielsweise mit Spaltenvektoren, während im entstandenen Algorithmus dieser Arbeit mit Zeilenvektoren gearbeitet wurde. Zunächst wird also weiter mit Spaltenvektoren verfahren. Zur Erinnerung, es gilt, dass $\beta = (\hat{b}_1, \hat{b}_2, \hat{b}_3, C)$ durch Rotation D aus $\delta = (\hat{d}_1, \hat{d}_2, \hat{d}_3)$ entstanden ist. Der Ursprung des Kamerakoordinatensystems ist bezüglich des Ursprungs des Weltkoordinatensystems verschoben. Zur Rotation des Koordinatesystems muss also auch noch eine Tranlation des Ursprungs durchgeführt werden. Der Translationsvektor ist gleich den Koordinaten des Ursprungs des Kamerakoordinatensystem $\vec{V} = C_\beta = (C_{\beta,1}, C_{\beta,2}, C_{\beta,3})^T$. Die Transformationsmatrix welche aus der Rotation D und dem Translationsvektor \vec{V} ensteht, wird im weiteren Verlauf mit R bezeichnet.

$$\begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ C_\beta \end{pmatrix} = \begin{bmatrix} b_{11} & b_{21} & b_{31} & 0 \\ b_{12} & b_{22} & b_{32} & 0 \\ b_{13} & b_{23} & b_{33} & 0 \\ C_{\beta,1} & C_{\beta,2} & C_{\beta,3} & 1 \end{bmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \\ \hat{d}_3 \\ O_\delta \end{pmatrix} \quad (3.16)$$

$$D^T = D^{-1} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \quad (3.17)$$

$$D = \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} \quad (3.18)$$

Für eine Rücktransformation von Kamera in Weltkoordinaten muss die Inverse von D gebildet werden. Außerdem muss der Translationsvektor \vec{V} ebenfalls mit dieser Inversen multipliziert werden, um diesen ebenfalls in das Zielkoordinatensystem zu überführen.

$$\rightsquigarrow \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \\ \hat{d}_3 \\ O_\delta \end{pmatrix} = \begin{bmatrix} b_{11} & b_{21} & b_{31} & 0 \\ b_{12} & b_{22} & b_{32} & 0 \\ b_{13} & b_{23} & b_{33} & 0 \\ -(C_{\beta,1}, C_{\beta,2}, C_{\beta,3})C^{-1} & 1 \end{bmatrix} \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ C_\beta \end{pmatrix} \quad (3.19)$$

Das selbe Verfahren mit Zeilenvektoren führt zu den Gleichungen 3.20 und 3.21.

$$(\hat{b}_1, \hat{b}_2, \hat{b}_3, C_\beta) = (\hat{d}_1, \hat{d}_2, \hat{d}_3, O) \cdot \begin{bmatrix} b_{11} & b_{21} & b_{31} & C_{\beta,1} \\ b_{12} & b_{22} & b_{32} & C_{\beta,2} \\ b_{13} & b_{23} & b_{33} & C_{\beta,2} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.20)$$

Daraus folgt, dass für den Fall der Rücktransformation gilt:

$$\rightsquigarrow (\hat{d}_1, \hat{d}_2, \hat{d}_3, O) = \begin{bmatrix} b_{11} & b_{12} & b_{13} & \\ b_{21} & b_{22} & b_{23} & - \begin{pmatrix} C_{\beta,1} \\ C_{\beta,2} \\ C_{\beta,3} \end{pmatrix} C^{-1} \\ b_{31} & b_{32} & b_{33} & \\ 0 & 0 & 0 & 1 \end{bmatrix} (\hat{b}_1, \hat{b}_2, \hat{b}_3, C_\beta) \quad (3.21)$$

3.2 Aufbau der Koordinatensysteme

Für die geplante Stereoanalyse werden insgesamt sieben verschiedene Koordinatensysteme definiert. Das Weltkoordinatensystem (O, δ) und zwei Kamerakoordinatensysteme (C, β) , welche alle drei zu den dreidimensionalen Koordinatensystemen gehören. Des Weiteren gibt noch vier zweidimensionale Koordinatensysteme. Zum einen die Bildebenenkoordinatensystem (I, τ) und (I', τ') und die Sensorkoordinatensystem (S, σ) und (S', σ') . Alle Koordinatensysteme sind innerhalb dieser Masterarbeit als kartesische, rechtsdrehende Koordinatensysteme festgelegt. Abbildung 3.2 zeigt als Übersicht die Koordinatensysteme von einer Kamera ausgehend im Überblick.

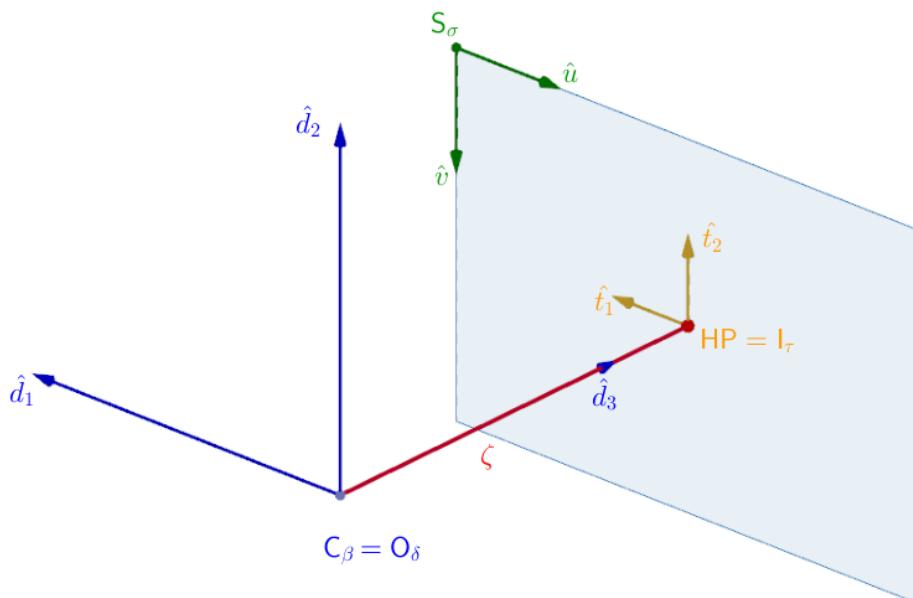


Abbildung 3.2: Koordinatensysteme von einer Kamera aus im Überblick. Das Weltkoordinatensystem (O, δ) ist Deckungsgleich mit dem Kamerakoordinatensystem (C, β)

Das Weltkoordinatensystem (O, δ) ist mit $\delta = (\hat{d}_1, \hat{d}_2, \hat{d}_3, O)$ definiert. Die Kamerakoordinatensysteme der Kameras werden mit (C, β) mit $(\beta = \hat{b}_1, \hat{b}_2, \hat{b}_3, C)$ und (C', β') mit $(\hat{b}_1', \hat{b}_2', \hat{b}_3', C')$ bezeichnet. Die Projektionszentren der Kameras entsprechen den Koordinatenursprüngen C und C' , wie aus dem Lochkameramodell bekannt. Das zweidimensionale Bildebenenkoordinatensystem wird im weiteren Verlauf mit (I, τ) mit $\tau = (\hat{t}_1, \hat{t}_2, I)$ bezeichnet und zu guter Letzt fehlt noch das Sensorkoordinatensystem (S, σ) mit $\sigma = (\hat{u}, \hat{v}, S)$. Der Ursprung des Bildkoordinatensystems ist Deckungsgleich mit dem Hauptpunkt HP auf der Bildebene. Der Ursprung des Sensorkoordinatensystems befindet sich in einer der Ecken der Bildebene. Das Bildebenenkoordinatensystem und das Sensorkoordinatensystem unterscheiden sich darin, dass das Sensorkoordinatensystem an die Geometrie der Pixel auf dem Sensor angepasst ist und somit die Werte der Koordinatenachsen in Pixeleinheiten angegeben werden. Währenddessen sind das Bildebenenkoordinatensystem, sowie auch die anderen hier vorgestellten Koordinatensysteme in dieser Arbeit immer in Millimeter definiert. Das Weltkoordinatensystem wird deckungsgleich mit dem Koordinatensystem von einer der beiden Kameras definiert, es gilt $(C, \beta) = (O, \delta)$.

Nachdem die einzelnen Koordinatensysteme definiert sind, soll symbolisch die Projektion eines Objekts aus dem 3D-Raum auf den 2D-Sensor aufgezeigt werden. Für die Transformation der Weltkoordinaten in Kamerakoordinaten gilt, dass $[\hat{d}_1 \hat{d}_2 \hat{d}_3 1] \cdot R = [\hat{b}_1 \hat{b}_2 \hat{b}_3 1]$, wobei R eine Transformationsmatrix ist, wie sie im vorherigen Abschnitt eingeführt wurde. $R = [D|V]$, wobei D eine 3×3 -Rotationsmatrix darstellt und V die Translationskomponente beschreibt. Der Translationsvektor \vec{V} besteht aus den

Koordinaten des Projektionszentrums C . Es gilt also $\vec{V} = C \rightsquigarrow \vec{v} = \begin{pmatrix} C_{\delta,1} \\ C_{\delta,2} \\ C_{\delta,3} \end{pmatrix}$.

$$(\hat{b}_1, \hat{b}_2, \hat{b}_3, 1) = (\hat{d}_1, \hat{d}_2, \hat{d}_3, 1) \cdot \begin{bmatrix} [D] & C_{\delta,1} \\ 0 & C_{\delta,2} \\ 0 & C_{\delta,3} \\ 0 & 1 \end{bmatrix} \quad (3.22)$$

Ein Punkt M_δ Mit $M_\delta = [X_\delta, Y_\delta, Z_\delta, 1]$ wird mit der soeben aufgestellten Transformationmatrix $R=[D—V]$ zu einem Punkt M_β transformiert.

$$M_\beta = M_\delta \cdot R \quad (3.23)$$

$$[X_\beta \ Y_\beta \ Z_\beta \ 1] = [X_\delta \ Y_\delta \ Z_\delta \ 1] \cdot \begin{bmatrix} [D] & C_{\delta,1} \\ 0 & C_{\delta,2} \\ 0 & C_{\delta,3} \\ 0 & 1 \end{bmatrix} \quad (3.24)$$

Nach der Transformation eines Punktes M_δ in 3D-Kamerakoordinaten zu M_β , erfolgt nun die Projektion des 3D-Kamerapunktes M_β in die 2D-Bildebene zu m_τ . Für die Projektion wird eine 3×3 -Kameramatrix K aufgestellt. Es wird angenommen, dass $\zeta_x = \zeta_y = \zeta$

$$K = {}_{I_\beta} [\pi]_{C_\beta} = \begin{bmatrix} \zeta & 0 & 0 & 0 \\ 0 & \zeta & 0 & 0 \\ 0 & 0 & \zeta & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (3.25)$$

$$\begin{bmatrix} \zeta & 0 & 0 & 0 \\ 0 & \zeta & 0 & 0 \\ 0 & 0 & \zeta & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_\beta \\ Y_\beta \\ Z_\beta \\ 1 \end{bmatrix} = \begin{pmatrix} \zeta X \\ \zeta Y \\ \zeta Z \\ Z \end{pmatrix} = \begin{bmatrix} \zeta \frac{X}{Z} \\ \zeta \frac{Y}{Z} \\ \zeta \\ 1 \end{bmatrix} \quad (3.26)$$

Für die Koordinaten der Bildebene I_β ergeben sich dann aus den Kamerakoordinaten $[\zeta \frac{X}{Z}, \zeta \frac{Y}{Z}, \zeta, 1]^T$. Da die Koordinaten momentan noch in homogenen 3D-Koordinaten angegeben sind, aber das Bildebenenkoordinatensystem ein 2D-Koordinatensystem muss der Tiefenwert noch eliminiert werden. ζ gibt

den Abstand von C zu I an, somit sind die 2D-Bildkoordinaten des die Bildebeneenkoordinatensystem $[\zeta \frac{X}{Z}, \zeta \frac{Y}{Z}, 1]^T = [X_\tau, Y_\tau, 1]$. Zuletzt folgt in der Transformationskette noch die Transformation der Bildebeneenkoordinaten in die Sensorkoordinaten mit dem Sensorkoordinatensystem (S, σ) mit $\sigma = (\hat{u}, \hat{v})$. \hat{u} und \hat{v} definieren die geometrische Beschaffung der Pixel, sie werden auch als Vektoren \vec{u} und \vec{v} angegeben, somit muss das Sensorkoordinatensystem nicht zwangsläufig ein kartesisches Koordinatensystem sein. Die Werte $V_{\sigma,1}$ und $V_{\sigma,2}$ stehen für die Verschiebung des Koordinatensystemursprungs Hauptpunkt auf der Bildebene in eine der Ecken der Bildebene.

$$\vec{u} = u_1 \hat{t}_1 + u_2 \hat{t}_2 \quad (3.27)$$

$$\vec{v} = v_1 \hat{t}_1 + v_2 \hat{t}_2 \quad (3.28)$$

$$S_\sigma = I_\tau + V_{\sigma,1} \hat{t}_1 + V_{\sigma,2} \hat{t}_2 \quad (3.29)$$

$$(\vec{u}, \vec{v}, 1) = (\hat{t}_1, \hat{t}_2, 1) \cdot \begin{bmatrix} u_1 & v_1 & V_{\sigma,1} \\ u_2 & v_2 & V_{\sigma,2} \\ 0 & 0 & 1 \end{bmatrix} \quad (3.30)$$

Der Koordinatenwechsel von Bildebeneenkoordinaten in Sensorkoordinaten wird noch einmal symbolisch veranschaulicht.

Es wird ein Punkt $X_\sigma = \begin{pmatrix} a \\ b \\ 1 \end{pmatrix}$ in Sensorkoordinaten definiert. Die Transfomationsmatrix $M = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix}$ beinhaltet statt einer Rotation eine Skalierung der Koordinatenwerte auf die Koordinaten des Sensorkoordinatensystems. Mit M lässt sich wieder eine Projektionsmatrix P aus einer Abbildungsmatrix K mit einer Transformationsmatrix M aufstellen. Ein Punkt X_τ wird in einen Punkt X_σ transformiert.

$$X_\sigma = \begin{bmatrix} \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix}^{-1} & -M^{-1} \begin{pmatrix} V_{\sigma,1} \\ V_{\sigma,2} \end{pmatrix} \\ 0 & 0 & 1 \end{bmatrix} \cdot X_\tau \quad (3.31)$$

Die Projektionsmatrix P , welche die Kamerakoordinaten in Sensorkoordinaten überführt, wird dann wie folgt aufgebaut:

$$P = \begin{bmatrix} M^{-1} & -M \begin{pmatrix} V_{\sigma,1} \\ V_{\sigma,2} \end{pmatrix}^{-1} \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -\zeta & 0 & 0 & 0 \\ 0 & -\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -\zeta M^{-1} & -M \begin{pmatrix} V_{\sigma,1} \\ V_{\sigma,2} \end{pmatrix}^{-1} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad (3.32)$$

Die symbolische Darstellung der Projektionsmatrix von Welt in Sensorkoordinaten $P = {}_{S_\sigma} [\pi]_{O_\delta}$ fügt sich somit aus dem vorhergehenden Abbildungsmatrix $K = {}_{S_\sigma} [\pi]_{C_\beta}$ mit der Tranformationsmatrix R zusammen.

$${}_{S_\sigma} [\pi]_{O_\delta} = {}_{S_\sigma} [\pi]_{C_\beta} \cdot \begin{bmatrix} [C]^{-1} & -[C]^{-1} Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.33)$$

Somit ist die Transformations-Pipeline eines Punktes im Raum auf den Sensor einer Kamera am Ende. Zum Schluss hier noch einmal zum Vergleich die Darstellung der Projektionsmatrix P aus *Hartley & Zisserman*[3]. Zu beachten ist hier, dass im *Hartley & Zisserman* R für $[R]^{-1}$ steht[3].

$$P = [K|0] \begin{bmatrix} R & -RZ \\ 0 & 0 & 0 & 1 \end{bmatrix} = [KR| -KRZ] = KR[I_{3x3}| -Z] \quad (3.34)$$

4 Homographien

Im vorherigen Kapitel wurde mathematisch die Projektion eines 3D-Punktes in einen 2D-Bildpunkt auf einem Kamerasensor durch Koordinatentransformation aufgezeigt. Im nächsten Schritt soll die Beziehung zwischen dem projizierten 3D-Punkt in zwei perspektivischen Kameras und den entstehenden 2D-Bildpunkten genauer erläutert werden. Die 2D-Bilder des projizierten 3D-Punktes sind abhängig vom Aufbau der abgebildeten Szene und den Positionen und Ausrichtungen der Kameras[5]. Die Beziehung zwischen einem 3D-Punkt und seinen Projektionen auf zwei Kameras, kann in einer sogenannten Homographie beschrieben werden. Die Homographie gilt jedoch nur für die Projektion einer planaren Szene im 3D-Raum auf auf zwei Kamerasensoren. Als Homographie wird eine projektive Transformation zwischen zwei Ebenen bezeichnet. Dabei bleiben Kollinearitäten und die Reihenfolge von Punkten auf Geraden, wie zum Beispiel Schnittpunkte mit anderen Geraden, erhalten. Bei den Ebenen handelt sich in diesem Fall um die Bildebene des Kameramodells. Aufgrund dieser Ebenenannahme, kann solch eine projektive Transformation durch eine 3x3-Homographiematrix ausgedrückt werden[12]. Die Homographiematrix H , projiziert jede Figur in eine Figur gleicher projektiver Entsprechung[3, 5]. Sind die Punkte A', B', C' und D' die projektiven Bilder eines Systems von vier kollinearen Punkten, so ist $(A', B', C', D') = (A, B, C, D)$ [13]. Der Grund warum Homographien in dieser Arbeit eingeführt werden ist, dass sie elementar sind wenn es darum geht Transformationen auf den Bildebene der Kamera durchzuführen, wie es im Kapitel Rektifizierung für die geometrische Entzerrung zweier Bilder von Nöten ist. Außerdem bildet ihre geometrische Herleitung eine gute Brücke zur Epipolar Geometrie, welche im nachfolgenden Kapitel vorgestellt wird und der Elementare Baustein für die Stereobildanalyse bildet. Mit Hilfe der Homography ist es möglich eine Beziehung zwischen den zwei projizierten 2D-Bildpunkten herzuleiten, ohne das intrinsische Parameter in Form einer Kameramatrix K noch die extrinsischen Parameter in Form der Matrix R bekannt sind[3, 5]. Diese können dann wiederum aus der, durch die Bildpunkte beider Kameras, entstandenen Homographiematrix geschätzt werden[3]. Die Homographie fasst die Transformationen wie Rotation und Translation, sowie die jeweiligen Abbildungsmatrizen der Kameras in einer 3x3-Matrix zusammen.[5, 13].

Es seien $m_\tau = \begin{pmatrix} m_{\tau,1} \\ m_{\tau,2} \\ m_{\tau,3} \end{pmatrix}$ die homogenen Koordinaten eines Punktes auf der Bildebene (I, τ) und $m'_{\tau'} = \begin{pmatrix} m'_{\tau',1} \\ m'_{\tau',2} \\ m'_{\tau',3} \end{pmatrix}$ ein Punkt der projektiv transformierten Bildebene (I', τ') . Dann gilt

$$m'_{\tau'} = Hm_\tau \quad (4.1)$$

$$Hm_\tau = \begin{bmatrix} h_1^T \cdot m_\tau \\ h_2^T \cdot m_\tau \\ h_3^T \cdot m_\tau \end{bmatrix} \quad (4.2)$$

$$\rightsquigarrow m'_{\tau'} = Hm_\tau = \begin{bmatrix} h_{11}m_{\tau,1} + h_{12}m_{\tau,2} + h_{13}m_{\tau,3} \\ h_{21}m_{\tau,1} + h_{22}m_{\tau,2} + h_{23}m_{\tau,3} \\ h_{31}m_{\tau,1} + h_{32}m_{\tau,2} + h_{33}m_{\tau,3} \end{bmatrix} \quad (4.3)$$

$$\rightsquigarrow H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \quad (4.4)$$

Dabei müssen die Koeffizienten so geartet sein, dass die zugehörige Transformation umkehrbar ist[3][13].

$$m'_{\tau'} = Hm_\tau \quad (4.5)$$

$$m_\tau = H^{-1}m'_{\tau'} \quad (4.6)$$

(GRAFIK NOCH ANPASSEN MIT m_β etc)

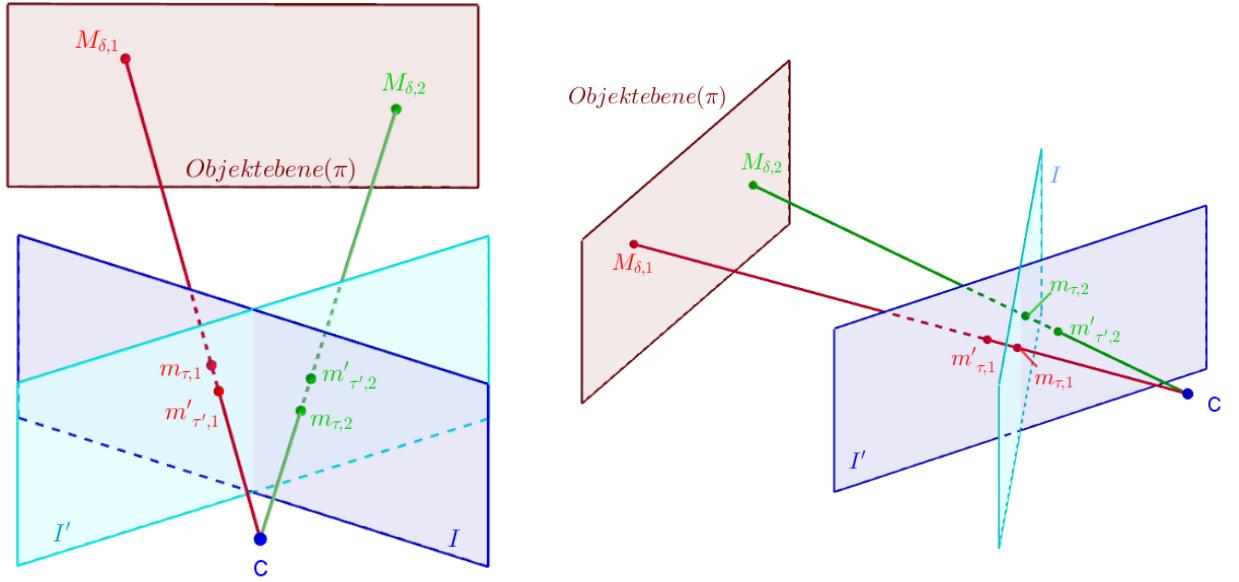


Abbildung 4.1: Die Abbildung zeigt zwei Aufnahmen einer um ihr Projektionszentrum C gedrehten Kamera. Die zwei blauen Ebenen stellen jeweils die Bildebene vor und nach der Drehung dar sie wird einmal als Bildebene (I, τ) und einmal als Bildebene (I', τ') bezeichnet. Die Objektebene befindet sich im 3D-Raum, dementsprechend sind die Punkte $M_{1,\delta}$ und $M_{2,\delta}$ in homogenen 3D-Koordinaten gegeben. Die Punkte $m_{1,\tau}$ und $m_{2,\tau}$ sind die projizierten Punkte auf der Bildebene (I, τ) und Die Punkte $m_{1,\tau'}$ und $m_{2,\tau'}$ sind die projizierten Punkte auf der Bildebene (I', τ')

Um die Herleitung der Homographiematrix zu veranschaulichen, werden zwei unterschiedliche Fälle betrachtet. Der erste Fall ist in Abbildung 4.1 gezeigt. Eine planare Szene im Raum soll auf die Bildebene (I, τ) einer Kamer projiziert werden, danach wird die Kamera um ihr Projektionszentrum gedreht die 3D-Objektpunkte somit auf die neu ausgerichtet Bildebene (I', τ') projiziert. Im zweiten Fall wird die Kamera nicht um ihr Projetionszentrum gedreht, sondern um einen definierten Drehpunkt, so dass zu der Rotation noch eine zusätzliche Translation der Kamera erfolgt. Für Fall eins wird zunächst angenommen, dass die Transformationsmatrizen R und R' sowie die Kameramatrizen K und K' bekannt sind. m_τ und $m'_{\tau'}$ sind die projizierten Punkte von M_δ auf den Bildebenen (I, τ) mit $\tau = (\hat{t}_1, \hat{t}_2, I)$ und (I', τ') mit $\tau' = (\hat{t}'_1, \hat{t}'_2, I')$. Die Kamerakoordinatensysteme werden definiert als (C, β) mit $\beta = (\hat{b}_1, \hat{b}_2, \hat{b}_3, C)$, wobei $\hat{b}_3 = \overline{CI}$ ist und (C', β') mit $\beta' = (\hat{b}'_1, \hat{b}'_2, \hat{b}'_3, C')$, wobei $\hat{b}'_3 = \overline{C'I'}$ [5]. Die folgenden Gleichungen 4.7 und 4.8 zeigen wieder die bereits bekannte Projektion eines 3D-Punktes auf die 2D-Bildebene, jedoch werden die projizierten Punkte m_β und $m'_{\beta'}$ noch bezüglich der dreidimensionalen Kamerakoordinatensysteme dargestellt. γ und γ' repräsentieren jeweils die Tiefe von m_β und $m'_{\beta'}$, welche gegen Ende eliminiert wird [5].

$$\gamma \vec{m}_\beta = P \begin{bmatrix} \vec{M}_\delta \\ 1 \end{bmatrix} = \left[KR | -KR\vec{C}_\delta \right] \cdot \begin{bmatrix} \vec{M}_\delta \\ 1 \end{bmatrix} = KR(\vec{M}_\delta - \vec{C}_\delta) \quad (4.7)$$

$$\gamma' \vec{m}'_{\beta'} = P' \begin{bmatrix} \vec{M}'_\delta \\ 1 \end{bmatrix} = \left[K'R' | -K'R'\vec{C}_\delta \right] \cdot \begin{bmatrix} \vec{M}_\delta \\ 1 \end{bmatrix} = K'R'(\vec{M}_\delta - \vec{C}_\delta) \quad (4.8)$$

Der Vektor $(\vec{M}_\delta - \vec{C}_\delta)$, beschreibt die Verbindung zwischen dem Objektpunkt M_δ und dem Projektionszentrum C .

onszentrum C_δ . Die Gleichungen 4.7 und 4.8 können auf Grund ihrer identischen Projektionszentren jeweils nach $(\vec{M}\delta - \vec{C}_\delta)$ aufgelöst werden

$$\gamma R K^{-1} \vec{m}_\beta = \vec{M}_\delta - \vec{C}_\delta \quad (4.9)$$

$$\gamma' R' K'^{-1} \vec{m}'_{\beta'} = \vec{M}_\delta - \vec{C}_\delta \quad (4.10)$$

Da die Rechten seiten der Gleichungen 4.9 und 4.10 identisch sind, können die linken Seiten gleichgesetzt werden.

$$\gamma' R' K'^{-1} \vec{m}'_{\beta'} = \gamma R K^{-1} \vec{m}_\beta \quad (4.11)$$

$$(4.12)$$

Um nun die gewünschte Form von $m'_{\tau'} = H m_\tau$ zu bekommen, wird die Gleichung nach $\frac{\gamma'}{\gamma} \vec{m}'_{\beta'}$ aufgelöst.

$$\frac{\gamma'}{\gamma} \vec{m}'_{\beta'} = K' R' R^T K^{-1} \vec{m}_\beta \quad (4.13)$$

$$\lambda \vec{m}'_{\beta'} = H \vec{m}_\beta \quad (4.14)$$

$$\leadsto m'_{\tau'} = H m_\tau \quad (4.15)$$

Es Resultiert also dass wenn $\lambda = \frac{\gamma'}{\gamma}$, dann gilt dass $H = K' R' R^T K^{-1}$ gleicht. Gilt nun noch der Fall, dass $K = K'$ ist, dann reduziert sich H zu $K^{-1} H K = R' R$, welches einer Rotation gleicht.[5].

4.1 Abbildungsunterschiede bei verschobenen Rotationsachsen

Wie am Anfang des Kapitels erwähnt, sollen zwei verschiedene Fälle für die Homographie betrachtet werden. Der Grund sind, die unterschiedlichen Abbildungsmerkmale die bei einer Drehung um das Projektionszentrum und einer Drehung um einen definierten Drehpunkt entstehen. Die Frage ist, ob und wie sich beim zweiten Fall eine Homographiematrix herleiten lässt und in wie fern sie sich von der bereits bekannten Herleitung unterscheidet. Um zu veranschaulichen, was genau sich bei einer Drehung um ein Drehpunkt und der Drehung um das Projektionszentrum ändert, wurde eine Simulation geschrieben, welche die Abbildungsunterschiede beider Drehungen aufzeigt. Abbildung 4.2 zeigt die Ausgangssituation der Simulation. In einem virtuellen Raum wurde eine Kamera platziert, welche auf ein Objekt gerichtet ist. Dieses Objekt besteht aus einem Quadrat, das durch die schwarzen Eckpunkte begrenzt ist. In der Mitte des Quadrats wurde mit Rot der Mittelpunkt des Quadrats gekennzeichnet. Dieser Mittelpunkt symbolisiert den Drehpunkt um welchem sich die Kamera später drehen soll.

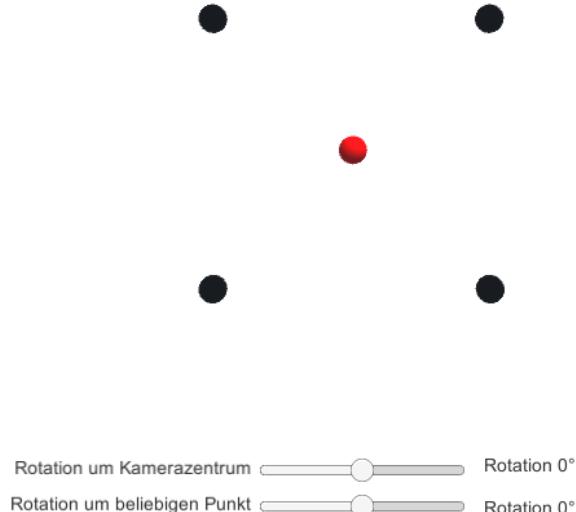


Abbildung 4.2: Ausgangssituation der Simulation

Für die Simulation der unterschiedlichen Drehungen wurden zwei Schieberegler implementiert mit denen sich die Kamera einmal um ihr Projektionszentrum, und einmal um den Drehpunkt drehen lässt. Abbildung 4.3 zeigt jeweils die entstehenden Bilder, wenn die Kamera um 20° beziehungsweise -20° um das Projektionszentrum gedreht wurde.

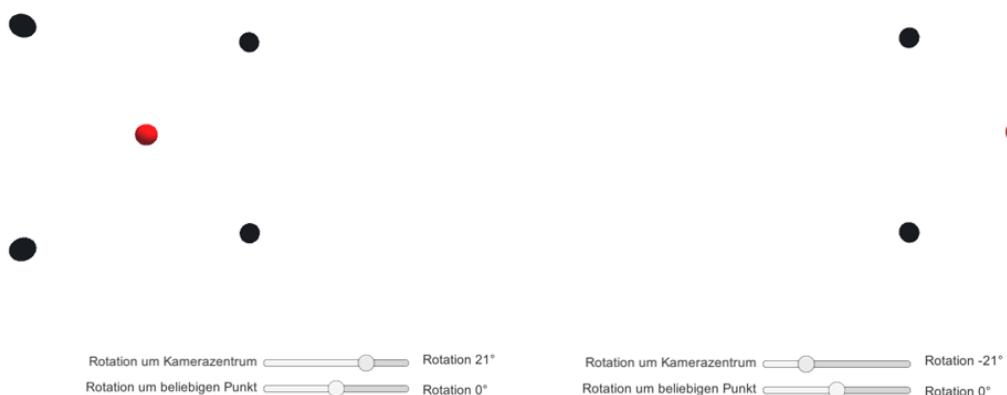


Abbildung 4.3: Drehung um das Projektionszentrum

Abbildung 4.4 zeigt die entstehenden Bilder, wenn die Kamera um 45° beziehungsweise -45° um den Drehpunkt gedreht wurde. Wie sich zeigt ist hier ein weiterer grüner Punkt zu sehen welcher zu Testzwecken hinter dem roten Punkt platziert wurde sichtbar.

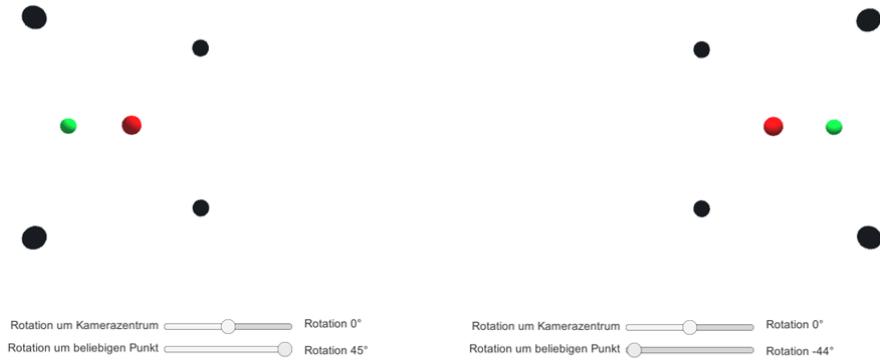


Abbildung 4.4: Drehung um einen Drehpunkt. In diesem Beispiel wurde der rote Punkt als Drehpunkt verwendet

Punkte die bei einer Drehung um das Projektionszentrum verdeckt bleiben, können bei einer Drehung um einen außerhalb der Kamera liegenden Drehpunkt sichtbar werden. Die Abbildungen 4.5 und 4.6 veranschaulichen anhand von Strahlenoptik, was der Grund für die Verdeckung ist.

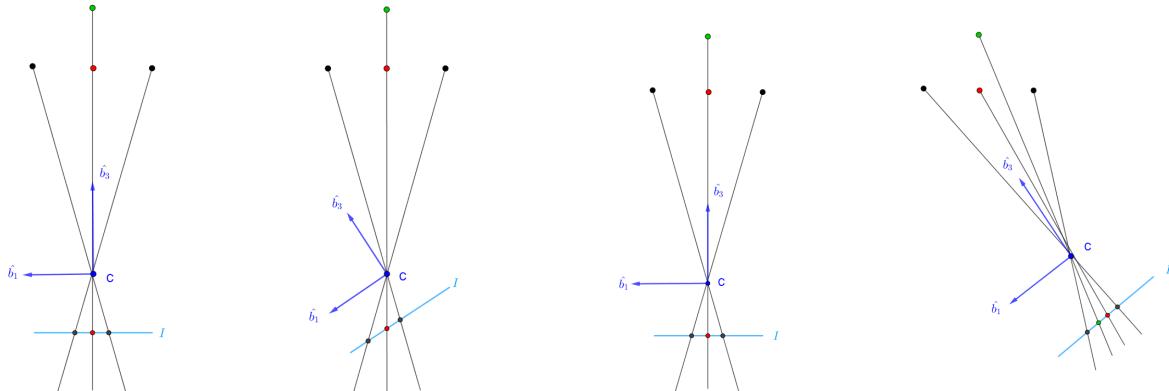


Abbildung 4.5: Strahlengang durch das Projektionszentrum. Auf der Abbildung ist erkennbar, dass der grüne Punkt auch nach der Drehung der Kamera um das Projektionszentrum vom roten Punkt verdeckt bleibt

Abbildung 4.6: Strahlengang durch das Projektionszentrum. Auf der Grafik ist erkennbar, dass der grüne Punkt nach der Drehung der Kamera um einen Drehpunkt, welcher in diesem Fall der rote Punkt darstellt, sichtbar wird.

Die Bildebene ist in diesem Beispiel hinter dem Projektionszentrum platziert. Die Lage der Bildebene ist in sofern egal, dass es sei sie im gleichen Abstand vor oder hinter dem Projektionszentrum, immer zur selben Abbildung kommt. Jedoch ist die Abbildung, bei hinter dem Projektionszentrum platzierten Bildebene, auf dem Kopf abgebildet. Die Positionierung hinter das Projektionszentrum dient allein der Übersichtlichkeit der Abbildungen, im restlichen Verlauf der Arbeit wird die Bildebene immer vor das

Projektionszentrum platziert um Verwirrungen zu vermeiden. Der grüne Punkt wurde nur als Beweis platziert, dass sich die entstehenden Abbildungen beider Fälle voneinander unterscheiden. Für den weiteren Verlauf wird er nicht weiter beachtet. Für die Homographie gelten nur die Punkte, welche sich im Raum auf einer Ebene befinden. Im folgenden wird eine Herleitung der Homographiematrix für den zweiten Fall aufgestellt.

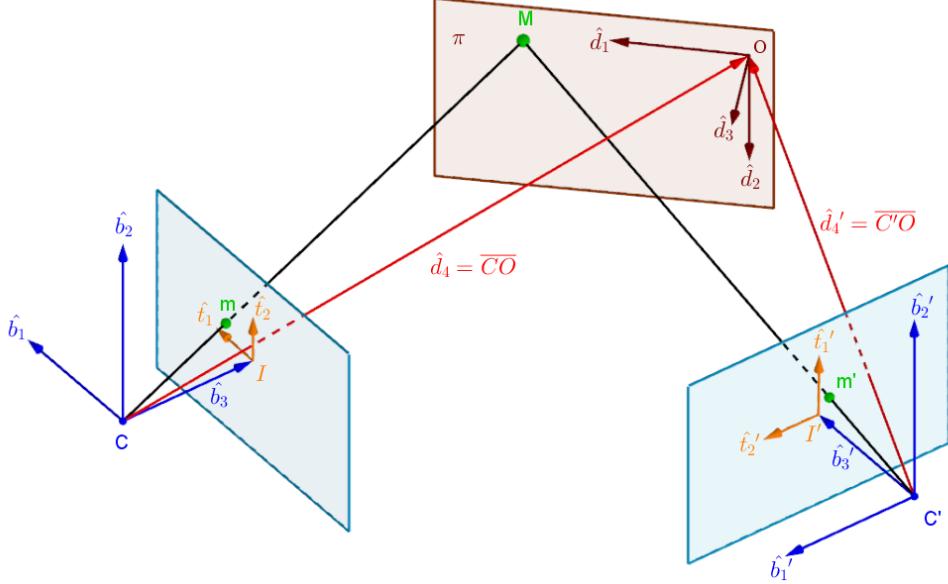


Abbildung 4.7: Veranschaulichung der Homographie bei zwei verschiedenen translatierten und rotierten Kameras.

Angenommen es existiert eine Ebene π im Raum mit einem Objektpunkt M_δ bezüglich eines Weltkoordinatensystems (O, δ) mit $\delta = (\hat{d}_1, \hat{d}_2, \hat{d}_3, O)$. \hat{d}_1 und \hat{d}_2 spannen die Ebene π auf. Der Punkt M_δ wird nun mit den Projektionsmatrizen $P = [KR] - KRC\vec{\delta}$ und $P' = [K'R'] - K'R'C\vec{\delta}$ in einen Punkt bezüglich der Bildebenen (I, τ) mit $\tau = (\hat{t}_1, \hat{t}_2, I)$ und (I', τ') mit $\tau' = (\hat{t}_1', \hat{t}_2', I')$ projiziert. Auch hier gilt wie im vorherigen Beispiel, dass die projizierten Bildkoordinaten als m_β und $m'_{\beta'}$ der jeweiligen Kamerakoordinatensysteme (C, β) mit $\beta = (\hat{b}_1, \hat{b}_2, \hat{b}_3, C)$, wobei $\hat{b}_3 = \overline{CI}$ und (C', β') mit $\beta' = (\hat{b}_1', \hat{b}_2', \hat{b}_3', \overline{C'I'})$, wobei $\hat{b}_3' = \overline{C'O}$, dargestellt werden. Da der Ursprung des Weltkoordinatensystems (O, δ) auf der Ebene π liegt und \hat{d}_1 und \hat{d}_2 diese aufspannen, gilt für den Punkt M_δ auf π

$$M_\delta = \begin{pmatrix} x_{\hat{d}_1} \\ y_{\hat{d}_2} \\ 0 \end{pmatrix}.$$

die Bezeichnungen p_i und p'_i stehen im folgenden für die einzelnen Spalten der Projektionsmatrizen P und P' . Für die Projektion des Punktes M_δ in die Bildpunkte m_β und $m'_{\beta'}$ bezüglich (C, β) und (C', β') gelten die folgenden beiden Gleichungen[5].

$$\gamma m_\beta = P \cdot \begin{bmatrix} M_\delta \\ 1 \end{bmatrix} = [p_1 \ p_2 \ p_3 \ p_4] \cdot \begin{bmatrix} x_\delta \\ y_\delta \\ 0 \\ 1 \end{bmatrix} = [p_1 \ p_2 \ p_4] \cdot \begin{bmatrix} x_\delta \\ y_\delta \\ 1 \end{bmatrix} = G \vec{m}_\tau \quad (4.16)$$

$$\gamma' m'_{\beta'} = P' \cdot \begin{bmatrix} M_\delta \\ 1 \end{bmatrix} = [p'_1 \ p'_2 \ p'_3 \ p'_4] \cdot \begin{bmatrix} x_\delta \\ y_\delta \\ 0 \\ 1 \end{bmatrix} = [p'_1 \ p'_2 \ p'_4] \cdot \begin{bmatrix} x'_\delta \\ y'_\delta \\ 1 \end{bmatrix} = G' \vec{m}'_{\alpha} \quad (4.17)$$

(4.18)

In diesem Prozess sind sowohl die zwei Matrizen G und G' , sowie zwei weitere Koordinatensysteme (C, α) mit $\alpha = (\hat{d}_1, \hat{d}_2, \hat{d}_4)$, wobei $\hat{d}_4 = \overline{CO}$ und (C', α') mit $\alpha' = (\hat{d}_1, \hat{d}_2, \hat{d}'_4)$, wobei $\hat{d}'_4 = \overline{C'O}$ entstanden[5]. Des Weiteren sind zwei neue Punkte m_α und m'_α entstanden, von denen behauptet wird, dass $\gamma m_\beta = G \cdot m_\alpha$ und $\gamma' m'_{\beta'} = G' \cdot m'_{\alpha'}$ gelte[5]. Um dies zu testen werden m_α und m'_α anders formuliert.

$$m_\alpha = M_\alpha + \overline{CO}_\alpha = M_{(\hat{d}_1, \hat{d}_2, \hat{d}_4)} + \overline{CO}_{(\hat{d}_1, \hat{d}_2, \hat{d}_4)} = \begin{pmatrix} x_\delta \\ y_\delta \\ 1 \end{pmatrix} \quad (4.19)$$

$$m'_{\alpha'} = M_{\alpha'} + \overline{CO}_{\alpha'} = M_{(\hat{d}_1, \hat{d}_2, \hat{d}'_4)} + \overline{CO}_{(\hat{d}_1, \hat{d}_2, \hat{d}'_4)} = \begin{pmatrix} x_\delta \\ y_\delta \\ 1 \end{pmatrix} \quad (4.20)$$

Aus Gleichungen 4.18 und 4.19 folgt dann, dass für alle Punkte in der Ebene π gilt, dass $m_\alpha = m'_\alpha$. Woraus für die Punkte γm_β und $\gamma' m'_{\beta'}$ folgt, dass $\gamma m_\beta = G \cdot m_\alpha$ und $\gamma' m'_{\beta'} = G' \cdot m'_{\alpha'}$ [5]. Somit kann für γm_β und $\gamma' m'_{\beta'}$ folgende Beziehungsgleichung aufgestellt werden[5].

$$\gamma' m'_{\beta'} = G' G^{-1} \gamma m_\beta \quad (4.21)$$

Für $\frac{\gamma'}{\gamma} = \lambda$, kann Gleichung 4.20 dann wieder umformuliert werden in:

$$\lambda m'_{\beta'} = H m_\beta \quad (4.22)$$

$$\rightsquigarrow m'_{\tau'} = H m_\tau \quad (4.23)$$

Zwei projizierte Bilder einer Ebene sind also über eine Homographie H miteinander verwandt, welche wiederum eine Kombination aus denjenigen Homographien ist, welche die jeweiligen Bilder mit der Ebene π in Beziehung setzten[5].

4.2 Berechnung der Homographiematrix aus Bildebenenkoordinaten

Nachdem nun bekannt ist, wie eine Homographie Beziehungen zwischen projizierten Punkten beschreiben kann, wird im nächsten Schritt davon ausgegangen, dass die Transformationsmatrizen R und R' sowie die Kameramatrizen K und K' nicht bekannt sind. Es wird schematisch aufgezeigt, wie die Homographiematrix aus 2D-Bildpunkten gewonnen werden kann. Um eine Homographiematrix mit

$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$ zu erhalten werden die Punkte beider Kameras in eine Koeffizientenmatrix A eingetragen, welche sich nach dem in den Gleichungen 4.25 bis 4.34 laufenden Schema aufstellen lässt.

$$H \cdot m_\tau = m'_{\tau'} \quad (4.24)$$

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} m_\tau \end{bmatrix} = \begin{bmatrix} m'_{\tau'} \end{bmatrix} \quad (4.25)$$

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad (4.26)$$

Aus Gleichung 4.24 lässt sich ein Gleichungssystem mit zwölf bekannten und neun unbekannten aufstellen.

$$h_{11}x + h_{12}y + h_{13}z = \lambda x' \quad (4.27)$$

$$h_{21}x + h_{22}y + h_{23}z = \lambda y' \quad (4.28)$$

$$h_{31}x + h_{32}y + h_{33}z = \lambda z' \quad (4.29)$$

Da mit homogenen Koordinaten gearbeitet wird und somit z und $z' = 1$ sind, ergibt sich für die letzte Zeile $h_{31}x + h_{32}y + h_{33}z = 1$. Dieser Ausdruck kann in den ersten beiden Gleichungen für λ eingesetzt werden. Pro Punktpaar m_τ und m'_τ , ergeben sich somit zwei Gleichungen.

$$h_{11}x + h_{12}y + h_{13}z = (h_{31}x + h_{32}y + h_{33}z) \cdot x' \quad (4.30)$$

$$h_{21}x + h_{22}y + h_{23}z = (h_{31}x + h_{32}y + h_{33}z) \cdot y' \quad (4.31)$$

Für den Aufbau von A werden beide Ausdrücke noch nach Null aufgelöst, so dass sich die Gleichungen 4.34 und 4.35 aus 4.32 und 4.33 ergeben.

$$h_{11}x + h_{12}y + h_{13}z - (h_{31}x + h_{32}y + h_{33}z) \cdot x' = 0 \quad (4.32)$$

$$h_{21}x + h_{22}y + h_{23}z - (h_{31}x + h_{32}y + h_{33}z) \cdot y' = 0 \quad (4.33)$$

$$\rightsquigarrow h_{11}x + h_{12}y + h_{13}z - h_{31}x \cdot x' - h_{32}y \cdot x' - h_{33}z \cdot x' = 0 \quad (4.34)$$

$$\rightsquigarrow h_{21}x + h_{22}y + h_{23}z - h_{31}x \cdot y - h_{32}y \cdot y - h_{33}z \cdot y = 0 \quad (4.35)$$

Die entstandenen Gleichungen werden jetzt pro Punktpaar m_τ und m'_τ , in die Koeffizientenmatrix A nach folgendem Schema eingetragen.[5, 3, 14, 8]

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & x_1x'_1 & y_1x'_1 & 1 \cdot x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & x_1y'_1 & y_1y'_1 & 1 \cdot y'_1 \\ & & & \ddots & & & & & \\ & & & \ddots & & & & & \\ & & & \ddots & & & & & \\ x_i & y_i & 1 & 0 & 0 & 0 & x_ix'_i & y_ix'_i & 1 \cdot x'_i \\ 0 & 0 & 0 & x_i & y_i & 1 & x_iy'_i & y_iy'_i & 1 \cdot y'_i \end{pmatrix} \cdot \begin{pmatrix} h1 \\ h2 \\ \vdots \\ hi \end{pmatrix} = 0 \quad (4.36)$$

Wenn ein nicht überbestimmter Fall vorliegt, sprich wenn der Matrixrang von $A \leq 8$ beträgt, kann aus A einfach der Kern bestimmt werden, um so die Einträge für die 3x3-Homographiematrix zu erhalten[3, 5, 14]. Gesucht wird also ein Vektor \vec{x} , für den gilt das $H \cdot x = 0$. Der gesuchte Vektor \vec{x} entspricht dem Kern der Koeffizientenmatrix und ist ein Spaltenvektor mit insgesamt neun Einträgen, welche in die 3x3-Homographiematrix eingetragen werden können[3, 14]. Tritt nun der Fall ein, dass es zu einem überbestimmten System kommt, sprich wenn der Rang von A größer 8 ist, so liefert die Bestimmung des Kerns kein eindeutiges Ergebnis mehr für den Vektor \vec{x} und somit kann die Homographiematrix nicht eindeutig bestimmt werden[3]. Für die Lösung überbestimmter Systeme wird die Singulärwertzerlegung angewandt[3][15]. Das bedeutet es wird nicht mehr derjenige Vektor \vec{x} gesucht für den gilt $H \cdot x = 0$ ergibt, sondern es wird derjenige Vektor \vec{x} gesucht, für den $\| H \cdot x \|$ minimal wird[3, 14]. Die Singulärwertzerlegung von A ist eine Faktorisierung der Matrix $A \in \mathbb{R}^{m \times n}$ der Form $A = U \cdot S \cdot V^T$ mit orthogonalen Matrizen $U \in \mathbb{R}^{m \times n}$ und $V \in \mathbb{R}^{m \times n}$ sowie mit einer Diagonalmatrix S .

$$S = \begin{pmatrix} s_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & \dots & & s_r & 0 & \dots & 0 \\ 0 & \dots & & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & & \vdots \\ 0 & \dots & & 0 & 0 & \dots & 0 \end{pmatrix} \quad (4.37)$$

Dabei soll für die diagonalen Singulärwerte in S mit s_1 bis s_r gelten, dass $s_1 \geq s_2 \geq \dots \geq s_r \geq 0$ [15]. Nach dem Schema der Singulärwertszerlegung, wird Matrix A in die Form USV^T zerlegt. Nach der Zerlegung sind die Diagonaleinträge von S in einer absteigenden Reihenfolge sortiert. Die Spalte von V^T , welche mit dem kleinsten Singulärwert von S korrespondiert, ergibt den Vektor \vec{x} , für den $\| H \cdot x \|$ minimal wird. Somit gleichen die neun Einträge der Homographiematrix gleich der letzten Spalte von V^T . Beispielrechnungen, in denen die Homographiematrix einmal für den Fall des rotierten Projektionszentrums und einmal mit der Drehung um einen Drehpunkt mit durchgerechnet wurde befindet sich im Appendix unter 11.1 und 11.2.

5 Epipolar Geometrie

Kurze zusammenfassung bis hier her... Bis hier her wurde Dann sagen für was wird jetzt die epipolar-geometrie benötigt

Die Epipolargeometrie beschreibt ähnlich der Homographie eine Beziehung der projektiven Geometrie zwischen zwei Bildern[3]. Sie dient insbesondere zur Korrespondenzanalyse von Punkten aus Bildern und zur Gewinnung von 3D-Informationen über die abgebildete Szene. Im Gegensatz zur Homographie ist die Epipolargeometrie unabhängig vom Szenenaufbau, sondern hängt nur von den intrinsischen Parametern der Kamera, sowie deren relative Position im Raum ab.[3]. Ohne Kenntnis der Kamera-positionen, kann mit Hilfe der Epipolargeometrie überprüft werden, ob eine Korrespondenz zwischen den Bildpunkten zweier Bilder besteht. Unter korrespondierenden Punkten versteht man Bildpunkte auf zwei Bildern, die Abbilder des selben Objektpunktes im 3D-Raum sind. Ein Beispiel für korrespondierende Punkte wären die Punkte m_τ und $m'_{\tau'}$ aus den Kapiteln davor. Diese entstanden aus der Projektion von M_δ auf zwei Bildebenen mit $m_\tau = PM_\delta$ und $m'_{\tau'} = P'M_\delta$. Die Verbindungsvektoren zwischen den Projektionszentren C und C' mit dem 3D-Szenenpunkt M_δ und die Basislinie bilden die Epipolarebene. Auf Abbildung 5.1 ist sie als Dreieck zwischen den Punkten zu sehen. Die Epipolargeometrie beschreibt in erster Linie die Geometrie zwischen den durch Schnitt der Bildebenen und der Epipolarebene entstehenden Schnittgeraden. Für jeden Objektpunkt im Raum, entsteht eine neue solche Ebene, die alle die Basislinie als Drehachse besitzen[3]. Die Schnittgeraden werden als Epipolarlinien l und l' bezeichnet. l' wird als die zu zum Bildpunkt m_τ korrespondierende Epipolarlinie bezeichnet und anders herum ist l die zum Bildpunkt $m'_{\tau'}$ korrespondierende Epipolarlinie. Die Epipolarlinien eines Bildes verlaufen alle durch einen Epipol e oder e' und den abgebildeten Bildpunkten und bilden somit ein Bündel. Epipol e ist die Abbildung von C' auf der Bildebene I von C und Epipol e' ist die Abbildung von C auf Bildebene I' von C' . Die Schnittpunkte der Basislinie mit den Bildebenen ergeben die jeweiligen Epipole. Ist eine der Bildebenen parallel zur Basislinie ausgerichtet, so kommt es zu keinem Schnittpunkt der Basislinie mit der parallelen Bildebene und der Epipol befindet sich somit im unendlichen[3, 1].

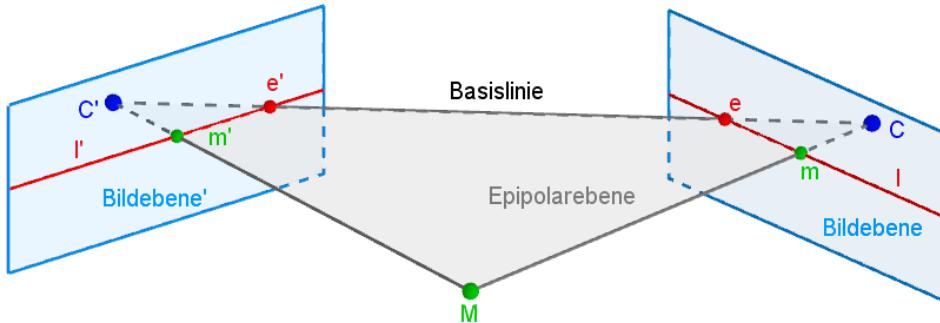


Abbildung 5.1: C und C' sind die Projektionszentren zweier Kameras. Beide Kameras besitzen jeweils eine Bildebene. Die Basislinien verbindet die Projektionszentren der Kameras. Der Punkt an welchem die Basislinie die Bildebenen schneidet, wird als Epipol bezeichnet. Durch den Epipol verlaufen alle Epipolarlinien des Bildes. M ist der Objektpunkt im 3D-Raum und m_1 und m_2 sind die jeweiligen Abbildungen dieses Punktes auf den Bildebenen. Die Verbindungsvektoren zwischen C, C' und M bilden die sogenannte Epipolarebene[5, 3, 1].

Ein Objektpunkt M_δ wird auf die Bildebenen I und I' der beiden Kameras C und C' projiziert. Es

entstehen die Bildpunkte m_τ und m'_τ . Die durch die Bildpunkte m_τ und m'_τ und den entsprechenden Epipolen e und e' verlaufenden Epipolarlinien, sind zueinander korrespondierende Epipolarlinien. Die zum Punkt m_τ korrespondierende Epipolarlinie l' beschreibt alle möglichen zu m_τ korrespondierenden Punkte, darunter ist auch der eindeutig korrespondierende Punkt m'_τ . Die Epipolargeometrie beschreibt eine Beziehung zwischen einem Bildpunkt m und dessen korrespondierender Epipolarlinie l' , welche wiederum alle möglichen zu m korrespondierenden Punkte m' beinhaltet. Diese Beziehung ist der sogenannte *Epipolar-Constraint* [3, 11, 1]. Abbildung 5.2 veranschaulicht die soeben genannte Beziehung eines Bildpunktes zu seiner korrespondierenden Epipolarlinie noch mal grafisch. Formal soll der Constraint aussagen, dass wenn ein 3D-Bildpunkt M_δ entlang seiner Verbindungslinie $\overline{CM_\delta}$ auf die Bildebene I zu bewegt wird, so entsteht auf dieser immer der selbe Bildpunkt, welcher in Abbildung 5.2 mit m_i bezeichnet wird. Auf der Bildebene I' von C' wiederum entstehen immer andere Abbildungen $m'_{\tau,1}$ bis $m'_{\tau,2}$ entlang der Epipolarlinie l' . Ist also nur Bildpunkt m_τ bekannt, so können alle Punkte auf dessen korrespondierender Epipolarlinie mögliche korrespondierende Punkte zu m_τ sein.

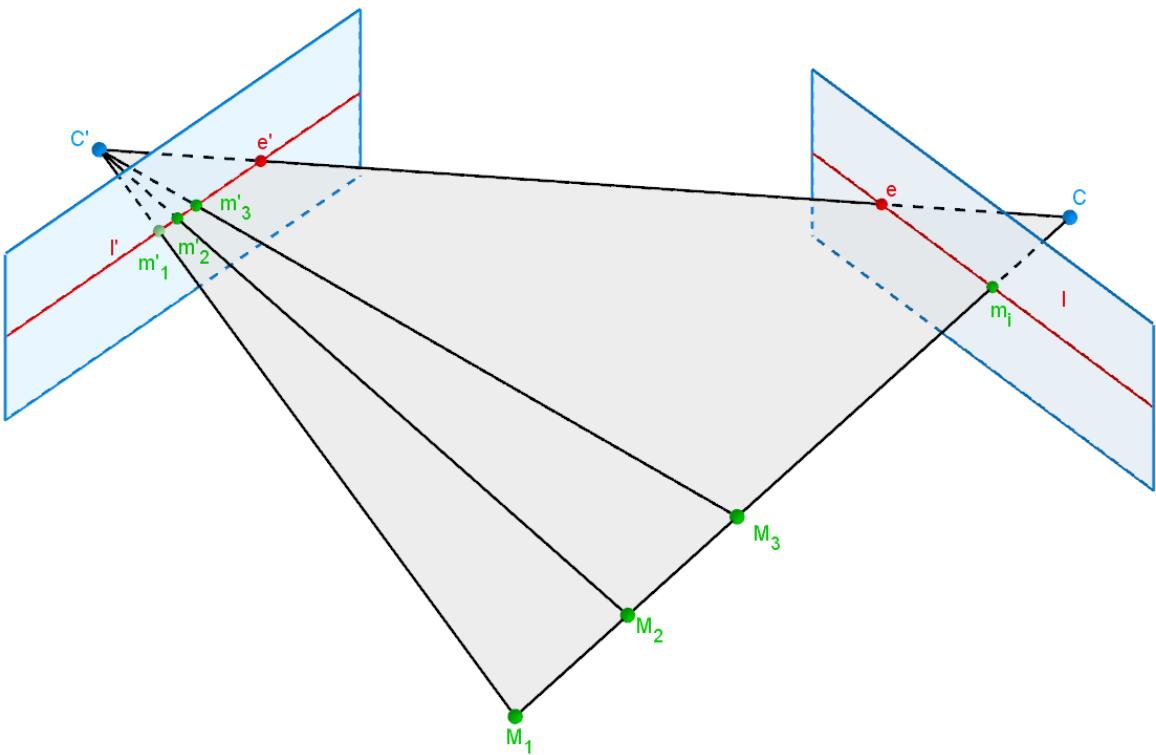


Abbildung 5.2: Die Objektpunkte M_1, M_2 und M_3 werden in I' als m'_1, m'_2 und m'_3 abgebildet, während sie in I immer den selben Bildpunkt m_1 ergeben.

(Ab hier weiter machen !!) Die Epipolargeometrie lässt sich, ähnlich wie eine Homographiematrix, in einer 3×3 -Matrix zusammenfassen. Diese ist jedoch singulär und besitzt somit nicht wie die Homographiematrix Rang 3 sondern Rang 2. Je nachdem ob ein ein kalibriertes oder unkalibriertes System vorliegen hat, handelt es sich entweder um die sogenannten Fundamental Matrix F oder die Essentielle Matrix E [5, 3, 16, 11, 1]. Von der Fundamentalmatrix F ist dann die rede, wenn die intrinsischen Kameraparameter nicht bekannt sind, sprich wenn das System unkalibriert ist. In diesem Fall wird mit Bildpixelkoordinaten gearbeitet. Sind die intrinsischen Parameter bekannt, so wird F zur essentiellen Matrix E und es wird mit sogenannten normalisierten Bildkoordinaten gearbeitet [16]. Was genau die unterschiedlichen Koordinaten ausmacht, wird in Kapitel Minimalbeispiel 3D-Stereokalibrierung und Szenenrekonstruktion bei Kameras gleicher Auflösung genauer erklärt. Mathematisch sagen die Matrizen F und E in Verbindung mit den korrespondierenden Punkten aus, ob für einen Bildpunkt m in einer Kamera, dessen korrespondierender Bildpunkt m' in der anderen Kamera auf der korrespon-

dierenden Epipolarlinie liegt. Das heißt, wenn $m'Fm = 0$ oder $\hat{m}'E\hat{m} = 0$, dann ist der sogenannte *epipolar-constraint* erfüllt. Der *epipolar-constraint* gibt somit Aufschluss darüber, ob m' ein möglicher korrespondierender Punkt von m ist. Dies ist nämlich genau dann der Fall, wenn $m'Fm = 0$ oder $\hat{m}'E\hat{m} = 0$ sind[3, 11]. Ist der *epipolar-constraint* erfüllt, so wird gleichzeitig der Suchaufwand nach weiteren Korrespondenzen reduziert, da somit nur noch eine eindimensionale Suche, entlang der Epipolarlinie, anstatt einer zweidimensionalen durchgeführt werden muss. Dieser neue *constraint* wird auch als *coplanarity constraint* oder Koplanaritätsbeschränkung bezeichnet. Dieser entsteht, da die Projektionszentren der Kameras und die korrespondierenden Bildpunkte auf ein und der selben Ebene liegen müssen [11]. Die Epipolargeometrie und die in ihr beinhalteten *constraints*, helfen bei der 3D-Szenenrekonstruktion. Szenenrekonstruktion ist dann möglich, wenn in einer Stereoaufnahme in beiden Bildern die zueinander gehörenden Bildpunkte lokalisiert wurden. Wird also eine Szene mit zwei Kameras aufgenommen, so liegen während der Aufnahme die aufgenommenen Objektpunkte, das Projektionszentrum und der zur Kamera gehörende Bildpunkt auf einer Linie. Wurde eine Objektpunkt nun zweimal aus verschiedenen Winkeln und/ oder Position aufgenommen, lassen sich nachdem die extrinsischen Parameter der Kameras ermittelt wurden, die Schnittpunkt der jeweiligen Linien aus Kamera eins und Kamera zwei berechnen. Diese Schnittpunkte ergeben den ursprünglichen Objektpunkt. Die Szene ist somit rekonstruiert[5, 1, 3].

5.1 Geometrische Erläuterung der Fundamentalmatrix und der Essentiellen Matrix

Nachdem die Theorie der geometrischen Hintergründe der Epipolargeometrie, bei der Stereokalibrierung und Szenenrekonstruktion, erläutert wurden, wird nun der mathematische Hintergrund genauer aufgezeigt. Vor allem soll auf die Herleitung der neu eingeführten Fundamental Matrix F und der essentiellen Matrix E eingegangen werden. Diese spielen nämlich eine entscheidende Rolle bei der Rekonstruktion der Kamerapose und der Szenenrekonstruktion[5, 3]. F und E bilden jeweils eine singuläre 3x3-Matrix, welche die Geometrie zwischen den Bildpunkten m_τ und m'_τ auf I und I' und dem Objektpunkt M_δ im Raum beschreibt. Die Vektoren $\overline{CM} = (\vec{M}_\delta - \vec{C}_\delta)$, $\overline{C'M} = (\vec{M}_\delta - \vec{C}'_\delta)$ und $\overline{CC'} = (\vec{C}'_\delta - \vec{C}_\delta)$ bilden das in Abbildung 5.1 sichtbare schwarze Dreieck. F und E fassen dieses Dreieck in ihren Matrizen zusammen. Um das ganze mathematisch zu erklären, wird ein Stereokamerabau definiert. Ein Objektpunkt M_δ in Weltkoordinaten(O, δ) wird von zwei Kameras C mit (C, β) und C' mit (C', β') aufgenommen und auf deren Bildebenen I und I' als m_β und m'_β abgebildet. C und C' besitzen jeweils eine eigene Projektionsmatrix P und P' . Anzumerken ist, dass die folgende Herleitung nach [5] aufgestellt wurde.

$$P = \begin{bmatrix} KR | -KR\vec{C}_\delta \end{bmatrix} \quad (5.1)$$

$$P' = \begin{bmatrix} K'R' | -K'R'\vec{C}'_\delta \end{bmatrix} \quad (5.2)$$

M wird mit P und P' auf die Bildebenen I und I' mit den jeweiligen Koordinatensystemen $I = (I, \tau)$ und $I' = (I', \tau')$ projiziert. Wichtig anzumerken, auch für den späteren Aufbau mit zwei Kameras unterschiedlicher Auflösung, ist, dass es sich bei den Koordinatensystemen von I und I' nicht um identische handeln muss.[5] Es entstehen die Bildpunkte γm_τ und $\gamma' m'_{\tau'}$ mit $\gamma \geq 0$ und $\gamma' \geq 0$. (gamma erklären)

$$\gamma \vec{m}_\tau = P \begin{bmatrix} \vec{M}_\delta \\ 1 \end{bmatrix} \quad (5.3)$$

$$\gamma \vec{m}_\tau = [KR| - KR\vec{C}_\delta] \begin{bmatrix} \vec{M}_\delta \\ 1 \end{bmatrix} \quad (5.4)$$

$$\gamma' \vec{m}'_\tau = P' \begin{bmatrix} \vec{M}_\delta \\ 1 \end{bmatrix} \quad (5.5)$$

$$\gamma' \vec{m}'_\tau = [K'R'| - K'R'\vec{C}'_\delta] \begin{bmatrix} \vec{M}_\delta \\ 1 \end{bmatrix} \quad (5.6)$$

(5.7)

$-KR\vec{C}_\delta$ und $-K'R'\vec{C}'_\delta$ verrechnet mit M sind gleich den Vektorausdrücken $(\vec{M}_\delta - \vec{C}_\delta)$ und $(\vec{M}_\delta - \vec{C}'_\delta)$, welche die Verbindungsgeraden der beiden Projektionszentren mit dem Objektpunkt M im Raum beschreiben.

$$\gamma \vec{m}_\tau = KR(\vec{M} - \vec{C}_\delta) \quad (5.8)$$

$$\gamma' \vec{m}'_\tau = K'R'(\vec{M} - \vec{C}'_\delta) \quad (5.9)$$

Gleichungen 4.8 und 4.9 werden nach $(\vec{M} - \vec{C}_\delta)$ und $(\vec{M} - \vec{C}'_\delta)$ aufgelöst.

$$\gamma R^T K^{-1} \vec{m}_\tau = (\vec{M} - \vec{C}_\delta) \quad (5.10)$$

$$\gamma R'^T K'^{-1} \vec{m}'_\tau = (\vec{M} - \vec{C}'_\delta) \quad (5.11)$$

Wie bereits erwähnt ergibt sich aus den Vektoren $(\vec{M}_\delta - \vec{C}_\delta)$, $(\vec{M}_\delta - \vec{C}'_\delta)$ und $(\vec{C}'_\delta - \vec{C}_\delta)$ das Dreieck aus Abbildung 5.1. Für das Dreieck kann, aus den drei Vektoren, die folgende Gleichung aufgestellt werden.

$$(\vec{C}'_\delta - \vec{C}_\delta) = (\vec{M}_\delta - \vec{C}_\delta) - (\vec{M}_\delta - \vec{C}'_\delta) \quad (5.12)$$

$(\vec{M} - \vec{C}_\delta)$ und $(\vec{M} - \vec{C}'_\delta)$ können durch die Ausdrücke in den Gleichungen 4.10 und 4.11 ersetzt werden.

$$(\vec{C}'_\delta - \vec{C}_\delta) = \gamma R^T K^{-1} \vec{m}_\tau - \gamma R'^T K'^{-1} \vec{m}'_\tau \quad (5.13)$$

Es gilt $\gamma \geq 0$ und $\gamma' \geq 0$, sie stehen für die Tiefe von m und m' und können mit Hilfe des Kreuzproduktes eliminiert werden[5]. Zunächst wird $(\vec{C}'_\delta - \vec{C}_\delta)$ auf die rechte Seite gebracht, so dass die Gleichung nach Null aufgelöst wird.

$$[\vec{C}'_\delta - \vec{C}_\delta] \times \gamma R^T K^{-1} \vec{m}_\tau - [\vec{C}'_\delta - \vec{C}_\delta] \times \gamma' R'^T K'^{-1} \vec{m}'_\tau = 0 \quad (5.14)$$

Gleichung 4.14 wird von links mit $\gamma' \vec{x}'_\tau^T K'^{-T} R'$. Somit kann eine der beiden Schiefsymmetrischen Matrizen aus der Gleichung eliminiert werden.

$$\gamma' \vec{m}'_\tau K'^{-T} R' [\vec{C}'_\delta - \vec{C}_\delta] \times \gamma R^T K^{-1} \vec{m}_\tau = 0 \quad (5.15)$$

Da $\gamma \geq 0$ und $\gamma' \geq 0$, kann für Gleichung 4.15 auch folgendes geschrieben werden.

$$\vec{m}'_{\tau'} K'^{-T} R' \left[\vec{C}'_{\delta} - \vec{C}_{\delta} \right]_{\times} R^T K^{-1} \vec{m}_{\tau} = 0 \quad (5.16)$$

Aus Gleichung 4.16 können nun die Matrizen F und E ausgelesen werden. Bei E handelt es sich um einen Kalibrierten Fall, dass bedeutet dass sowohl K als auch K' bekannt sind und die normalisierten Bildkoordinaten \vec{m} und \vec{m}' durch multiplikation mit K und K' entstehen. E selbst fasst die Schiefsymmetrische Matrix $[\vec{C}'_{\delta} - \vec{C}_{\delta}]_{\times}$ und die beiden Transformationsmatrizen R und R' zusammen.

$$\vec{m}'_{\tau'}^T K'^{-T} E K^{-1} \vec{m}_{\tau} = 0 \quad (5.17)$$

$$\vec{m}_{\tau}^T E \vec{m}'_{\tau'} = 0 \quad (5.18)$$

Matrix E wird zu F , wenn es sich um einen unkalibrierten Fall handelt. Unkalibriert bedeutet, dass K und K' nicht bekannt sind, die Informationen zu K und K' in F befinden. Werden K und K' zu E' multipliziert, wird E zu F .

$$\vec{m}'_{\tau'}^T K'^{-T} E K^{-1} \vec{m}_{\tau} = 0 \quad (5.19)$$

$$\vec{m}'_{\tau'}^T F \vec{m}_{\tau} = 0 \quad (5.20)$$

F und E , fassen die komplette Epipolarageometrie, sprich externe und interne Parameter, sowie die geometrische Beziehung der jeweiligen Bildpunkte zu den 3-D Objektpunkten in einer 3x3-Matrix zusammen. Für F und E gibt es nicht nur eine Lösung. Werden F oder E Beispielsweise über den *eight-Point-Algorithm* ermittelt, so sind die entstehenden 3x3-Matrixen und jedes vielfache von diesen gültige Lösungen für F und E [3, 17]. **Noch herausfinden ob das mit den Tiefen γ und γ' zusammenhängt!!**. Mit F und E kann wie bereits nachgeprüft werden, ob der *epipolar-constraint* $m'^T F m = 0$ oder $\hat{m}'^T E \hat{m} = 0$ zwischen zwei Bildpunkte gilt. Des weiteren können Epipole e und e' und Epipolarlinien l und l' ausfindig gemacht werden, sobald E oder F bekannt ist[3, 5, 17, 1]. Um die zu m oder m' korrespondierende Epipolarlinie l' oder l zu berechnen gilt:

$$l' = Fm \quad (5.21)$$

$$l = F^T m' \quad (5.22)$$

Um die Epipole e und e' zu berechnen die Gleichungen 3.103 und 3.104 erfüllt sein. Für e reicht es also den rechten Nullraum von F zu bestimmen und für e' muss dementsprechend der linke Kern von F gefunden werden.

$$Fe = 0 \quad (5.23)$$

$$F^T e' = 0 \quad (5.24)$$

Die Matrizen F und E sind die ausschlaggebenden Elemente, wenn es um die Rekonstruktion der Kameraorientierungen und der Rekonstruktion der Szene geht. In beiden folgenden Kapitel werden zwei Beispiele zur Findung der externen Kameraparameter und der Szenenrekonstruktion aufgezeigt. Beim ersten Beispiel handelt es sich um ein Minimalbeispiel mit synthetisch erzeugten reinen Daten, um die theoretische Funktionalität des Algorithmus zu beweisen. Im zweiten Beispiel, wird der Algorithmus, mit einigen Anpassungen an die Realverhältnisse, auf Stereobildpaare, aufgenommen von zwei verschiedenen Kameras, angewandt. Die implementierten Algorithmen ermitteln aus einem Satz korrespondierender Bildpunkte die Fundamental Matrix und die essentielle Matrix mit Hilfe des sogenannten *eight-Point-Algorithm*, Im Anschluss werden dann die externen Kameraparameter ermittelt und die Szene mit einem Triangulationsverfahren rekonstruiert.

6 Minimalbeispiel 3D-Stereokalibrierung und Szenenrekonstruktion bei Kameras gleicher Auflösung

Um den Mathematischen Vorgänge der stereoskopischen Szenenrekonstruktion zu veranschaulichen und besser nachvollziehen zu können, wurde ein Minimalszenario erstellt. Mit Hilfe dieses Minimalszenarios, können Theorien besser getestet und bestimmte Situationen simuliert werden. Des Weiteren kann es von Nutzen sein, wenn es darum geht, Fehler in Realbeispielen nachzustellen oder eine Lösung für diese zu finden. Für dieses Minimalszenario wurde ein Objekt, in diesem Falle ein Quader, in ein zuvor definiertes Weltkoordinatensystem (O, δ) mit $\delta = (d_1, d_2, d_3)$ platziert. Außerdem wurden zwei Kameras (C, β) mit $\beta = (b_1, b_2, b_3)$ und (C', β') mit $\beta' = (b'_1, b'_2, b'_3)$ definiert und in (O, δ) platziert. Kamera eins (C, β) ist Deckungsgleich mit (O, δ) . Kamera zwei (C', β') wurde von C in positive d_1 -Richtung, verschoben und um einen Winkel α zu C um die eigene b'_3 -Achse rotiert. Die verwendeten kartesischen Koordinatensysteme sind in diesem Minimalbeispiel alle rechtshändig orientiert. Die äußeren und inneren Kameraparameter wurden für den Aufbau der Szene festgelegt. Dies hat den positiven Effekt, dass somit die späteren Ergebnisse besser validiert werden können. In den Abbildung 6.1 bis 6.3 wird der Aufbau noch einmal genauer veranschaulicht.

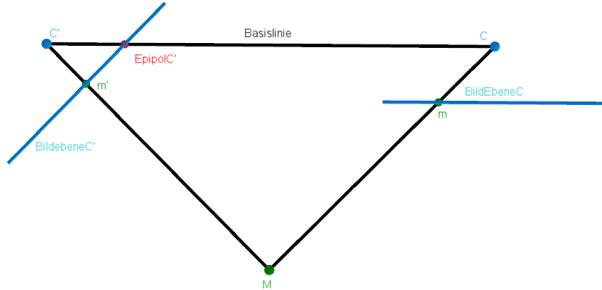


Abbildung 6.1: vereinfachte Top-Down-Ansicht des Szenenaufbaus des Minimalbeispiele

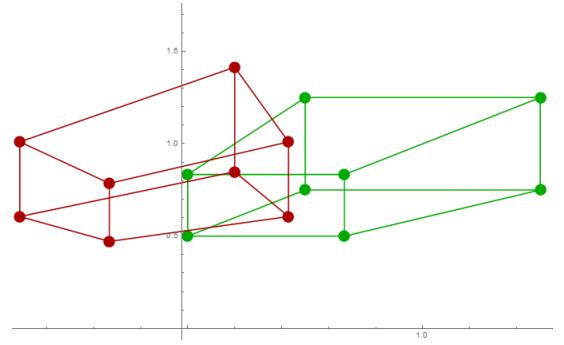


Abbildung 6.2: In Grün ist die Abbildung auf der Bildebenen I von C und in rot ist die Abbildung auf der Bildebenen I' von C'

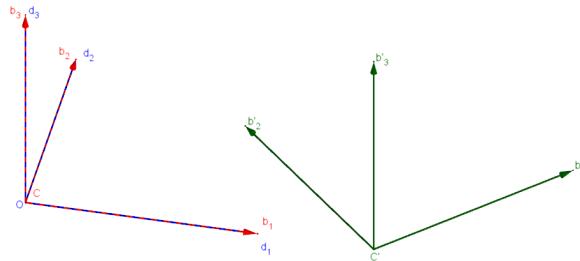


Abbildung 6.3: In Blau und Rot sind jeweils das Welt- und Kamerakoordinatensystem von Kamera eins zu sehen. In grün ist das gedrehte Koordiantensystem von Kamera 2 zu sehen.

6.1 Vorgehen: Projektion eines Quaders in zwei verschiedene transformierte Kameras

Für die Stereokamerakalibrierung wird C' relativ zur zu C um einen Vektor \vec{V}' verschoben und anschließend um einen Winkel α um die b'_3 Achse gedreht. Für die Rotation um b'_3 wird eine Drehmatrix D' aufgestellt.

$$D' = \begin{pmatrix} \cos(\alpha) & 0 & \sin(\alpha) \\ 0 & 1 & 0 \\ -\sin(\alpha) & 0 & \cos(\alpha) \end{pmatrix} \quad (6.1)$$

Um die Transformationsmatrix R' , welche die Rotation und die Translation von C' beinhaltet, zu erhalten werden D' und \vec{V}' miteinander verrechnet.

$$\vec{V}' = \begin{pmatrix} 1 & 0 & 0 & -v'_1 \\ 0 & 1 & 0 & -v'_2 \\ 0 & 0 & 1 & -v'_3 \end{pmatrix} \quad (6.2)$$

$$R' = D'^T \cdot \vec{V}' \quad (6.3)$$

$$R' = \begin{pmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & -v'_1 \\ 0 & 1 & 0 & -v'_2 \\ 0 & 0 & 1 & -v'_3 \end{pmatrix} \quad (6.4)$$

$$R' = \begin{pmatrix} \cos(\alpha) & 0 & -\sin(\alpha) & -v'_1 \cos(\alpha) + v'_3 \sin(\alpha) \\ 0 & 1 & 0 & -v'_2 \\ \sin(\alpha) & 0 & \cos(\alpha) & -v'_1 \sin(\alpha) - v'_3 \cos(\alpha) \end{pmatrix} \quad (6.5)$$

Die entstandene Matrix R' beschreibt die Transformation C' und somit auch die Transformation von Punkten des Koordinatensystems (C, β) in (C', β') . Da $(C, \beta) = (O, \delta)$ ist, beinhaltet die Transformationsmatrix R für C weder eine Translation noch eine Rotation.

$$V = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (6.6)$$

$$R = [I|V] \quad (6.7)$$

$$R = \begin{bmatrix} 1 & 0 & 0 & v_1 \\ 0 & 1 & 0 & v_2 \\ 0 & 0 & 1 & v_3 \end{bmatrix} \quad (6.8)$$

6.2 Berechnung der Projektionsmatrizen

Die Eckpunkte des Quaders $A_\delta, B_\delta, C_\delta, D_\delta, A'_\delta, B'_\delta, C'_\delta, D'_\delta$ sind bekannt. Neben den Eckpunkten des Quaders wird noch ein neunter Punkt E_δ außerhalb des Quaders platziert und zwar so, dass es zu keinen linearen Abhängigkeiten zwischen E_δ und den anderen Punkten kommt. So kann vermieden werden, dass die später aufgestellt Koeffizientenmatrix, zum berechnen der Fundamentalmatrix, einen Rang kleiner als acht bekommt und somit zwei linear unabhängige Lösungen ausgibt[3]. Im Unterkapitel Fundamentalmatrix und der eight-Point-Algorithmus wird nochmal genauer drauf eingegangen, was das für F bedeutet. Fürs erste wird festgelegt, dass insgesamt neun Punkte sich in der Szene befinden. Um diese Punkte auf die Bildebenen (I, τ) und (I', τ) zu projizieren, müssen neben den Transformationsmatrizen R und R' noch die Kameramatrizen K und K' festgelget werden.

$$K = \begin{bmatrix} \zeta_C & 0 & 0 & 0 \\ 0 & \zeta_C & 0 & 0 \\ 0 & 0 & \zeta_C & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (6.9)$$

$$K' = \begin{bmatrix} \zeta_{C'} & 0 & 0 & 0 \\ 0 & \zeta_{C'} & 0 & 0 \\ 0 & 0 & \zeta_{C'} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (6.10)$$

Sind R, R', K und K' bekannt, können die Projektionsmatrizen P und P' berechnet werden. Vorher werden R und R' noch projektiv erweitert.

$$R = \begin{bmatrix} 1 & 0 & 0 & v_1 \\ 0 & 1 & 0 & v_2 \\ 0 & 0 & 1 & v_3 \end{bmatrix} \quad (6.11)$$

$$P = K \cdot R \quad (6.12)$$

$$P = \begin{bmatrix} \zeta_C & 0 & 0 & 0 \\ 0 & \zeta_C & 0 & 0 \\ 0 & 0 & \zeta_C & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (6.13)$$

$$P' = K' \cdot R' \quad (6.14)$$

$$P' = \begin{bmatrix} \zeta_{C'} \cos(\alpha) & 0 & \zeta_{C'} \sin(\alpha) & -\zeta_{C'}(v'_1 \cos(\alpha) + v'_3 \sin(\alpha)) & 0 \\ 0 & 1 & 0 & \zeta_{C'} - v'_2 & 0 \\ \zeta_{C'} \sin(\alpha) & 0 & \zeta_{C'} \cos(\alpha) & -\zeta_{C'}(v'_1 \sin(\alpha) + v'_3 \cos(\alpha)) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (6.15)$$

6.3 Transformation der Objektpunkte von Weltkoordinaten in Kamerakoordinaten

Die 3D-Punkte $A_\delta, B_\delta, C_\delta, D_\delta, A'_\delta, B'_\delta, C'_\delta, D'_\delta, E_\delta$, werden um eine Homogene Komponenten erweitert und mit den Projektionsmatrizen P und P' zu, auf die Bildebenen I und I' projizierten, Punkten der Kamerakoordinatensysteme (C, β) und C', β' transformiert.

$$Q_\beta = P \cdot \begin{bmatrix} \begin{pmatrix} A_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} B_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} C_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} D_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} A'_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} B'_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} C'_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} D'_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} E_\delta \\ 1 \end{pmatrix} \end{bmatrix} \quad (6.16)$$

$$Q_{\beta'} = P' \cdot \begin{bmatrix} \begin{pmatrix} A_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} B_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} C_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} D_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} A'_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} B'_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} C'_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} D'_\delta \\ 1 \end{pmatrix} & \begin{pmatrix} E_\delta \\ 1 \end{pmatrix} \end{bmatrix} \quad (6.17)$$

Die entstanden Bildebenenkoordinaten zur Basis der Kamerakoordinatensysteme müssen dann wieder auf eine homogene Form gebracht werden, indem sie durch ihre jeweils letzte Komponenten dividiert werden.

$$A_\beta = \begin{pmatrix} b_1 \\ b_3 \\ b_2 \\ \gamma \end{pmatrix} \rightsquigarrow A_\beta = \begin{pmatrix} \frac{b_1}{\gamma} \\ \frac{b_3}{\gamma} \\ \frac{b_2}{\gamma} \\ \frac{\gamma}{\gamma} \end{pmatrix} \rightsquigarrow A_\beta = \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ 1 \end{pmatrix} \quad (6.18)$$

$$\simeq \begin{pmatrix} t_1 \\ t_2 \\ \zeta \\ 1 \end{pmatrix} \quad (6.19)$$

Der in Gleichung 4.18 verwendete Kamerapunkt A_β liegt direkt auf der Bildebene I , was durch die Projektionsmatrix P bedingt wurde. Die Tiefenkomponente entspricht deshalb nach der homogenisierung des Punktes dem Wert ζ , welcher den Abstand \overline{IC} beschreibt. Um die 3D-Kamerapunkte in 2D-Bildebenepunkte zur Basis (I, τ) mit $\tau = (t_1, t_2)$ und für C' entsprechend (I', τ') mit $\tau' = (t'_1, t'_2)$ umzuwandeln, muss lediglich der Tiefenwert ζ entfernt und durch die Homogene Komponenten ersetzt werden. In Abbildung 4.4, sind die entstehenden Abbildilder auf die jeweiligen Bildebene in 2D-Bildebenekoordinaten (I, τ) und (I', τ') zu sehen.

$$A_\beta = \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ 1 \end{pmatrix} \simeq \begin{pmatrix} t_1 \\ t_2 \\ \zeta \\ 1 \end{pmatrix} \quad (6.20)$$

$$\rightsquigarrow A_\tau = \begin{pmatrix} t_1 \\ t_2 \\ 1 \end{pmatrix} \quad (6.21)$$

(6.22)

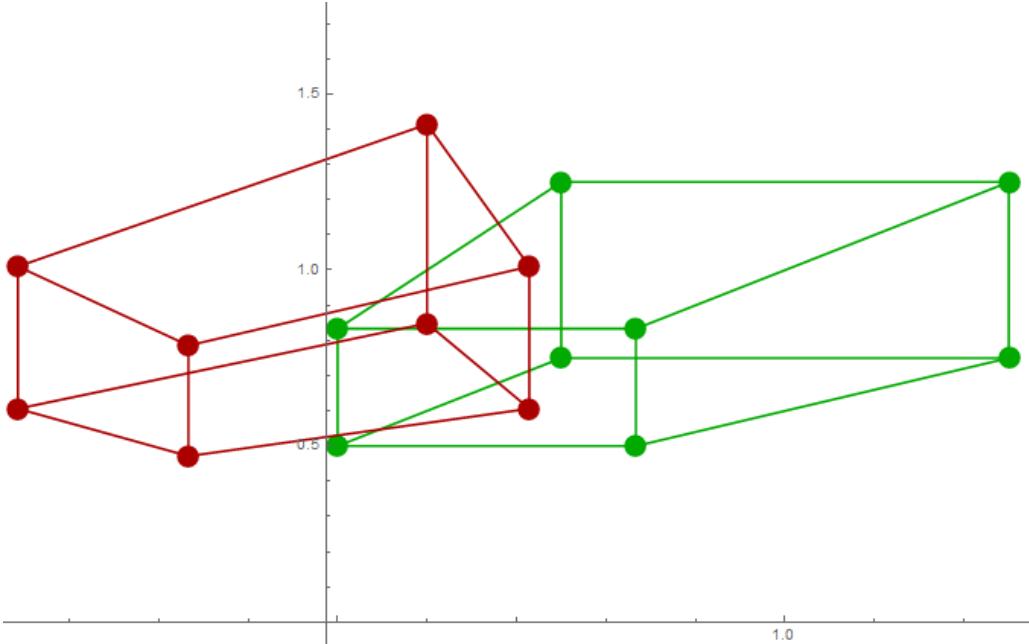


Abbildung 6.4: Grün zeigt den Quader welcher auf I von C abgebildet wird. Das größere Quadrat sind die vorderen Punkte A, B, C, D , das kleinere Quadrat sind die hinteren Punkte A', B', C', D' . Der Punkt E ist weiter weg von den Abbildungen und deshalb auf dieser Abbildung momentan nicht zu sehen. Rot zeigt denselben Quader auf I' von C' abgebildet.

6.4 Umrechnung von Bildebenenkoordinaten in Sensorkoordinaten

Für die Umrechnung der Bildebenenkoordinaten $A_\tau, B_\tau, C_\tau, D_\tau, A'_\tau, B'_\tau, C'_\tau, D'_\tau, E_\tau$ und $A_{\tau'}, B_{\tau'}, C_{\tau'}, D_{\tau'}, A'_{\tau'}, B'_{\tau'}, C'_{\tau'}, D'_{\tau'}, E_{\tau'}$ in Sensorkoordinaten (S, σ) mit $\sigma = (\vec{u}, \vec{v})$, muss der sogenannte Pixelpitch des Sensors bekannt sein. Für das Minimalbeispiel wird ein PixelPitch von 1 angenommen. Die Bildebenenkoordinaten werden so 1:1 in die Sensorkoordinaten umgesetzt. In Realbeispielen ist dies aber eher selten der Fall. Die Sensorkoordinaten sind im normalfall in Pixeleinheiten gegeben und werden für die Berechnung der Fundamentalmatrix benötigt.

$$\sigma = (\vec{u}, \vec{v}) \quad (6.23)$$

$$\vec{u} = u_1 t_1 + u_2 t_2 \quad (6.24)$$

$$\vec{v} = v_1 t_1 + v_2 t_2 \quad (6.25)$$

$$S_\sigma = I_\tau + p_{x\tau} + p_{y\tau} \quad (6.26)$$

$$(\vec{u}, \vec{v}, S_\sigma) = (t_1, t_2, I_\tau) \cdot \begin{bmatrix} u_1 & u_2 & z_1 \\ v_1 & v_2 & z_2 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.27)$$

$$M = \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \quad (6.28)$$

Mit dem Berechnen der Sensorkoordinaten ist der Aufbau der Szene vollendet. Das Ergebnis sind zwei Bilder des Quaders von zwei verschiedenen Kameras mit momentan noch gleicher Auflösung. In den nächsten Schritten soll nun anhand der Bildpunkte eine Rekonstruktion der externen Kameraparameter und eine Rekonstruktion der synthetischen Szene folgen. Die intrinsischen Parameter werden in dieser Arbeit immer als bereits bekannt festgelegt. (Hier vllt noch erwähnen, dass die calibrierung der inneren Kameraparameter in einer anderen Arbeit behandelt wurde, außerdem kann hier auch eine getrennt kalibrierung beider kameras durchgeführt werden)

6.5 Fundamentalmatrix und der eight-Point-Algorithmus

Nachdem nun die Bildkoordinaten in Pixel des Quaders auf den Bildern der beiden Kameras bekannt sind, kann nun die eigentliche Rekonstruktion der externen Kameraparameter und die 3D-Szenenrekonstruktion beginnen. Der erste Schritt beinhaltet die Ermittlung der Fundamentalmatrix F mit den korrespondierenden Bildpunkten der beiden Kamerabilder mit Hilfe des sogenannten *8-Point-Algorithmus*. Der *8-Point-Algorithm* ist eine lineare Technik die angewandt wird, um die essentielle Matrix und die Fundamentalmatrix aus $n \geq 8$ Punkten zu schätzen [11, 3]. Der Algorithmus benötigt $n \geq 8$ Punkte, um ein valides eindeutiges Ergebnis zu liefern [3, 4]. Es besteht noch die Möglichkeit den *7-Point-Algorithm* anzuwenden, welcher nach dem selben Prinzip wie sein Verwandter *8-Point-Algorithm* verfährt, jedoch kein eindeutiges Ergebnis liefert. Mit sieben Punkten bekommen wir als Lösung zwei linear unabhängige Kerne als Lösung der zuvor aufgestellten Koeffizientenmatrix mit welchen für ein eindeutiges Ergebnis noch weiter verfahren werden muss [3, 4]. Die aufgestellte Koeffizientenmatrix, hat im Fall des *7-Point-Algorithm* nämlich nur den Rang 7 [3]. In der Schätzung von F wurde in diesem Minimalbeispiel und auch später in Realbeispiel der *8-Point-Algorithm* verwendet. Im Falle des Minimalbeispiels, wurde mit einem zusätzlichen neunten Punkt auch noch dafür gesorgt, dass die aufgestellte Koeffizientenmatrix auch den Rang 8 besitzt. Somit kann, wie bei der Bestimmung der Homographie, einfach der Nullraum der Koeffizientenmatrix bestimmt werden, welcher die Einträge der 3x3-Matrix von F liefert [3]. Besitzt die Koeffizientenmatrix A einen Rang größer 8, so wird auch hier mit einem *least-square*-Verfahren, mit Hilfe der Singulärwertszerlegung, angewendet werden, so dass $\|A \cdot f\|$ minimal wird. f ist der Singulärvektor, welcher mit dem kleinsten Singulärwert korrespondiert [3]. Die genaue Abfolge des Verfahrens, wird im Kapitel Singularity-Constraint der Fundamentalmatrix genauer aufgezeigt. Da im Minimalbeispiel der Rang 8 der Koeffizientenmatrix erzwungen wurde, reicht es vorerst den Nullraum zu bestimmen. Die Koeffizientenmatrix wird aus

dem Ausdruck *epipolar-constraint* in Gleichung 4.29 aufgestellt. Der ermittelte Kern und alle seine Vielfache sind mögliche Lösungen für $F[3, 4]$.

$$m'^T_{\sigma'} \cdot F \cdot m_{\sigma} = 0 \quad (6.29)$$

$$F = \begin{bmatrix} f_{11} & f_{122} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \quad (6.30)$$

$$\begin{bmatrix} x'_n & y'_n & 1 \end{bmatrix} \cdot \begin{bmatrix} f_{11} & f_{122} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \cdot \begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} = 0 \quad (6.31)$$

$$f_{11}x_nx'_n + f_{12}y_nx'_n + f_{13}x'_n + f_{21}x_ny'_n + f_{22}y_ny'_n + f_{23}y'_n + f_{31}x_n + f_{32}y_n + f_{33} = 0 \quad (6.32)$$

$$(x_nx'_n, y_nx'_n, x'_n, x_ny'_n, y_ny'_n, y'_n, x_n, y_n, 1) \cdot f = 0 \quad (6.33)$$

$$\begin{bmatrix} x_1x'_1 & y_1x'_1 & x'_1 & x_1y'_1 & y_1y'_1 & y'_1 & x_1 & y_1 & 1 \\ x_2x'_2 & y_2x'_2 & x'_2 & x_2y'_2 & y_2y'_2 & y'_2 & x_2 & y_2 & 1 \\ \vdots & \vdots \\ x_nx'_n & y_nx'_n & x'_n & x_ny'_n & y_ny'_n & y'_n & x_n & y_n & 1 \end{bmatrix} \cdot \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = 0 \quad (6.34)$$

6.6 Epipole und Epipolargeraden

Mit Hilfe der Fundamentalmatrix und dem Wissen über die Epipolargerometrie, kann man die Epipole e und e' , sowie die Epipolargeraden l und l' ermitteln.

$$l' = F \cdot m \quad (6.35)$$

$$l = F^T \cdot m' \quad (6.36)$$

l' ist die zu m korrespondierende Epipolargerade. l ist die zu m' korrespondierende Epipolargerade. Zu Berechnung des Epipols e muss der Rechte Kern von F ermitteln werden und für den Epipol e' brauchen wir den linken Kern. Die Abbildung 4.5 zeigt, dass Ergebnis der Epipole im Minimalbeispiel.

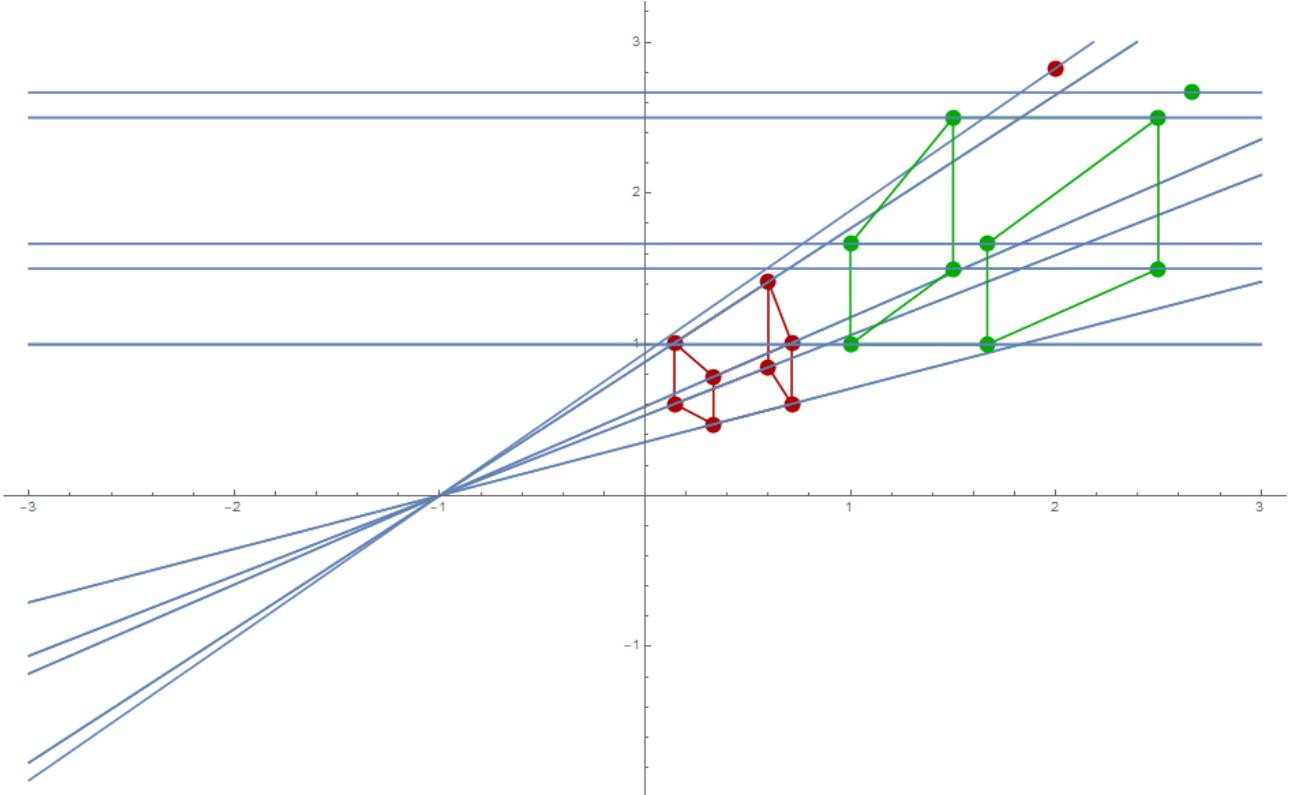


Abbildung 6.5: Die blauen Geraden zeigen die jeweiligen Epipolgeraden. Die vom roten Quader schneiden sich bei -1 im Epipol e' . Die Epipolgeraden vom grünen Quader schneiden sich im Epipol e' im Unendlichen, weshalb die Epipolarlinien Parallel zueinander verlaufen.

6.6.1 Geometrische Konstruktion der Epipole und der Epipolarlinien

Die Vorgehensweise zur Ermittlung der Epipole und der Epipolgeraden ist zwar schnell aber zur Verdeutlichung der geometrischen Beziehungen untereinander wird nochmal eine rein geometrische Konstruktion der Epipole und der Epipolgeraden durchgeführt. Im nachfolgenden Beispiel wird der Epipol e' geometrisch konstruiert. Umrechnung von allen Punkten in ein gemeinsames Koordinatensystem, zum Beispiel in (C, β) . Für die Konstruktion der Epipole wird die BasisLinie $B = C'_\beta - C_\beta$, ein Bildebenenpunkt $A_{C'_\beta}$ auf I' und die Position von I' selbst benötigt.

$$B = C'_\beta + t \cdot (C'_\beta - C_\beta) \quad (6.37)$$

Um den Epipol zu ermitteln wird der Schnittpunkt der Gerade B mit der Bildebenen I' von C' benötigt. Die Bildebene I' wird in einer Ebenen Normalenform dargestellt. Als Bildebenenpunkt wird $A_{C'_\beta}$ von I' von C' gewählt.

$$\vec{n}_0 \cdot [\vec{x} - A_{C'_\beta}] \quad (6.38)$$

Um den Schnittpunkt der Gerade B mit der Ebene I zu berechnen, wird die Gerade B in die Ebenengleichung eingesetzt und es wird ein Wert für t ermittelt.

$$\vec{n}_0 \cdot \vec{x} - \vec{n}_0 \cdot \vec{a} \quad (6.39)$$

$$(6.40)$$

Danach wird die Geradengleichung *BaseLine* in die Ebenengleichung *ImagePlane* eingesetzt und es wird ein Wert für t ermittelt. Der berechnete Wert für t wird dann wiederum in G eingesetzt und das

Ergebnis ist der Epipol e'_β . Dieser kann dann wieder über bereits bekannte Transformationen in die Sensorkoordinaten von (S', σ) bezüglich C' überführt werden.

$$e'_\sigma = P' \cdot e'_\beta \quad (6.41)$$

Für die Epipolarlinien $l'_{\sigma'}$ durch $e'_{\sigma'}$, müssen einfach Geradengleichungen eines Sensorpunktes, beispielsweise $A_{\sigma'}$, durch $e'_{\sigma'}$ aufgestellt werden. Die entstandenen Linie ist dann die Epipolarlinie zum Punkt $A_{\sigma'}$ und gleichzeitig die korrespondierende Epipolarlinie zum Sensorpunkt A_σ von I von C .

$$g := A_{\sigma'} + t \cdot (A_{\sigma'} - e'_{\sigma'}) \quad (6.42)$$

6.7 Essentiellen Matrix

Für die Rekonstruktion der externen Kameraparameter, gibt es verschiedene Ansätze [3]. Es ist zum einen Möglich, ohne Vorwissen der Kameramatrizen oder Transformationsmatrizen, aus der Fundamentalmatrix über den sogenannten *Stratified approach* die komplette Projektionsmatrix P und P' zu ermitteln, jedoch ohne weitere Informationen über die 3S-Szene nur bis zu einem Projektiven Vieldeutigkeit[3]. Da in dieser Arbeit davon ausgegangen wird, dass die Kameras zuvor einzeln Kalibriert wurden und die intrinsischen Kameraparameter bereits bekannt sind(hier vllt vorherige Arbeit erwähnen), wird für die Rekonstruktion der externen Kameraparameter ein Ansatz verfolgt, in welchen die essentielle Matrix zum Einsatz kommt. Die essentielle Matrix ist eine Spezialform der Fundamentalmatrix und beschreibt den *epipolar constraint*, im kalibrierten Fall, zwischen den normalisierten Bildkoordinaten \hat{m}_σ und $\hat{m}'_{\sigma'}$ [3, 5, 1, 11, 4].

$$\hat{m}'_{\sigma'}^T \cdot E \cdot \hat{m}_\sigma = 0 \quad (6.43)$$

Ist die Fundamentalmatrix F bereits ermittelt worden und die Kameramatrizen K und K' bekannt, kann die essentielle Matrix wie im Kapitel Epipolar Geometrie aus F berechnet werden. Die essentielle Matrix ist eine Fundamentalmatrix, welche zu einem paar normalisierten Projektionsmatrizen P mit $P = [I|0]$ und P' mit $P' = [R'|V']$ korrespondierend ist[3, 11, 1, 4]. Um eine Projektionsmatrix $P' = K'[R'|V']$ auf die normalisierte Form zu bringen, müssen die intrinsischen Parameter für K' bekannt sein.

$$m'_\sigma = P' \cdot M_\delta \quad (6.44)$$

$$m'_\sigma = K'[R'|V'] \cdot M_\delta \mid \cdot K'^{-1} \quad (6.45)$$

$$K'^{-1} \cdot m'_\sigma = K'^{-1} \cdot K'[R'|V'] \cdot M_\delta \quad (6.46)$$

$$\hat{m}'_\sigma = [R'|V']M_\delta \quad (6.47)$$

M_δ ist der ein Objektpunkt im 3D-Raum, welcher mit P' zum Sensorpunkt $m'_{\sigma'}$ abgebildet wird. Nach der Normalisierung, wird \hat{m}'_σ als normalisierte Koordinat bezeichnet und die entstandene Projektionsmatrix $P' = [R'|V']$ als normalisierte Projektionsmatrix[3, 4]. Die essentielle Matrix E beschreibt dann die Korrespondenz zwischen den korrespondierenden normalisierten Koordinaten \hat{m}_σ und $\hat{m}'_{\sigma'}$, wie in Gleichung 4.43 gezeigt. Um E aus F zu gewinnen, werden die Kameramatrizen K und K' mit F multipliziert. Zu Erinnerung im Kapitel Epipolar Geometrie wurde genau diese Beziehung zwischen E und F in den Gleuchungen 3.98 bis 3.101 aufgeführt.

$$E = K'^T \cdot F \cdot K \quad (6.48)$$

Das Ergebnis dieser Gleichung 4.48 und alle deren Vielfache, sind mögliche Lösungen für die essentielle Matrix. Wie die Fundamentalmatrix muss auch die essentielle Matrix einen Rang von 2 haben. Des Weiteren darf eine 3x3-Matrix nur dann als essentielle Matrix bezeichnet werden, wenn die Singulärwerte, welche durch eine Singulärwertszerlegung ersichtlich werden, bestimmte Merkmale aufweisen. So müssen zwei der drei Singulärwerte gleich und die dritte null sein[3]. Des Weiteren muss für die Determinante gelten, dass $\det(E) = 0$ ist und die Quadratwurzel der Eigenwerte müssen wieder die Singulärwerte ergeben. Die essentielle Matrix kann, wie die Fundamentalmatrix auch, über den *8-Point-Algorithm* ermittelt werden. Jedoch müssen auch hier die intrinsischen Kameraparameter K und K' bekannt sein. Die korrespondierenden Punkte, werden in normalisierte Bildkoordinaten umgerechnet und dann in eine Koeffizientenmatrix gleich der Matrix in Gleichung 4.34 eingetragen. Danach kann E entweder über die Bestimmung des Kerns oder, im Falle von Realen Bilddaten, dem *least-square-* Verfahren geschätzt werden[3, 4].

6.8 Exteren Kameraparameter mit essentieller Matrix

Mit der essentiellen Matrix ist es möglich die Transformationsmatrix R' zu ermitteln. Es wird davon ausgegangen, dass für R gilt $R = [I|0]$. Die aus E ermittelte Matrix R' beschreibt dann die Transformation von C' relativ zu C [3, 4]. Um die externen Kameraparameter zu bestimmen, wird zunächst die essentielle Matrix E mit Hilfe der Singulärwertszerlegung in drei Matrizen zerlegt.

$$E = U\Sigma V^T \quad (6.49)$$

Die Singulärwerte befinden sich in der mittleren Matrix $\Sigma = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$ wieder. Damit die Matrix E sich essentielle Matrix nennen darf, muss für zwei der Diagonaleinträge für die Singulärwerte gelten das $\sigma_1 = \sigma_2$ und für σ_3 muss dann dementsprechend gelten, dass $\sigma_3 = 0$. Sollten die Singulärwerte diesen Bedingungen nicht entsprechen, so wird ein so genannten *singularity-constraint*, der Form $\Sigma = \text{diag}(\frac{\sigma_1+\sigma_2}{2}, \frac{\sigma_1+\sigma_2}{2}, 0)$, erzwungen[3]. Im Minimalbeispiel kommt es auf Grund der eigens erstellten Bildpunkte zu keinen Bildfehlern, wie im Realbeispielen. Deswegen ist der *singularity-constraint* nicht notwendig, da die Bedingungen für E erfüllt sind.

Ein wichtiges Detail, was sich später auch auf die vier möglichen Ergebnisse für P auswirkt, ist noch zu beachten. Zum einen sollte einem bewusst sein, dass die Zerlegung der essentiellen Matrix keine eindeutige Lösung hervorbringen muss. Für E muss gelten, dass $\det(E) = 0$ ist. Es wird also vorausgesetzt, dass die Determinante der aus der *SVD* gewonnenen Matrizen $UV^T = 1$ ist. Angenommen aus der *SVD* von E ergibt sich für die Determinanten von $UV^T = -1$, so können die Determinanten der Matrizen U und V getrennt voneinander bestimmt werden. Sollte $\det(U) = -1$ oder $\det(V) = -1$ oder beides zusammen der Fall sein, so kann die jeweilige Matrix einfach mit -1 multipliziert werden. E setzt sich zusammen aus einer Rotationsmatrix R und einer schiefsymmetrischen Matrix S .

$$E = [v]_x R \quad (6.50)$$

$$S = [v]_x \quad (6.51)$$

$$E = SR \quad (6.52)$$

Zur Schätzung von S und R werden des Weiteren die schiefsymmetrische Matrix W und die Blockdiagonale Matrix Z eingeführt[3]. Mit diesen Matrizen lassen sich die gesuchten R und S für C' rekonstruieren, jedoch nur bis zu einer gewissen Skaleninvarianz[3, 4].

$$W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad Z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (6.53)$$

Mit dem, je nach dem leicht modifizierten, Ergebnis der *SVD* von E lassen sich nun jeweils zwei mögliche Lösungen für S und R aufstellen.

$$S_1 = -UZU^T \quad R_1 = UW^TV^T \quad (6.54)$$

$$S_2 = UZU^T \quad R_2 = UWV^T \quad (6.55)$$

Um sicher zu gehen, dass es sich bei R_1 und R_2 auch um gültige Rotationsmatrizen handelt, kann eine Probe durchgeführt werden. Zum einen muss $R \cdot R^T = I_{3x3}$ sein. I_{3x3} bezeichnet in diesem Fall die 3x3-Einheitsmatrix. Des Weiteren kann man überprüfen, ob die Determinante $\det(R_1) = 1$ ist. Ist die Determinante = -1, so ist die Rotation für ein Rechtsdrehendes System, ist die Determinante = 1, so handelt es sich um ein linksdrehendes System[3]. S_1 und S_2 sind jeweils schiefsymmetrische Matrizen aus welchen, jetzt noch zwei Lösungen für v finden lassen müssen[3].

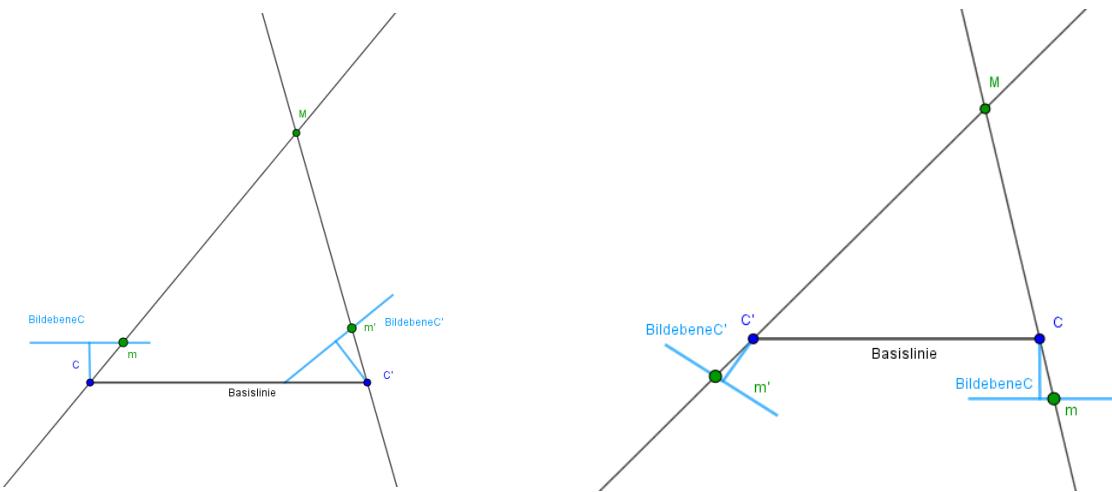
$$St = [v]_\times \cdot v = v \times v \quad (6.56)$$

Um v zu ermitteln muss also lediglich der Kern von S_1 und S_2 gefunden werden. Die jeweiligen Ergebnisse für v_1 und v_2 sind bis auf ihre Vorzeichen die selben.

Die externen Kameraparameter lassen sich wie gesagt nur bis zu einem Skalierungsfaktor genau bestimmen. Die Rotationen R ist von dieser Skaleninvarianz nicht betroffen, es betrifft nur den Translationsvektor \vec{v} . Wie mit der Skaleninvarianz weiter Verfahren wird, wird im nachfolgenden Abschnitt der Szenenrekonstruktion durch Triangulierung noch aufgeführt. Letztendlich können, für die Rekonstruktion der externen Kameraparameter vier Lösungen für P gefunden werden.[3, 4]. λv heißt dabei, dass sowohl v also auch alle Vielfache von v , Lösungen sein können, was durch die Skaleninvarianz der Lösung bedingt ist[3, 4].

$$P = [UWV^T| + \lambda v] \quad \text{oder} \quad [UW^TV^T| + \lambda v] \quad (6.57)$$

$$\text{oder} \quad [UWV^T| - \lambda v] \quad \text{oder} \quad [UW^TV^T| - \lambda v] \quad (6.58)$$



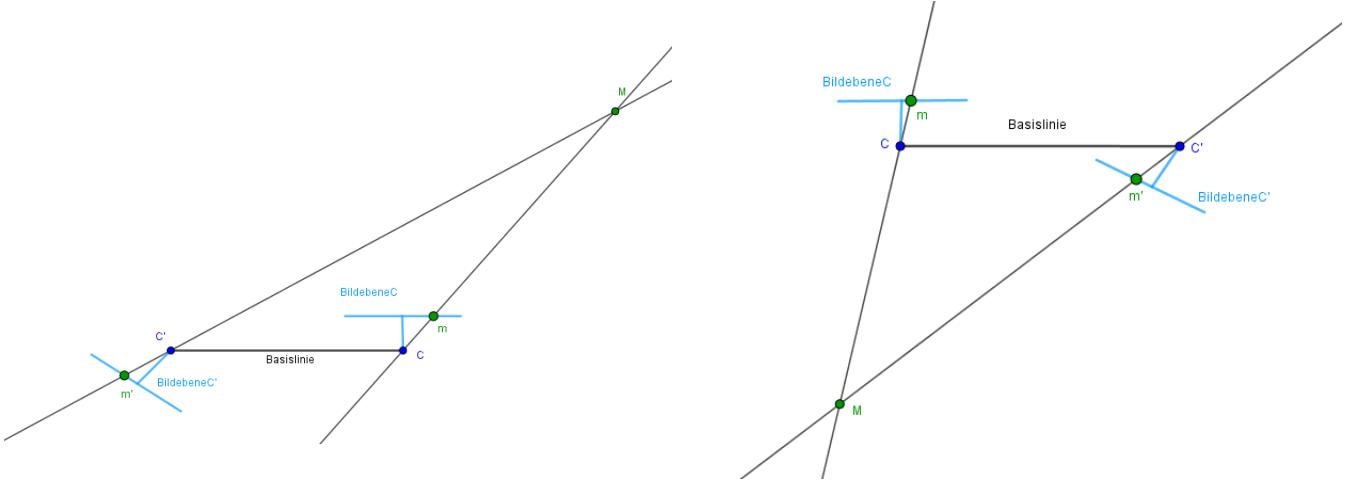


Abbildung 6.6: Anordnung der Kameras bei den vier verschiedenen Lösungen für P

6.9 Szenenrekonstruktion durch Triangulation

Unter Triangulierung versteht man das Rekonstruieren der 3D-Szenenpunkte durch Schnittpunktberechnung derjenigen Geraden, welche durch die jeweiligen Projektionszentren der Kameras und deren korrespondierenden Bildpunkten auf deren Bildebenen gehen, so dass sich ein Dreick bildet.

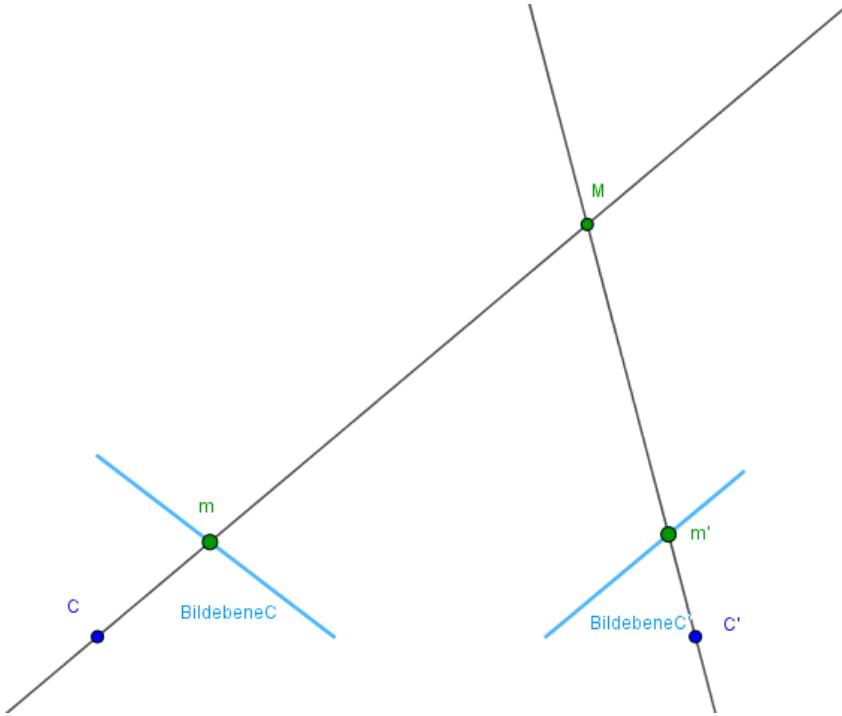


Abbildung 6.7: Optimale Triangulierung: Beide Geraden Treffen sich in einem Punkt im 3D-Raum

Da das Minimalbeispiel mit reinen Daten arbeitet, ist damit garantiert, dass die Linien sich in einem Schnittpunkt treffen. Anders hingegen wäre es in einem realen Beispiel mit korrespondierenden Punkten, welche Beispielsweise über einen SURF-Algorithmus detektiert wurden[18]. In Realbildern, können Bildfehler wie beispielsweise Rauschen nicht vermieden werden, des Weiteren können korrespondierende Punkte nicht immer exakt auf den Pixel genau bestimmt werden. Diese Fehler führen dazu, dass wenn ein Schnittpunkt der Geraden durch die vermeintlichen korrespondierenden Punkte nicht gefunden werden kann, da die Geraden sich sehr wahrscheinlich nicht in einem Punkt treffen werden[18, 3].

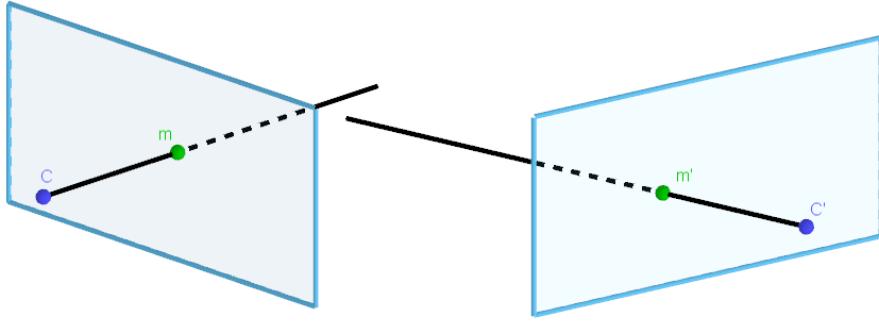


Abbildung 6.8: Durch Ungenauigkeiten in der korrespondierenden Punkte, verfehlten sich die Linien und es kommt zu keinem Schnittpunkt

Für diese Fälle gibt es mehrere Näherungsverfahren, wovon eines, im Kapitel Minimieren der Kostenfunktion durch Sampson-approximation, im Realbeispiel eingeführt wird[3]. In diesem Minimalbeispiel tritt der optimale Fall ein. Das bedeutet, dass ein zu Bildpunkt m korrespondierender Bildpunkt m' auf der zu m korrespondierenden Epipolarlinie l' liegt und somit garantiert ist, dass sich die Geraden mm' und $m'm$ auf jedenfall in einem gemeinsamen Punkt schneiden. Um den Schnittpunkt beider Geraden zu berechnen, werden zum einen die Bildpunkte $A_\tau, B_\tau, C_\tau, D_\tau, A2_\tau, B2_\tau, C2_\tau, D2_\tau$ und $A_{\tau'}, B_{\tau'}, C_{\tau'}, D_{\tau'}, A2_{\tau'}, B2_{\tau'}, C2_{\tau'}, D2_{\tau'}$, sowie die korrekt ermittelte Projektionsmatrix P' von zuvor benötigt. Bei der Berechnung der externen Kameraparameter wurde festgesetzt, dass die Projektionsmatrix von Kamera eins $P = [I|0]$ und dementsprechend $P' = [R^T | -R^T V]$ die Transformation für Kamera zwei, ausgehend von Kamera eins ist. Also ist die Position von Kamera eins in Weltkoordinaten $C_\delta = [0 \ 0 \ 0 \ 1]^T$. Was jetzt noch fehlt ist die Position von Kamera zwei in Weltkoordinaten. Da die Szene dieses Minimalbeispiels durchkonstruiert wurde, ist C' eigentlich bekannt. Es soll nun aber angenommen werden, dass die Position von Kamera zwei nicht bekannt ist und diese wie im realen Fall erst einmal aus P berechnet werden muss. Um C' zu ermitteln, wird der Translationsvektor V aus $P' = [R^T | -R^T V]$ benötigt.

$$C'_\delta = -R * (-R^T V) \quad (6.59)$$

Im nächsten Schritt, müssen die Bildebenepunkte von C' noch in das selbe Koordinatensystem wie die Bildebenepunkte von C transformiert werden. Da Kamera eins Deckungsgleich mit dem Weltkoordinatensystem ist, sind die Bildpunkt der Bildebene I von C bereits in Weltkoordinaten gegeben, diese müssen nur noch mit, ζ als dritte Tiefenkomponenten z , erweitert werden. Die Bildpunkte von C' , werden ebenfalls mit ihrem ζ erweitert und danach mit der Projektionsmatrix P , welche als V die Koordinaten von C' besitzt, in das Weltkoordinatensystem überführt. Sind die Basen der Bildpunkte von C_δ und C'_δ , mit $\delta = \beta$ im selben Koordinatensystem, so können nun die Gleichungen durch die Projektionszentren C und C' und den entsprechenden Bildebene koordinaten $A_\delta, B_\delta, C_\delta, D_\delta, A2_\delta, B2_\delta, C2_\delta, D2_\delta$ und den neu umgerechneten Punkten $A'_\delta, B'_\delta, C'_\delta, D'_\delta, A2'_\delta, B2'_\delta, C2'_\delta, D2'_\delta$ gezogen. Beispielhaft wird dies Anhand der korrespondierenden Punkte A_τ und dem umgerechneten A'^τ aufgezeigt.

$$\begin{aligned} A_\delta + t * (A_\delta - C_\delta) &= 0 \\ A'_\delta + t' * (A'_\delta - C'_\delta) &= 0 \\ \text{Geraden gleichsetzen:} \\ A_{\delta x} - t_x - C_{\delta x} &= A'_{\delta x'} - t'_x - C'_{\delta' x} \\ A_{\delta y} - t_y - C_{\delta y} &= A'_{\delta y'} - t'_y - C'_{\delta' y} \end{aligned}$$

Nun muss für jedes Linienpaar eine Lösung für t und t' gefunden werden und die Lösungen in die Gleichungen 4.65 und 4.66 eingesetzt werden. Es sollte für beide Gleichungen die selbe Lösung für den rekonstruierten Punkt A im Raum ergeben. Die Lösung entspricht meist noch nicht exakt dem eigentlichen Ergebnis, das liegt an der zuvor erwähnten Skaleninvariants der Rekonstruktion der externen Kameraparameter. Bei den zuvor ermittelten externen Kameraparametern, ist der Translationsvektor Skaleninvariant, was dazu führt, dass die rekonstruierten Objekte nach der Szenenrekonstruktion noch nicht ihrer Originalgröße entsprechen. Es wird noch ein Skalierungsfaktor benötigt, welcher die Szene auf Originalgröße skaliert. Hierfür ist es in einer Realsszene ratsam, wenn man zuvor von zwei Punkten in Szene den Abstand zueinander abmisst, um anhand dessen einen Skalierungsfaktor zu berechnen. Im hier beschriebenen Minimalbeispiel sind die Originalkoordinaten der Objektpunkte im Raum bekannt, weshalb hier nach dem passenden Vielfachen der Rekonstruierten Punkte gesucht werden kann. Da die Verhältnisse der Abstände der Punkte zueinander bei der Skalierung beibehalten wird, kommt zu keinen Verfälschungen des Objektes, da die Rotationen der beiden Kameras unangetastet bleibt. Abbildung 4.6 zeigt, dass sich durch verändern des Translationsvektors nur die Größe des Objektes ändert nicht aber seine Orientierung im Raum.

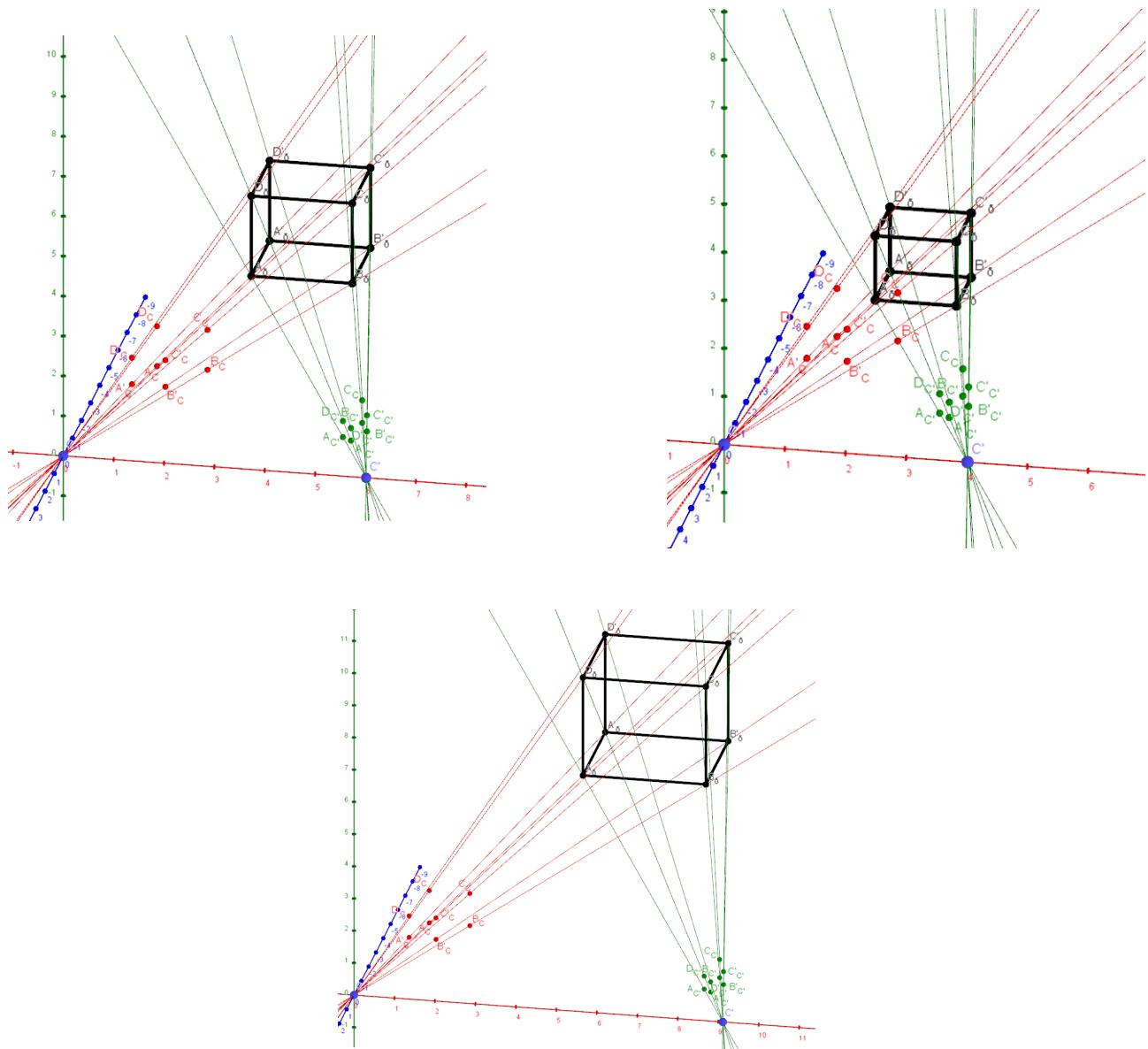


Abbildung 6.9: Veranschaulichung der Skaleninvarianz und dessen Auswirkung auf die geometrische Form der Objekte

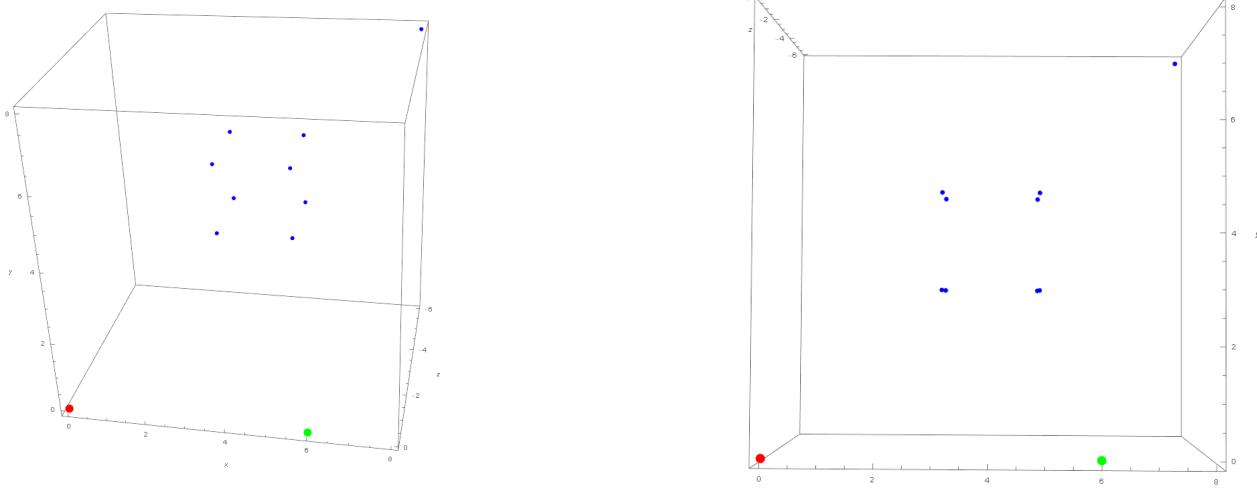


Abbildung 6.10: auf Originalgröße skalierte rekonstruierte Szene

Abbildung 4.7 zeigt die Komplett rekonstruierte Szene des Minimalbeispiels, welche beweist, dass die beschriebenen Methoden für das Minimalbeispiel mit reinen Daten und auch bei Kameras mit unterschiedlichen Auflösungen, funktioniert hat.

6.10 Rektifizierung

Im Minimalbeispiel ist die Triangulierung und Rekonstruktion der 3D-Weltpunkte ohne Kommunikationen durchführbar. Das liegt daran, dass mit reinen Werten gerechnet wird und Fehler wie Beispielsweise Bildrauschen nicht vorkommen. Im Minimalbeispiel kann davon ausgegangen werden, dass die Linien zweier korrespondierender Punkte sich, welche durch die jeweiligen Kamerazentren und Bildpunkte gehen, sich ziemlich sicher in einem Punkt treffen werden. In einem Beispiel mit realen Daten ist es nicht unwahrscheinlich, dass die herausgefilterten korrespondierenden Punkte, nicht zu hundert Prozent stimmen. Es kommt immer zu kleineren Abweichungen, was dazu führen kann, dass wenn die Linien der korrespondierenden Punkte im Realbild sich nicht treffen. Ein Grund dafür ist, dass sie nicht hundertprozentig auf der selben Höhe im Bild liegen und die Linien sich somit verfehlten

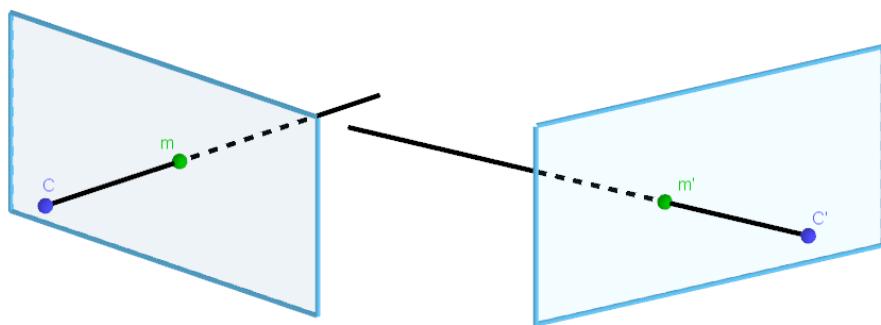


Abbildung 6.11: Durch Ungenauigkeiten in der korrespondierenden Punkte, verfehlen sich die Linien und es kommt zu keinem Schnittpunkt

Im Realbeispiel dieser Arbeit wird das Auftreten solcher Fehler durch das sogenannte *Sampson-Approximation* - Verfahren behoben, welches bei kalibrierten Fällen zum Einsatz kommt[3]. Mehr zu diesem Verfahren wird im Kapitel Minimieren der Kostenfunktion durch Sampson-approximation

aufgeführt. Das ist eine Möglichkeit um eine Szenerekonstruktion trotz Fehlerhafter korrespondierender Punkte zu ermöglichen. Ein weiteres weit verbreitetes Verfahren, ist cor der Szenenrekonstruierung durch Triangulierung eine Rektifizierung beider Bilder vorzunehmen[19, 20, 21, 22]. Da bestimmte Formen der Rektifizierung keine vorherige Kalibrierung der Kameras benötigen, wird diese Methode in den meisten gängigen Echtzeit-Szenenrekonstruktionen eingesetzt. [22, 21, 23]. Rektifizierte Bilder müssen zwei Eigenschaften erfüllen. Zum einen müssen alle Epipolareraden parallel zur x-Koordinatenachse verlaufen und zweitens müssen alle korrespondierenden Punkte die selben y-Koordinaten besitzen[20]. Mit Hilfe dieser Eigenschaften ist es somit möglich die entstandenen korrespondierenden Epipolarlinien als horizontale Scanlinien zu benutzen[21, 20]. Mit hilfe dieser Scanlinien und den darauf sich befindenden korrespondierenden Punkten ist es zum Beispiel Möglich eine Tiefenkarte des Bildes zu berechnen allein durch die Differenz der horizontalen Lage der korrespondierenden Punkte[21, 20].

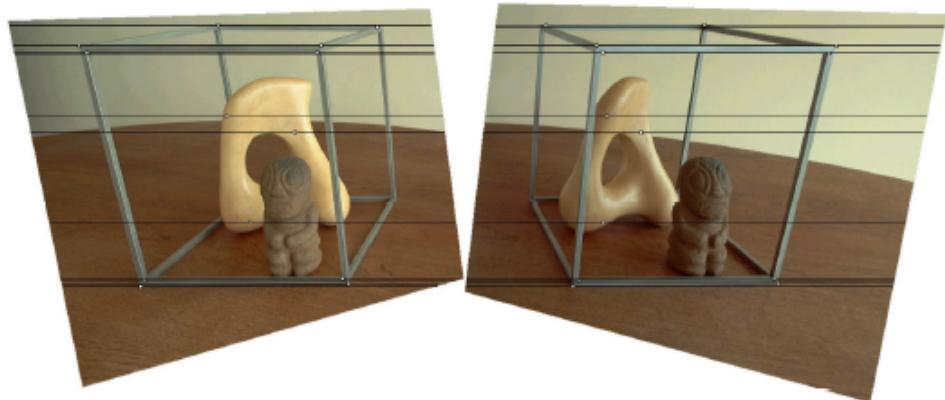


Abbildung 6.12: Beispiel eines rektifizierten Bildes. Quelle: [20]



Abbildung 6.13: Beispiel einer einfachen Tiefenkarte eines Stereobildpaars nach der Rektifizierung. Quelle: [21]

Die Rektifizierung, allem voraus vor allem die Optimierung des Rektifizierungsvorgangs, von Stereo- oder auch multplen- Kamerasystemen, wird heutzutage von vielen Entwicklergruppen der Computer Vision untersucht(Der satz ist mist!). Es gibt mittlerweile viele Ansätze, jedoch funktionieren nicht alle bei den selben Fällen. So setzten zum Beispiel manche Rektifizerungsalgorithmen voraus, dass die Bilder von Kameras mit selber Auflösung aufgenommen wurden. Ein Beispiel ist die Rektifizierung welche in *Matlab* verwendet wird [19]. Die Rektifizierung wurde anhand einer Methode implementiert, welcher sich ähnlich verhält wie in [24] beschrieben. Die Grundidee hier hinter ist, dass die Kameramatrizen von zwei Kameras so aufgebaut sind dass die intrinsischen Parameter die selben sind, sie sich aber in ihren Rotationen und Translationen voneinander unterscheiden. Die extrinsischen Kameraparameter werden dann dementsprechend so manipuliert, dass die Bildebenen Achsenparallel zueinander stehen[24, 22]. Um horizontale Epipolarlinien zu erhalten muss gleichzeitig die Basislinie zwischen den zwei Kamerazentren parallel zur neuen x-Achse beider Kameras sein. Zudem soll, um eine angemessene Rektifizierung zu gewährleisten, müssen konjugierende Punkte die selbe vertikale Koordinate haben. Dies wird hier durch die die Bedingung gewährleistet, dass beide Kameras die selben intrinsischen Parameter haben[24]. Eine Frage welche mit unter in dieser Arbeit beantwortet werden sollte, war, ob es möglich ist, ohne deutlich größeren Aufwand eine Kamerakalibrierung und Szenerekonstruktion mit Kameras unterschiedlicher Auflösung zu gewährleisten. Im Kapitel Epipolar Geometrie, in welchem ausführlich die Epipolargeometrie vorgestellt wurde, wurde bereits bezug auf die unterschiedlichen Auflösungen genommen. Prinzipiell spielen unterschiedliche intrinsische Kameraparameter keine Rolle, wenn es um die Rekonstruktion der Kameraposen geht, da die Fundamental Matrix und die essentielle Matrix die Information über die intrinsischen und extrinsischen Kameraparameter besitzen und es klar gestellt wurde, dass die Bildkoordinatensysteme der Kameras nicht identisch sein müssen. [5]. In dieser Arbeit wurde ein Rektifizierungsalgorithmus nach *Zhang*[20] implementiert, welcher sich die Fundamentalmatrix zu nutzen macht. *Loop* und *Zhang* zerlegen jede Kollinearität in eine Ähnlichkeitstransformation, eine Schertransformation und eine projektive Transformation. Die projektive Komponenten wird dabei in einem nichtlinearen Optimierungsprozess so affin wie möglich gemacht.[22, 20]. Im folgenden wird zunächst der genaue Vorgang des implementierten Algorithmus genauer erklärt und **des Weiteren werden zwei Beispiele vorgestellt, welche die Bilder des Minimalbeispiels einmal mit gleichen intrinsischen Parametern und einmal mit unterschiedlichen intrinsischen Parametern der Kamera aufzeigt. Es wird sich Herausstellen, dass beide Beispiele eine gelungene Rektifizierung der Bilder aufweisen.(Nochmal genau nachprüfen ob das geht!!!)**. Während sich einige Rektifizierungsverfahren im 3D-Raum abspielen, wird beim Verfahren nach *Zhang*, hauptsächlich im 2D-Raum gearbeitet. Des Weiteren wird vorausgesetzt, dass die Fundamental Matrix F und somit auch korrespondierende Punkte bereits bekannt sind. Sind die intrinsischen Kameraparameter bekannt, so wird aus der Fundamentalmatrix die Essentielle Matrix. Das Verfahren kann sowohl in einem kalibrierten als auch in einen unkalibrierten Fall angewendet werden[20]. Im Algorithmus wurde der unkalibrierte Fall implementiert und somit wird in der Erläuterung und in den danach folgenden Beispielen die Fundamentalmatrix F verwendet. Die korrespondierenden Punkte werden mit x für das erste beziehungsweise x' für das zweite Bild definiert, die Kamerazentren dementsprechend mit C und C' . Bildebene der ersten Kamera wird mit I definiert und die Bildebene von Kamera zwei mit I' , die entsprechenden Epipole mit e und e' . Der Prozess der im Algorithmus erfolgt kann quasi als eine Transformation der Epipolar Geometrie eines Bildpaars in eine kanonische Form angesehen werden. Diese Transformation wird durch eine Homographiematrix durchgeführt, welche sich aus den bereits erwähnten drei Komponenten zusammenstellt. Zu Beginn sei noch erwähnt dass wir pro Bild zwei unterschiedliche Homographien H und H' brauchen. Die Fundamentalmatrix liefert, die Epipolarbedingung, dass $x'^T F x = 0$ ergibt, wenn x' auf der zu x korrespondierenden Epipolarlinie liegt. Die korrespondierenden Punkte x und x' werden, für die Rektifizierung, jeweils mit den Homographien H und H' verrechnet.

$$\bar{x} = Hx \quad (6.60)$$

$$\bar{x}' = Hx' \quad (6.61)$$

Die Fundamentalmatrix, welche sich aus durch die Rektifizierten korrespondierenden Punkte resultiert, wird mit \bar{F} bezeichnet. Daraus folgt für die Fundamentalmatrix folgendes:

$$\bar{x}'^T \bar{F} \bar{x} = 0 \quad (6.62)$$

$$\rightsquigarrow x'^T H' T \bar{F} H x = 0 \quad (6.63)$$

$$\rightsquigarrow F = H'^T [i]_x H \quad (6.64)$$

Das Ziel ist es diese zwei Homographien in deren bereits erwähnten projektiven und affinen Komponenten zu zersetzten, wobei diese die jeweils entstehenden Bildverzerrungen minimieren sollen. Die Homographiematrizen bestehen aus drei Linien, welche jeweils durch den Epipol verlaufen. Des Weiteren werden noch ein paar weitere Bedingungen für die jeweils drei Linien festgelegt. So müssen die Linien v und v' sowie w und w' korrespondierende Epipolarlinien sein. Diese Bedingung schafft eine geometrische Verbindung beider Bilder zueinander und ist gerade bei der Minimierung der durch die Rektifizierung entstehenden Bildverzerrung von Bedeutung.

$$H = \begin{bmatrix} u^T \\ v^T \\ w^T \end{bmatrix} = \begin{bmatrix} u_a & u_b & u_c \\ v_a & v_b & v_c \\ w_a & w_b & w_c \end{bmatrix} \quad (6.65)$$

$$H' = \begin{bmatrix} u'^T \\ v'^T \\ w'^T \end{bmatrix} = \begin{bmatrix} u'_a & u'_b & u'_c \\ v'_a & v'_b & v'_c \\ w'_a & w'_b & w'_c \end{bmatrix} \quad (6.66)$$

Für die Bestimmung der einzelnen Komponenten von H und H' werden diese in ihre projektiven und affinen Teilstücke zerlegt. Davor wird noch die letzte Komponente w_c raus dividiert, um somit skaleninvariante Matrizen H und H' zu bekommen.

$$H = \begin{bmatrix} u^T \\ v^T \\ w^T \end{bmatrix} = \begin{bmatrix} u_a & u_b & u_c \\ v_a & v_b & v_c \\ w_a & w_b & 1 \end{bmatrix} \quad (6.67)$$

$$H' = \begin{bmatrix} u'^T \\ v'^T \\ w'^T \end{bmatrix} = \begin{bmatrix} u'_a & u'_b & u'_c \\ v'_a & v'_b & v'_c \\ w'_a & w'_b & 1 \end{bmatrix} \quad (6.68)$$

Beide Matrizen werden nun auf die selbe Weise in ihre projektiven und affinen Bestandteile zerlegt.

$$H = H_p \cdot H_a \quad (6.69)$$

$$H' = H'_p \cdot H'_a \quad (6.70)$$

H_p ist die projektive Komponente, sie bezieht sich nur auf die letzte Zeile der Matrix H und wirkt sich somit auch nur auf die homogenen Komponenten der mit ihr verrechneten Punkte aus.

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ w_a & w_b & 1 \end{bmatrix} \quad (6.71)$$

Die affine Komponenten H_a lässt sich aus H und H_p konstruieren. Es gilt:

$$H_a = H \cdot H_p^{-1} = \begin{bmatrix} u_a - v_c w_b & v_c w_a - v_a & 0 \\ v_a - v_c w_a & v_b - v_c w_b & v_c \\ 0 & 0 & 1 \end{bmatrix} \quad (6.72)$$

Für die Matrizen H'_p und H'_a gilt das selbe nur mit den Epipolarlinien u' , v' und w' . Die projektive Matrix sorgt dafür, dass die Epipole beider Bilder ins unendliche gesetzt werden und die Epipolarlinien der Bilder jeweils parallel zueinander verlaufen. Zu Beginn wurde erwähnt, dass es eine Zerlegung in eine projektive, eine Ähnlichkeits- und eine Scherungstransformation gibt. Die projektive Komponente ist mit H_p und H'_p bereits vollständig definiert. Was nun noch fehlt ist die Zerlegung der affinen Matrizen H_a und H'_a in ihre jeweiligen Ähnlichkeits- und Scherungstransformationen.

$$H_a = H_s \cdot H_r \quad (6.73)$$

$$H_r = \begin{bmatrix} v_b - v_c w_b & v_a - v_c w_a & 0 \\ v_a - v_c w_a & v_b - v_c w_b & v_c \\ 0 & 0 & 1 \end{bmatrix} \quad (6.74)$$

$$H_s = \begin{bmatrix} u_a & u_b & u_c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.75)$$

H_r und auch H'_r definieren eine Rotation und auch eine Verschiebung, welche die bereits parallelen Epipolarlinien beider Bilder zueinander parallel und horizontal ausrichtet. Durch die Verschiebung werden die korrespondierenden Epipolarlinien noch auf die selbe Höhe verschoben. Somit entstehen die gewünschten Scanlinien in den Bildern. Die Matrix H_s und H'_s wirken sich nur auf die u -Elemente der Matrix H und H' aus und definieren eine Scherung. Sie haben keine Auswirkung auf die Rektifizierung an sich, aber sorgen dafür, dass die horizontale Verzerrung der beiden Bilder zueinander reduziert wird.

6.10.1 Projektive Transformation

Die projektiven Matrizen H_p und H'_p werden von den Linien w und w' bestimmt. w und w' sind dabei jedoch nicht unabhängig. Definiert werden sie durch einen Punkt $z = [\lambda \ \mu \ 0]^T$, welche die, durch die Rektifizierung entstehende, Bildverzerrung minimieren soll. Für beide Bilder werden w und w' folgendermaßen gewählt

$$w = [e]_x \cdot z \quad (6.76)$$

$$w' = F \cdot z \quad (6.77)$$

Jedes beliebige z würde zwei korrespondierende Epipolarlinien definieren, um ein z zu finden, welches die Verzerrung der Bilder minimiert, wird ein Kriterium aufgestellt, welches ein z finden soll, dass

die Verzerrung minimal halten wird. Minimierung bedeutet in diesem Falle, dass versucht wird die Matrizen H_p und H'_p so affin wie möglich zu machen. So affin wie möglich bedeutet, dass die Werte von w_a und w_b so nah wie möglich an den Wert 0 gebracht werden sollen.

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ w_a & w_b & 1 \end{bmatrix} \quad (6.78)$$

Jedoch sollen sie nicht ganz null werden, da die projektive Matrix dann keine projektive mehr wäre, sondern eine affine. Deswegen heißt es auch sie soll so affin wie möglich gemacht werden. Das selbe gilt natürlich auch für w'_a und w'_b aus H'_p . Wäre das der Fall, so wären die beiden Epipole e und e' bereits im unendlichen und die Matrizen H_p und H'_p hätten keine Auswirkungen auf die Punkte. Für die Minimierung wird die Methode des *least-square-fitting*, also die Anpassung des kleinsten Quadrats, genutzt[25]. Es werden also die Gewichtungen der Punkte in beiden Bildern in der Methode der Anpassung der kleinsten Quadrate verbaut, welche versucht eine Funktion zu finden, die einen Wert für z berechnen soll welcher die Bildverzerrung minimal hält. **Anders ausgedrückt man sucht einen Wert für z , welcher am nächsten an den gegebenen Punktesammlungen der jeweiligen Bildern dran liegt, wobei für z bereits gilt, dass es sich um einen Punkt im Unendlichen handeln soll**[20, 25]. Angenommenem, dass die Annäherungsfunktion $g(x)$ eine Funktion $f(x)$, mit $x \in [a, b]$, annähern soll, dann versucht die Methode, die Summe der Quadrate der ordinatischen Differenzen, welche zwischen den von der Funktion generierten Punkten und den Punkten aus den Daten gewonnen wird, zu minimieren[25, 26]. Zum Beispiel werden n Datenpunkte angenommen, dann gilt:

$$e = \sum_{i=1}^n [f(x_i) - g(x_i)]^2 \quad (6.79)$$

Für die Minimierung der Bildverzerrung werden die Gewichtungen der Punkte beider Bilder benötigt. p_i beinhaltet alle Punkte von Bild eins und p_j beinhaltet alle Punkte von Bild zwei. Angenommen wir nehmen einen Punkt aus Bild eins $p_{i1} = [p_{i1,u} \ p_{i1,v} \ 1]^T$, so soll dieser Punkt mit der Matrix H_p zu einem Punkt der Form $p_{i1} = [\frac{p_{i1,u}}{w_i} \ \frac{p_{i1,v}}{w_i} \ 1]^T$ transformiert werden. w_i ist die Gewichtung welche durch die Verrechnung von w mit p_i zustande kommt.

$$w_i = w^T p_i \quad (6.80)$$

Ist die Gewichtung der Punkte identisch gibt es keine projektive Verzerrung und die Homographie ist eine affine Transformation. Jedoch wenn die Epipole der Bilder ins Unendliche transformiert werden sollen, so können H_p und H'_p keine affine Homographien sein. Sonst könnte man die Epipole nur innerhalb der affinen Ebenen, sprich den Bildebenen, verschieben. Also bildet der Versuch H_p und H'_p so affin wie möglich zu machen die Basis für die Minimierung. Im Realbeispiel werden alle Pixel des Bildes verwendet. Die Rektifizierung wurde im aufgeführten Beispiel anhand des erstellten Minimalbeispiels durchgeführt, somit wurden die Eckpunkte des Quaders des jeweiligen Bildes für die das Minimierungskriterium verwendet. Es wird eine Funktion nach dem Prinzip der Anpassung der kleinsten Quadrate aufgestellt, welche die Abweichung der Gewichtung der Punkte in Bezug auf die Gewichtung des Bildzentrums p_c berechnet. p_c ergibt sich aus der Mittelung aller verwendeten Punkte eines Bildes $p_c = \frac{1}{n} \sum_{i=1}^n p_i$, dessen Gewichtung ergibt sich aus $w_c = w^T p_c$. Die gesuchte Abweichung ausgedrückt in der Anpassung der kleinsten Quadrate ergibt dann die folgende Formel.

$$\sum_{i=1}^n \left[\frac{w_i - w_c}{w_c} \right]^2 \quad (6.81)$$

$$\rightsquigarrow \sum_{i=1}^n \left[\frac{w^T(p_i - p_c)}{w^T p_c} \right]^2 \rightsquigarrow \sum_{i=1}^n \left[\frac{w^T(p_i - p_c)(p_i - p_c)^T w}{w^T p_c p_c^T w} \right] \quad (6.82)$$

Vereinfacht lässt sich das auch in einer Matrixgleichung angeben

$$\frac{w^T P P^T w}{w^T p_c p_c^T w} \quad (6.83)$$

in welcher für P gilt:

$$P = \begin{bmatrix} p_{1,u} - p_{c,u} & p_{2,u} - p_{c,u} & \dots & p_{i,u} - p_{c,u} \\ p_{1,v} - p_{c,v} & p_{2,v} - p_{c,v} & \dots & p_{i,v} - p_{c,v} \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad (6.84)$$

die Gleichungen 4.79 bis 4.83 werden ebenfalls für die Punkte p_j in Bild zwei aufgestellt. So ergibt sich für das zweite Bild die Matrixgleichung:

$$\frac{w'^T P' P'^T w'}{w'^T p'_c p'_c^T w'} \quad (6.85)$$

Das Ziel ist es einen Wert für z zu finden, welches bis jetzt noch nicht ersichtlich in den Gleichungen vorkommt. Also werden w und w^T noch mit ihren Definitionen aus den Gleichungen 4.75 und 4.76 ersetzt. Gleichzeitig werden die Gleichungen 4.82 und 4.84 summiert um die Gleichung zu erhalten, welche sich auf beide Bilder gleichzeitig bezieht und somit eine Lösung für z , das für beide Bilder gilt, gesucht werden kann.

$$\frac{z^T [e]_x^T P P^T [e]_x z}{z^T [e]_x^T p_c p_c^T [e]_x z} + \frac{z^T F^T P' P'^T F z}{z^T F^T p'_c p'_c^T F z} \quad (6.86)$$

Für den weiteren Verlauf werden die Ausdrücke noch durch die Variablen A, B, A' und B' vereinfacht.

$$A = [e]_x^T P P^T [e]_x \quad (6.87)$$

$$B = [e]_x^T p_c p_c^T [e]_x \quad (6.88)$$

$$A' = F^T P' P'^T F \quad (6.89)$$

$$B' = F^T p'_c p'_c^T F \quad (6.90)$$

$$\rightsquigarrow \frac{z^T A z}{z^T B z} + \frac{z^T A' z}{z^T B' z} \quad (6.91)$$

Da die dritte Komponente von z laut definition null sein soll, wird zu $z = \begin{pmatrix} \lambda \\ \mu \end{pmatrix}$ umgeschrieben. A, B, A' und B' sind 3x3-Matrizen, von welchen uns dann nur noch der erste 2x2-Block interessiert. Bei dem somit aufgestellten Minimalisierungs Kriterium, handelt es sich um ein nicht lineares optimierungs Problem. Die Gleichung 4.90 ist dann minimiert, wenn die erste Ableitung dieser Funktion nach $\lambda =$ gleich null ist. Es entsteht also ein Polynom mit dem Grad sieben, da die 4.90 die Summe zweier rationaler Funktionen ist, welche jeweil das Verhältnis von quadratischen Polynomen darstellt.

$$Hier soll das Polynom aufgestellt werden, ist aber nicht mehr klar wie das ging!!! \quad (6.92)$$

Für die nicht lineare Optimierung wird das gesamte Polynom aufgeteilt, so minimieren wir zunächst $\frac{z^T A z}{z^T B z}$ und danach $\frac{z^T A' z}{z^T B' z}$. So entstehen für z zunächst zwei Lösungen \hat{z}_1 und \hat{z}_2 , welche über eine Mittelung eine ersten Schätzung für z geben, welche schon ziemlich nah an den optimalen Wert heranreicht.

$$z = \frac{\frac{\hat{z}_1}{\|\hat{z}_1\|} + \frac{\hat{z}_2}{\|\hat{z}_2\|}}{2} \quad (6.93)$$

Da es sich um eine nicht lineare Optimierung handelt ist die Minimierung von $\frac{z^T A z}{z^T B z}$ gleichzusetzen mit der Maximierung von $\frac{z^T B z}{z^T A z}$. Beide als eine Funktion von $f(z)$. Matrix A wird mit der Cholesky-Zerlegung in zwei höhere Dreiecksmatrizen zerlegt $A = D^T D$ [27]. Dies geht nur da A nachweislich eine symmetrische und positiv-definite Matrix ist.[27] positiv-Definite bedeutet, dass die Singulärwerte von A immer positiv bleiben, egal mit welchem Vektor z diese multipliziert wird. (**HIER NOCH LITERATUR FINDEN UND NOCHMAL PRÜFEN OB DEFINITION SO STIMMT**). Des Weiteren wir definiert, dass $y = Dz$ ist und $f(z)$ wird dann zu einen $\hat{f}(y)$

$$A = D^T D \quad (6.94)$$

$$y = Dz \rightsquigarrow z = D^{-1}y \quad (6.95)$$

$$f(z) = \frac{z^T B z}{z^T A z} \quad (6.96)$$

$$\rightsquigarrow f(z) = \frac{z^T B z}{z^T D^T D z} \quad (6.97)$$

$$\hat{f}(y) = \frac{y^T D^{-T} B D^{-1} y}{y^T y} \quad (6.98)$$

Durch die Defintion von $y = Dz$ ist y bis auf einen Skalierungsfaktor definiert. $\hat{f}(y)$ ist maximiert, wenn y gleich dem Eigenvektor von $D^{-T} B D - 1$ ist, welcher mit dem größten Eigenwert assoziiert wird. Zum Schluss erhalten wir dann einen Wert für \hat{z}_1 mit $\hat{z}_1 = D^{-1}y$. Exakt das selbe Verfahren wird für die Findung von z_2 mit $\frac{z^T B' z}{z^T A' z}$ angewandt. Sind z_1, z_2 und eine erste Schätzung für z gefunden, so kann ein Wert für z gesucht werden, welcher noch näher an ein optimales Ergebnis heranreicht. Beide Lösungen z_1 und z_2 , werden in die Funktion $f(z)$ eingesetzt und es jeweils ein wert ermittelt, welcher am nächsten an einem Nullpunkt sich befindet. So kann iterativ eine optimale Lösung für z gefunden werden. Ist der Wert für z bestimmt, so kann dieser die Gleichungen 4.75 und 4.76 eingesetzt werden und w beziehungsweise w' bestimmt werden, welche die Elemente für die Matritzen H_p und H'_p bereitstellen. Abbildung 4.7 zeigt in rot den Quader des Minimalbeispiels wie er in Kamera zwei abgebildet ist und in grün wie er in Kamera eins abgebildet ist. Kamera zwei ist horizontal zu Kamera eins verschoben und um 45° zu Kamera eins um die eigene vertikale Achse gedreht. Die Auflösungen beider Kameras sind identisch, sprich die intrinsischen Kameraparameter sind die selben. Abbildung 4.8 zeigt die momentanen Epipolarlinien. Die Epipolarlinien von Bild eins, also dem grünen Abbild, sind bereits Parallel, was aber keine Voraussetzung für die Funktion des Rektifizierungsalgorithmus ist. Der Schnittpunkt der Epipolarlinien von Bild zwei, also dem Roten Abbild, treffen sich in einem Punkt und bilden somit den Epipol von Bild zwei.

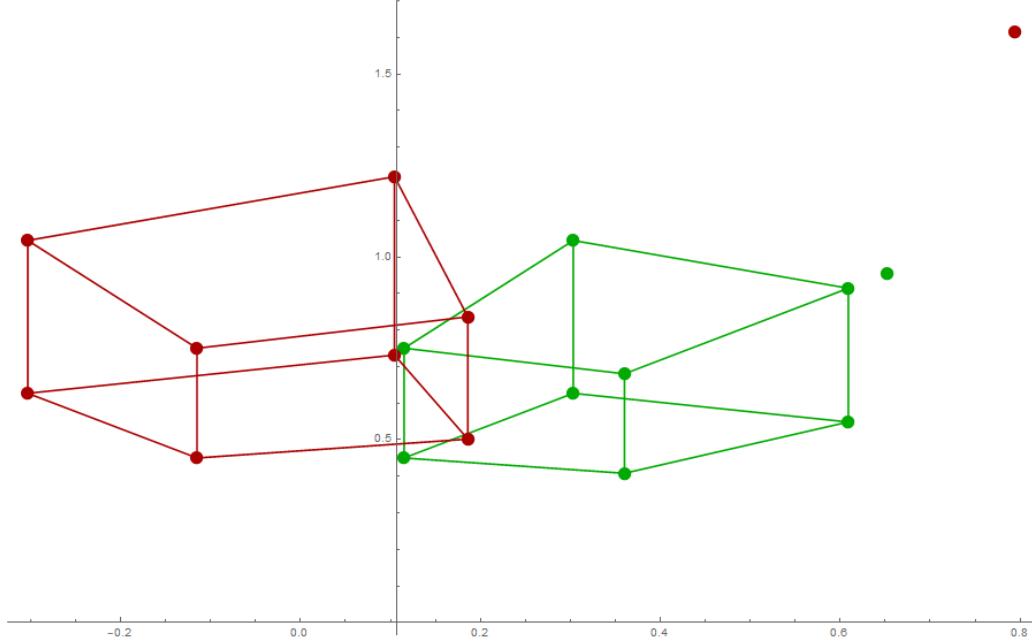


Abbildung 6.14: Aufnahmen zweier Kameras mit den selben Auflösungen, Kamera eins(Grün) und Kamera(rot) zwei gelten jeweils $\zeta = 1$

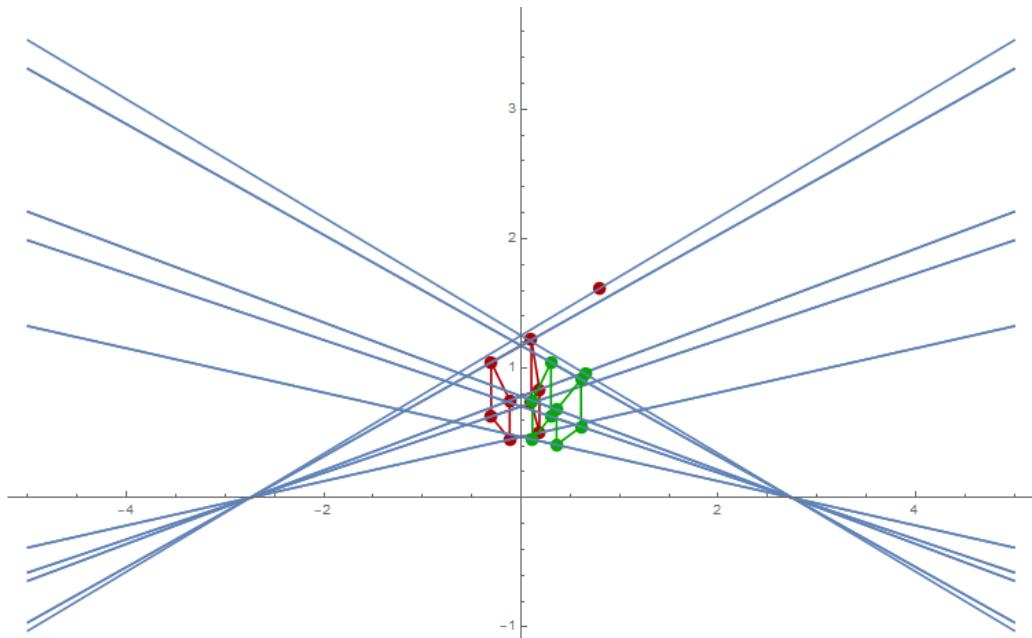


Abbildung 6.15: Epipole für Kamera eins und Kamera zwei vor der Rektifizierung

Werden nun die Matritzen H_p und H'_p auf die jeweiligen Punkte der Bilder, p_i für Bild eins und p_j für Bild zwei, angewandt, so kann man eine erste Veränderung beobachten. Abbildung 4.9 zeigt beide Quader aus Abbildung 4.7 nachdem die jeweiligen Bildpunkte mit den projektiven Matrizen multipliziert wurden. Der Epipol in Bild eins bleibt natürlich wie zuvor im unendlichen, jedoch kann man erkennen, dass der rote Quader aus Bild zwei sich verändert hat. Sein Epipol wurde ins Unendliche transformiert und parallele Linien sind nun auch auf dem Bild parallel. Das die Epipolarlinien bereits horizontal parallel zur x-Achse verlaufen ist Zufall und ist nach der Anwendung der projektiven Matrizen auch noch nicht verlangt. Das Anpassen der Epipolarlinien, dazu gehört sie zunächst von beiden Bildern aus parallel zur x-Achse verlaufen zu lassen und dann noch sie so zueinander anzupassen, dass sie zu Scanlinien über beide Bilder verlaufen, vergleiche Abbildung 4.12, folgt im nächsten Schritt.

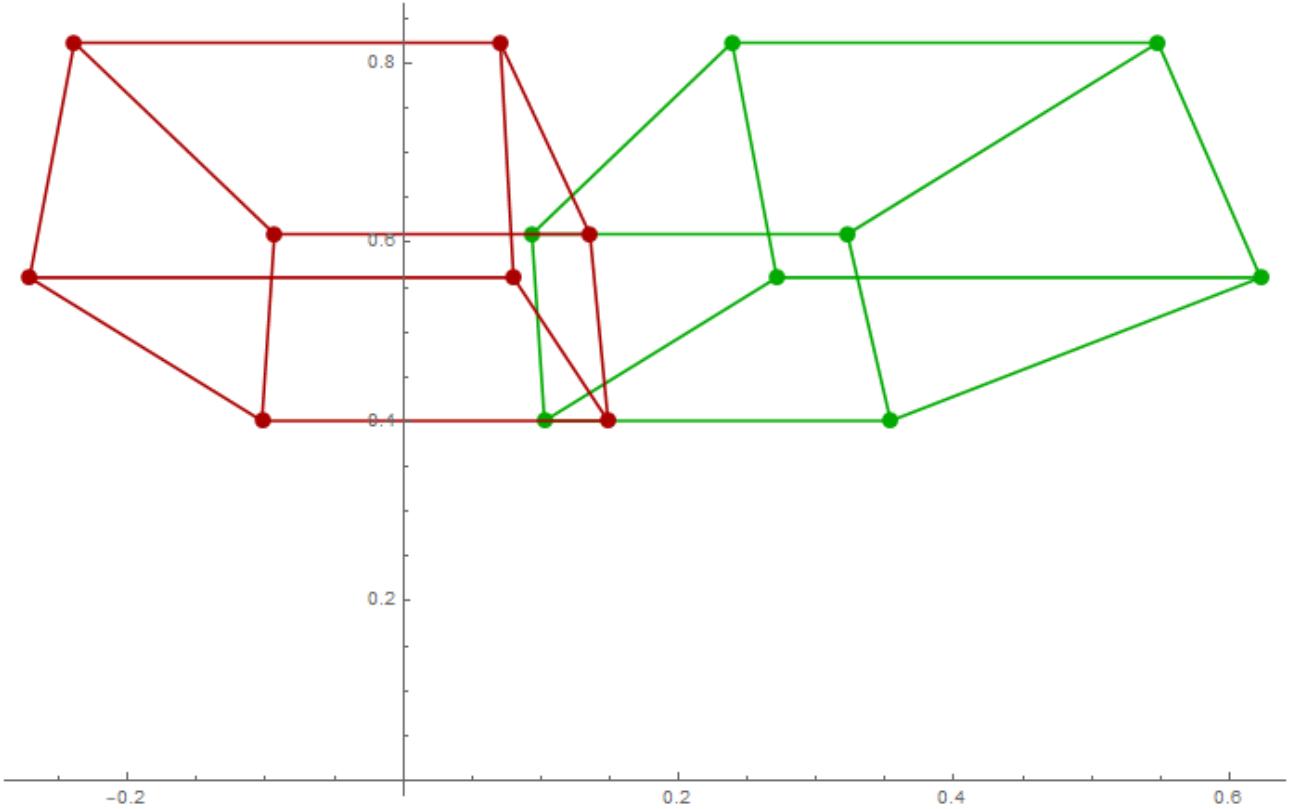


Abbildung 6.16: Abbildung beider Bilder nach anwenden der Matrizen H_p und H'_p . Die Epipole beider Bilder sind nun im unendlichen. Das die entstehenden parallelen Epipolarlinien auch hier schon horizontal ausgerichtet sind ist Zufall. Die Epipolarlinien sind immer parallel nach dieser Transformation aber die Richtung ist nicht immer automatisch bereits $i = [1,0,0]$.

6.10.2 Ähnlichkeitstransformation

Nachdem die Epipole ins Unendliche verschoben wurden, müssen diese nun so rotiert und verschoben werden, dass die Epipolarlinien als Richtung $i = [1 \ 0 \ 0]$ haben und die Epipolarlinien beider Bilder zu einheitlichen Scanlinien werden. Für die Ähnlichkeitstransformation wird davon ausgegangen, dass w und w' bereits bekannt sind. H_r und H'_r wurden bereits aus der Zerlegung von H_a und H'_a gewonnen.

$$H_r = \begin{bmatrix} v_b - v_c w_b & v_a - v_c w_a & 0 \\ v_a - v_c w_a & v_b - v_c w_b & v_c \\ 0 & 0 & 1 \end{bmatrix} \quad (6.99)$$

$$H'_r = \begin{bmatrix} v'_b - v'_c w'_b & v'_a - v'_c w'_a & 0 \\ v'_a - v'_c w'_a & v'_b - v'_c w'_b & v'_c \\ 0 & 0 & 1 \end{bmatrix} \quad (6.100)$$

$$(6.101)$$

w und w' sind bereits bekannt. Mit Hilfe von F , können v_a und v_b ersetzt werden. Dazu kann die letzte Zeile von F nach v_a, v_b und v_c aufgelöst werden. Für v'_a, v'_b und v'_c wird die letzte Spalte von F verwendet. So können folgende Gleichungen für $v_a, v'_a, v_b, v'_b, v_c$ und v'_c gewonnen werden.

$$F = H'^T[i]_x H \quad (6.102)$$

$$F = \begin{bmatrix} v_a w'_a - v'_a w_a & v_b w'_a - v'_a w_b & v_c w'_a - v'_a \\ v_a w'_b - v'_b w_a & v_b w'_b - v'_b w_b & v_c w'_b - v'_b \\ v_a - v'_c w_a & v_b - v'_c w_b & v_c - v'_c \end{bmatrix} \quad (6.103)$$

$$v_a = F_{31} + v'_c w_a \quad (6.104)$$

$$v_b = F_{32} + v'_c w_b \quad (6.105)$$

$$v_c = F_{33} + v'_c \quad (6.106)$$

$$v'_a = v_c w'_a - F_{13} \quad (6.107)$$

$$v'_b = v_c w'_b - F_{23} \quad (6.108)$$

$$v'_c = v_c - F_{33} \quad (6.109)$$

Eingesetzt in die jeweiligen Matrizen H_r und H'_r , entstehen die folgenden Matrizen in Gleichungen 4.114 und 4.115, welche nur noch die unbekannte v'_c beinhalten. Die gemeinsame Variable v'_c zeigt die geometrische Verbindung beider Bilder in ihrer Verschiebung entlang ihrer v-Richtung. Es wird also ein Offset von F_{33} benötigt, um die Epipolarlinien horizontal zu Scanlinien auszurichten. Den Wert für v_c wird so ermittelt, dass das Minimum einer v-Koordinaten eines Pixel als minimum den Wert null besitzt

$$H_r = \begin{bmatrix} F_{32} - w_b F_{33} & w_a F_{33} - F_{31} & 0 \\ F_{31} - w_a F_{33} & F_{32} - w_b F_{33} & F_{33} + v'_c \\ 0 & 0 & 1 \end{bmatrix} \quad (6.110)$$

$$H'_r = \begin{bmatrix} w'_b F_{33} - F_{23} & F_{13} - w'_a F_{33} & 0 \\ w'_a F_{33} - F_{13} & w'_b F_{33} - F_{23} & v'_c \\ 0 & 0 & 1 \end{bmatrix} \quad (6.111)$$

Das Ergebnis der Bildpunkte p_i und p_j multipliziert mit den Matrizen $H_r H_p$ und $H'_r H'_p$ mit ist in Abbildung 4.10 zu sehen. Als letztes folgt noch die Scherungstransformation H_s und H'_s für die horizontale Entzerrung beider Bilder.

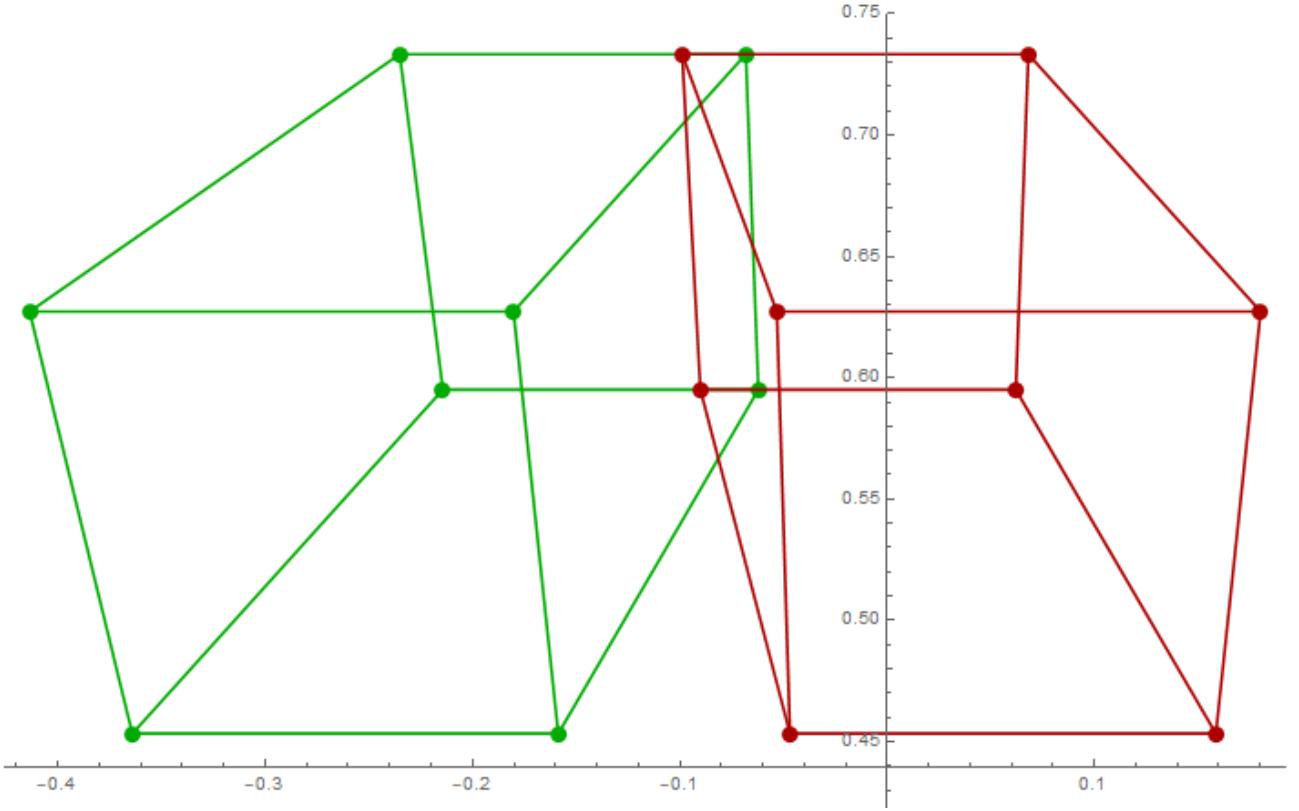


Abbildung 6.17: Abbildung beider Bilder nach anwenden der Matrizen $H_r \cdot H_p$ und $H'_r \cdot H'_p$. Die Epipolarlinien sind nun horizontal zueinander ausgerichtet

6.10.3 Scherungstransformation

Die letzte Transformation, welche an den Bilder durchgeführt werden soll, ist die sogenannten Scherungstransformation. Sie soll vor allem dazu dienen, die horizontale Verzerrung der Bilder zueinander nochmal weiter zu minimieren. Die Matrizen H_s und H'_s wirken sich hauptsächlich auf die u und u' Komponenten aus.

$$H_s = \begin{bmatrix} u_a & u_b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.112)$$

$$H'_s = \begin{bmatrix} u'_a & u'_b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.113)$$

Um die richtigen Werte für a, a', b und b' zu bekommen, werden zunächst Punkte an den jeweiligen gegenüberliegenden Kanten der Bilder definiert. Da die Bilder des Quaders nicht aus tausenden von Pixeln bestehen, wie ein reales Bild, sondern nur über dessen Eckpunkte bestimmt ist, wird eine Bildbreite w und w' und eine Bildhöhe h und h' definiert. Die Höhen und Breiten der Bilder rahmen die abgebildeten Quader ein, somit wurde quasi eine Bildgröße für beide Bilder definiert. Nun können die Punkte an den Kantenhalbierenden $a = [\frac{w-1}{2} \ 0 \ 1]^T, b = [w-1 \ \frac{h-1}{2} \ 1]^T, c = [\frac{w-1}{2} \ h-1 \ 1]^T, d = [0 \ \frac{h-1}{2} \ 1]^T$ gebildet werden. Der Gedanke, der damit verfolgt wird ist, dass die Punkte der jeweiligen gegenüberliegenden Kanten mit einander verbunden werden können und dann so ausgerichtet werden sollen, dass sie sich wieder direkt gegenüber liegen. Schematisch wird as in Abbildung ???? aufgezeigt.

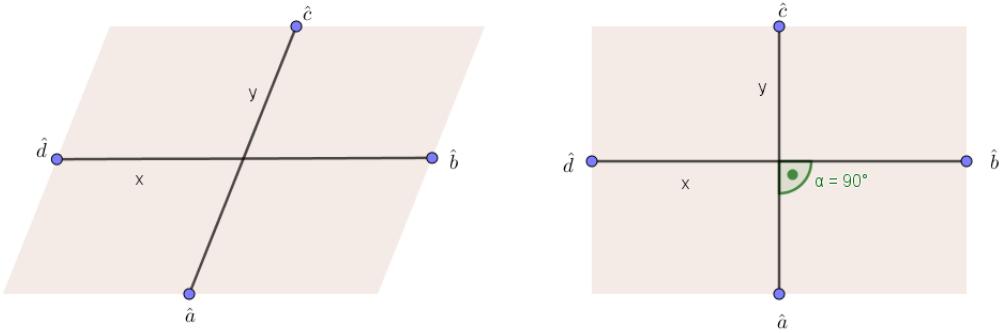


Abbildung 6.18: Die Abbildung verdeutlicht noch mal Schematisch, wie sich die Punkte ausrichten sollen. Bild a) zeigt die durch die Rektifizierung verschobenen Bildkantenmitten. Bild b= zeigt, wie sich die Bildkantenmitten durch die Scherungstransformation wieder ausrichten sollen.

Die Punkte a, b, c, d und auch a', b', c', d' geben die Bildbreiten der noch unberührten Bilder an. Nach der Rektifizierung sind die Bilder so verzerrt, dass die Kanten mitten sich meistens nicht mehr direkt gegenüber von einander befinden. Die Punkte a, b, c, d und a', b', c', d' werden mit den Matrizen H_p, H'_p, H_r und H'_r verrechnet, so dass man die genaue neue Position der Kanten Mitten nach der Rektifizierung hat.

$$\hat{a} = H_r \cdot H_p \cdot a$$

$$\hat{b} = H_r \cdot H_p \cdot b$$

$$\hat{c} = H_r \cdot H_p \cdot c$$

$$\hat{d} = H_r \cdot H_p \cdot d$$

$$\hat{a}' = H'_r \cdot H'_p \cdot a'$$

$$\hat{b}' = H'_r \cdot H'_p \cdot b'$$

$$\hat{c}' = H'_r \cdot H'_p \cdot c'$$

$$\hat{d}' = H'_r \cdot H'_p \cdot d'$$

Um aus $\hat{a}, \hat{b}, \hat{c}, \hat{d}$ und auch $\hat{a}', \hat{b}', \hat{c}', \hat{d}'$ wieder Punkte der affinen Ebene zu machen werden sie jeweils durch ihre dritte Komponenten geteilt, so das $\hat{a}_w, \hat{b}_w, \hat{c}_w, \hat{d}_w$ und $\hat{a}'_w, \hat{b}'_w, \hat{c}'_w, \hat{d}'_w$ jeweils den Wert eins besitzen. Danach können die Vektoren \vec{x} und \vec{y} aus den Differenzen der sich ursprünglich gegenüberliegenden Punkte gebildet werden.

$$x = \hat{b} - \hat{d} \quad (6.114)$$

$$y = \hat{c} - \hat{a} \quad (6.115)$$

$$x' = \hat{b}' - \hat{d}' \quad (6.116)$$

$$y' = \hat{c}' - \hat{a}' \quad (6.117)$$

x und y sind Vektoren der euklidischen Bildebene. Die Rechtwinkligkeit beider wird also erhalten, wenn gilt:

$$(H_s x)^T (H_s y) = 0 \quad (6.118)$$

$$(H'_s x')^T (H'_s y') = 0 \quad (6.119)$$

Die Seitenverhältnisse der Bilder werden beibehalten, wenn gilt:

$$\frac{(H_s x)^T (H_s x)}{(H_s y)^T (H_s y)} = \frac{w^2}{h^2} \quad (6.120)$$

$$\frac{(H'_s x')^T (H'_s x')}{(H'_s y')^T (H'_s y')} = \frac{w'^2}{h'^2} \quad (6.121)$$

Für u_a, u'_a, u_b und u'_b jeweils Gleichungen auf Basis der jeweiligen Bild Höhen und Breiten w, w', h, h' und x, x', y und y' und unter einhaltung der Aussagen der Gleichungen 5.118 bis 5.121, aufgestellt werden[20, 28].

$$u_a = \frac{h^2 x_v^2 + w^2 + y_v^2}{h w (x_v y_u - x_u y_v)} \quad (6.122)$$

$$u_b = \frac{h^2 x_u x_v + w^2 y_u y_v}{h w (x_u y_v - x_v y_u)} \quad (6.123)$$

Selbe Gleichungen werden auch für u'_a und u'_b aufgestellt. Das Ergebnis der Scherungstransformation ist in Abbildung 5.17 dargestellt. Wie zu sehen ist, ist die Minimierung noch nicht zu hindert prozent perfekt, hierfür müsste man noch ein paar mehr Interationsschritte bei finden von z einfügen.(ICH WEIß GANZ EHRLICH NICHT WORAN ES LIEGT...)

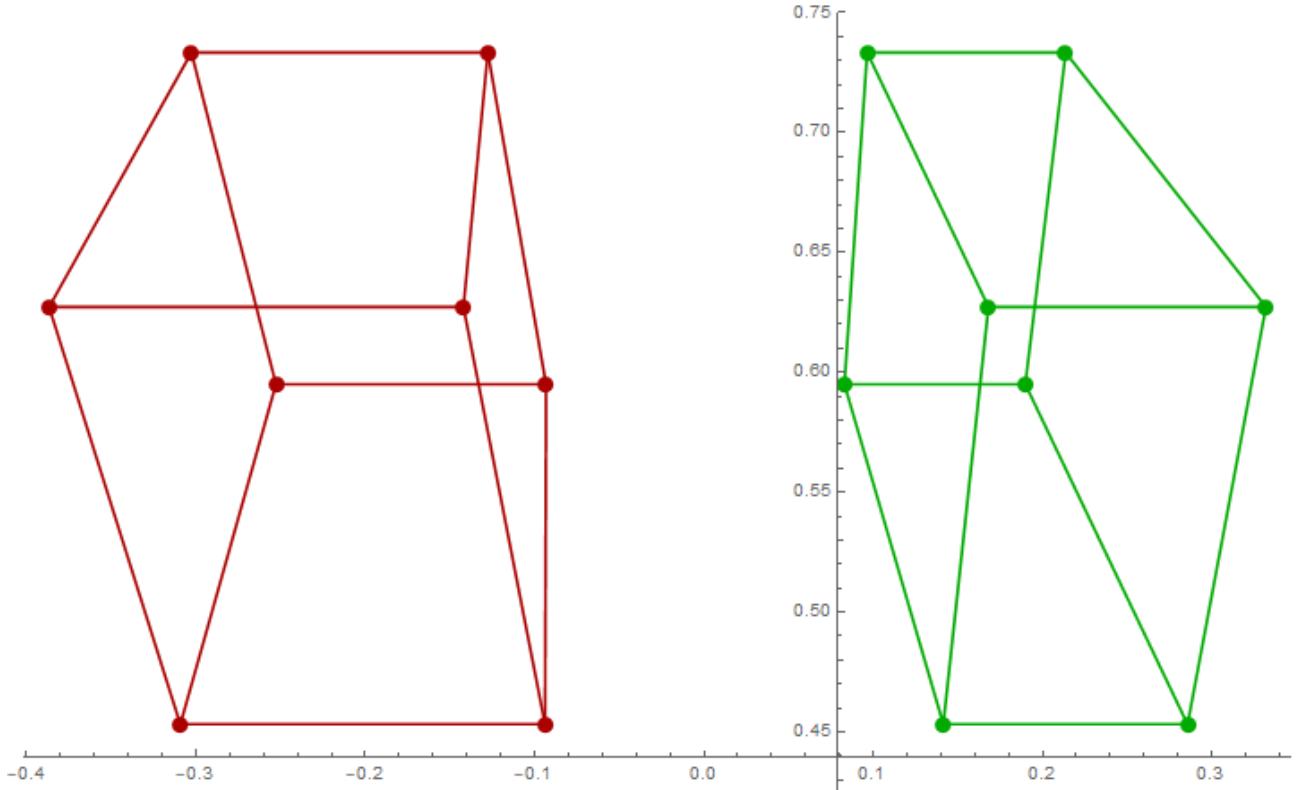


Abbildung 6.19: Abbildung beider Bilder nach anwenden der Matrizen $H_s \cdot H_r \cdot H_p$ und $H'_s \cdot H'_r \cdot H'_p$. Die horizontale Verzerrung wurde reduziert.

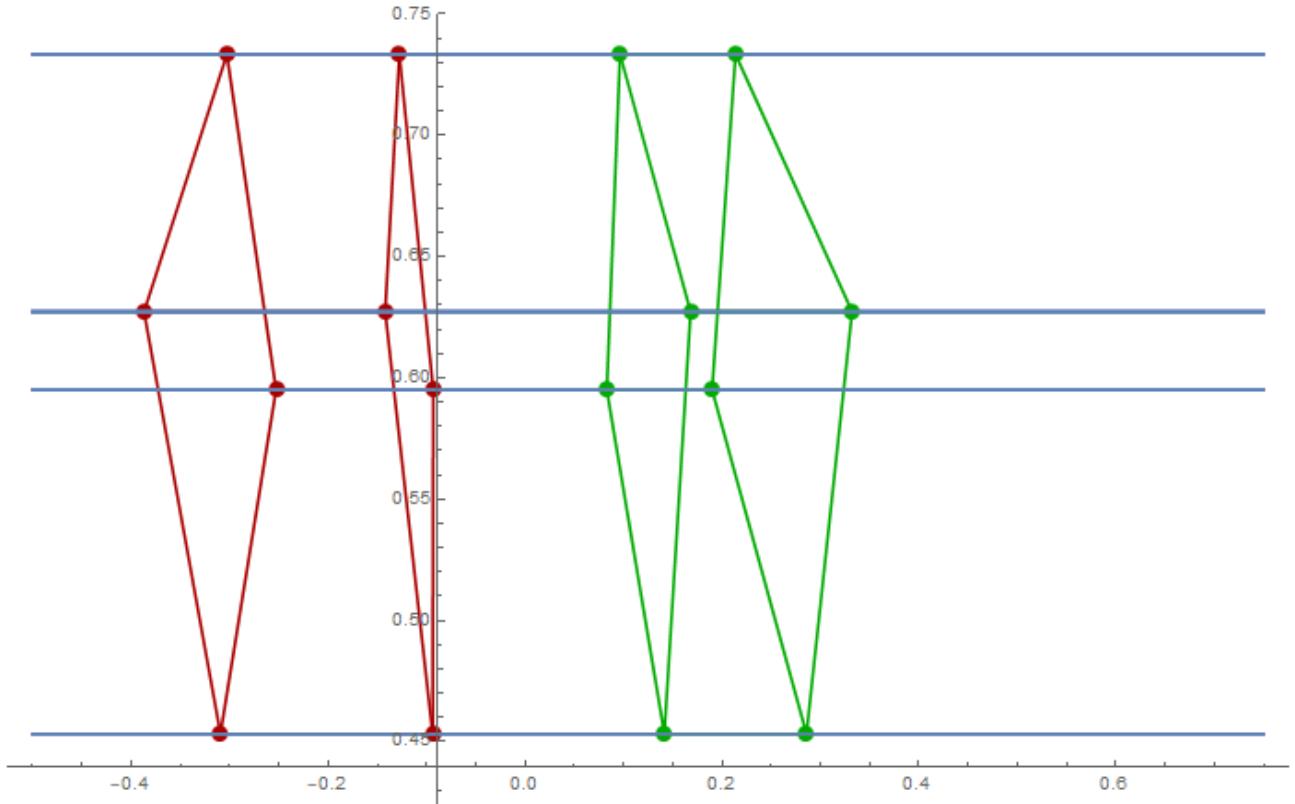


Abbildung 6.20: In dieser Abbildung wurden die Epipolarlinien noch in den Grafikplot mit eingebaut

7 Auswirkungen von unterschiedlichen Kameraauflösungen

Wenn man sich mit digitalen Bilddaten auseinandersetzt, so kommt man nicht drum herum sich auch mit den verschiedenen Auflösungsarten beschäftigen zu müssen. Im vorherigen Kapitel wurde gezeigt, wie eine 3D-Szene aus zwei heterogenen Bildquellen, welche von zwei Kameras gleicher Auflösung aufgenommen wurden, rekonstruiert werden konnte und gleichzeitig noch die externen Kameraparameter von C' in Relation zu C geschätzt werden konnten. Dies wirft die Frage auf, welche Auswirkungen Bilder zweier Kameras mit unterschiedlichen Auflösungen für die Rekonstruktionen haben können. Aus der Stereokalibrierungsapp, welche *MatLab* anbietet, ist bekannt, dass diese nicht mit Bildern unterschiedlicher Auflösung eine Szeneriekonstruktion durchführen kann. Der erste Schritt bestand erstmal darin zu überprüfen, warum zwei unterschiedliche Auflösungen in *MatLab* Probleme machen. *MatLab* verfolgt einen etwas anderen Rekonstruktionsansatz. Zu aller erst werden die Kameras kalibriert. Dies geschieht über die Matlab-Funktion `estimateCameraParameters`[29]. Diese Funktion funktioniert auch bei Bildern unterschiedlicher Auflösung noch ohne Probleme. Das Problem, welches sich als eigentlich minimales Problem herausstellt, ist, dass die darauf folgenden Rektifizierung der Stereobilder nicht mit zwei Bildern unterschiedlicher Auflösung funktioniert. In den *MatLab* references, steht es nicht explizit drin [19]. Die Rektifizierung in Matlab funktioniert nach einem Schema, welches ähnlich dem aufgezeigten Beispiel im Kapitel Rektifizierung bereits erwähnt wurde und in [22, 24] nochmal genau aufgeführt wird. Das Problem liegt also nicht daran, dass Bilder unterschiedlicher Auflösung nicht rektifiziert werden können, sondern das Problem liegt an dem in *MatLab* verwendeten Algorithmus für die Rektifizierung zweier Stereobilder (Foren Zitieren????). Warum *MatLab* überhaupt rektifiziert, liegt daran, dass ein Ansatz der Szeneriekonstruktion gewählt wurde, welcher die essentielle Matrix nicht benötigt. In diesem Falle, werden zwei Stereobilder aufgenommen, danach rektifiziert und anschließend über eine sogenannte *Disparity-Map*, die Szenen Punkte rekonstruiert[30, 6, 22, 21]. Der in dieser Arbeit gewählte Rektifizierungsalgorithmus, ist nicht auf gleiche Kameraauflösungen beschränkt. Mittlerweile gibt es natürlich deutlich fortgeschrittenere Rektifizierungsansätze, jedoch wurde für diese Arbeit der Ansatz von [20] gewählt, um ein Gefühl zu vermitteln, dass wenn man sich auf die Epipolargeometrie bezieht, die Auflösungen der Kameras keine Rolle spielen[5]. Was genau unterschiedliche Auflösungen der Kameras für die einzelnen Bilder bedeutet und was genau sich bei der Aufnahme mit dem Sensor dabei ändert, soll im folgenden Unterkapitel kurz erläutert werden. Danach wird aufgezeigt, was genau diese Veränderungen für die Epipolargeometrie bedeuten und letztendlich wird das Minimalbeispiel so angepasst, als hätten zwei Kameras unterschiedlicher Auflösung die selbe Szene wie davor aufgenommen und die Resultate miteinander verglichen.

7.1 Abbildungsunterschiede

Das Herz einer modernen Kamera ist der Halbleiter-Bildsensor. Die Bildsensoren spielen in der optischen Messtechnik und industriellen Bildverarbeitung eine große Rolle[10]. Die in dieser Arbeit verwendeten Kameras für die Aufnahme der Stereobildpaare, waren zum einen die Vollformatkamera Canon EOS 6D und die Halbformatkamera Canon EOS 60D. Beide besitzen einen CMOS-Sensorchip, welcher zu den Halbleiterbildsensoren gehört[31]. Die Geometrie dieser Sensoren, ist grob gesprochen als eine $M \times N$ - Matrix, bestehend aus $M \times N$ Pixeln. Die Pixel bedecken nur einen Teil der Sensorfläche und können in ihrer geometrischen Beschaffenheit von der Form des Sensors unterschieden werden[10]. Abbildung 5.1 zeigt einen schematischen Aufbau eines Halbleiterbildsensors

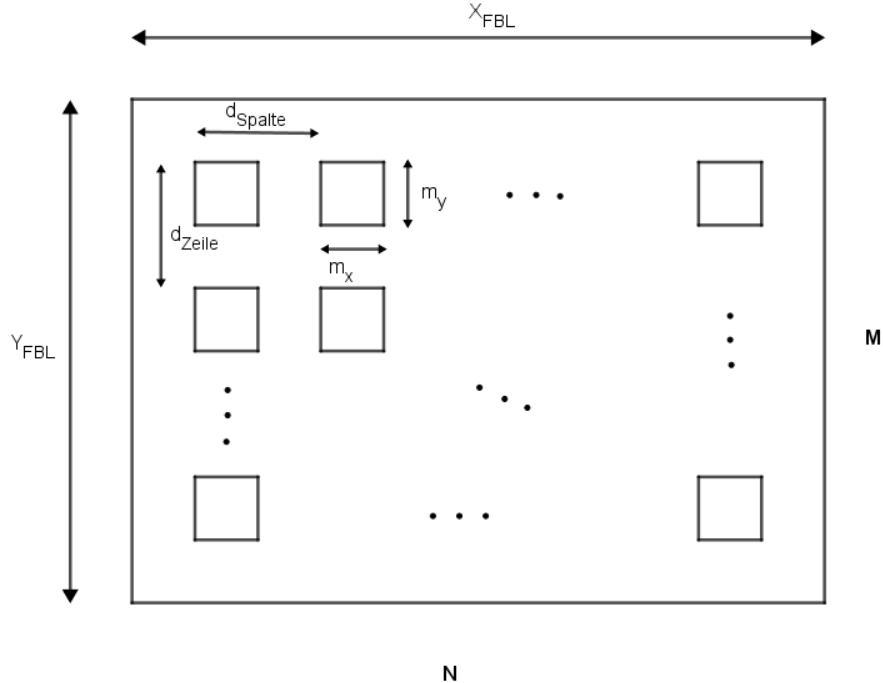


Abbildung 7.1: Rechteckiger Bildsensor mit darauf sich befindenden quadratischen Pixeln. Vlg [10]

Die Auflösung des Sensors hängt von den horizontalen und vertikalen Pixelabständen ab und gibt wieder, welche Objektdetails mit dem Sensor gerade noch erkannt werden können.[10]. Ein Sensor hat eine maximale Auflösung, welche durch die Anzahl seiner fest installierten Pixel bestimmt wird. Die Bildqualität ist abhängig von der Größe des Sensorchips und der Menge der sich darauf befindenden Pixel. Der CMOS-Sensor einer Canon 6D hat beispielsweise eine Größe von $36 \times 24\text{mm}$ und eine maximale Auflösung von 5.472×3.648 Pixel. Jedoch ist das nicht die einzige Auflösung, welche beim photographieren oder filmen mit der Kamera genutzt werden kann. Es können sowohl die Pixelanzahl als auch das Seitenverhältnis der entstehenden Bilder eingestellt werden. Bei der Canon EOS 6D können insgesamt vier verschiedene Seinsverhältnisse eingestellt werden [3 : 2], [4 : 3], [16 : 9] und [1 : 1][31]. Die Bildauflösungen unterscheiden sich pro Seitenverhältnis in sechs Auflösungseinstellungen L , M , $S1$, $S2$, $S3$.

Tabelle 7.1: Auflösungen Canon EOS 6D

	3:2	4:3	16:9	1:1
L	5478×3648 px	4864×3648 px	5472×3072 px	3648×3648 px
M	4104×2736 px	3648×2736 px	4104×2310 px	2736×2736 px
S1	2736×1824 px	2432×1824 px	2736×1536 px	1824×1824 px
S2	1920×1280 px	1696×1280 px	1920×1080 px	1280×1280 px
S3	720×480 px	640×480 px	720×408 px	480×480 px

Tabelle 7.2: Vgl [31]

Je geringer die Auflösung, desto geringer ist die Anzahl der Pixel. Die Anzahl der Pixel auf einem Sensorchip kann natürlich nicht variieren. Eine geringere Pixelanzahl bei gleichbleibender Bildgröße, bedeutet, dass sich ein Pixel mit den um sich befindenden Pixeln interpoliert, so dass ein neuer Pixel bestehend aus mehreren kleinen Pixeln entsteht. Dieser Prozess wird Nachbarschaftsoperation genannt. Für die Berechnung des neuen Bildpixels px' an der Stelle (m, n) wird nicht nur das Bildpixel p des Originalbildes an der Stelle (m, n) verwendet, sondern auch einige seiner Nachbarpunkte[10]. Bei

der Canon 6D bietet das Seitenverhältnis [3 : 2] die Möglichkeit die maximale Pixelanzahl des Sensors zu verwenden, vergleiche hierzu Tabelle 5.1. Bei einem Seitenverhältnis von [4 : 3] ist die Anzahl der maximal möglichen Pixel nur noch 4864×3648 . Anders sich das Seitenverhältnis des Bildausschnitts, so wird auch nicht mehr die gesamte Fläche des Sensors benutzt.

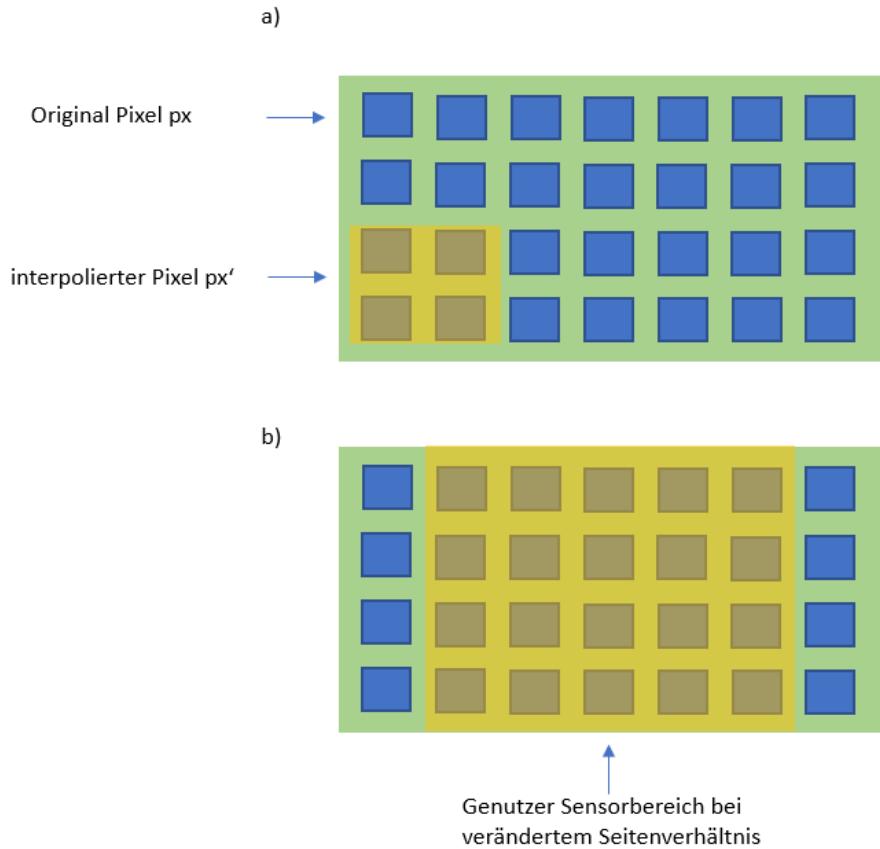


Abbildung 7.2: Bild a) zeigt die Interpolation von Pixeln, wenn bei gleichbleibenden Seitenverhältnissen weniger Pixel für das Bild verwendet werden sollen. Die interpolierten Pixel leiten dann alle das selbe Signal weiter. Bild b) zeigt in gelb markiert, den verwendeten Bereich des Sensors, wenn sich die Seitenverhältnisse ändern und nicht mehr der volle Sensor genutzt wird.

7.2 Auswirkungen auf die Epipolargeometrie

Um nun auf die Fragestellung der Auswirkung unterschiedlicher Auflösungen auf die Szenenrekonstruktion zu kommen, kann nun folgende Behauptung aufgestellt werden. Beide Kameras besitzen einen Bildsensor, welcher fest in der Kamera installiert ist und sich weder in Position noch seiner Form ändern kann. Dieser Bildsensor beinhaltet sowohl das Bildebenenkoordinatensystem, bei welchem der Hauptpunkt den Koordinatenursprung bildet als auch das Sensorkoordinatensystem, dess Ursprung leicht außerhalb einer der Ecken des Sensors sich befindet. Abbildung ??? zeigt einen stereoskopischen Szenenaufbau mit Kameras gleicher Auflösung. Die Sensorkoordinatensysteme besitzen die selbe Skalierung.

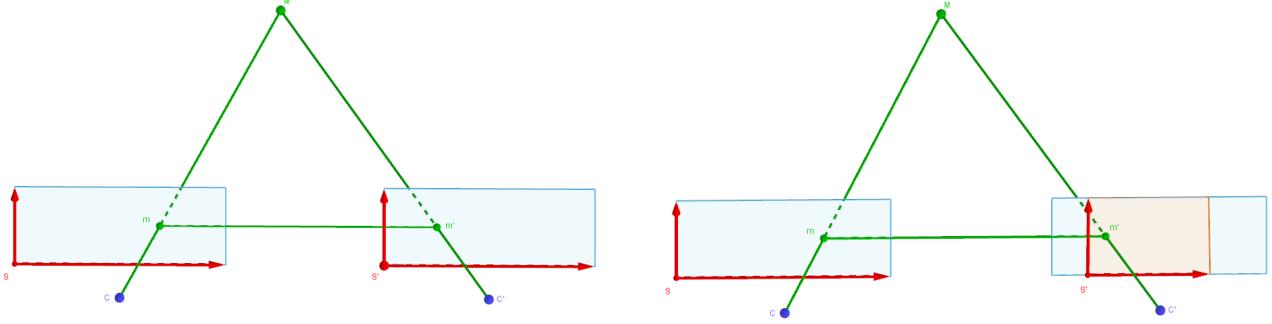


Abbildung 7.3: C und C' haben die selbe Auflösung
Abbildung 7.4: C und C' haben unterschiedliche
eingestellt Auflösungen eingestellt

Es bildet sich wieder das bekannte Dreieck zwischen den Bildpunkten m_σ und m'_σ und dem Objekt-
punkt M_δ . Das in diesem Falle einen korrekten Szenenrekonstruktion funktioniert, ist im Kapitel Minimalbeispiel 3D-Stereokalibrierung und Szenenrekonstruktion bei Kameras gleicher Auflösung anhand des Minimalbeispiels aufgezeigt worden. Wird die Auflösung auf einer der beiden Kameras verringert und damit einhergehend auch noch das Seitenverhältnis geändert, so ändert sich zum einen der aktive Teilbereich des Sensorschips, sowie die Skalierung der Werte auf den Koordinatenachsen des Sensor-koordinatensystems. Die Skalierung der Koordinatenachsen hängt mit der Interpolation der mehrerer Pixel zu einem neuen Pixel zusammen. Abbildung 5.3 zeigt schematisch, was sich nach veränderten Auflösungseinstellungen am Sensor verändert. Epipolare Geometrie ändert sich wie man in Abbildung ??? sehen kann nichts. Das zuvor erwähnte Dreieck zwischen den Bildpunkten und dem Objektpunkt bleibt unverändert. Wie in Kapitel Epipolar Geometrie bereits erwähnt, dürfen die Bildkoordinatensysteme und somit auch die Sensorkoordinatensysteme unterschiedlich voneinander sein [5]. Für die relative Position des Bildpunktes auf dem Sensor ändert sich nichts, dieser bleibt statisch, einzige seine Koordinatenwerte ändern sich im Bezug auf das Sensorkoordinatensystems. Für die Fundamentalmatrix und die Essentielle Matrix ergeben sich lediglich andere Vielfache voneinander, welche wie erwähnt ebenfalls gültige Lösungen sind [3, 4].

7.3 Minimalbeispiel mit unterschiedlichen Kameraauflösungen

Als Beweise der aufgestellten Behauptung wurde im Minimalbeispiel die Kameramatrix einer der beiden Kameras unterschiedlich modifiziert. Während für die Kameramatrix von C der Wert $\zeta = -1$ in der Kameramatrix K bleibt, wurde das ζ in C' verändert, so dass sie drei verschiedene neue Kameramatrizen K'_1, K'_2 und K'_3 ergeben. Die resultierenden Abbildungen des Quaders sind in den Abbildungen ??? bis ??? zu sehen.

$$K = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (7.1)$$

$$K'_1 = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (7.2)$$

$$(7.3)$$

$$K'_2 = \begin{bmatrix} -3.2 & 0 & 0 & 0 \\ 0 & -1.2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (7.4)$$

$$K'_3 = \begin{bmatrix} -0.5 & 0 & 0 & 0 \\ 0 & -4.3 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (7.5)$$

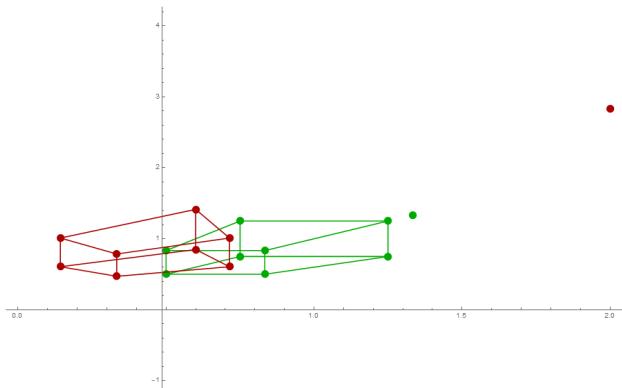


Abbildung 7.5: C und C' haben die selbe Auflösung eingestellt

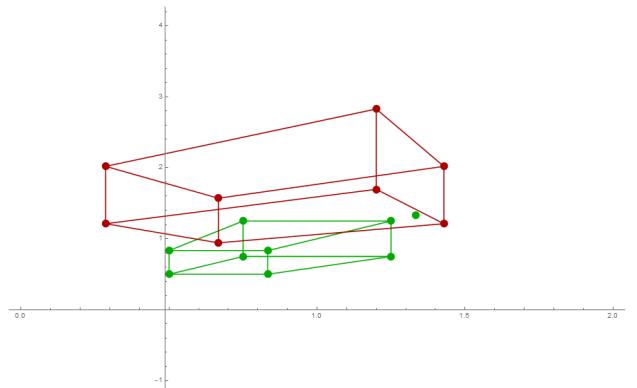


Abbildung 7.6: C und C' haben unterschiedliche Auflösungen eingestellt

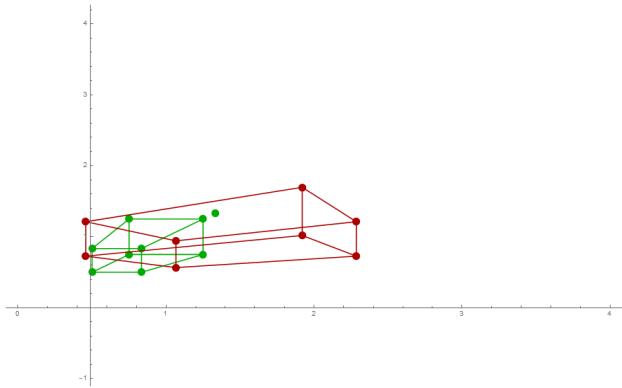


Abbildung 7.7: C und C' haben die selbe Auflösung eingestellt

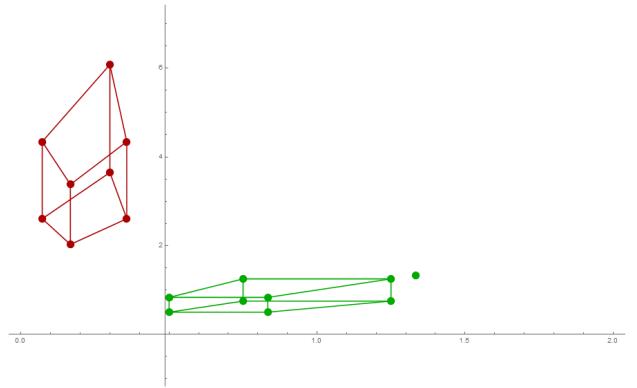


Abbildung 7.8: C und C' haben unterschiedliche Auflösungen eingestellt

Während die Abbildung von C Unverändert bleibt, wird in Abbildung ??? Die Abbildung des Quaders in C' "vergrößert", was für eine höhere Anzahl an verwendeten Pixeln steht. In Abbildung ??? werden die Pixel in horizontaler Richtung um das 3.2-fache und in vertikaler Richtung um das 1.2-fache erweitert und in Abbildung ??? wird in horizontaler Richtung die Anzahl der Pixel um das 0.5-fache und in vertikaler Richtung um das 4.3-fache skaliert. Für die Fundamentalmatrix und die essentielle Matrix ergeben sich verglichen mit denen aus dem Beispiel mit gleicher Abbildung folgende Matrizen.

$$\begin{aligned}
\zeta_x = 1, \zeta_y = 1 : \quad F &= \begin{pmatrix} 0 & -0.5 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -0.5 & 0 \end{pmatrix} | : -0.5 \rightsquigarrow F = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \\
\zeta_x = 2, \zeta_y = 2 : \quad F &= \begin{pmatrix} 0 & 0.378 & 0 \\ 0 & 0 & -0.534 \\ 0 & 0.756 & 0 \end{pmatrix} | : 0.756 \rightsquigarrow F = \begin{pmatrix} 0 & 0.5 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \\
&\quad \frac{1}{\zeta_x} = 0.5, \quad \frac{\sqrt{2}}{\zeta_y} = \frac{1}{\sqrt{2}} \\
\zeta_x = 3.2, \zeta_y = 1.2 : \quad F &= \begin{pmatrix} 0 & 0.198 & 0 \\ 0 & 0 & -0.747 \\ 0 & 0.634 & 0 \end{pmatrix} | : 0.634 \rightsquigarrow F = \begin{pmatrix} 0 & 0.312 & 0 \\ 0 & 0 & -1.178 \\ 0 & 1 & 0 \end{pmatrix} \\
&\quad \frac{1}{\zeta_x} = 0.312, \quad \frac{\sqrt{2}}{\zeta_y} = -1.178 \\
\zeta_x = 0.5, \zeta_y = 4.3 : \quad F &= \begin{pmatrix} 0 & 0.885 & 0 \\ 0 & 0 & -0.145 \\ 0 & 0.442 & 0 \end{pmatrix} | : 0.442 \rightsquigarrow F = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & -0.328 \\ 0 & 1 & 0 \end{pmatrix} \\
&\quad \frac{1}{\zeta_x} = 2, \quad \frac{\sqrt{2}}{\zeta_y} = -0.328
\end{aligned}$$

(hier können vllt noch die verschiedenen rotationen gezeigt werden ,welche immer die selbe rotation bewirken) Die Werte für ζ_x wirken sich auf die erste Zeile der Fundamentalmatrix aus, während die Werte von ζ_y sich auf die zweite Zeile auswirken. Bei der nachfolgenden Umrechnung der Fundamentalmatrix in die essentielle Matrix mit Hilfe der Kameramatrizen K und K' , kann gestellt werden, dass die Ergebnisse jeweils Vielfache voneinander sind.

$$\begin{aligned}
\zeta_x = 1, \zeta_y = 1 : \quad E &= \begin{pmatrix} 0 & -0.5 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0.5 & 0 \end{pmatrix} | : 0.5 \rightsquigarrow E = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \\
\zeta_x = 2, \zeta_y = 2 : \quad E &= \begin{pmatrix} 0 & 0.756 & 0 \\ 0 & 0 & 1.069 \\ 0 & -0.756 & 0 \end{pmatrix} | : -0.756 \rightsquigarrow E = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \\
\zeta_x = 3.2, \zeta_y = 1.2 : \quad E &= \begin{pmatrix} 0 & 0.634 & 0 \\ 0 & 0 & 1.069 \\ 0 & -0.634 & 0 \end{pmatrix} | : -0.634 \rightsquigarrow E = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \\
\zeta_x = 0.5, \zeta_y = 4.3 : \quad E &= \begin{pmatrix} 0 & 0.442 & 0 \\ 0 & 0 & 1.069 \\ 0 & -0.442 & 0 \end{pmatrix} | : -0.442 \rightsquigarrow E = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix}
\end{aligned}$$

Bei der Rekonstruktion der externen Kameraparameter ergibt sich daraus stehts die selbe Matrix für P' . Was wie gezeigt daran liegt, dass sich geometrisch nichts ändert, sondern lediglich die Skalierung der Koordinatenwerte der Bildpunkte und somit auch eine Skalierung der Einträge in F und E , welche aber ebenfalls als Skaleninvariant definiert sind[3]. Die Ergebnisse der darauffolgenden Szenenrekonstruktionen, der einzelnen Szenen zeigt, dass sich immer die selbe Szene ergibt, welche mit der eigens aufgebauten Szene übereinstimmen.

Reconstructed scaled Points 3D =

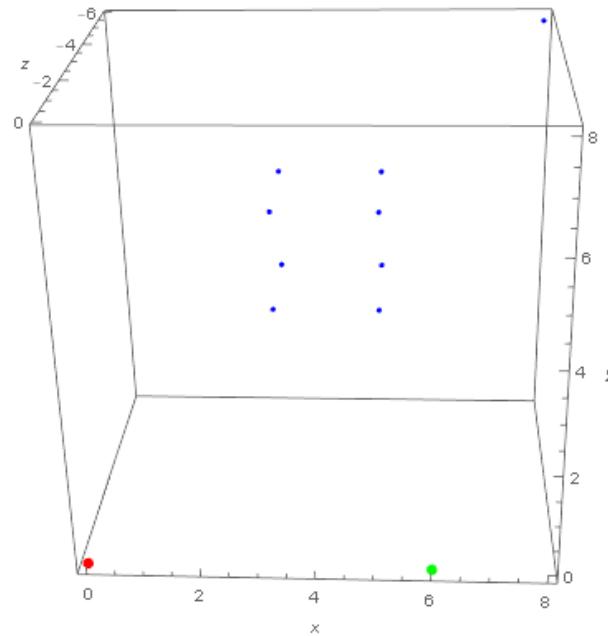
$$\begin{pmatrix} 3. & 3. & -4. \\ 5. & 3. & -4. \\ 5. & 5. & -4. \\ 3. & 5. & -4. \\ 3. & 3. & -6. \\ 5. & 3. & -6. \\ 5. & 5. & -6. \\ 3. & 5. & -6. \\ 8. & 8. & -6. \end{pmatrix}$$


Abbildung 7.9: Die rekonstruierten Szenenpunkte und Kamerapositionen bleibt auch bei unterschiedlichen Auflösungen die selben

Die Behauptung, dass die Auflösung der Kamera bei dem in dieser Arbeit gewählten Workflow für die Rekonstruktion der Szene keine Auswirkung hat, kann für das Minimalbeispiel bestätigt werden.

(HIER VLLT NOCH ZEIGEN, DAS DER REKTIFIZIERUNGSANSATZ AUCH MIT KAMERA UNTERSCHIEDLICHER AUFLÖSUNG FUNKTIONIERT)

8 3D-Stereokalibrierung und Szenenrekonstruktion mit reellen Daten und Kameras gleicher Auflösung

Durch im Minimabeispiel durchgeföhrten Stereoanalyse steht das Grundgerüst für die Stereoanalyse mit realen Bilddaten. Der selbe Arbeitsprozess soll nun auch auf ein Realbeispiel mit Bildern zweier Kameras durchgeföhrten werden. Die exterenen Kameraparameter, werden in extern beschaffen. Über *Matlab* wird für jede Kamera einzeln mit Hilfe der *Single-camera-acibration* App eine Kalibrierung durchgeföhr. Da diese auch bei exteren Parameter liefert, können diese später als Vergleich mit den eigens berechneten Parameter dienen. Im fortschreitenden Arbeitsprozess der Stereoanalyse, werden des öfteren ungereimtheiten im Vergleich mit dem Minimalbeispiel auftreten, welche behoben werden müssen um weiter zu verfahren. Diese Fehler entstehen durch die Ungenauigkeit, der zuvor detektierten korrespondierenden Punkte oder durch Bildfehler wie rauschen oder unschärfe, welche in der Photographie nicht immer vermeiden werden können. Um diese Fehler zu beheben, werden einige zwischenschritte benötigt, welche im Minimalbeispiel durch die Reinheit der Daten und der selbst aufgebauten Szenen nicht nötig waren. Im Minimalbeispiel wurden einige solcher möglichen Fehler bereits erwähnt. Für die Stereoaufnahmen wurden die halbformat Kamera Canon 60 D und die Vollformatkamera 6D verwendet. Die Bilder wurden mit beiden Kameras mit selber Auflösung und Seitenverhältnis augenommen. Später wurden auch noch aufnahmen mit unterschiedlichen Auflösungen gemacht und ebenfalls die Stereoanalyse auf die Bildpaare angewandt.

8.1 Arbeitsprozess

Die beiden Kameras wurden so im Raum platziert, dass sie leicht zueinander hin gedreht waren. Da die Canon 60D eine halbformatkamera ist, wurde sie weiter hinten als die Canon 6D positioniert. Somit konnte ungefähr der selbe Bildausschnitt der Szene mit beiden Kameras aufgenommen werden.

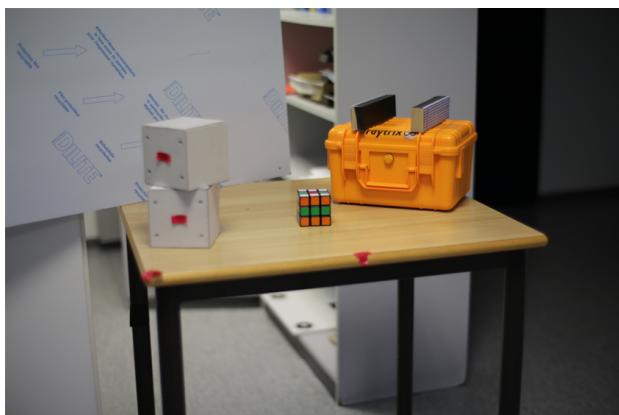


Abbildung 8.1: Aufnahme der Canon 6D von links

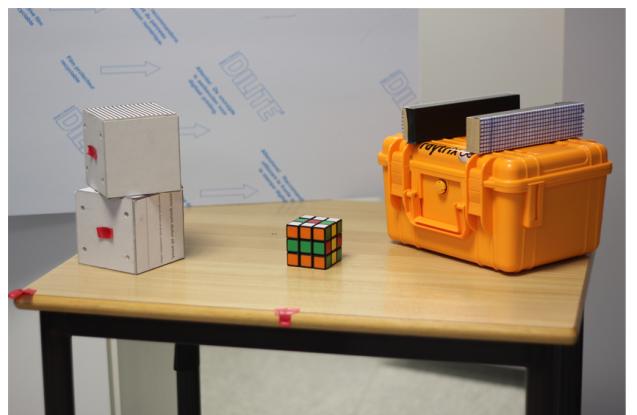


Abbildung 8.2: Aufnahme der Canon 60D von rechts



Abbildung 8.3: Szenenaufbau: Die Canon 60D befindet sich in dieser Abbildung auf der linken Seite, die Canon 60 D auf der rechten. Auf dem Tisch zwischen den Kamera sind die in den Abbildungen 6.1 und 6.1 abgebildete Szene zu sehen. Die Canon 60D ist etwas hinter der Canon 6D positioniert. Beide Kamera sind zu Szene hin gedreht und auch leicht nach unten geneigt.

Die Canon 6D wird als primär Kamera definiert und bekommt somit die Bezeichnung C , während die Canon 60D ab jetzt mit C' gekennzeichnet wird. Die räumliche Orientierung und Position von C' wird also relativ zu C berechnet und auch die Szene wird davon ausgehend, dass C als Projektionsmatrix $P = [I|0]$ besitzt.

$$P = (I|0) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (8.1)$$

Die Koordinatensysteme von C mit (C, β) und das Weltkoordinatensystem (O, δ) werden deckungsgleich definiert. Das Koordinatensystem von C' wird mit (C', β') definiert. Für die spätere Skalierung der rekonstruierten Szene ist es ratsam innerhalb der Szene einen Abstand zweier Punkte zueinander abzumessen, um einen Referenzwert für die Skalierung zu bekommen. Der Stereoaufbau ist somit fertig installiert. Der nächste Schritt ist mit Hilfe von *Matlab* und dem mitgelieferten Kalibrierungs-Tool namens *Single-camera-calibration* für jede Kamera eine Kalibrierung durchzuführen, um so die intrinsischen Parameter K und K' zu bekommen. Diese werden für den verwendeten Ansatz für die Schätzung der extrinsischen Parameter benötigt. Nach der Kalibrierung werden Stereoaufnahmen der Szene gemacht und anschließend korrespondierende Bildpunkte aus den beiden Bildern gesucht. Für den ersten Versuch wurden diese von Hand ausgelesen, da dies jedoch sehr Zeitaufwändig ist, wurden andere Möglichkeiten zur Detektion von korrespondierenden Bildpunkten überlegt. Wie in Abbildung (Einleitung Workflow) ersichtlich ist, wurde zwei Ansätze verfolgt. Fürs erste, wurde überlegt, ähnlich wie in *Matlab* anhand von 2D-Schachbrettern eine Stereoanalyse durchzuführen. [6]. Ein 2D-Schachbrett wird von beiden Kameras positioniert und es werden Stereoaufnahmen des selben gemacht. Mit Hilfe eines Algorithmus, welcher in einer vorherigen Arbeit angefertigt wurde, können die Eckpunkte des Schachbretts der jeweiligen Bilder bestimmt werden. Wichtig für die Funktion des Algorithmus ist, dass die Schachbretter vor einem einfarbigen Hintergrund aufgenommen wurden, so dass sich die Eckpunkt-Bestimmung auch nur auf das Schachbrett bezieht und keine weiteren Punkte außerhalb mit in die entstehende Punkteliste aufgenommen werden. Wenn die Koordinaten der Eckpunkte des

Schachbrettmusters bekannt sind, folgt ein weiterer Algorithmus, welcher im Zuge dieser Arbeit implementiert wurde. Dieser Algorithmus nimmt die Liste mit den Koordinaten der Eckpunkte entgegen und sortiert und nummeriert diese Zeilen- und Spaltenweise durch. Jeder Punkt ist somit über zwei Indizes codiert und enthält die Information, in welcher Zeile und in welcher Spalte des Schachbrettmusters er sich befindet. Dieser Algorithmus wird auf beiden Schachbrettern angewandt. Nach der Sortierung der Punkte auf beiden Bildern, ist es möglich, korrespondierende Punkte der Bilder anhand gleicher Indizes der Eckpunkte zu bestimmen. Da nicht immer garantiert ist, dass alle Punkte innerhalb des Schachbretts zuvor gefunden worden, enthält der Sortierungsalgorithmus eine Funktion, in welchem er Lücken des innerhalb ausfindig macht und synthetische Eckpunkte setzt. Diese synthetisch gesetzten Punkte, werden markiert, so dass sie nicht in die Liste der möglichen korrespondierenden Punkte fallen. Wie genau der Sortierungsalgorithmus implementiert wurde wird im Kapitel Punktesortierung in Schachbrettmustern genauer beschrieben. Danach wird wie in Abbildung ???(Einleitung) gezeigt weiter verfahren. Um eine 3D Szene rekonstruieren zu können, wurde mit Hilfe eines in *Mathematica* bereits implementierten Verfahren zu Findung korrespondierender Punkte in zwei Bildern genutzt. Die benutzte Funktion basiert auf dem sogenannten *SURF*-Algorithmus. *SURF* steht für *Speeded Up Robust Features* und ist ein Rotations- und Skaleninvarianter Punkte Detektor und Deskriptor[32]. Die Suche nach diskreten Bildkorrespondenzen kann in drei Schritte eingeteilt werden. Im ersten Schritt werde sogennaten *Point of interest* an markanten Stellen im Bild detektiert. Darunter fallen zum Beispiel Eckpunkte-Erkennung, "Blob"- Erkennung oder "T-Junctions"- Erkennung[32]. Diese Aufgabe wird dem Detektor zugeordnet. Der wohl am meisten genutzte Detektor in heutigen Computer Vision Applikationen ist der sogenannte *Harris corner detector*[32]. Die Umgebung eines jeden gefundenen Punktes wird durch einen Merkmalsvektor beschrieben, den Deskriptor[32]. Dieser Deskriptor muss unverwechselbar und zu gleichen Zeit robust gegenüber Bildrauschen, Detektionsfehlern und geometrischen Deformationen sein. Im letzten Schritt müssen die Deskriptoren zwischen den Bildern abgestimmt werden. Meistens wird dieses *matching* über die Distanzen der Vektoren betrieben. In Abbildung 6.4 ist eines der Ergebnisse des verwendeten *SURF*-Algorithmus zu sehen. Das linke Bild wurde mit der Canon 6D aufgenommen, das rechte mit der Canon 60D. Die gelben Ziffern in den Bildern markieren die jeweiligen korrespondierenden Punkte in den Bildern.

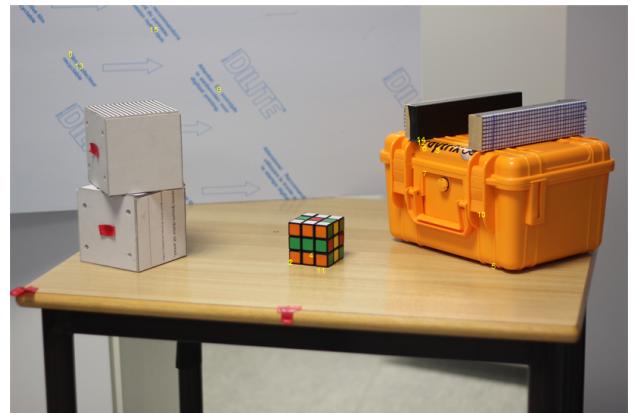
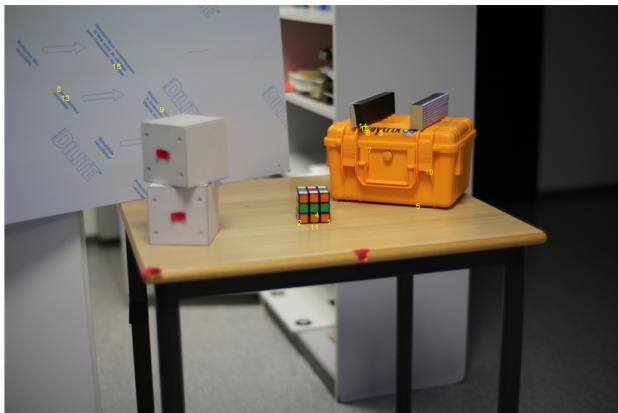


Abbildung 8.4: Die mit dem *SURF*-Algorithmus gefundenen Punkte sind mit den gelben Ziffern im Bild gekennzeichnet

8.2 Normalized-eight-Point-Algorithm

Nachdem die korrespondierenden Punkte in den Bilder gefunden wurden, wird nun der Arbeitsprozess wie in den Abbildungen ??? und ??? in der Einleitung und wie bereits aus dem Minimalbespiel bekannt, weiterverfolgt. In den einzelnen Schritten müssen jedoch ein paar Änderungen vorgenommen werden, um die durch die ungenauen Bilddaten entstehenden Fehler im Verlauf des Arbeitsprozesses zu minimieren.

Als erstes erfolgt die Schätzung der Fundamentalmatrix. Für die Schätzung wurde eine leicht abgeänderte Form des *eight-point-algorithms* namens *normalized-eight-point-algorithm* angewandt[3, 17, 4]. Zur Durchführung des *normalized-eight-point-algorithm* wird eine vorherige Normierung der eingehenden Bildpunkte pro Bild verlangt. Diese sollen so normiert werden, dass ihr durchschnittlicher Abstand zu ihrem den Koordinatenursprung verschobenenen Schwerpunkt $\sqrt{2}$ beträgt[3, 17, 4]. Zu aller erst werden die jeweiligen Schwerpunkte der Punkte in den einzelnen Bildern gesucht und dieser dann in den jeweiligen Sensorskoordinatenursprung verschoben. Die Bildpunkte werden unter Beibehaltung ihres momentanen Abstandes zum Schwerpunkt mit verschoben. Danach werden die Anständer der Bildpunkten zum Schwerpunkt so skaliert, dass der Durchschnittsabstand der Punkte zu Schwerpunkt $\sqrt{2}$ beträgt. Die so skalierten Bildpunkte befinden sich nun in einem deutlich kleineren Zahlenbereich von circa -1 bis 1[3, 17, 4]. Die Transformation der Bilddpunkte für beide Bilder wird jeweils einer Matrix T und T' vollzogen. Diese Matrix ist wichtig, um nach dem schätzen einer auf den normalisierten Koordinaten basierten Fundamentalmatrix \hat{F} wieder eine denormalisierte F zu generieren. Die Normierung der Bildkoordinaten ist wichtig, um die Auswirkung der Bildfehler auf das Endergebnis zu minimieren. Die Entscheidung den normalized-8-Point-Algorithm zu benutzen fiel als festgestellt wurde, dass ohne vorherige Normalisierung der ausgelesenen Punkte es zu größeren Fehlern in den weiteren Berechnungen kam. Zur Erklärung dieser Fehler kann zum einen die *Condition-Number* betrachten. Als *Condition Number*, Kondition im deutschen, wird die Abhängigkeit der Lösung eines Problems von der Störung der Eingangsdaten beschrieben. Die Kondition lässt sich durch Bestimmung des kleinsten Eigenvektors der Matrixmultiplikation der Koeffizientenmatrix A mit ihrer Transponierten A^T herausfinden. Die Matrix AA^T wird in die Matrizen UDU^T , wobei U eine orthogonale und D eine diagonale Matrix ist, zerlegt. Die diagonaleinträge von D sind in einer nicht ansteigenden Reihenfolge, woraus resultiert, dass der kleinste Singulärwert von D mit der letzten Spalte von U korrespondiert und somit ist die letzte Spalte von U gleich dem kleinsten Eigenvektor von AA^T [17]. Wird angenommen, dass AA^T eine 9×9 -Matrix ist, so ergeben die Einträge d_1/d_9 die gesuchte *Condition Number*. Je größer die *Condition-Number* ist, desto größer wirken sich auch kleinste Abweichungen, wie Bildrauschen, auf die Resultate aus. Da sich die original Bildkoordinaten in diesem Beispiel in einem Zahlenbereich von 0 bis 5478 befinden, sind auch die Werte innerhalb der Koeffizientenmatrix in einem sehr großen Zahlenbereich, was zu Folge hat, dass schon kleinste Abweichungen in den Bilddaten, große Auswirkungen auf die daraus resultierende Fundamentalmatrix haben kann, in Bezug darauf, dass die Werte der Einträge innerhalb von F , sehr große Ungleichgewichte aufweisen. Anders im Fall von Normierten Koordinaten, deren Zahlen sich in einem Bereich zwischen circa -1 und 1 befinden.

$$F = \begin{pmatrix} 10^{-4} & 10^{-4} & 10^{-2} \\ 10^{-4} & 10^{-4} & 10^{-2} \\ 10^{-2} & 10^{-2} & 1 \end{pmatrix} \quad (8.2)$$

$$F = \begin{pmatrix} -10^{-9} & 10^{-6} & -10^{-4} \\ -10^{-7} & 10^{-4} & 10^{-3} \\ 10^{-4} & -10^{-3} & -10^{-2} \end{pmatrix} \quad (8.3)$$

Gleichung 6.2 zeigt schematisch was unter einer Gleichgewichtigen Fundamentalmatrix zu verstehen ist, welche bei einer sehr geringen *Condition-number* resultieren kann. Gleichung 6.3 wiederum zeigt schematisch das Resultat einer Ungleichgewichteten Fundamentalmatrix, dessen *Condition-Number* sehr groß ausfällt[17]. Durch normieren der Bildkoordinaten, kann die *Condition-Number* kleiner und damit einhergehend die entstehenden Fehler minimiert werden. Nachdem die Fundamentalmatrix aus den normierten Koordinaten geschätzt wurde, wird sie anschließend mit den beiden aufgestellten Matrizen T und T' wieder denormalisiert, so dass sie wieder als *epipolar – constraint* zwischen die original Koordinaten geschalten werden kann. Für normierte Koordinaten \hat{m} und \hat{m}' gilt $\hat{m}'^T \cdot \hat{F} \cdot \hat{m} = 0$ und für die ursprünglichen Bildkoordinaten gilt, dass $m'^T \cdot T'^T \cdot \hat{F} \cdot T \cdot m = 0$ und somit wieder $m'^T \cdot F \cdot m = 0$ [3, 17, 4]. Die Normierung der Koordinaten für die Verwendung des *eight-point-algorithms* darf auf keinen Fall mit der Normierung der Koordinaten für die essentielle Matrix E verglichen werden. Die Normierung der Koordinaten für die Schätzung von F , soll die Auswirkungen von Fehler auf die Resultate minimieren, während die Normierung der Koordinaten durch deren Verrechnung mit den

Kameramatrizen K und K' dafür sorgt, dass daraus normierte Bildkoordinaten entstehen, dessen Koordinatenursprung nicht mehr in einer Bildecke sondern in der Bildmitte sich befindet[3]. Um zurück zum Arbeitsprozess zu kommen, sind die Koordinaten normiert, so wird der im Kapitel Einleitung aufgezeigte Verfahren zur Schätzung der Fundamentalmatrix gleichermaßen wie in den Gleichungen 4.29 bis 4.34 aufgebaut. Durch die möglichen Ungenauigkeiten wie Bildrauschen oder dem detektieren der korrespondierenden Punkte, ist der Rang der aufgestellten Koeffizientenmatrix A in den meisten Fällen größer als acht, was bedeutet, dass hier nicht einfach der Kern mit $A \cdot f = 0$ gesucht werden kann, um eine Lösung zu finden. Im Falle eines höhren Ranges als 8 muss ein Verfahren, ähnlich wie dem, welches angewandt wurde um überbestimmte Systeme zu Lösen um eine Homographiematrix zu erhalten. Es wird also derjenige Vektor für f gesucht, welcher $\|A \cdot f\|$ minimiert. Hierzu wird eine Singulärwertszerlegung von A in $A = UDV^T$ durchgeführt. die Lösung für f , welche $\|A \cdot f\|$ minimiert ist dann diejenige Spalte von V^T , welche mit dem kleinsten Singulärwert von D korrespondiert. Da die Singulärwerte eine absteigende Reihenfolge besitzen, bildet die letzte Spalte von V^T den Vektor f [3].

8.2.1 Singularity-Constraint der Fundamentalmatrix

Die Fundamentalmatrix ist eine singuläre-Matrix und ist somit eine Matrix von Rang zwei. Die Singularität der Fundamentalmatrix sorgt zum einen dafür das ihr rechter und linker Kern jeweils den Epipol des jeweiligen Bildes ergibt und die Epipolarlinien auch alle durch eben diese Epipole verlaufen. wird die Fundamentalmatrix durch eine Singulärwertszerlegung von A geschätzt, ist die Chance sehr hoch, dass das Ergebnis für \hat{F} eine Matrix von Rang 3 ist. Sollte dies der Fall sein gehen die Epipolarlinien der Bilder nicht mehr durch genau einen Punkt, wie man in den Abbildungen 6.5 und 6.6 erkennen kann. Diese bilden Epipolarlinien in einem Stereobildpaar ab.

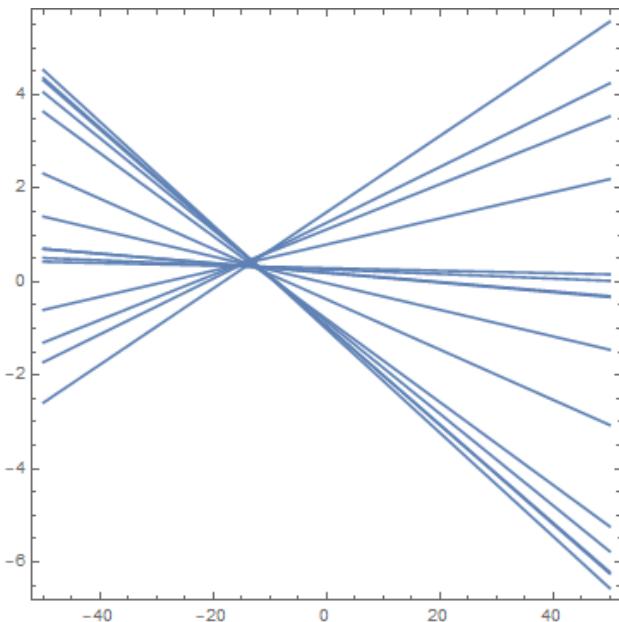


Abbildung 8.5: Epipolarlinien ohne *Epipolar-constraint* im Bild der Canon 6D

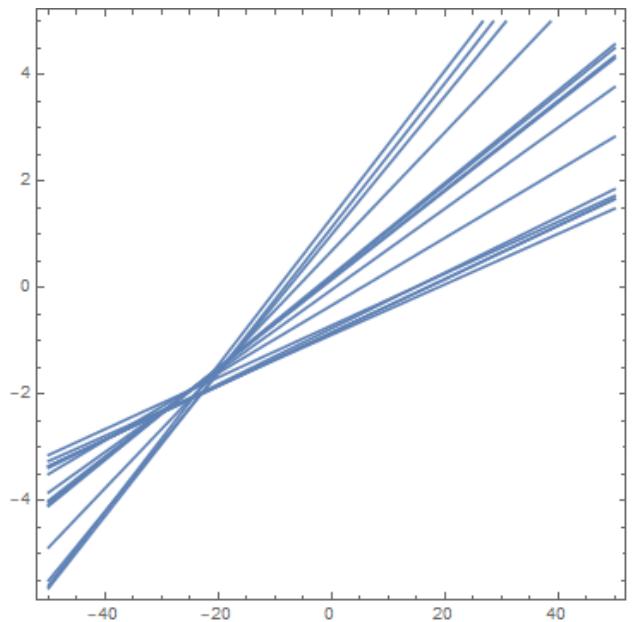


Abbildung 8.6: Epipolarlinien ohne *Epipolar-constraint* im Bild der Canon 60D

Um eine gültige Fundamentalmatrix für den weiteren Arbeitsprozess zu generieren, kommt hier ein sogenannter *singularity constraint* zum Einsatz. Dieser erzwingt in \hat{F} eine Singularität. Zu aller erst wird eine Singulärwertszerlegung an F durchgeführt, so dass \hat{F} in $\hat{F} = UDV^T$ zerlegt wird. D beinhaltet in einer diagonalen Matrix die Singulärwerte $D = \text{diag}(r, s, t)$, welche die Bedingung $r \leq s \leq t$ erfüllen. Um nun den *singularity-constraint* in \hat{F} zu erzwingen, wird der letzte Singulärwert $t = 0$

gesetzt, so dass am Ende dasteht $D = \text{diag}(r, s, 0)$. Danach werden die Matrizen UDV^T , wobei D nun die modifizierten Singulärwerte beinhaltet, wieder zu \hat{F} multipliziert. Die jetzt resultierende Fundamentalmatrix \hat{F} besitzt den Rang 2. Der rechte und linke Kern ergeben wieder die Epipole und die Epipolarlinien verlaufen wieder durch eben diese Epipole. Die Abbildungen 6.7 und 6.8 zeigen die selben Epipolarlinien wie in 6.5 und 6.6 nachdem der *singularity-constraint* in \hat{F} erzwungen wurde. Die somit entstandene Matrix \hat{F} , ist die laut Frobenius norm, nächste zum ursprünglichen \hat{F} [3]. Jetzt erst erfolgt die zuvor erwähnte Denormierung von \hat{F} durch T und T' . Die Abbildungen 6.9 und 6.10 zeigen die Epipolarlinien im Originalbild mit denormalisierten Koordinaten.

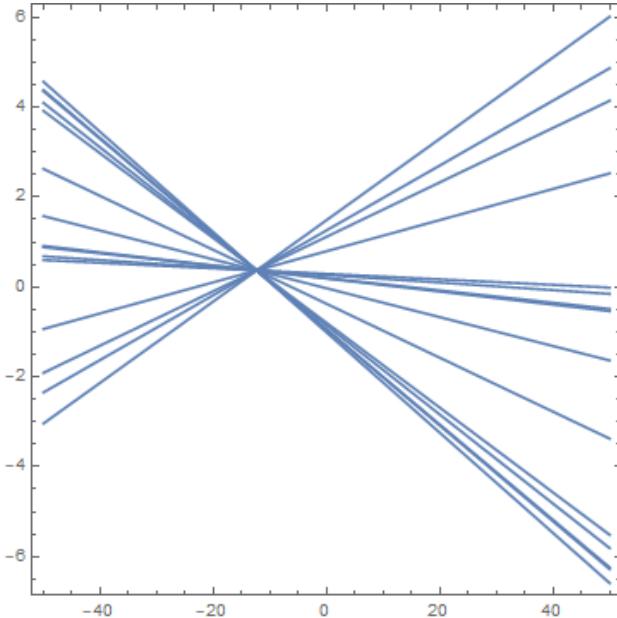


Abbildung 8.7: Epipolarlinien mit *Epipolar-constraint* im Bild der Canon 6D

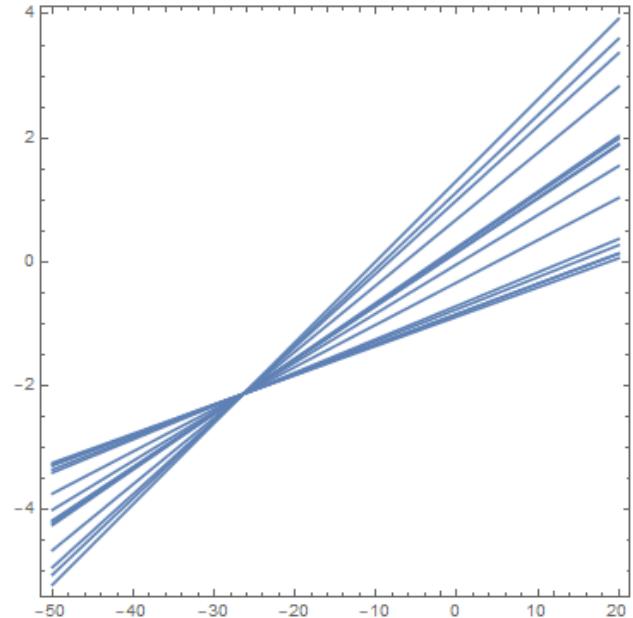


Abbildung 8.8: Epipolarlinien ohne *Epipolar-constraint* im Bild der Canon 60D

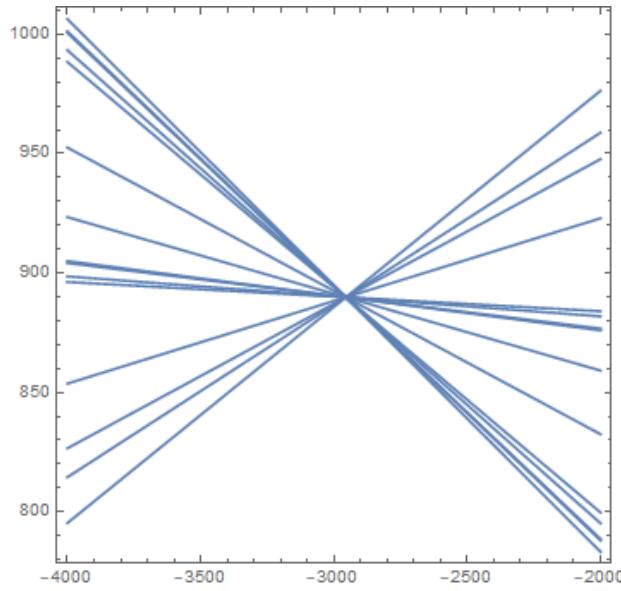


Abbildung 8.9: Epipolarlinien mit *Epipolar-constraint* im Bild der Canon 6D, nach der denormalisierung von F

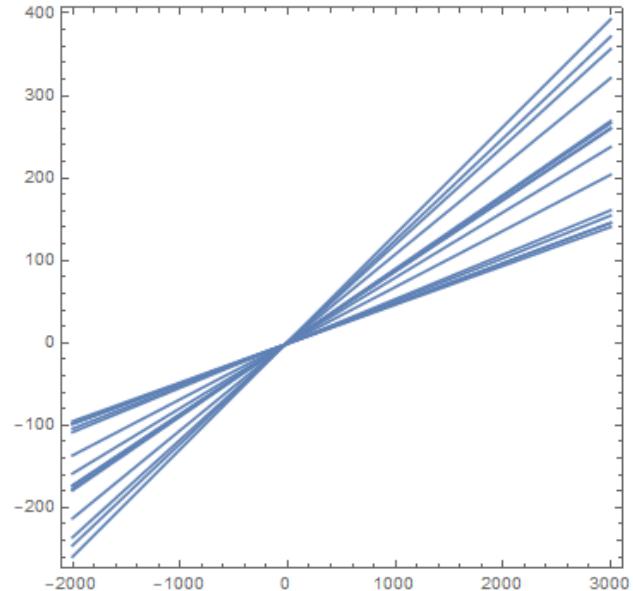


Abbildung 8.10: Epipolarlinien ohne *Epipolar-constraint* im Bild der Canon 60D, nach der denormalisierung von F

8.2.2 Singularity- Constraint der essentiellen Matrix

Die essentielle Matrix E kann wenn sie über den *eight-point-algorithm* ermittelt wird, auch eine Rang 3 Matrix anstelle einer Rang 2 Matrix sein. Eine essentielle Matrix wird darüber definiert, dass sie eine Matrix mit Rang 2 ist und ihre Singulärwerte in D von $E = UDV^T$ die Eigenschaft besitzen, dass $D = \text{diag}(a, b, c)$ mit $a = b$ und $c = 0$. Sind die Singulärwerte nicht in der gezeigten Form vorhanden und E hat den Rang 3, so ist sie keine gültige essentielle Matrix[3]. Im implementierten Algorithmus, welcher in dieser Arbeit vorgestellt wird, wird die essentielle Matrix über die Fundamentalmatrix F und den intrinsischen Kameraparameter K und K' gewonnen. Da im vorherigen Schritt für die Matrix F schon der *singularity-constraint* erwirkt wurde, ist dadurch dass F nun eine Matrix von Ran 2 ist auch versichert, dass E ebenfalls von Rang 2 ist. Jedoch bedeutet das nicht gleichzeitig, dass auch die Bedingungen für die Singulärwerte von E erfüllt sind. Wird E in UDV^T zerlegt und die Singulärwerte in D haben beispielsweise die Form $D = \text{diag}(a, b, c)$ mit $a \geq b \geq c$, so muss auch hier die für E typische Singularität erzwungen werden. Die laut Frobenius Norm nächste Matrix E zur momentanen E kann durch modifizieren der Singulärwerte von D mit $D = \text{diag}(\frac{a+b}{2}, \frac{a+b}{2}, 0)$ erzwungen werden[3]. Mit der neuen essentiellen Matrix können dann genau wie im Kapitel Minimalbeispiel 3D-Stereokalibrierung und Szenenrekonstruktion bei Kamera gleicher Auflösung auch wieder die vier möglichen Lösungen der externen Kameraparameter ermittelt werden.

8.3 Szenenrekonstruktion

Im letzten Schritt des Arbeitsprozesses, wird nun noch die Szenen mit Hilfe eines Triangulationsverfahrens rekonstruiert. Wie bereits im Kapitel Minimalbeispiel 3D-Stereokalibrierung und Szenenrekonstruktion bei Kamera gleicher Auflösung erwähnt wurde, ist es bei den Fehlerhaften Bildkoordinaten nicht möglich die 3D-Szenenpunkte durch eine einfache Rückprojektion der Bildpunkte zu einem Punkt im 3D-Raum zu erhalten. liegt nur einer der beiden Bildpunkte m oder m' nicht hundert prozentig auf der jeweiligen korrespondierenden Epipolarlinie, so liegen die rückprojizierten Strahlen windschief im Raum. Das liegt daran, dass die Bildpunkte m und m' nicht den *Epipolar-Constraint* $m'^T F m = 0$ erfüllen. Sprich die Gleichungen $m = PM$ und $m' = P'M$ können nicht erfüllt werden, da es kein M gibt, dass für beide Gleichungen mit den momentanen m und m' gilt. Abbildung 6.11 veranschaulicht die Rückprojektion der Kamerazentren durch zwei Fehlerhafte Bildpunkte.

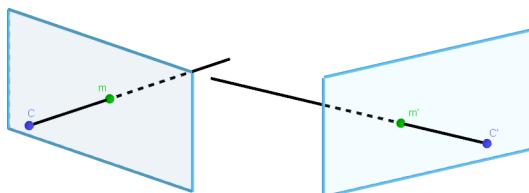


Abbildung 8.11: a)

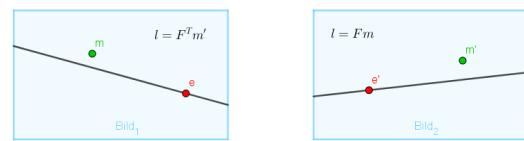


Abbildung 8.12: b)

Abbildung 8.13: a) Die rückprojizierten Strahlen der ungenauen korrespondierenden Punkte m und m' sind schief und treffen sich nicht in einem Punkt im 3D-Raum. b) The epipolar geometry for m, m' . The measured points do not satisfy the epipolar constraint. The epipolar line $l' = Fm$ is the image of the ray through n , and $l = F^T m'$ is the image of the ray through m' . Since the rays do not intersect, m' does not lie on l' , and m does not lie on l .

Um eine Triangulation zu ermöglichen, muss eine Methode gefunden werden, welche diesen Fehler so weit minimiert, dass es zu einer erfolgreichen Rückprojektion kommt. Die verwendete Methode zur Rekonstruktion der Szene wurde nach der Vorlage von Hartley & Zisserman[3] implementiert und wird im folgenden Schritt für Schritt beschrieben. Voraussetzung ist, dass die Projektionsmatrizen P und P' , sowie die Fundamentalmatrix F bekannt sein müssen. Sind die Projektionsmatrizen P

und P' bis auf eine projektive oder affine Komponenten bekannt, so ist es wünschenswert, wenn die Triangulierung auf einem affinen und projektiv invarianten Verfahren funktioniert[3]. Die hier verwendeten Projektionsmatrizen sind bis auf eine Skaleninvarianz genau bestimmt, was unter den Fall der affinen Invarianz Fällt. Wären die intrinsischen Kameraparameter K und K' nicht bekannt gewesen, gibt es die Möglichkeit die Projektionsmatrizen über die Fundamentalmatrix F mit dem, im Buch von *Hartley & Zisserman* beschriebenen *Stratified-Approach* bis auf eine projektive Invarianz genau zu bestimmen[3]. Die hier verwendete Triangulierung ist nur projektiv invariant, kann aber trotzdem genutzt werden. Die rekonstruierte Szene ist, dann wie im Minimalbeispiel auch, nicht auf ihre Originalmaße skaliert, was aber nach der Triangulierung noch getan werden kann. Die Triangulierung ist deshalb projektiv invariant, weil alle Rechenoperationen, wie die Minimierungen von Distanzen, sich nur auf die 2D-Bildern beziehen und sich nicht in den projektiven 3D-Raum erstreckt[3]. Der Grundgedanke der Triangulation ist, dass zwei Punkte \hat{m} und \hat{m}' gefunden werden sollen, die möglichst nah an den ursprünglichen m und m' sind und gleichzeitig den *Epipolar-Constraint* $\hat{x}'^T F \hat{x} = 0$ erfüllen. Dies erfolgt durch die Minimierung einer Kostenfunktion C . In vielen bekannten Computer Vision Applikationen wird für diese Minimierung eine numerische Lösung gewählt, die wohl bekannteste Methode ist der *Levenberg-Marquardt* Algorithmus[3]. Jedoch hat sich gezeigt, dass ein nahezu optimales Minimum der geometrischen Kostenfunktion C auch durch eine Annäherung ersten Grades finden lässt. Die Annährung um die es sich handelt ist die sogenannten *Sampson-approximation*.

8.3.1 Minimieren der Kostenfunktion durch Sampson-approximation

Es sollen zwei Punkte \hat{m} und \hat{m}' gefunden werden, welche nahe an den Ursprünglichen m und m' liegen und gleichzeitig den *Epipolar-Constraint* erfüllen. \hat{m} und \hat{m}' sollen durch Minimierung einer Kostenfunktion C ermittelt werden, welche die Distanz d zwischen m und \hat{m} und m' und \hat{m}' minimiert.

$$C(m, m') = d(m, \hat{m})^2 + d(m', \hat{m}')^2 \quad (8.4)$$

Die projizierten Punkte \hat{m} und \hat{m}' eines 3D-Objektpunktes \hat{M} liegen auf einem paar korrespondierender Epipolarlinien. Jedes Punktpaar, welches den *Epipolar-Constraint* erfüllt, liegt auf einem paar korrespondierender Epipolarlinien. Die optimalen \hat{m} und \hat{m}' liegen am Fuße des Lots, welches von den ursprünglich projizierten Punkten m und m' ausgeht. Zusätzlich liegen \hat{m} und \hat{m}' auf korrespondierenden Epipolarlinien l und l' [3].

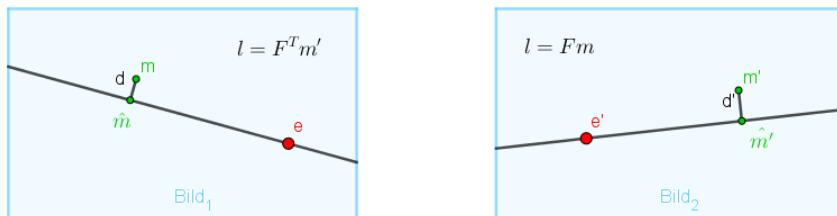


Abbildung 8.14: Grafische Darstellung der optimalen Punkte \hat{m} und \hat{m}'

Jedes Punktpaar auf l und l' würde den *Epipolar-Constraint* erfüllen, jedoch minimieren nur x_\perp und x'_\perp die quadratischen Distanzen in der Kostenfunktion C . Ausgehend von dieser Aussagen wird C so umformuliert, dass gilt $d(m, \hat{m}) = d(\hat{m}, l)$ und $d(m', \hat{m}') = d(\hat{m}', l')$, was jeweils den senkrechten Abstand m zu l und m' zu l' beschreibt. Werden l und l' frei aus allen möglichen Epipolarlinien gewählt, so wird immer der senkrechte Abstand von x zu dieser gewählten l berechnet, entsprechend gilt das auch für m' und irgendeine l' . Nun muss der Abstand $d(\hat{m}, l)^2 + d(\hat{m}', l')$ minimiert werden, da natürlich nicht einfach jede beliebigen Epipolarlinien genutzt werden können. Es wird eine Strategie

mit insgesamt vier Schritten für die Minimierung verfolgt[3]. Zuerst werden die Epipolarlinienbündel pro Bild so Parametrisiert, dass beispielsweise eine Epipolarlinie im ersten Bild als $l(t)$ geschrieben werden kann. Danach wird die Fundamentalmatrix F dazu benutzt, die entsprechend korrespondierende Epipolarlinie l' zu berechnen. Die Kostenfunktion C kann somit als eine Funktion von t definiert werden. Schlussendlich muss ein Wert für t gefunden werden, welcher C minimal werden lässt.

$$C(m, m') = d(m, \hat{m})^2 + d(m', \hat{m}')^2 \quad (8.5)$$

$$\rightsquigarrow C(m, m') = d(m, l(t))^2 + d(m', l'(t))^2 \quad (8.6)$$

Es kann passieren, dass ein Bildpunkt korrespondierend zum jeweiligen Epipol des anderen Bildes ist, der Rückprojizierte Punkt im 3D-Raum würde sich dann auf der Basislinie der zwei Projektionszentren befinden und es ist somit nicht möglich ihn zu rekonstruieren. Um solche Fälle zu vermeiden, wird eine Transformation der Punkte m und m' in den Ursprung $(0, 0, 1)^T$ zu verschieben.

$$T = \begin{bmatrix} 1 & 0 & -x \\ 0 & 1 & -y \\ 0 & 0 & 1 \end{bmatrix} \rightsquigarrow \bar{m} = T \cdot m \quad (8.7)$$

$$T' = \begin{bmatrix} 1 & 0 & -x' \\ 0 & 1 & -y' \\ 0 & 0 & 1 \end{bmatrix} \rightsquigarrow \bar{m}' = T' \cdot m' \quad (8.8)$$

Die Fundamentalmatrix F wird dann wieder an die neu translatierten Punkte \bar{m} und \bar{m}' angepasst.

$$\bar{F} = T'^{-T} F T^{-1} \quad (8.9)$$

Als nächstes wird F mit T und T' so Transformiert, dass den *Singularity-Constraint* zwischen Des Weiteren sollen die Epipole auf die x-Achse an die Punkte $\hat{e} = (1, 0, f)^T$ und $\hat{e}' = (1, 0, f')^T$, wobei f und f' nahezu null sein werden. Die Epipole lassen sich durch den rechten und linken Kern der neuen \bar{F} berechnen. Angenommen f und f' seien genau 0, so lauten die Koordinaten der Epipole $e = (1, 0, f)^T$ und $e' = (1, 0, f')^T$. Ist dies der Fall so hat \bar{F} für welche dann gilt, dass $\bar{F}(1, 0, f)^T = (1, 0, f')\bar{F} = 0$ eine spezielle Form.

$$\bar{F} = \begin{pmatrix} ff'd & -f'c & -f'd \\ -fb & a & b \\ -fd & c & d \end{pmatrix} \quad (8.10)$$

$$\begin{pmatrix} ff'd & -f'c & -f'd \\ -fb & a & b \\ -fd & c & d \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ f \end{pmatrix} = \begin{pmatrix} ff'd + (-ff'd) \\ -fb + fb \\ -fd + fd \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (8.11)$$

$$(1 \ 0 \ f') \cdot \begin{pmatrix} ff'd & -f'c & -f'd \\ -fb & a & b \\ -fd & c & d \end{pmatrix} = \begin{pmatrix} ff'd + (-ff'd) \\ -f'c + f'c \\ -f'd + f'd \end{pmatrix} = (0 \ 0 \ 0) \quad (8.12)$$

Im Realfall sind die Werte der Epipole e und e' nicht so rein wie im Beispiel gezeigt. Aufgrund dessen, werden zwei Rotationsmatrizen aufgestellt, welche die Epipole e und e' auf $Re = (1, 0, e_3)$ und $R'e' = (1, 0, e'_3)$ rotiert.

$$R = \begin{bmatrix} e_1 & e_2 & 0 \\ -e_2 & e_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (8.13)$$

$$R' = \begin{bmatrix} e'_1 & e'_2 & 0 \\ -e'_2 & e'_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (8.14)$$

\bar{F} wird dann nochmals mit $R'FR^T$ ersetzt. Die Einträge in \bar{F}_{Rot} haben nun die Form wie in Gleichung 7.10, mit $f = e_3$, $f' = e'_3$, $a = \bar{F}_{Rot,22}$, $b = \bar{F}_{Rot,23}$, $c = \bar{F}_{Rot,32}$ und $d = \bar{F}_{Rot,33}$. Angenommen eine Epipolarlinie verläuft nun durch einen Punkt $(0, t, 1)^T$ und dem Epipol $e = (1, 0, f)^T$, wird diese Epipolarlinie mit $l(t)$ bezeichnet. Das Kreuzprodukt dieser beiden Punkte beschreibt die Gerade.

$$\begin{pmatrix} 0 \\ t \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ f \end{pmatrix} = \begin{pmatrix} tf \\ 1 \\ -t \end{pmatrix} \quad (8.15)$$

Die quadratische Distanz dieser Linie zum Ursprung wird dann bezeichnet mit:

$$d(m', l'(t))^2 = \frac{t^2}{1 + (tf)^2} \quad (8.16)$$

Für die Herleitung von Gleichung 7.16 wird angenommen die Gleichung einer Gerade sei zunächst in Koordinatenform Dargestellt

$$Ax + By - C = 0 \quad (8.17)$$

Die Selbe Gerade kann auch in Normalform ausgedrückt werden

$$\vec{n} \cdot (\vec{x} - \vec{p}) = 0 \quad (8.18)$$

$$\begin{pmatrix} A \\ B \end{pmatrix} \cdot (\vec{x} - \vec{p}) = 0 \quad (8.19)$$

Der Abstand eines Punktes zu einer Geraden kann folgend ausgedrückt werden.

$$\vec{v} = \frac{\vec{p} \cdot \vec{n}}{\vec{n} \cdot \vec{n}} \cdot \vec{n} \rightsquigarrow \frac{-C}{A^2 + B^2} \cdot \begin{pmatrix} A \\ B \end{pmatrix} \quad (8.20)$$

$$\|\vec{v}\| = \frac{|\vec{p} \cdot \vec{n}|}{\|\vec{n}\|^2} \cdot \|\vec{n}\| \rightsquigarrow \|\vec{v}\| = \frac{|\vec{p} \cdot \vec{n}|}{\|\vec{n}\|} \quad (8.21)$$

$$\Rightarrow |C| = |\vec{p} - \vec{n}| \quad (8.22)$$

$$\Rightarrow |\sqrt{A^2 + B^2}| = \|\vec{n}\| \quad (8.23)$$

$$\|\vec{v}\| = \frac{|C|}{\sqrt{A^2 + B^2}} \quad (8.24)$$

Werden nun A , B und C mit den Werten der Geraden $(tf, 1, -t)^T$ ersetzt, kann Gleichung 7.16 rekonstruiert werden.

$$A = tf, B = 1, C = -t, \vec{v} = d \quad (8.25)$$

$$d^2 = \frac{t^2}{\sqrt{((tf)^2 + 1^2)^2}} = \frac{t^2}{(tf)^2 + 1^2} = \frac{t^2}{1 + (tf)^2} \quad (8.26)$$

Für korrespondierende Epipolarlinie $l'(t)$ wird der Punkt $(0, t, 1^T)$ und die Fundamentalmatrix \bar{F}_{Rot} multipliziert.

$$l'(t) = F(0, t, 1)^T = (-f'(ct + d), at + b, ct + d)^T. \quad (8.27)$$

Für die quadratische Distanz $d(m', l'(t))^2$ ergibt sich dann:

$$d(m', l'(t))^2 = \frac{(ct + d)^2}{(at + b)^2 + f'^2(ct + d)^2} \quad (8.28)$$

Für die ursprüngliche Kostenfunktion C kann jetzt eine Funktion $s(t)$ überestzt werden.

$$C(m, m') = d(m, \hat{m})^2 + d(m', \hat{m}')^2 \quad (8.29)$$

$$\rightsquigarrow C(m, m') = d(m, l(t))^2 + d(m', l'(t))^2 \quad (8.30)$$

$$\rightsquigarrow s(t) = \frac{t^2}{1 + (tf)^2} + \frac{(ct + d)^2}{(at + b)^2 + f'^2(ct + d)^2} \quad (8.31)$$

$s(t)$ ist dann Minimal, wenn für dessen Ableitung gilt $s'(t) = 0$. Werden die beiden Terme in $s(t)$ Nennergleich gemacht und der Nenner gleich Null gesetzt, ergibt sich der folgende Ausdruck $g(t)$

$$g(t) = t((at + b)^2 + f'^2(ct + d)^2)^2 - (ad - bc)(1 + f^2t^2)^2(at + b)(ct + d) \quad (8.32)$$

Funktion $g(t)$ ist ein Polynom vom Grad 6. Das Minimum von $s(t)$ ergibt sich aus einer der 6 möglichen Lösungen für t aus $g(t)$. Für die Ermittlung des Minimums werden nur die reellen Lösungen in betracht gezogen, die nicht-reellen Lösungen können ignoriert werden. Die reellen Lösungen für t aus $g(t)$, werden dann wieder in $s(t)$ eingesetzt. Das t , welches durch einsetzte in $s(t)$ den kleinsten Wert ergibt, ist das gesuchte Minimum. Ist t_{min} gefunden, können die Epipolarlinien $l = (tf, 1, -t)$ und l' durch einsetzen von t_{min} berechnet werden. Nun müssen nur noch die zwei Punkte \hat{m}_{Rot} und \hat{m}'_{Rot} gefunden werden, welche dieser Epipolarlinien vom Ursprung aus am nächsten sind. Der Punkt vom Ursprung aus mit dem geringsten Abstand zu einer Linie (λ, μ, v) berechnet sich durch $(-\lambda \cdot v, -\mu \cdot v, \lambda^2 + \mu^2)$

$$l = (tf, 1, -t) \quad (8.33)$$

$$\hat{m}_{Rot} = (-(tf) \cdot v, -1 \cdot v, (tf)^2 \cdot 1^2) \quad (8.34)$$

Nachdem zu beiden Linien l und l' der jeweils nächste Punkte \hat{m}_{Rot} und \hat{m}'_{Rot} vom Ursprung aus gefunden wurden, müssen diese nun wieder mit an ihre Ausgangsposition verschoben werden.

$$\hat{m} = T^{-1}R^T \hat{m}_{Rot} \quad (8.35)$$

$$\hat{m}' = T'^{-1}R'^T \hat{m}'_{Rot} \quad (8.36)$$

Vergleicht man die Punkte m und \hat{m} und die Punkte m' und \hat{m}' , so kann die minimalen Abweichungen der Punkte voneinander sehen. Um nun noch den Punkt \hat{M} im 3D-Raum zu rekonstruieren, kann nun

jegliche bekannte Methode für die Triangulierung verwendet werden. Durch die zuvorigen Rechenoperationen ist nun gewährleistet, dass sich die Gerade der Projektionszentren C und C' durch ihre jeweiligen Bildpunkte \hat{m} und \hat{m}' auf jeden Fall im Raum treffen[3]. Für die Rückprojektion der Punkte \hat{m} und \hat{m}' zu \hat{M} wurde ebenfalls sich wieder auf ein Verfahren von *Hartley & Zisserman* berufen. Es handelt sich um eine lineare Triangulierungsmethode. Die Gleichungen $\hat{m} = P\hat{M}$, $\hat{m}' = P'\hat{M}$ werden in eine Gleichung der Form $AX = 0$ zusammengeschrieben. Durch die Verwendung des Kreuzproduktes, wird die Homogene Komponente eliminiert.

$$\hat{m} \times (P\hat{M}) = 0 \quad (8.37)$$

$$\hat{m}' \times (P'\hat{M}') = 0 \quad (8.38)$$

Was ausgeschrieben für \hat{m} und \hat{m}' zu den folgenden drei Gleichungen führt.

$$x(p^{3T}X) - (p^{1T}X) = 0 \quad (8.39)$$

$$y(p^{3T}X) - (p^{2T}X) = 0 \quad (8.40)$$

$$x(p^{3T}X) - y(p^{1T}X) = 0 \quad (8.41)$$

p^{iT} bezeichnet hier jeweils die Reihen der Projektionsmatrix P beziehungsweise P' . Die Matrix A stellt sich, aufgrund der Tatsache, dass die Komponenten der Gleichungen 7.37 bis 7.39 linear zu \hat{M} sind, wie folgt zusammen.

$$A = \begin{bmatrix} xp^{3T} - p^{1T} \\ yp^{3T} - p^{2T} \\ x'p'^{3T} - p'^{1T} \\ y'p'^{3T} - p'^{2T} \end{bmatrix} \quad (8.42)$$

Die zwei Wege eine solche Matrix zu lösen sind bereits bekannt, so kann zum einen wieder die Inhomogenene Methode angewandt werden und Kern dieser Koeffizientenmatrix berechnet werden, oder es kann das homogene Verfahren verfolgt werden. Hier wird die Singulärwertzerlegung an A durchgeführt und derjenige Vektor gesucht werden, welcher mit dem kleinsten Singulärwert korrespondiert[3]. Das Ergebnis ist jeweils \hat{M} im 3D-Raum. Da die vorherige P und P' nur bis zu einem Skalierungsfaktor genau bestimmt wurden, muss nachdem die Punkte rekonstruiert wurden noch die Skalierung auf ihre ursprüngliche Größe erfolgen. Dies ist am einfachsten, wenn eine Referenzgröße zuvor in der Originalszene gemessen wurde. Die Abbildungen 7.15 und 7.16 zeigen die Rekonstruierte Szene des Beispiels, jedoch noch nicht skaliert auf ihre Ursprungsgrößen. Abbildung 7.15 zeigt die 3D Szene. Der Rote Punkt symbolisiert die Position von C also der Canon 6D und der grüne Punkt symbolisiert die Position von C' also der Canon 60D. Die Blauen Punkte sind die durch den *SURF*-Algorithmus detektierten Punkte der Szene. Abbildung 7.16 zeigt die rekonstruierten Objektpunkte als 2D-Punkte, hierfür wurden ihre Koordinaten einfach durch ihren Tiefenwert geteilt.

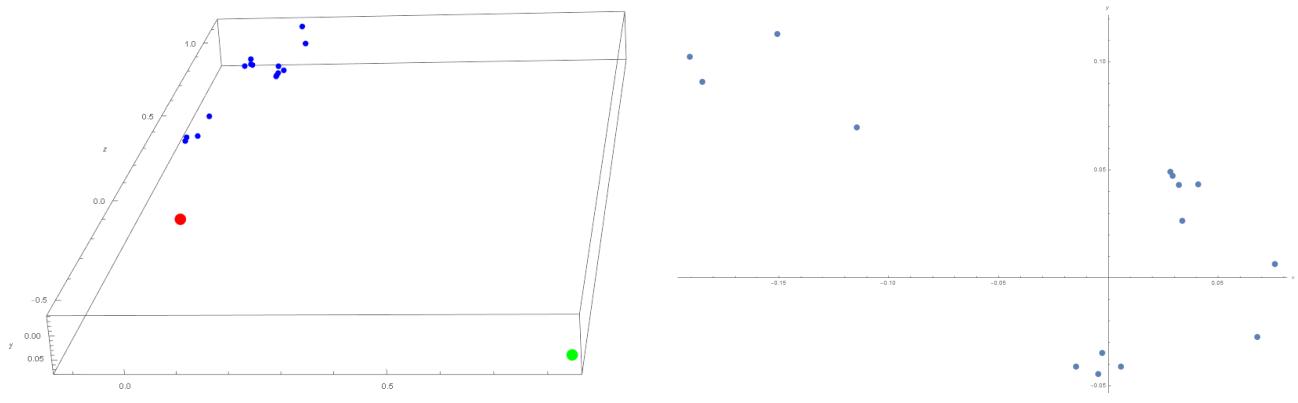


Abbildung 8.15: Rekonstruierte Szene, unskaliert
in Pixeleinheiten

Abbildung 8.16: Rekonstruierte Szene, unskaliert, in Pixeleinheiten und in einem 2D-Plot geschrieben

9 3D-Stereokalibrierung und Szenenrekonstruktion mit reellen Daten und Kameras unterschiedlicher Auflösung

9.1 Ergebnisse einer Stereoanalyse mit Kameras unterschiedlicher Auflösung

Für den Test, ob Szeneriekonstruktion im Realbeispiel auch mit unterschiedlichen Kameraauflösungen funktioniert, wurde eine der von Matlab ermittelten Kameramatrizen K' und auch die durch den Surf Algorithmus detektierten Punkte jeweils skaliert. Im Kapitel Basis Transformationen wurden die einzelnen Bauteile der Kameramatrix genau beschrieben. Die Kameramatrix K' aus Matlab für die Canon 60D gegeben.

$$K' = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (9.1)$$

α_x und α_y setzen sich auch dem Abstand des Kamerazentrums zum Hauptpunkt zusammen, welcher in dieser Arbeit als mit ζ bezeichnet wurde, und den Kantenlängen der Pixel auf dem Sensor m_x und m_y . Um die Auflösung der Kamera zu verändern, wird auf α_x und α_y jeweils ein beliebiger Faktor dazu multipliziert. Zum Beweis, dass die Rekonstruktion der externen Kameraparameter und die Szeneriekonstruktion, bei egal welcher Skalierung, die ähnlichen Ergebnisse liefern, wurde die Kameramatrix K' mit den Verhältnissen [2 : 2], [5 : 2], [2 : 1], [1 : 2] und [1.2 : 2.3] skaliert.

$$\begin{aligned} K'_{[2:2]} &= \begin{bmatrix} \alpha_x \cdot 2 & s & x_0 \cdot 2 \\ 0 & \alpha_y \cdot 2 & y_0 \cdot 2 \\ 0 & 0 & 1 \end{bmatrix} \\ K'_{[5:2]} &= \begin{bmatrix} \alpha_x \cdot 5 & s & x_0 \cdot 5 \\ 0 & \alpha_y \cdot 2 & y_0 \cdot 2 \\ 0 & 0 & 1 \end{bmatrix} \\ K'_{[2:1]} &= \begin{bmatrix} \alpha_x \cdot 2 & s & x_0 \cdot 2 \\ 0 & \alpha_y \cdot 1 & y_0 \cdot 1 \\ 0 & 0 & 1 \end{bmatrix} \\ K'_{[1:2]} &= \begin{bmatrix} \alpha_x \cdot 1 & s & x_0 \cdot 1 \\ 0 & \alpha_y \cdot 2 & y_0 \cdot 2 \\ 0 & 0 & 1 \end{bmatrix} \\ K'_{[1.2:2.3]} &= \begin{bmatrix} \alpha_x \cdot 1.2 & s & x_0 \cdot 1.2 \\ 0 & \alpha_y \cdot 2.3 & y_0 \cdot 2.3 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Die Formulierung, dass die jeweils neu rekonstruierten Szenen ähnlich sind, wurde deshalb verwendet, da durch die zuvorigen Fehler der korrespondierenden Punkte und später, bei der Triangulierung,

durch die *Sampson-Approximation* Abweichungen auftreten können. Als Beweise werden im folgenden vier Beispiele für die vier Lösungen der rekonstruierten Translationsmatrizen R' aufgezeigt. Des Weiteren werden die 3D-Plots und 2D-Plots der rekonstruierten Szenen bei unterschiedlich Auflösungen im Vergleich mit der Szene bei gleichen Auflösungen gezeigt. Die Koordinaten sind in unskalierten Pixel-einheiten gegeben. Die Originalszene ist in Abbildung 7.15 und 7.16 zu sehen. Zu beachten ist, das die Ausgabe des 3D Plots in *Mathematica* manchmal rechtsdrehend, manchmal linksdrehend dargestellt sind, weshalb der Eindruck aufkommt, die Szene und die Kameraposition seien gespiegelt dargestellt. Dies leidet auf ein generellen Darstellungsproblem von 3D-Plots in *Mathematica* zurückzuführen. Dieses kann mit zusätzlichem Code bereinigt werden, wurde aber zu diesem Zeitpunkt noch nicht implementiert.

$$\begin{aligned} P1 &= \begin{pmatrix} 0.991092 & 0.120891 & 0.0558752 & -0.812277 \\ -0.120366 & 0.992649 & -0.0126719 & 0.0389145 \\ -0.0569964 & 0.00583352 & 0.998357 & 0.581973 \end{pmatrix} \\ P2 &= \begin{pmatrix} 0.991092 & 0.120891 & 0.0558752 & 0.812277 \\ -0.120366 & 0.992649 & -0.0126719 & -0.0389145 \\ -0.0569964 & 0.00583352 & 0.998357 & -0.581973 \end{pmatrix} \\ P3 &= \begin{pmatrix} 0.378236 & -0.0296342 & -0.925235 & -0.812277 \\ 0.0547644 & -0.997021 & 0.0543212 & 0.0389145 \\ -0.924088 & -0.0712161 & -0.375487 & 0.581973 \end{pmatrix} \\ P4 &= \begin{pmatrix} 0.378236 & -0.0296342 & -0.925235 & 0.812277 \\ 0.0547644 & -0.997021 & 0.0543212 & -0.0389145 \\ -0.924088 & -0.0712161 & -0.375487 & -0.581973 \end{pmatrix} \end{aligned}$$

Abbildung 9.1: Zeigt die rekonstruierte Matrix R' bei unveränderter Auflösung. Die Auflösungen von C_δ und C'_δ sind die selben.

$$\begin{aligned} P1 &= \begin{pmatrix} 0.376619 & -0.0296193 & -0.925895 & 0.812113 \\ 0.0551443 & -0.996999 & 0.0543246 & -0.0388039 \\ -0.924725 & -0.0715175 & -0.373856 & -0.582208 \end{pmatrix} \\ P2 &= \begin{pmatrix} 0.376619 & -0.0296193 & -0.925895 & -0.812113 \\ 0.0551443 & -0.996999 & 0.0543246 & 0.0388039 \\ -0.924725 & -0.0715175 & -0.373856 & 0.582208 \end{pmatrix} \\ P3 &= \begin{pmatrix} 0.991142 & 0.121017 & 0.0546967 & 0.812113 \\ -0.120498 & 0.992632 & -0.0126975 & -0.0388039 \\ -0.0558303 & 0.00599422 & 0.998422 & -0.582208 \end{pmatrix} \\ P4 &= \begin{pmatrix} 0.991142 & 0.121017 & 0.0546967 & -0.812113 \\ -0.120498 & 0.992632 & -0.0126975 & 0.0388039 \\ -0.0558303 & 0.00599422 & 0.998422 & 0.582208 \end{pmatrix} \end{aligned}$$

Abbildung 9.2: Zeigt die rekonstruierte Matrix R' wenn K' mit einem Verhältnis von [5 : 2] skaliert wurde

$$\begin{aligned} P1 &= \begin{pmatrix} 0.990947 & 0.120272 & 0.0596553 & 0.810768 \\ -0.119751 & 0.992728 & -0.0122401 & -0.0387536 \\ -0.0606937 & 0.0049855 & 0.998144 & -0.584083 \end{pmatrix} \\ P2 &= \begin{pmatrix} 0.990947 & 0.120272 & 0.0596553 & -0.810768 \\ -0.119751 & 0.992728 & -0.0122401 & 0.0387536 \\ -0.0606937 & 0.0049855 & 0.998144 & 0.584083 \end{pmatrix} \\ P3 &= \begin{pmatrix} 0.37685 & -0.0292569 & -0.925812 & 0.810768 \\ 0.0543725 & -0.997079 & 0.0536412 & -0.0387536 \\ -0.924677 & -0.0705534 & -0.374158 & -0.584083 \end{pmatrix} \\ P4 &= \begin{pmatrix} 0.37685 & -0.0292569 & -0.925812 & -0.810768 \\ 0.0543725 & -0.997079 & 0.0536412 & 0.0387536 \\ -0.924677 & -0.0705534 & -0.374158 & 0.584083 \end{pmatrix} \end{aligned}$$

Abbildung 9.3: Zeigt die rekonstruierte Matrix R' wenn K' mit einem Verhältnis von [1 : 2] skaliert wurde

$$\begin{aligned} P1 &= \begin{pmatrix} 0.990963 & 0.12034 & 0.0592445 & -0.81095 \\ -0.119818 & 0.99272 & -0.0122887 & 0.0387766 \\ -0.060292 & 0.0050791 & 0.998168 & 0.583829 \end{pmatrix} \\ P2 &= \begin{pmatrix} 0.990963 & 0.12034 & 0.0592445 & 0.81095 \\ -0.119818 & 0.99272 & -0.0122887 & -0.0387766 \\ -0.060292 & 0.0050791 & 0.998168 & -0.583829 \end{pmatrix} \\ P3 &= \begin{pmatrix} 0.377058 & -0.0293027 & -0.925726 & -0.81095 \\ 0.0544044 & -0.997073 & 0.0537206 & 0.0387766 \\ -0.92459 & -0.0706194 & -0.37436 & 0.583829 \end{pmatrix} \\ P4 &= \begin{pmatrix} 0.377058 & -0.0293027 & -0.925726 & 0.81095 \\ 0.0544044 & -0.997073 & 0.0537206 & -0.0387766 \\ -0.92459 & -0.0706194 & -0.37436 & -0.583829 \end{pmatrix} \end{aligned}$$

Abbildung 9.4: Zeigt die rekonstruierte Matrix R' wenn K' mit einem Verhältnis von [1.2 : 2.3] skaliert wurde

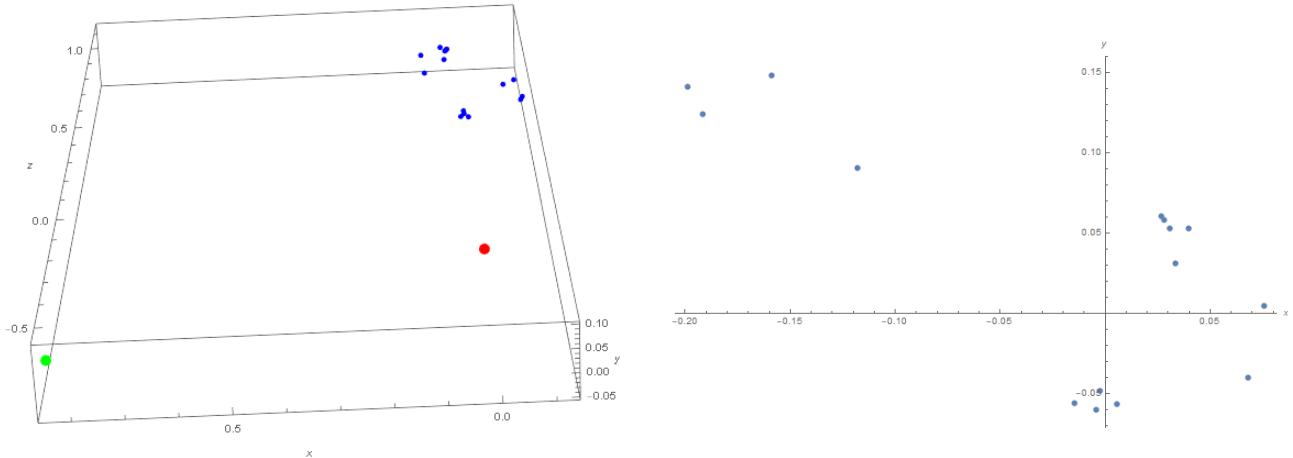


Abbildung 9.5: Rekonstruierte Szene, wenn K' mit einem Verhältnis von $[5 : 2]$ skaliert wurde

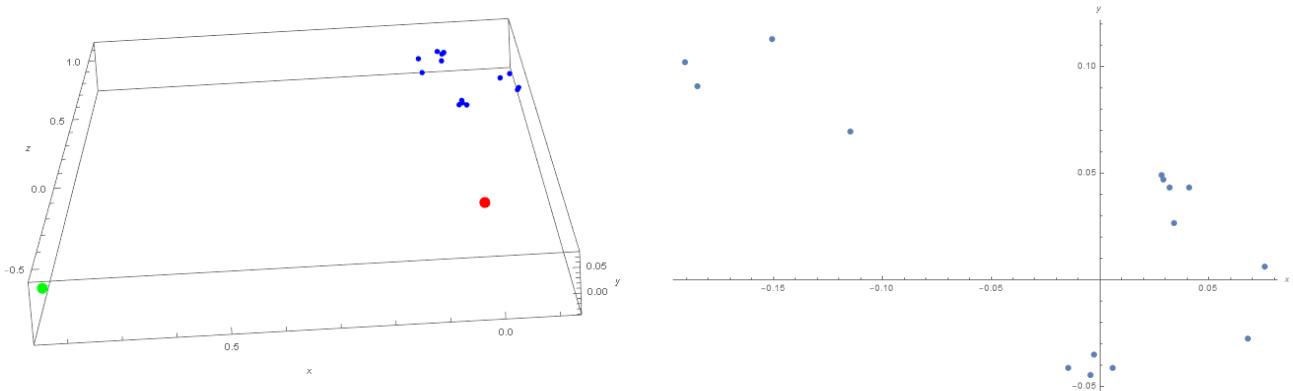


Abbildung 9.6: Rekonstruierte Szene, wenn K' mit einem Verhältnis von $[2 : 1]$ skaliert wurde

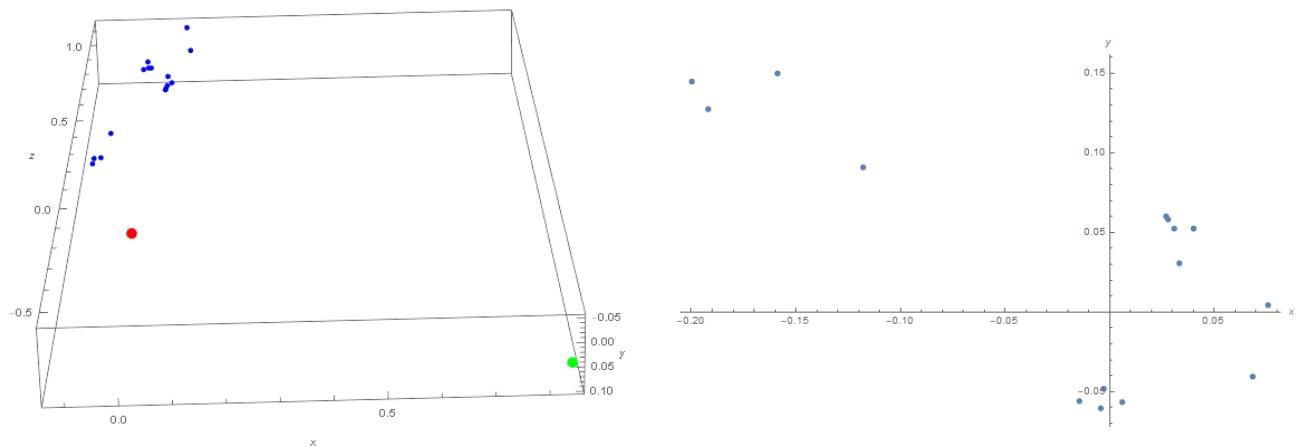


Abbildung 9.7: Rekonstruierte Szene, wenn K' mit einem Verhältnis von $[1 : 2]$ skaliert wurde

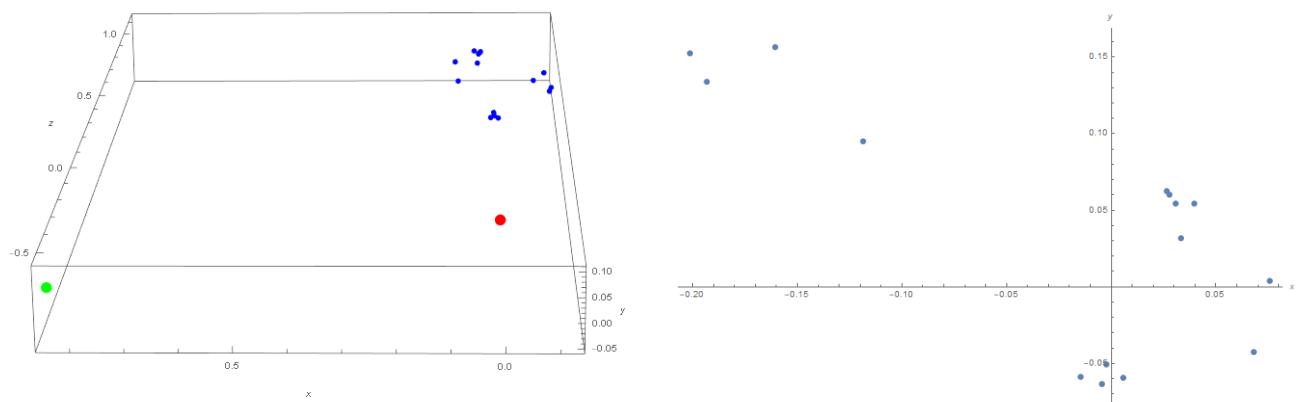


Abbildung 9.8: Rekonstruierte Szene, wenn K' mit einem Verhältnis von $[1.2 : 2.3]$ skaliert wurde

10 Punktesortierung in Schachbrettmustern

10.1 Algorithmus zur Punktesortierung in verzeichneten Schachbrettbildern

In diesem Teil der Masterthesis soll am Ende ein Algorithmus entstehen, welcher durch einen bereits bestehenden Algorithmus zur Detektion von Eckpunkten eines Schachbretts, eine Liste an Eckpunkten bekommt und diese auf deren Nachbarschaftsverhältnisse prüft. Die Schachbretter können dabei sowohl Kissen- als auch Tonnennverzeichnungen aufweisen und oder perspektivisch verzerrt sein. Mit den Algorithmus sollen Punkte wissen in welchen Reihen sie sich sowohl in x- als auch y-Richtung befinden. Jeder Punkt bekommt also eine Indexnummer in x-, sowie y-Richtung beziehungsweise in unserem Beispiel wird die y-Koordinate als j bezeichnet und die x-Koordinate als i , zugewiesen. Jeder Punkt bekommt mit Hilfe von den Mathematica eigenen *Associations* einen *Key* mit *NeighbourJ* und *NeighbourI* zugeteilt. Mit Hilfe dieser *Keys* kann dann später bei einem Stereobildpaar zum Beispiel die Korrespondierenden Eckpunkte der Schachbretter rausgesucht werden, was vielleicht genauere Ergebnisse liefert also die Suche von Hand. Des weiteren kann dieser Algorithmus in späteren Projekten vielleicht bei der Rausrechnung von Verzeichnungen hilfreich sein.

10.1.1 Vorläufiges Klassendiagramm

Module	Parameter	Lokale Variablen	Funktion
FindMinMax	Pointlist	Imin, imax, jmin, jmax, iSplits, jSplits, iDistance, jDistance	<ul style="list-style-type: none"> Die minimas und maximas der i und j-Werte der Koordinaten werden gesucht, um den „Rahmen“ des Gitters um das Schachbrett festzulegen In den ConstantArrays JSplits und ISplits werden die Zellen des Gitters gespeichert. Diese werden über die Distanz der jeweiligen Minimalwerte und Maximalwerte geteilt durch die gewünschte Anzahl an Zellen geteilt.
SortPointList	iSplits, jSplits, Pointlist	pi,pj	<ul style="list-style-type: none"> Die Eckpunkte werden zunächst der Größe nach nach ihren i-Werten Sortiert. Die sortierte Liste wird dann durchgezählt, so dass jeder Punkt seinen Indexwert in I-Richtung bekommt Danach werden die Eckpunkte der Größe nach nach ihren J-Werten sortiert und bekommen hier ebenfalls einen Index zugeordnet (Diese Sortierung ist nach jetzigem Stand des Algorithmus vllt nicht mehr zwingend notwendig)
GoThroughConvex Hulls	iSplits, jSplits, pj	ConvexHull	<ul style="list-style-type: none"> Nun wird herausgefiltert, welcher Punkt in welche Zelle des erstellten Gitters gehört, somit wird eine grobe Vorsortierung der Punkte für den weiteren Verlauf vorgenommen. In einer For-Schleife welche alle iSplits durchzählt wird die Funktion FindPointsInConvexHull bei jedem Durchgang aufgerufen welche eine Liste mit Associations in die Liste ConvexHull hinzufügt. Der Funktion werden die momentanen iSplits der Durchzählung übergeben und alle Jsplits. Des Weiteren wird die nach J sortierte Punktliste übergeben
FindPointsInConvexHull	iSplits[[1,ii]], iSplits[[1,ii+1]], jSplits, pj	ConvexHullCell={}, ConvexHullList, ConvexHullCellKeys = < >	<ul style="list-style-type: none"> Eine Liste namens ConvexHullCell und eine Association nach dem ConvexHullKeys wird angelegt Zwei For-Schleifen werden gestartet. Die erste läuft durch alle Jsplits, die zweite geht alle Punkte von pj durch. Innerhalb der For-Schleife wird dann überprüft, welche Punkte aus pj sich innerhalb der übergebenen iSplits und den dazugehörigen jSplits befinden. Die Koordinaten, die Indizes und die Zellenbezeichnung werden dann in Keys in die Association ConvexHullCellKeys gespeichert und er Liste ConvexHullCell angehängt. Diese Liste wird dann und die Liste ConvexHull angehängt Wiederholung des Vorganges mit neuen iSplits.

Abbildung 10.1: Klassendiagramm

FindStartVectors	ConvexHull	PointCloud={}, StartPointCloudKeys=< >, VecI,VecJ,countI,countJ, Start, nextI,nextJ,	<ul style="list-style-type: none"> Die Punkte der Zellen ($i = 1, j = \text{All}$) und ($i = \text{all}, j = 1$) werden in eine neue Liste namens StartPointCloud gespeichert. Die Liste wird zweimal durchlaufen <ul style="list-style-type: none"> Alle Punkte welche sich in den Zellen $j = 1$ und $i = \text{all}$ aufhalten. Aus ihnen wird der geringste i-Wert ermittelt (VecI) Alle Punkte welche sich in den Zellen $j = \text{all}$ und $i = 1$ aufhalten. Aus ihnen wird der geringste j-Wert ermittelt (VecJ) Die Punkte mit den geringsten Werten werden in VecI und VecJ gespeichert. Jetzt wird die Liste nochmals zweimal durchgegangen. <ul style="list-style-type: none"> Alle Punkte welche sich in den Zellen $j = 1$ und $i = \text{all}$ aufhalten werden durchgegangen. Aus ihnen wird derjenige Wert ermittelt, welcher einen Wert für j besitzt der kleiner ist als der momentane j-Wert von VecI und dessen i-Wert kleiner ist als der i-Wert von VecI plus einem Offset. Dieser Wert ist das neue VecI Alle Punkte welche sich in den Zellen $j = \text{all}$ und $i = 1$ aufhalten. Aus ihnen wird derjenige Wert ermittelt, welcher einen Wert für i besitzt der kleiner ist als der momentane i-Wert von VecJ und dessen j-Wert kleiner ist als der j-Wert von VecI plus einem Offset. Dieser Wert ist das neue VecJ VecI und VecJ ergeben den gleichen Punkt und somit ist der Startwert gesetzt. Nun sollen die ersten Punkte in i- und j-Richtung vom Startpunkt aus gefunden werden. Es wird ein nexti und ein nextj definiert, dessen Koordinaten sehr groß anfangen Es wird wieder die StartPointListe zweimal durchlaufen <ul style="list-style-type: none"> Es werden Punkte gesucht, welche sich in der selben Zelle i wie der Startpunkt befinden und auch die Zellen +1 und -1 drum herum. Sollte es ein Punkt geben, der kleiner ist als das momentane nexti und größer als der Startpunkt, jedoch nicht gleich dem Startpunkt. So nimmt nexti dessen Wert an. Danach muss geprüft werden, ob das potentielle nexti auch wirklich das richtige nexti ist. Hierzu wird eine neue For-Schleife gestartet, welche wieder die StartPointCloud durchgeht und überprüft ob es einen Punkt gibt dessen j-Koordinatenabstand zum Startpunkt kleiner ist also der j-Koordinatenabstand des momentanen nexti zum Startpunkt und ob dessen i-Koordinatenabstand zum Startpunkt kleiner ist als der momentane i-Koordinatenabstand von nexti zum Startpunkt.
------------------	------------	---	---

Abbildung 10.2: Klassendiagramm

			<ul style="list-style-type: none"> Ist dies der Fall so wird dieser Punkt zum neuen nexti. Mit dem potentiellen nextj wird ebenso verfahren.
CreatePossiblePoint - ListsIAndJ	nextI, nextJ, Start, ConvexHull	IList= {}, JList= {}, IDir, JDir, distance, cache, PotNextI, PotNextJ	<ul style="list-style-type: none"> IDir und JDir sind die Richtungsvektoren vom Startpunkt aus in beide Kantenrichtungen des Schachbretts. Danach werden die ersten beiden Spalten in I- und J-Richtung jeweils durchlaufen, und in IList und JList gespeichert. Diese Listen enthalten weitere potentielle Punkte entlang der gesuchten Kante. Die Kanten können natürlich durch die perspektivische Verzerrung mancher Bilder auch noch weiter in die Zellen hineinragen. Hierum kümmert sich dann im späteren Algorithmus die SaftyJList[] und SaftyIList[] Funktionen
FindNeighbours	IList, JList ,Start, nextI, nextJ, ConvexHull	SortedPointsKeys = <>, Sortedpoints = {}, proportionJ, proportionI, Jtemp, itemp, PotNextJDir, distanceNextPotPointJ, PotNextIDir, distanceNextPotPointI, NeighbourNumberJ, NeighbournumberI, distanceJ, distanceI, NextJDir, NextIDir, StartPropJForFirstCompleteGridJ	<ul style="list-style-type: none"> StartPoint und NextPointI und NextPointJ werden die Keys NeighbourI und NeighbourJ gegeben mit StartPoint(NeighbourJ → 1, NeighbourI → 1), NextPointI(NeighbourJ → 1, NeighbourI → 2) und NextPointJ(NeighbourJ → 2, NeighbourI → 1). Diese drei bereits bekannten Punkte werden dann auch in eine angelegte CheckPointList gespeichert, diese wird für das spätere Prüfen von weiteren Punkten benötigt. Nun wird zunächst in einer For-Schleife die Punkte von StartPoint und NextPointJ aus gesucht. <ul style="list-style-type: none"> Anmerkung: Für die Punkte in I-Richtung des Schachbretts wird das selbe Verfahren angewandt. Benötigt wird die Distanz zwischen dem momentanen StartPointJ, welcher nach jedem Durchlauf der Schleife den Wert des momentanen NextPointJ bekommt und einem momentanen NextPointJ, welcher nach jedem Durchlauf der Schleife den Wert des gerade neu gefundenen nächsten Punktes bekommt. Die Schleife selbst durchläuft alle Punkte, welche in der für die Richtung entsprechenden Richtung Liste sind. In diesem Fall die JList Es wird außerdem bei der Suche den nächsten Punktes in j-Richtung eine Distanz namens proportion berechnet, welcher die Distanz i zwischen StartPoint und Nextpoint beinhaltet. Innerhalb der durchlaufenden Liste wird derjenige Punkt gesucht welcher zum NextPointJ den geringsten Abstand in J-Richtung hat und dessen Abstand in I-Richtung <= der i-Koordinate des NextPointJ + proportion+noch einen Puffer ist und >= der i-Koordinate des NextPointJ – proportion+noch einen Puffer.

Abbildung 10.3: Klassendiagramm

			<ul style="list-style-type: none"> Ist der nächste Punkt gefunden, so wird dieser der SortedPointsList und der der CheckPointsList übergeben mit den passenden NeighbourI und NeighbourJ associationKey. Des Weiteren bekommt für den nächsten Schleifendurchlauf StartPointJ die Werte von NextPointJ und NextPointJ' in der neu gefundenen Punkt aus der JList gespeichert. Im Anschluss werden noch in AppendTo[SortedPoints, SaftyListJ[Start, CheckPointJ, proportionY, CheckCellForJ, ConvexHull, distanceJ]], AppendTo[SortedPoints, CompleteJGrid[nextI, ConvexHull, StartDistanceJ, StartProportionJ, Start, Jp, al]] Weitere Punkte zur SortedList in J-Richtung hinzugefügt, bei ersterem nur in bestimmten Fällen. Mehr zu den Funktionen folgt. Nicht zu vergessen: selbiges wie oben wird auch mit den Punkten in I-Richtung vollzogen, bis auf die CompleteGrid Funktion
SaftyList	Start, CheckLastPointJ , proportionJ, LastJPointsCell, ConvexHull, NextJDir	SaftyList = {}, SaftyKeys =<>, SaftyKeysList = {}, propJ, lastDir, lastdistanceJ	<ul style="list-style-type: none"> Der Funktion werden die Parameter CheckLastPointJ und LastJPointCell mitgegeben. Diese stammen aus der Funktion FindNeighbours und es handelt sich um den letzten Punkt der innerhalb der JListe ermittelt wurde und dessen i-Zelle in welcher sich dieser befindet. Da die I- bzw die JListe in jede Richtung nur die Punkte der ersten beiden Zellen beinhaltet, kann es bei einem rotierten Schachbrett sein, dass sich noch weitere Punkte in Zellen weiter oben/unten befinden Die Funktion SaftyList, erstellt eine Liste aus möglichen weiteren Punkten, indem sie die in diesem Falle I-Zelle des letzten Punktes nimmt und diese so wie die unter und oberhalb dieser Zelle und alle deren J-Zellen aufwärts auf einen möglichen nächsten Punkt untersucht. → Dies geschieht nach dem selben Verfahren wie in FindNeighbours. Sollte es noch einen geben wird dieser ebenfalls der CheckPointList und der SortedPointsList zugewiesen, ansonsten passiert nichts.
CompleteJGrid	StartPointI, ConvexHull, StartDistanceJ, proportionJ, Start,	PossiblePointsList = {}, SortedPointsKeys = <>, SaftyPossiblePointsListJ = {}, propJ, StartPointForJGrid, distanceJ,	<ul style="list-style-type: none"> Nachdem die äußersten Punkte der linken und unteren Kante des Schachbretts gefunden wurden, muss nun das restliche Grid des Schachbretts detektiert und mit den richtigen NeighbourI und NeighbourJ Werten versehen werden. Jeder Punkt der in I-Richtung als „Rahmenpunkt“ detektiert wurde, wird einmal als Startpunkt gesetzt, von ihm aus wird dann in einem sehr ähnlichen Verfahren wie schon zuvor der nächste Punkt in J-Richtung gesucht und wenn nötig tritt auch hier

Abbildung 10.4: Klassendiagramm

NeighbourNumberJ, aI	NextNeighbourNumberJ, distanceNextPotGridPointJ, tempJ, NextPointJDir, NextJDir, CheckPointJ, CheckCellForJ	nochmal die SaftyList Funktion in Kraft um auch wirklich alle Punkte jeder Reihe ausfindig zu machen
----------------------	--	--

Abbildung 10.5: Klassendiagramm

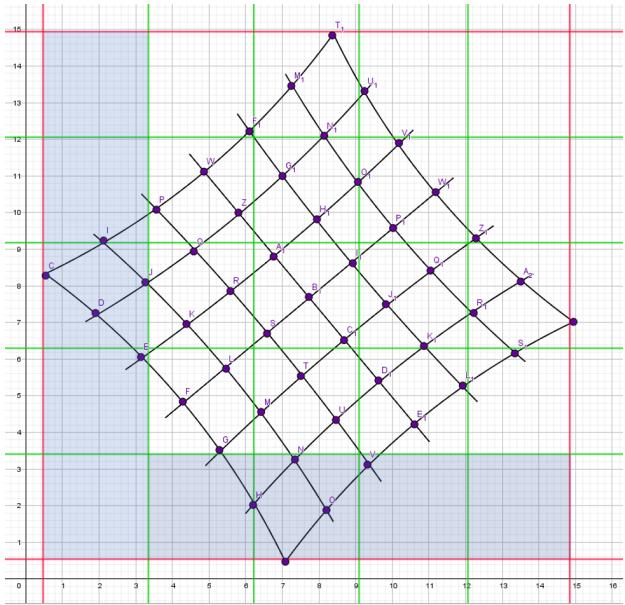


Abbildung 10.6

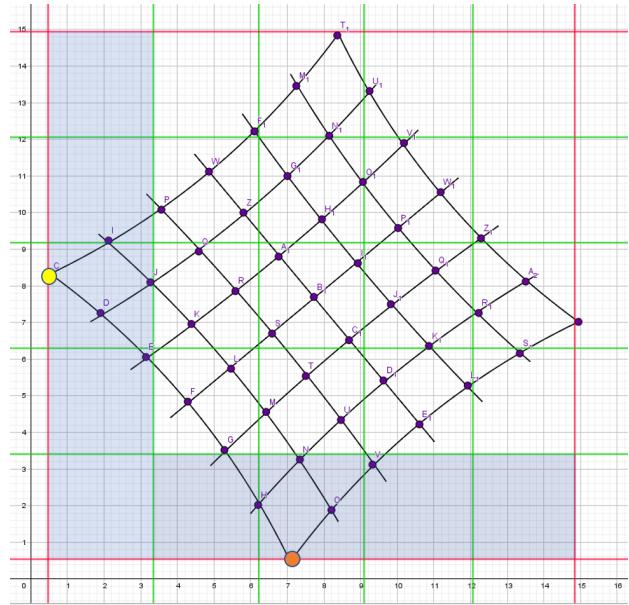


Abbildung 10.7

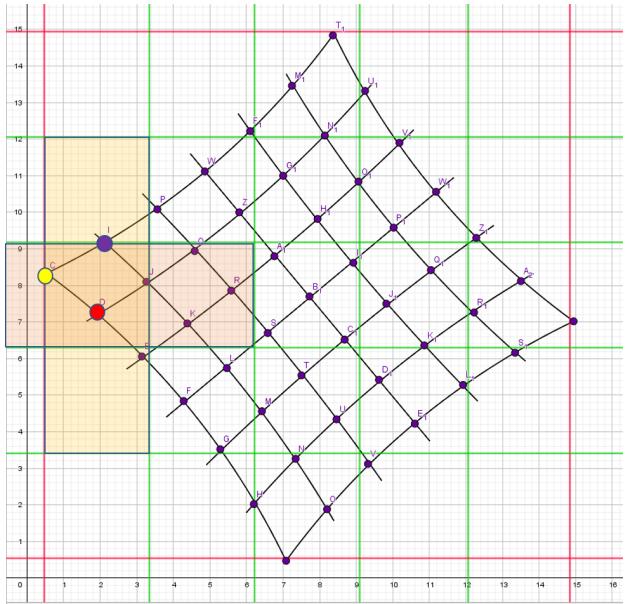


Abbildung 10.8

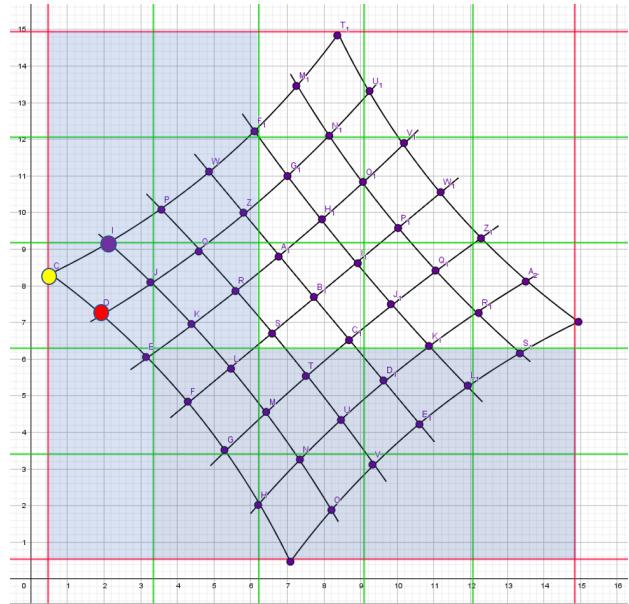


Abbildung 10.9

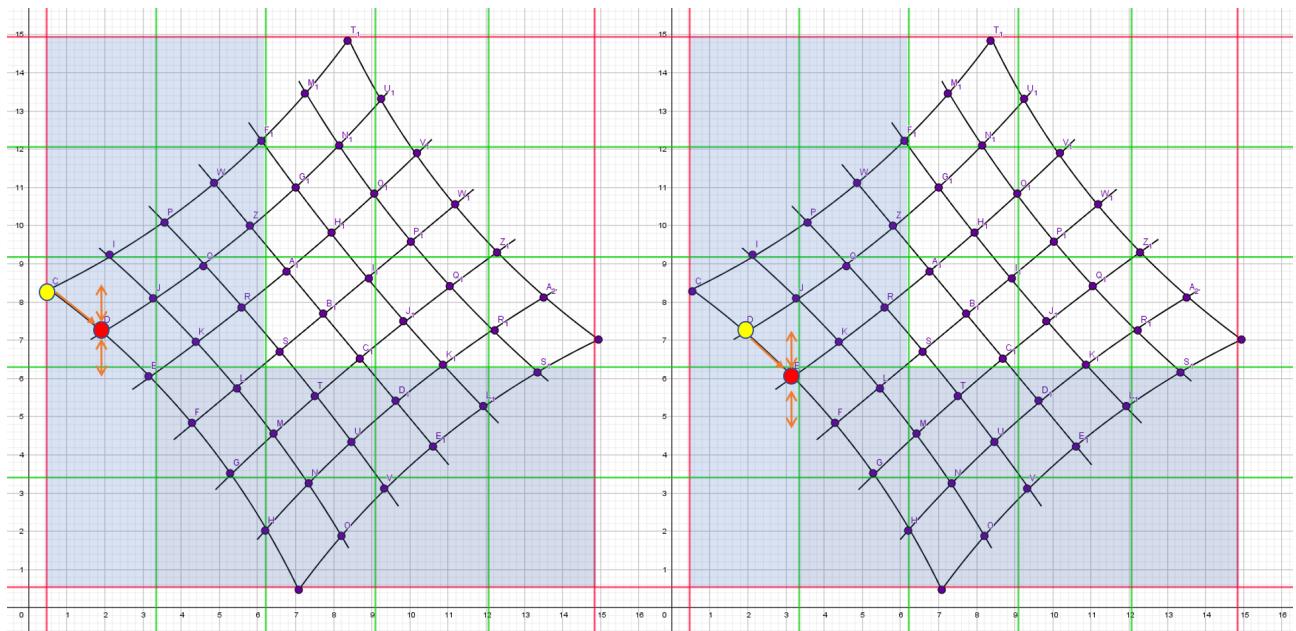


Abbildung 10.10: Klassendiagramm

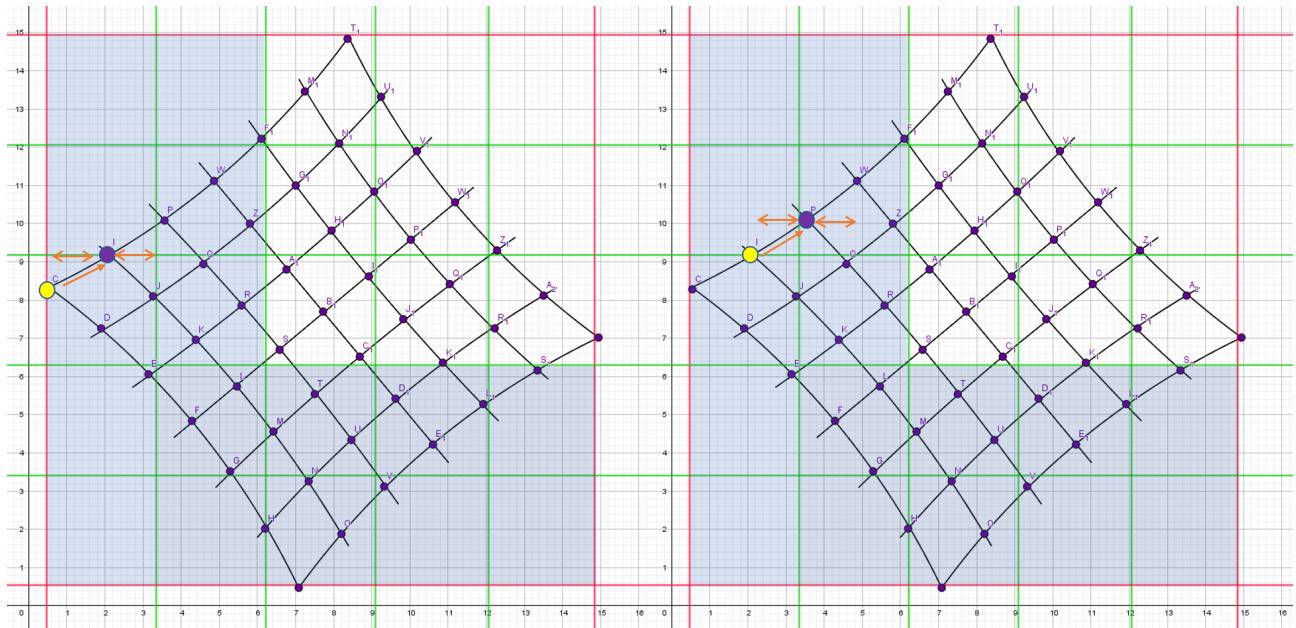


Abbildung 10.11: Klassendiagramm

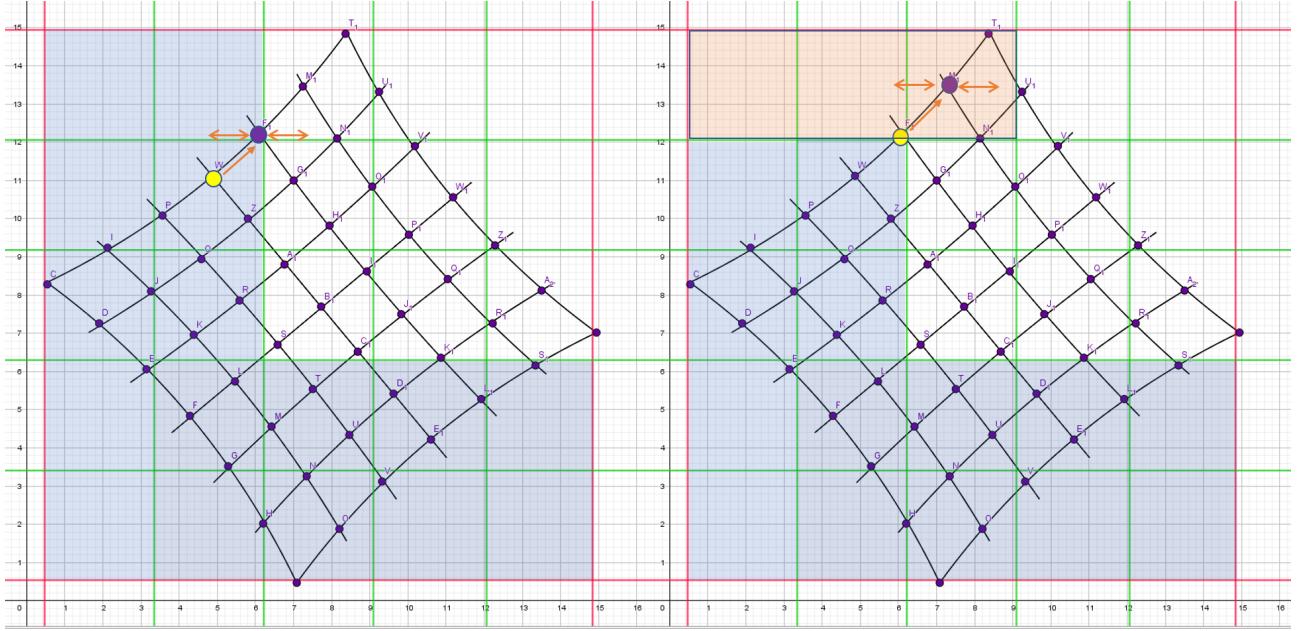


Abbildung 10.12: Klassendiagramm

10.1.2 Beispiele

In den folgenden Beispielen sieht man jeweils das Originalbild und ein Bild welches die durch den Algorithmus sortierten Punkte farbig ausgibt. Die grünen eingefärbten Punkte sind in den Bildern des Algorithmus die Nachbarn, welche sich in i-Richtung an der dritten Stelle befinden. Natürlich können auch andere Reihen oder auch einzelne Punkte abgefragt werden.

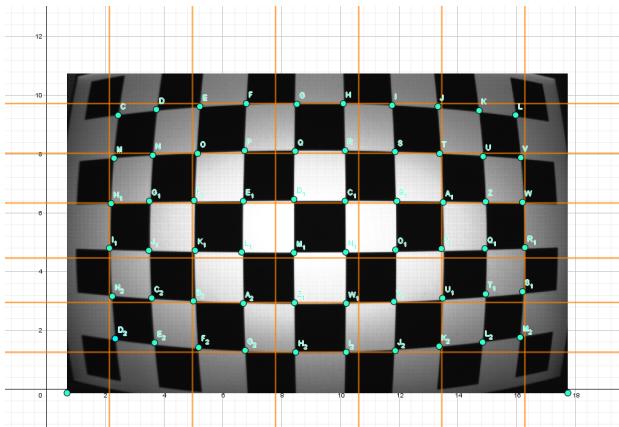


Abbildung 10.13: Bild eines Tonnenförmig verzeichneten Schachbretts

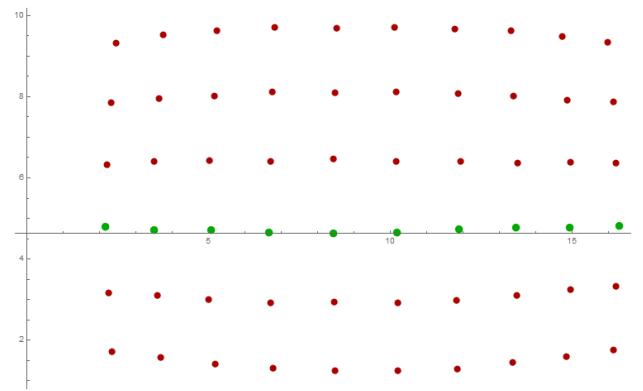


Abbildung 10.14: Algorithmisch detektierte Linie der dritten i-Reihe

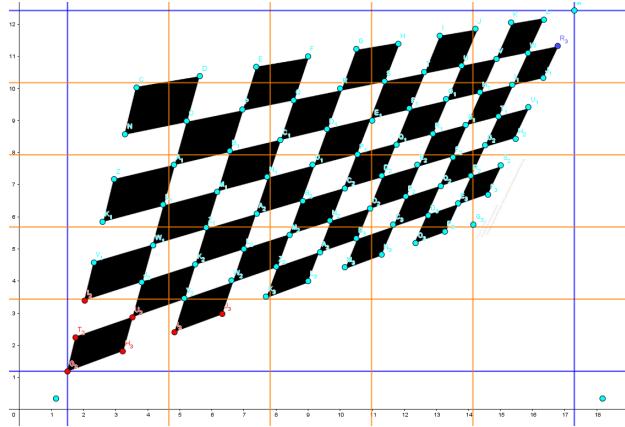


Abbildung 10.15: Bild eines perspektivisch verzerrtem Schachbretts

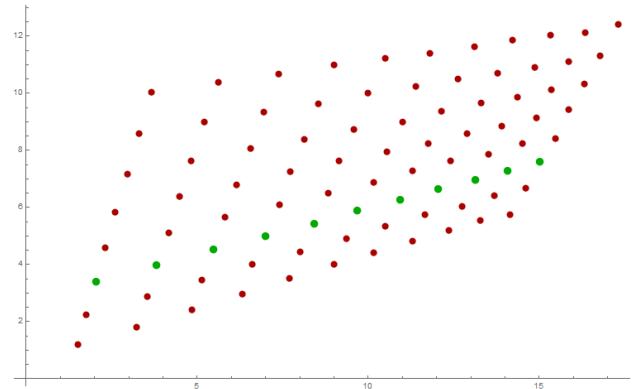


Abbildung 10.16: Algorithmisch detektierte Linie der dritten i-Reihe

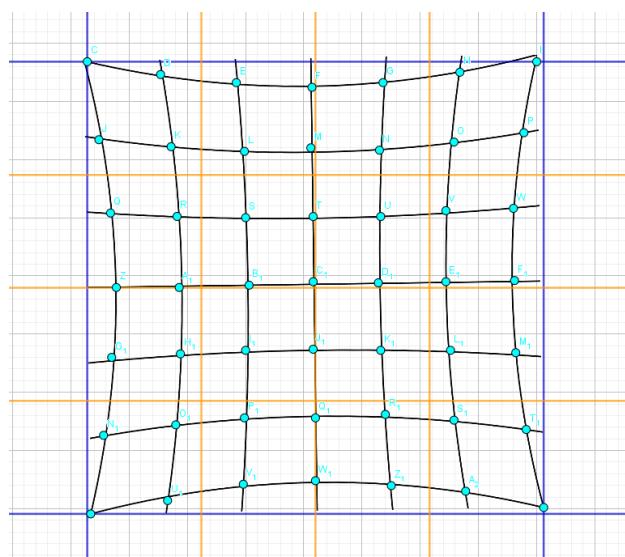


Abbildung 10.17: Bild eines Kissenförmig verzeichnetem Schachbretts

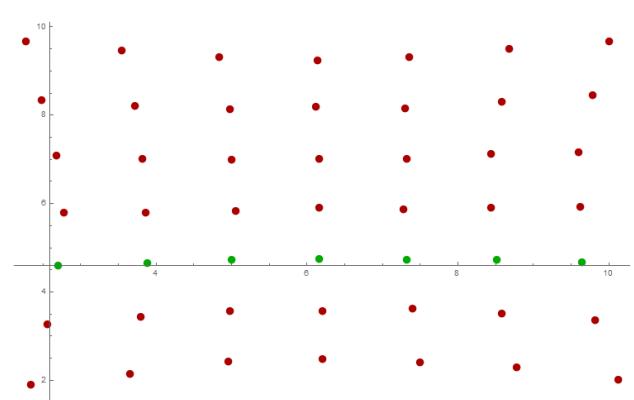


Abbildung 10.18: Algorithmisch detektierte Linie der dritten i-Reihe

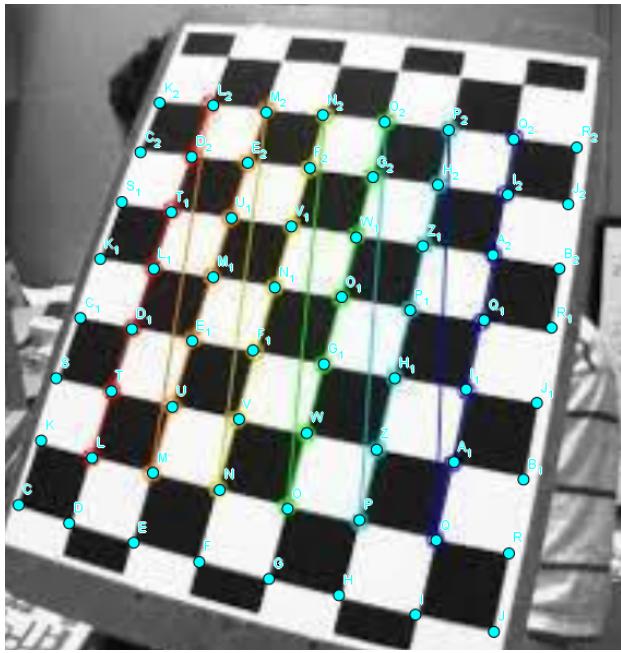


Abbildung 10.19: Bild eines Tonnenförmig verzeichnetem leicht perspektivisch verzerrtem Schachbretts(GRFIK AUSTAUSCHEN BILD IS KACKE)

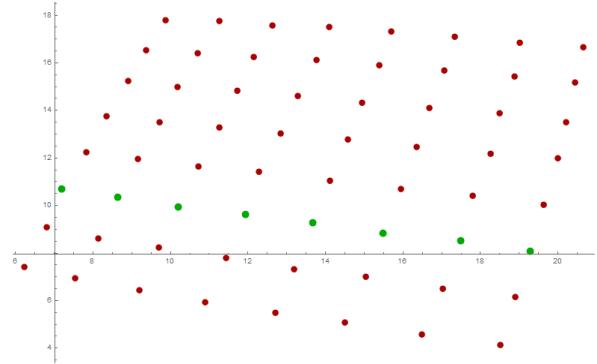


Abbildung 10.20: Algorithmisch detektierte Linie der dritten i-Reihe

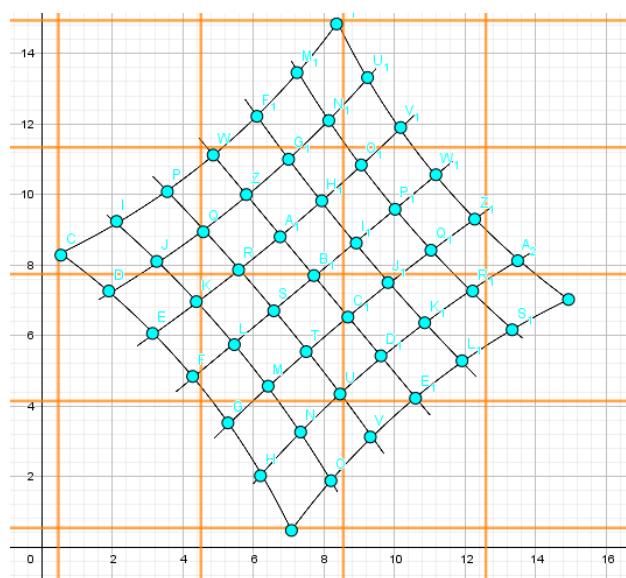


Abbildung 10.21: Bild eines Tonnenförmig verzeichnetem leicht perspektivisch verzerrtem Schachbretts

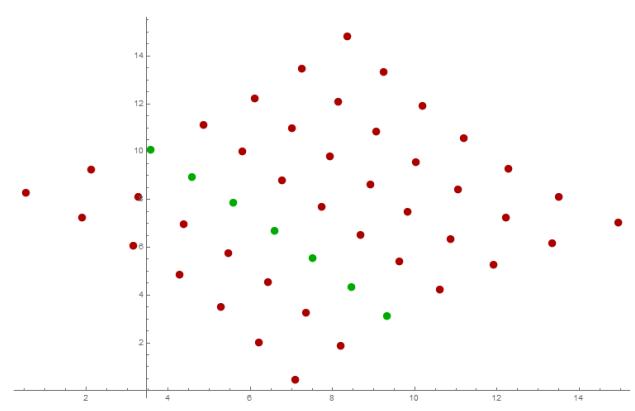


Abbildung 10.22: Algorithmisch detektierte Linie der dritten i-Reihe

11 Appendix

11.1 Homographien bei Drehungen um ein Projektionszentrum

In diesem Beispiel werden die abgebildeten Punkte eines Quadrats einer Kamera (C_β) mit $\beta = [\vec{b}_1, \vec{b}_2, \vec{b}_3, C]$ in die einer anderen Kamera (C', β') mit $\beta' = [\vec{b}'_1, \vec{b}'_2, \vec{b}'_3, C']$ mit Hilfe einer selbst aufgestellten Homographiematrix überführt. Hierzu wird zunächst eine Szene mit den Eckpunkten des Quadrates $A_\delta, B_\delta, C_\delta, D_\delta$ und dessen Mittelpunkt mit E_δ und zwei Kameras C_δ und C'_δ definiert. Für Die Projektionszentren der Kameras gilt hier, dass $C'_\delta = C_\delta$. Die Eckpunkte so wie die Kameras werden in Weltkoordinaten mit den Basen (O, δ) mit $\delta = [\vec{d}_1, \vec{d}_2, \vec{d}_3, O]$ angegeben. Mit Hilfe der gelernten Transformationsmethoden, werden die Eckpunkte in ihre jeweiligen 2D-Bildebenenkoordinaten (I, τ) und (I', τ') umgerechnet. Nachdem die Szene vollständig definiert und die auf den Kameras abgebildeten Punkte berechnet sind, werden zwei Methoden gezeigt, mit welchem die Homographiematrix ermittelt werden kann. Da es sich in diesem Beispiel um ein selbst aufgestelltes Minimalbeispiel handelt, kommt es zu keinem überbestimmten Fall bei der Berechnung der Homographiematrix. Da in Realbeispielen, jedoch immer damit gerechnet werden muss, dass es zu einem überbestimmten Fall kommt, wird das Verfahren, welches diesen Fall abdeckt ebenfalls aufgezeigt. Der Szenenaufbau des Minimalbeispiels soll folgendermaßen aussehen. Die Kamerakoordinatensysteme unterscheiden sich vom Weltkoordinatensystem durch eine Drehung um 180° um die \vec{d}_1 -Achse. Der Ursprung beider Kamerakoordinatensysteme entspricht dem Projektionszentrum C . Des Weiteren ist Kamera eins noch um 45° um die \vec{b}_2 zu Kamera eins eingedreht. Es werden zwei Bilder der selben Szene mit diesen Kameras aufgenommen. Die Behauptung ist, dass sich die beiden entstandenen Bilder mit einer Homographie ineinander überführen lassen. In Abbildung 11.1 ist der Aufbau nochmal grafisch veranschaulicht.

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \quad (11.1)$$

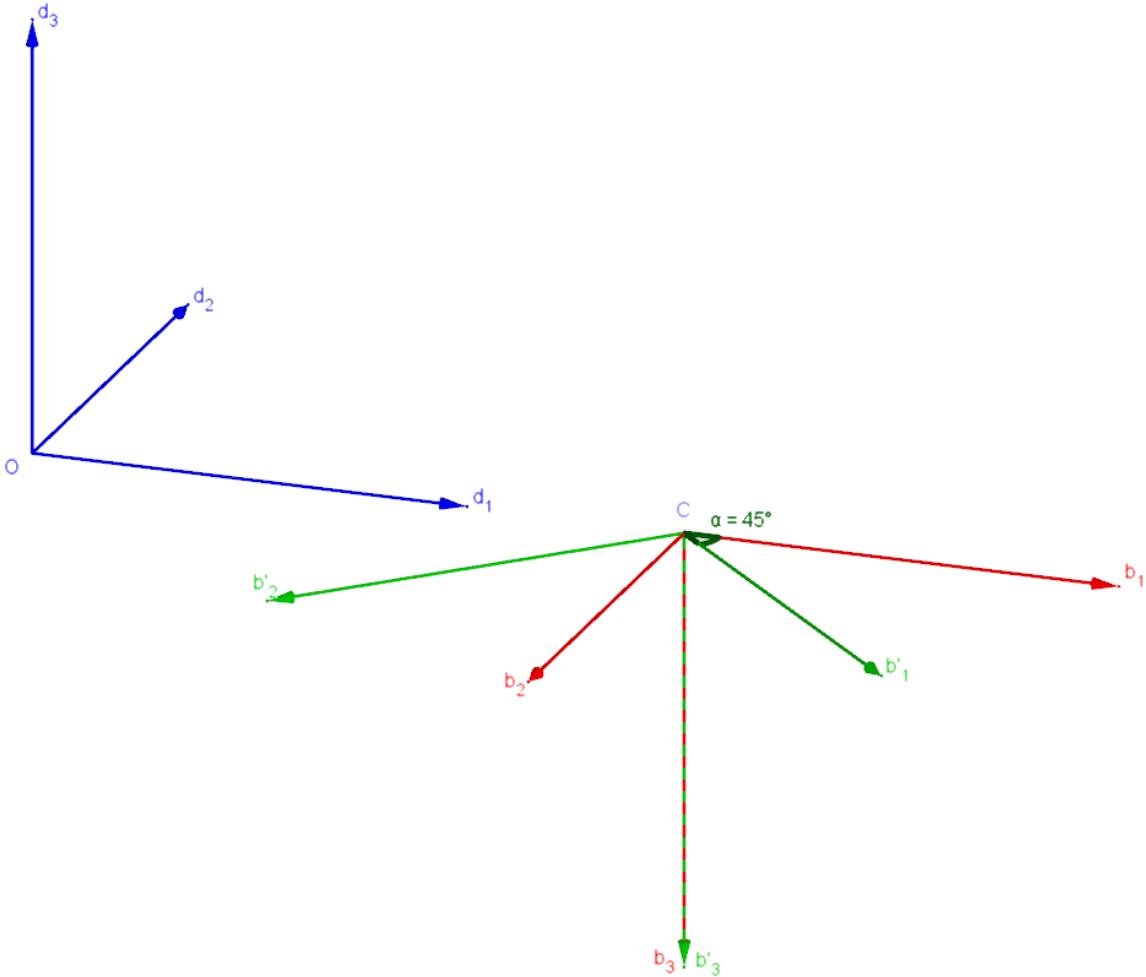


Abbildung 11.1: Weltkoordinatensystem (O, δ) mit $\delta = [\vec{d}_1, \vec{d}_2, \vec{d}_3, O]$ und Kamerakoordinatensysteme $\beta = [\vec{b}_1, \vec{b}_2, \vec{b}_3, C]$ und (C', β') mit $\beta' = [\vec{b}'_1, \vec{b}'_2, \vec{b}'_3, C']$.

Als nächstes werden die intrinsischen Kameraparameter K und K' für beide Kameras festgelegt. Für das Beispiel

$$K = K' = \begin{pmatrix} \zeta & 0 & 0 & 0 \\ 0 & \zeta & 0 & 0 \\ 0 & 0 & \zeta & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (11.2)$$

$$\rightsquigarrow \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = K \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} \zeta X \\ \zeta Y \\ \zeta Z \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\zeta}{Z} X \\ \frac{\zeta}{Z} Y \\ Z \\ 1 \end{pmatrix} \quad (11.3)$$

Für das Beispiel gilt die Bedingung, dass $\zeta \neq 0$ sein soll. Des Weiteren soll gelten, dass \vec{b}_1 gleich der Lotgeraden vom Objektpunkt zum Projektionszentrum entspricht und somit folgt, dass \vec{b}_1 in Sensorebene liegt und $\vec{b}_2 = \vec{b}_3 \times \vec{b}_1$ ist. Der Quader besteht aus den Punkten $A_\delta, B_\delta, C_\delta, D_\delta$ und $E\delta$ in homogenen Weltkoordinaten. Die Punkte in Weltkoordinaten bekommen die Koordinatenwerte zugewiesen.

$$A_\delta = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix}, B_\delta = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}, C_\delta = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}, D_\delta = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix}, E_\delta = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 2 \\ 1 \end{pmatrix} \quad (11.4)$$

Für die Transformation der Weltkoordinatentupel in die Bildebenen I und I' werden die Projektionsmatrizen $P = [KR| - KRC_\delta]$ und $P' = [K'R'| - K'R'C_\delta]$ aufgestellt werden. R soll eine Drehung der Punkte um 180° um die \vec{d}_1 -Achse durchführen. R' muss zusätzlich im Anschluss noch eine Drehung, um 45° um die neue \vec{b}_3' -Achse von Kamera zwei, an die vorherige Drehung anhängen. Die so erhaltenen Matrizen R und R' können nun dazu verwendet werden, die Objektpunkte bezüglich des Weltkoordinatensystems (O, δ) in Punkte bezüglich der jeweiligen Kamerakoordinatensysteme (C, β) und (C', β') zu transformieren.

$$\begin{bmatrix} \cos(180) & -\sin(180) & 0 & 0 \\ \sin(180) & \cos(180) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} (\vec{d}_1, \vec{d}_2, \vec{d}_3, O) = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} (\vec{d}_1, \vec{d}_2, \vec{d}_3, O) \quad (11.5)$$

$$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^T = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = R \quad (11.6)$$

Im nächsten Schritt wird Kamera zwei noch um 45° um ihre \vec{b}_3' -Achse gedreht. Diese Transformation wird mit der vorherigen Transformation multipliziert und es entsteht P' für Kamera zwei.

$$\cos(45) = \sin(45) = \frac{1}{\sqrt{2}} = a \quad (11.7)$$

$$\begin{bmatrix} a & 0 & a & 0 \\ 0 & 1 & 0 & 0 \\ -a & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -a & 0 & a & 0 \\ 0 & -1 & 0 & 0 \\ a & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} (\vec{d}_1, \vec{d}_2, \vec{d}_3, O) \quad (11.8)$$

$$= \begin{bmatrix} -a & 0 & a & 0 \\ 0 & -1 & 0 & 0 \\ a & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^T = \begin{bmatrix} -a & 0 & a & 0 \\ 0 & -1 & 0 & 0 \\ a & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = R' \quad (11.9)$$

Handelt es sich bei den Kamerakoordinatensystemen nicht um kartesische Koordinatensysteme, so muss von den Rotationsmatrizen jeweils die Inverse genommen werden, um R und R' zu erhalten. Die Inversen Rotationsmatrizen werden in den folgenden Gleichungen nicht mit R^{-1} , sondern R bezeichnet. So entsprechen sie auch den Notationen der Literatur[5, 3]. Sind die Matrizen R und R' ermittelt, können nun die Punkte des Quadrats im Raum in Punkte der jeweiligen Kameras umgerechnet werden. Für Punkte spezifisch Kamera eins ergeben sich die in den Gleichungen 3.23 bis 3.27 aufgezeigten Werte.

$$\begin{pmatrix} A_\beta \\ 1 \end{pmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} \quad (11.10)$$

$$\begin{pmatrix} B_\beta \\ 1 \end{pmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} \quad (11.11)$$

$$\begin{pmatrix} C_\beta \\ 1 \end{pmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 2 \\ 1 \end{pmatrix} \quad (11.12)$$

$$\begin{pmatrix} D_\beta \\ 1 \end{pmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 2 \\ 1 \end{pmatrix} \quad (11.13)$$

$$\begin{pmatrix} E_\beta \\ 1 \end{pmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 2 \\ 1 \end{pmatrix} \quad (11.14)$$

Für Punkte spezifisch Kamera zwei ergeben sich die Punkte aus den Gleichungen 3.28 bis 3.32.

$$\begin{pmatrix} A'_{\beta'} \\ 1 \end{pmatrix} = \begin{bmatrix} -a & 0 & a & 0 \\ 0 & -1 & 0 & 0 \\ a & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2a \\ 0 \\ 2a \\ 1 \end{pmatrix} \quad (11.15)$$

$$\begin{pmatrix} B'_{\beta'} \\ 1 \end{pmatrix} = \begin{bmatrix} -a & 0 & a & 0 \\ 0 & -1 & 0 & 0 \\ a & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ 0 \\ 3a \\ 1 \end{pmatrix} \quad (11.16)$$

$$\begin{pmatrix} C'_{\beta'} \\ 1 \end{pmatrix} = \begin{bmatrix} -a & 0 & a & 0 \\ 0 & -1 & 0 & 0 \\ a & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ -1 \\ 3a \\ 1 \end{pmatrix} \quad (11.17)$$

$$\begin{pmatrix} D'_{\beta'} \\ 1 \end{pmatrix} = \begin{bmatrix} -a & 0 & a & 0 \\ 0 & -1 & 0 & 0 \\ a & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2a \\ -1 \\ 2a \\ 1 \end{pmatrix} \quad (11.18)$$

$$\begin{pmatrix} E'_{\beta'} \\ 1 \end{pmatrix} = \begin{bmatrix} -a & 0 & a & 0 \\ 0 & -1 & 0 & 0 \\ a & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{2}a \\ -\frac{1}{2} \\ \frac{5}{2}a \\ 1 \end{pmatrix} \quad (11.19)$$

Die entstandenen Punkte sind nun entsprechend der Kamerakoordinatensysteme definiert. Jetzt müssen noch die jeweiligen Kameramatrizen K und K' mit diesen verrechnet werden, um so die 2D-Bildebenenpunkte der entsprechenden Kameras zu erhalten. ζ bekommt in diesem Beispiel den Wert -1, das bedeutet laut der Definition der Kamerakoordinatensysteme, dass der Sensor sich hinter dem Projektionszentrum befindet. Das entstehende Bild ist somit um 180° gedreht auf dem Sensor abgebildet.

$$K = K' = \begin{pmatrix} \zeta & 0 & 0 & 0 \\ 0 & \zeta & 0 & 0 \\ 0 & 0 & \zeta & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (11.20)$$

Man kann natürlich wie bereits gewohnt die Rotationsmatrizen und die Kameramatrizen bereits im Vorhinein miteinander multiplizieren und dann auf einen 3D-Weltpunkt anwenden. Auf diese Weise bekommt man die fertigen Projektionsmatrizen $P = [KR| - KRC_\delta]$ und $P' = [K'R'| - K'R'C_\delta]$. In diesem Beispiel werden sie nochmal getrennt voneinander auf die Punkte angewandt. Die fünf Punkte werden jeweils in einer Matrix zusammengeschrieben und mit K beziehungsweise K' verrechnet.

$$K \cdot \begin{pmatrix} 0 & -1 & -1 & 0 & -\frac{1}{2} \\ 0 & 0 & -1 & -1 & -\frac{1}{2} \\ 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & -1 & 0 & -\frac{1}{2} \\ 0 & 0 & -1 & -1 & -\frac{1}{2} \\ 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (11.21)$$

$$= \begin{pmatrix} 0 & 1 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 1 & -\frac{1}{2} \\ -2 & -2 & -2 & -2 & -2 \\ 2 & 2 & 2 & 2 & 2 \end{pmatrix} \quad (11.22)$$

$$\simeq \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \\ -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \simeq \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (11.23)$$

$$K' \cdot \begin{pmatrix} 2a & a & a & 2a & \frac{3}{2}a \\ 0 & 0 & -1 & -1 & -\frac{1}{2} \\ 2a & 3a & 3a & 2a & \frac{5}{2}a \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2a & a & a & 2a & \frac{3}{2}a \\ 0 & 0 & -1 & -1 & -\frac{1}{2} \\ 2a & 3a & 3a & 2a & \frac{5}{2}a \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (11.24)$$

$$= \begin{pmatrix} -2a & -a & -a & -2a & -\frac{3}{2}a \\ 0 & 0 & 1 & 1 & \frac{1}{2} \\ -2a & -3a & -3a & -2a & -\frac{5}{2}a \\ 2a & 3a & 3a & 2a & \frac{5}{2}a \end{pmatrix} \quad (11.25)$$

$$\simeq \begin{pmatrix} -1 & -\frac{1}{3} & -\frac{1}{3} & -1 & \frac{3}{5} \\ 0 & 0 & \frac{1}{3a} & \frac{1}{2a} & \frac{1}{5}a \\ -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \simeq \begin{pmatrix} -1 & -\frac{1}{3} & -\frac{1}{3} & -1 & \frac{3}{5} \\ 0 & 0 & \frac{1}{3a} & \frac{1}{2a} & \frac{1}{5}a \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (11.26)$$

Die entstandenen Punkte beider Kameras ergeben die in Abbildung 11.2 veranschaulichten Abbildungen, des Quadrats mit seinem Mittelpunkt, auf den jeweiligen Kamerabildebenen. Die aus den Objektpunkten M_δ entstandenen 2-D-Bildebenenkoordinaten m_τ und $m'_{\tau'}$ werden in den Koordinatensystemen (I, τ) mit $(\tau = \vec{b}_1, \vec{b}_2, \vec{b}_4)$ und (I', τ') mit $(\tau = \vec{b}'_1, \vec{b}'_2, \vec{b}'_4)$ dargestellt.

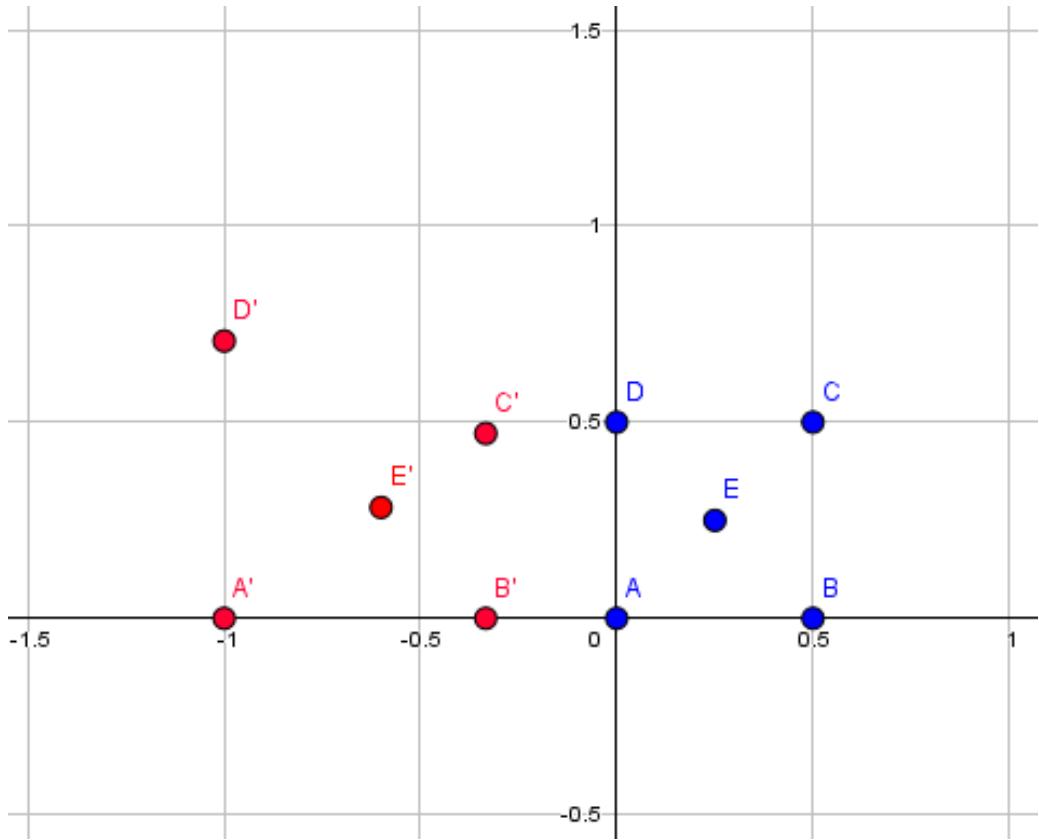


Abbildung 11.2: in blau ist die Abbildung des Quaders von Kamera eins und in rot die Abbildung des selben Quaders in Kamera zwei

Die fertigen Bildebenekoordinaten beider Kameras, sollen nun durch eine eigens aufgestellte Homographiematrix H ineinander überführt werden. Um eine Homographiematrix mit $H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$ zu erhalten werden die Punkte beider Kameras in eine Koeffizientenmatrix eingetragen, welche sich nach dem in den Gleichungen 4.40 bis 3.52 laufenden Schema ergibt. Das Verfahren wird anhand der Bildebenekoordinaten aufgezeigt.

$$H \cdot m_\tau = m'_\tau \quad (11.27)$$

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} m_\tau \end{bmatrix} = \begin{bmatrix} m'_{\tau'} \end{bmatrix} \quad (11.28)$$

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad (11.29)$$

Aus Gleichung 3.42 lassen sich ein Gleichungssystem mit zwölf bekannten und neun unbekannten aufstellen.

$$h_{11}x + h_{12}y + h_{13}z = \lambda x' \quad (11.30)$$

$$h_{21}x + h_{22}y + h_{23}z = \lambda y' \quad (11.31)$$

$$h_{31}x + h_{32}y + h_{33}z = \lambda z' \quad (11.32)$$

Da mit homogenen Koordinaten gearbeitet wird und somit z und $z' = 1$ sind, ergibt sich für die letzte Zeile $h_{31}x + h_{32}y + h_{33}z = 1$. Dieser Ausdruck kann in den ersten beiden Gleichungen für λ eingesetzt werden. Pro Punktpaar m_τ und $m'_{\tau'}$ ergeben sich somit zwei Gleichungen.

$$h_{11}x + h_{12}y + h_{13}z = (h_{31}x + h_{32}y + h_{33}z) \cdot x' \quad (11.33)$$

$$h_{21}x + h_{22}y + h_{23}z = (h_{31}x + h_{32}y + h_{33}z) \cdot y' \quad (11.34)$$

Für den Aufbau der nötigen Koeffizientenmatrix werden beide Ausdrücke noch nach Null aufgelöst, so dass sich die Gleichungen 3.48 und 3.49 aus 3.46 und 3.47 ergeben.

$$h_{11}x + h_{12}y + h_{13}z - (h_{31}x + h_{32}y + h_{33}z) \cdot x' = 0 \quad (11.35)$$

$$h_{21}x + h_{22}y + h_{23}z - (h_{31}x + h_{32}y + h_{33}z) \cdot y' = 0 \quad (11.36)$$

$$\rightsquigarrow h_{11}x + h_{12}y + h_{13}z - h_{31}x \cdot x' - h_{32}y \cdot x' - h_{33}z \cdot x' = 0 \quad (11.37)$$

$$\rightsquigarrow h_{21}x + h_{22}y + h_{23}z - h_{31}x \cdot y' - h_{32}y \cdot y' - h_{33}z \cdot y' = 0 \quad (11.38)$$

Die entstandenen Gleichungen werden jetzt pro Punktpaar m_τ und $m'_{\tau'}$ in eine Matrix nach folgendem Schema eingetragen.[5, 3, 14, 8]

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & x_1x'_1 & y_1x'_1 & 1 \cdot x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & x_1y'_1 & y_1y'_1 & 1 \cdot y'_1 \\ & & & \ddots & & & & & \\ & & & & \ddots & & & & \\ x_i & y_i & 1 & 0 & 0 & 0 & x_ix'_i & y_ix'_i & 1 \cdot x'_i \\ 0 & 0 & 0 & x_i & y_i & 1 & x_iy'_i & y_iy'_i & 1 \cdot y'_i \end{pmatrix} \cdot \begin{pmatrix} h1 \\ h2 \\ \vdots \\ \vdots \\ hi \end{pmatrix} = 0 \quad (11.39)$$

Wenn ein nicht überbestimmter Fall vorliegt, sprich wenn der Rang der Koeffizientenmatrix genau acht und nicht höher beträgt, kann aus der Koeffizientenmatrix einfach der Nullraum berechnet werden, um so die Einträge für die 3x3-Homographiematrix zu erhalten[3, 5, 14]. Gesucht wird also ein Vector \vec{x} , für den gilt das $H \cdot \vec{x} = 0$. Der gesuchte Vektor \vec{x} entspricht dem Kern der Koeffizientenmatrix und ist ein Spaltenvektor mit 9 Einträgen, welche in die 3x3-Homographiematrix eingetragen werden können[3, 14]. Tritt nun der Fall ein, dass es zu einem überbestimmtes System kommt, **was Beipelsweise auftritt wenn mehr als neun Punktpaare durch eine Homographie ineinander überführt werden sollen**, so kann nicht mehr die Ermittlung des Nullraums für die Berechnung der Homographiematrix genutzt werden. Für die Lösung überbestimmter Gleichungssysteme bietet sich die Singulärwertzerlegung an[3][15]. Das bedeutet es wird nicht derjenige Vektor \vec{x} gesucht für den gilt $H \cdot \vec{x} = 0$, sondern es wird derjenige Vektor \vec{x} gesucht, für den $\| H \cdot \vec{x} \|$ minimal wird[3, 14]. Die Singulärwertzerlegung von zum Beispiel der Koeffizientenmatrix ist eine Faktorisierung einer beliebigen Matrix $A \in \mathbb{R}^{m \times n}$ der Form $A = U \cdot S \cdot V^T$ mit orthogonalen Matrizen $U \in \mathbb{R}^{m \times m}$ und $V \in \mathbb{R}^{n \times n}$ sowie mit einer Diagonalmatrix.

$$S = \begin{pmatrix} s_1 & & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & & \ddots & \ddots & & \ddots \\ & & \ddots & \ddots & \ddots & & \ddots \\ & & & \ddots & \ddots & & \ddots \\ 0 & & \dots & & s_r & 0 & \dots & 0 \\ 0 & & \dots & & 0 & 0 & \dots & 0 \\ & & & & \ddots & \ddots & & \ddots \\ & & & & \ddots & \ddots & & \ddots \\ 0 & & \dots & & 0 & 0 & \dots & 0 \end{pmatrix} \quad (11.40)$$

Dabei soll für die Singulärwerte s_1 bis s_r gelten, dass $s_1 \geq s_2 \geq \dots \geq s_r \geq 0$ [15]. Entsprechend dieser Methode wird eine Singulärwertszerlegung, kurz *SVD* der entstandenen Koeffizientenmatrix

durchgeführt. Wir erhalten 3 Matrizen $U \cdot S \cdot V^T$. Durch die Zerlegung sind die diagonaleinträge von S in einer absteigenden Reihenfolge sortiert. Die Spalte von V^T , welche mit dem kleinsten Singulärwert von S korrespondiert, ergibt den Vektor \vec{x} , für den $\| H \cdot x \|$ minimal wird. Somit gleichen die neun Einträge der Homographiematrix gleich der letzten Spalte von V . Das Ergebnis für H hat dann die folgende Form

$$H = \begin{pmatrix} v_{19} & v_{29} & v_{39} \\ v_{49} & v_{59} & v_{69} \\ v_{79} & v_{89} & v_{99} \end{pmatrix} \quad (11.41)$$

Für das Minimalbeispiel mit reinen Punkten, welches in diesem Kapitel erstellt wurde, würde die Herleitung der Homographiematrix über die Ermittlung des Nullraumes der Koeffizientenmatrix genügen. Für die Matrix H ergibt sich aus den Werten der im Beispiel verwendeten Punkte

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad (11.42)$$

Nun werden die Punkte aus Kamera eins s mit Hilfe von H in die Punkte von Kamera zwei überführt und umgekehrt werden die Punkte aus Kamera zwei mit der Inversen H^{-1} in die Punkte von Kamera eins überführt.

$$x' = H \cdot x \rightsquigarrow \begin{pmatrix} -1 & -\frac{1}{3} & -\frac{1}{3} & -1 \\ 0 & 0 & \frac{1}{3a} & \frac{1}{2a} \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad (11.43)$$

$$x = H^{-1} \cdot x' \rightsquigarrow \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} -1 & -\frac{1}{3} & -\frac{1}{3} & -1 \\ 0 & 0 & \frac{1}{3a} & \frac{1}{2a} \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad (11.44)$$

11.2 Realbeispiel der Homographieberechnung bei Drehung um einen Drehpunkt

Im folgenden wird das breits bekannte Beipiel für die Drehung um das Projektionszentrum, für die Drehung um einen Drehpunkt umgebaut. Danach soll nach dem gleichen Verfahren wie im vorherigen Beispiel aufgezeigt wurde eine Homographiematrix ermittelt werden, welche die Bildebenenkoordinaten der einen Kamera in die der anderen Kamera überführen kann. Die Herangehensweise für dieses Beispiel größtenteils die selbe wie in dem Beispiel zuvor. Der unterschied liegt lediglich in der Transformation der Punkte in as Koordinatensystem von der gedrehten Kamera zwei. Diese besteht, im Gegensatz zum vorherigen Beispiel, nicht nur aus hintereinander geschalteten Rotationen, sondern beinhaltet des Weiteren noch zwei Translationen. Die komplette Transformationsmatrix wird, um der vorherigen Notation gerecht zu bleiben, wieder mit R beziehungsweise R' notiert, auch wenn es sich hier nicht mehr um eine Reine Rotation handelt. In der Literatur wird die gesamte Transformationsmatrix ebenfalls mit R notiert[3, 5, 14]. R und R' bestehen aus insgesamt drei Transformationsmatrizen $R = \text{Verschiebung}_1 \cdot \text{Rotation} \cdot \text{Verschiebung}_2$. Da die Position des Projektionszentrums von Kamera zwei nicht mehr mit dem von Kamera eins übereinstimmt, muss diese erst mit Hilfe einer Tranformationsmatrix T berechnet werden. $C_\delta = [0 \ 0 \ 0 \ 1]$ sind die Koordinaten von Kamera eins in Weltkoordinaten. Da in diesem Beispiel das Kamerakoodrinatensystem von Kamera eins und das Weltkoordinatensystem deckungsgleich sind heißt das, $C_\beta = [0 \ 0 \ 0 \ 1]$. somit ist Matrix R , welche die Transformation der ersten Kamera im Bezug auf das Weltkoordinatensystem beschreibt gleich der Einheitsmatrix.

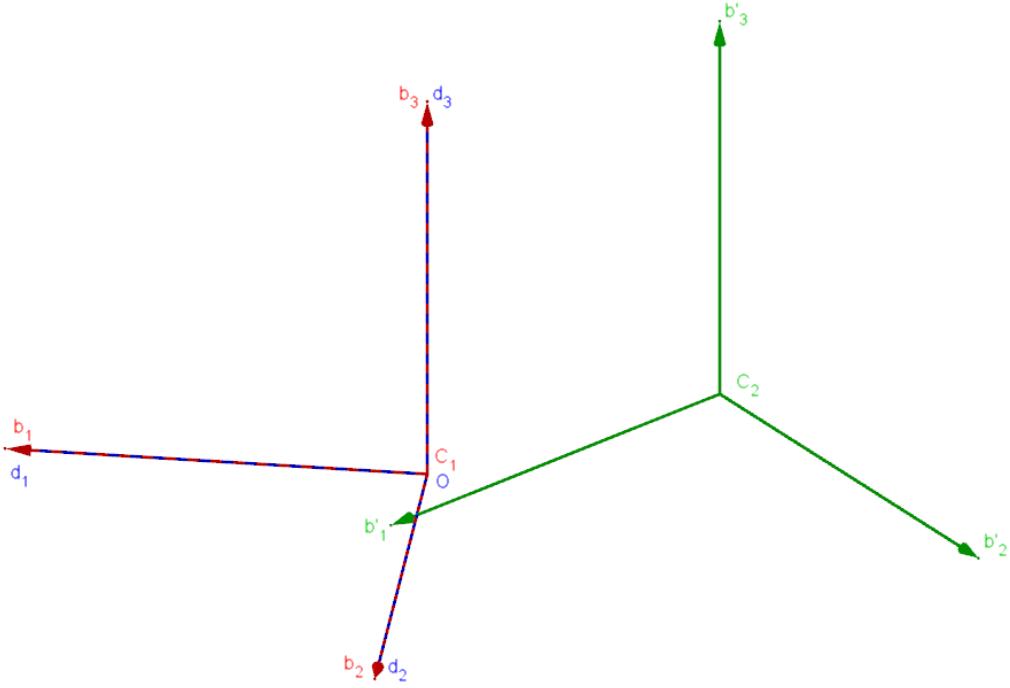


Abbildung 11.3: Weltkoordinatensystem (O, δ) mit $\delta = [\vec{d}_1, \vec{d}_1, \vec{d}_1, O]$ und Kamerakoordinatensysteme $\beta = [\vec{b}_1, \vec{b}_2, \vec{b}_3, C]$ und (C', β') mit $\beta' = [\vec{b}'_1, \vec{b}'_2, \vec{b}'_3, C']$.

Mit C' wird der Ursprung des Koordinatensystems von Kamera zwei bezeichnet. Als Objekte in der Ebene im \mathbb{R}^3 -Raum, werden die selben Punkte $A_\delta, B_\delta, C_\delta, D_\delta$ und E_δ wie im vorherigen Beispiel verwendet.

$$A_\delta = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix}, B_\delta = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}, C_\delta = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix}, D_\delta = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix}, E_\delta = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 2 \\ 1 \end{pmatrix} \quad (11.45)$$

Punkt E_δ bildet den Mittelpunkt des Quadrates und wird als Drehpunkt gewählt und mit *PivotPoint* bezeichnet. Als nächstes wird Matrix T aus drei Transformationsmatrizen zusammengestellt. Diese Matrix verschiebt das Projektionszentrum von C' an den gewünschten Ort im Weltkoordinatensystem. Bezeichnet werden die drei Matrizen mit T_1, T_2 und T_3 . T_1 beinhaltet die Verschiebung von des Ursprungs von Kamera eins zum *PivotPoint*, T_2 bildet die Rotationsmatrix, welche den verschobenen Punkt um die gewünschten 45° dreht. Die letzte Matrize T_3 beinhaltet wieder eine Translation, welche den Punkt vom *PivotPoint* zurück verschiebt.

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & -PivotPoint_x \\ 0 & 1 & 0 & -PivotPoint_y \\ 0 & 0 & 1 & -PivotPoint_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11.46)$$

$$T_2 = \begin{bmatrix} \cos(45^\circ) & 0 & \sin(45^\circ) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(45^\circ) & 0 & \cos(45^\circ) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11.47)$$

$$T_3 = \begin{bmatrix} 1 & 0 & 0 & PivotPoint_x \\ 0 & 1 & 0 & PivotPoint_y \\ 0 & 0 & 1 & PivotPoint_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11.48)$$

$$T = T_3 \cdot T_2 \cdot T_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & -1.26777 \\ 0 & 1 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0.93934 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11.49)$$

$$C'_\delta = T \cdot C_\delta = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & -1.26777 \\ 0 & 1 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0.93934 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1.27 \\ 0 \\ 0.94 \\ 1 \end{bmatrix} \quad (11.50)$$

Der Ursprung des Koordinatensystems von Kamera zwei befindet sich, angegeben in Weltkoordinaten, bei $C'_\delta = [-1.26777 \ 0 \ 0.93934 \ 1]^T$. Da nun die C'_δ bekannt ist, können die 3D-Objektpunkte in das Kamerakoordinatensystem von Kamera zwei transformiert werden. Hierzu wird die Matrix R' aufgestellt.

$$R' = \begin{bmatrix} [T_2]^{-1} & -[T_2]^{-1} \cdot V \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11.51)$$

aufgestellt werden. V ist der Translationsvektor welcher den Wert von C'_δ bekommt. Da es sich wieder um kartesische Koordinatensysteme handelt gilt wieder $[T_2]^{-1} = [T_2]^T$.

$$-[T_2]^T \cdot C'_\delta = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \cdot \begin{pmatrix} -1.27 \\ 0 \\ 0.94 \end{pmatrix} = \begin{pmatrix} 1.56 \\ 0 \\ 0.23 \end{pmatrix} \quad (11.52)$$

Für die Transformationsmatrizen R und R' gilt dann jeweils:

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11.53)$$

$$R' = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 1.56 \\ 0 & 1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0.23 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11.54)$$

Nachdem R und R' bestimmt sind, müssen die Punkte in Weltkoordinaten noch in die entsprechenden Kamerakoordinatensysteme und mit den Projektionsmatrizen K und K' in deren Bildkoordinatensysteme transformiert werden. Es gilt wieder wie im Beispiel zuvor, dass $K = K'$ ist.

$${}_{K_{c1}} [\pi]_{K_{c1}} = {}_{K_{c2}} [\pi]_{K_{c2}} = \begin{pmatrix} \zeta & 0 & 0 & 0 \\ 0 & \zeta & 0 & 0 \\ 0 & 0 & \zeta & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (11.55)$$

Es entstehen die folgenden beiden Punktematrizen pC für Kamera eins und pC' für Kamera zwei. Die Koordinaten der jeweiligen Punkte $pC = [A_\tau \ B_\tau \ C_\tau \ D_\tau \ E_\tau]$ und $pC' = [A_{\tau'} \ B_{\tau'} \ C_{\tau'} \ D_{\tau'} \ E_{\tau'}]$ aus Sicht der beiden Kameras befinden sich der Reihe nach in den Spalten der Punktematrix.

$$pC = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (11.56)$$

$$pC' = \begin{bmatrix} 0.09 & 0.36 & 0.36 & 0.09 & \frac{1}{4} \\ 0 & 0 & 0.42 & 0.61 & \frac{1}{4} \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (11.57)$$

$$(11.58)$$

Die Homographiematrix wird durch aufstellen der Koeffizientenmatrix und anschließendes Bestimmen von $H \cdot x = 0$ oder durch findes desjenigen Vektors \vec{x} für den $\|H \cdot x\|$ minimal wird. Da es sich auch hier nicht um einen überbestimmten Fall handelt, kann die homographiematrix entweder durch die Bestimmung des Kerns oder durch anwenden der Singulärwertszerlegung SVD , gewonnen werden. Die resultierende Homographiematrix sieht folgendermaßen aus.

$$H = \begin{bmatrix} 0.43 & 0 & 0.43 \\ 0 & 0.6 & 0 \\ 0.4 & 0 & 0.5 \end{bmatrix} \quad (11.59)$$

Wird H auf die Punkte pC angewandt, so liefert das Ergebnis die Punkte von pC' und wird die Inverse H^{-1} auf die Punkte pC' angewandt, so erhält man die Punkte von pC . Somit wurde bewiesen, dass Homographiematrizen immer zur Transformation von Punkten genutzt werden können, solange sich diese Punkte auf einer Ebene im Raum befinden. Die Definition der Homographie sagt aus, dass sowohl Translationen und Rotationen in der Homographiematrix vorkommen dürfen [12] [13]. Die Drehung um einen Drehpunkt ist nicht weiter als die Hintereinaderschaltung verschiedener Transformationsmatrizen. Das wichtigste Kriterium welches erfüllt sein muss, damit Homographien angewendet werden können ist, dass die Abbildungen der Punkte in allen Kameras auf einer Ebene sich befinden müssen [5].

In der Abbildung 4.4 ist ein grüner Punkt zu sehen welcher sich in Gegensatz zu den anderen Punkten nicht auf der selben Ebene befindet. Dieser Punkt lässt sich nicht mit der errechneten Homographiematrix ineinander überführen. Mit Hilfe von Homographien, können die Positionen und Orientierung der jeweiligen Kameras zueinander ermittelt werden, jedoch nur wenn sich die Szenepunkte auf einer Ebene befinden. Ein Beispiel hierfür wäre zum Beispiel die Aufnahme einer Gebäudefassade aus unterschiedlichen Kamerawinkeln und Positionen [5]. Für eine Szenenrekonstruktion einer kompletten 3D-Szene reichen pure Homographien nicht aus, hierfür muss sich der geometrischen Eigenschaften der Epipolaregeometrie bedient werden.

12 Fazit - Conclusion

13 Nächste Schritte - next steps

14 Protocol - 10.11.2015

15 Abkürzungsverzeichnis - List of Abbreviations

Abbildungsverzeichnis

1.1	Arbeitsprozesse der Stereoanalyse bei Verwendung von eigens erstellten synthetischen Bilddaten	6
1.2	Arbeitsprozesse der Stereoanalyse bei Verwendung von realen Bilddaten in einem unkalibrierten Fall	7
1.3	Arbeitsprozesse der Stereoanalyse bei Verwendung von realen Bilddaten in einem kalibrierten Fall	8
2.1	C ist die Kamera beziehungsweise das Projektionszentrum der Kamera. In rot ist die Strecke ζ markiert. M ist der Dreidimensionale Objektpunkt in Weltkoordinaten und m sein projizierter Punkt auf der 2D-Bildebene	10
2.2	Lochkameramodell: C ist das Kamerazentrum	11
2.3	Lochkameramodell [9]	11
2.4	Steroskopischer Aufbau mit Lochkameramodellen. Der Punkt M im Raum kann mit den jeweiligen Projektionsmatrizen P und P' der Kameras C und C' in die Bildpunkte m und m' projiziert werden. $m = PM$ und $m' = P'M$. Eine Rückprojektion der Punkte m und m' zum Punkt M ist mit bekannten Projektionsmatrizen genauso möglich.	12
2.5	In blau ist die Bildebene dargestellt auf ihr befinden sich die Punkte Q und HP . HP steht in dieser Abbildung für den Hauptpunkt. Z liegt nicht auf der Ebene, Das Projektionszentrum C liegt hinter der Bildebene und somit auch hinter dem Sensor. n ist die Normale der Bildebene	13
2.6	ζ_x und ζ_y beschreiben den Abstand des Kamerazentrums zu den jeweiligen Pixelkanten	14
3.1	Weltkoordinatensystem (O, δ) mit $\delta = (\hat{d}_1, \hat{d}_2, \hat{d}_3)$ und Kamerakoordinatensystem (C, β) mit $\beta = (\hat{b}_1, \hat{b}_2, \hat{b}_3)$	15
3.2	Koordinatensysteme von einer Kamera aus im Überblick. Das Weltkoordinatensystem (O, δ) ist Deckungsgleich mit dem Kamerakoordinatensystem (C, β)	18
4.1	Die Abbildung zeigt zwei Aufnahmen einer um ihr Projektionszentrum C gedrehten Kamera. Die zwei blauen Ebenen stellen jeweils die Bildebene vor und nach der Drehung dar sie wird einmal als Bildebene (I, τ) und einmal als Bildebene (I', τ') bezeichnet. Die Objektebene befindet sich im 3D-Raum, dementsprechend sind die Punkte $M_{1,\delta}$ und $M_{2,\delta}$ in homogenen 3D-Koordinaten gegeben. Die Punkte $m_{1,\tau}$ und $m_{2,\tau}$ sind die projizierten Punkte auf der Bildebene (I, τ) und Die Punkte $m_{1,\tau'}$ und $m_{2,\tau'}$ sind die projizierten Punkte auf der Bildebene (I', τ')	23
4.2	Ausgangssituation der Simulation	25
4.3	Drehung um das Projektionszentrum	25
4.4	Drehung um einen Drehpunkt. In diesem Beispiel wurde der rote Punkt als Drehpunkt verwendet	26
4.5	Strahlengang durch das Projektionszentrum. Auf der Abbildung ist erkennbar, dass der grüne Punkt auch nach der Drehung der Kamera um das Projektionszentrum vom roten Punkt verdeckt bleibt	26
4.6	Strahlengang durch das Projektionszentrum. Auf der Grafik ist erkennbar, dass der grüne Punkt nach der Drehung der Kamera um einen Drehpunkt, welcher in diesem Fall der rote Punkt darstellt, sichtbar wird.	26
4.7	Veranschaulichung der Homographie bei zwei verschieden translatierten und rotierten Kameras.	27

5.1	C und C' sind die Projektionszentren zweier Kameras. Beide Kameras besitzen jeweils eine Bildebene. Die Basislinien verbindet die Projektionszentren der Kameras. Der Punkt an welchem die Basislinie die Bildebenen schneidet, wird als Epipol bezeichnet. Durch den Epipol verlaufen alle Epipolarlinien des Bildes. M ist der Objektpunkt im 3D-Raum und m_1 und m_2 sind die jeweiligen Abbildungen dieses Punktes auf den Bildebenen. Die Verbindungsvektoren zwischen C, C' und M bilden die sogenannte Epipolarebene[5, 3, 1].	31
5.2	Die Objektpunkte M_1, M_2 und M_3 werden in I' als m'_1, m'_2 und m'_3 abgebildet, während sie in I immer den selben Bildpunkt m_1 ergeben.	32
6.1	vereinfachte Top-Down-Ansicht des Szenenaufbaus des Minimalbeispiels	36
6.2	In Grün ist die Abbildung auf der Bildebenen I von C und in rot ist die Abbildung auf der Bildebenen I' von C'	36
6.3	In Blau und Rot sind jeweils das Welt- und Kamerakoordinatensystem von Kamera eins zu sehen. In grün ist das gedrehte Koordinatensystem von Kamera 2 zu sehen.	36
6.4	Grün zeigt den Quader welcher auf I von C abgebildet wird. Das größere Quadrat sind die vorderen Punkte A, B, C, D , das kleinere Quadrat sind die hinteren Punkte A', B', C', D' . Der Punkt E ist weiter weg von den Abbildungen und deshalb auf dieser Abbildung momentan nicht zu sehen. Rot zeigt denselben Quader auf I' von C' abgebildet.	39
6.5	Die blauen Geraden zeigen die jeweiligen Epipolgeraden. Die vom roten Quader schneiden sich bei -1 im Epipol e' . Die Epipolgeraden vom grünen Quader schneiden sich im Epipol e' im Unendlichen, weshalb die Epipolarlinien Parallel zueinander verlaufen.	42
6.6	Anordnung der Kameras bei den vier verschiedenen Lösungen für P	46
6.7	Optimale Triangulierung: Beide Geraden Treffen sich in einem Punkt im 3D-Raum	46
6.8	Durch Ungenauigkeiten in der korrespondierenden Punkte, verfehlten sich die Linien und es kommt zu keinem Schnittpunkt	47
6.9	Veranschaulichung der Skaleninvarianz und dessen Auswirkung auf die geometrische Form der Objekte	48
6.10	auf Originalgröße skalierte rekonstruierte Szene	49
6.11	Durch Ungenauigkeiten in der korrespondierenden Punkte, verfehlten sich die Linien und es kommt zu keinem Schnittpunkt	49
6.12	Beispiel eines rektifizierten Bildes. Quelle: [20]	50
6.13	Beispiel einer einfachen Tiefenkarte eines Stereobildpaars nach der Rektifizierung. Quelle: [21]	50
6.14	Aufnahmen zweier Kameras mit den selben Auflösungen, Kamera eins(Grün) und Kamera(rot) zwei gelten jeweils $\zeta = 1$	57
6.15	Epipole für Kamera eins und Kamera zwei vor der Rektifizierung	57
6.16	Abbildung beider Bilder nach anwenden der Matrizen H_p und H'_p . Die Epipole beider Bilder sind nun im unendlichen. Das die entstehenden parallelen Epipolarlinien auch hier schon horizontal ausgerichtet sind ist Zufall. Die Epipolarlinien sind immer parallel nach dieser Transformation aber die Richtung ist nicht immer automatisch bereits $i = [1, 0, 0]$	58
6.17	Abbildung beider Bilder nach anwenden der Matrizen $H_r \cdot H_p$ und $H'_r \cdot H'_p$. Die Epipolarlinien sind nun horizontal zueinander ausgerichtet	60
6.18	Die Abbildung verdeutlicht noch mal Schematisch, wie sich die Punkte ausrichten sollen. Bild a) zeigt die durch die Rektifizierung verschobenen Bildkantenmitten. Bild b= zeigt, wie sich die Bildkantenmitten durch die Scherungstransformation wieder ausrichten sollen.	61
6.19	Abbildung beider Bilder nach anwenden der Matrizen $H_s \cdot H_r \cdot H_p$ und $H'_s \cdot H'_r \cdot H'_p$. Die horizontale Verzerrung wurde reduziert.	62
6.20	In dieser Abbildung wurden die Epipolarlinien noch in den Grafikplot mit eingebaut .	63
7.1	Rechteckiger Bildsensor mit darauf sich befindenden quadratischen Pixeln. Vlg [10]	65

7.2	Bild a) zeigt die Interpolation von Pixeln, wenn bei gleichbleibenden Seitenverältnissen weniger Pixel für das Bild verwendet werden sollen. Die interpolierten Pixel leiten dann alle das selbe Signal weiter. Bild b) zeigt in gelb markiert, den verweneten Bereich des Sensors, wenn sich die Seitenverältnisse ändern und nicht mehr der volle Sensor genutzt wird.	66
7.3	C und C' haben die selbe Auflösung eingestellt	67
7.4	C und C' haben unterschiedliche Auflösungen eingestellt	67
7.5	C und C' haben die selbe Auflösung eingestellt	68
7.6	C und C' haben unterschiedliche Auflösungen eingestellt	68
7.7	C und C' haben die selbe Auflösung eingestellt	68
7.8	C und C' haben unterschiedliche Auflösungen eingestellt	68
7.9	Die rekonstruierten Szenenpunkte und Kamerapositionen bleibt auch bei unterschiedlichen Auflösungen die selben	70
8.1	Aufnahme der Canon 6D von links	71
8.2	Aufnahme der Canon 60D von rechts	71
8.3	Szenenaufbau: Die Canon 60D befindet sich in dieser Abbildung auf der linken Seite, die Canon 60 D auf der rechten. Auf dem Tisch zwischen den Kameras ist die in den Abbildungen 6.1 und 6.1 abgebildete Szene zu sehen. Die Canon 60D ist etwas hinter der Canon 6D positioniert. Beide Kameras sind zu Szene hin gedreht und auch leicht nach unten geneigt.	72
8.4	Die mit dem <i>SURF</i> -Algorithmus gefundenen Punkte sind mit den gelben Ziffern im Bild gekennzeichnet	73
8.5	Epipolarlinien ohne <i>Epipolar-constraint</i> im Bild der Canon 6D	75
8.6	Epipolarlinien ohne <i>Epipolar-constraint</i> im Bild der Canon 60D	75
8.7	Epipolarlinien mit <i>Epipolar-constraint</i> im Bild der Canon 6D	76
8.8	Epipolarlinien ohne <i>Epipolar-constraint</i> im Bild der Canon 60D	76
8.9	Epipolarlinien mit <i>Epipolar-constraint</i> im Bild der Canon 6D, nach der denormalisierung von F	76
8.10	Epipolarlinien ohne <i>Epipolar-constraint</i> im Bild der Canon 60D, nach der denormalisierung von F	76
8.11	a)	77
8.12	b)	77
8.13	a) Die rückprojizierten Strahlen der ungenauen korrespondierenden Punkte m und m' sind schief und treffen sich nicht in einem Punkt im 3D-Raum. b) The epipolar geometry for m , m' . The measured points do not satisfy the epipolar constraint. The epipolar line $l' = Fm$ is the image of the ray through n , and $l = F^T m'$ is the image of the ray through m' . Since the rays do not intersect, m' does not lie on l' , and m does not lie on l	77
8.14	Frafische Darstellung der optimalen Punkte \hat{m} und \hat{m}'	78
8.15	Rekonstruierte Szene, unskaliert in Pixeleinheiten	83
8.16	Rekonstruierte Szene, unskaliert, in Pixeleinheiten und in einem 2D-Plot geschrieben	83
9.1	Zeigt die Die rekonstruierte Matrix R' bei unveränderter Auflösung. Die Auflösungen von C_δ und C'_δ sind die selben.	85
9.2	Zeigt die rekonstruierte Matrix R' wenn K' mit einem Verhältnis von $[5 : 2]$ skaliert wurde	85
9.3	Zeigt die rekonstruierte Matrix R' wenn K' mit einem Verhältnis von $[1 : 2]$ skaliert wurde	85
9.4	Zeigt die rekonstruierte Matrix R' wenn K' mit einem Verhältnis von $[1.2 : 2.3]$ skaliert wurde	85
9.5	Rekonstruierte Szene, wenn K' mit einem Verhältnis von $[5 : 2]$ skaliert wurde	86
9.6	Rekonstruierte Szene, wenn K' mit einem Verhältnis von $[2 : 1]$ skaliert wurde	86
9.7	Rekonstruierte Szene, wenn K' mit einem Verhältnis von $[1 : 2]$ skaliert wurde	87
9.8	Rekonstruierte Szene, wenn K' mit einem Verhältnis von $[1.2 : 2.3]$ skaliert wurde	87

10.1 Klassendiagramm	89
10.2 Klassendiagramm	89
10.3 Klassendiagramm	90
10.4 Klassendiagramm	90
10.5 Klassendiagramm	90
10.6	91
10.7	91
10.8	91
10.9	91
10.10Klassendiagramm	92
10.11Klassendiagramm	92
10.12Klassendiagramm	93
10.13Bild eines Tonnenförmig verzeichneten Schachbretts	93
10.14Algorithmisch detektierte Linie der dritten i-Reihe	93
10.15Bild eines perspektivisch verzerrtem Schachbretts	94
10.16Algorithmisch detektierte Linie der dritten i-Reihe	94
10.17Bild eines Kissenförmig verzeichnetem Schachbretts	94
10.18Algorithmisch detektierte Linie der dritten i-Reihe	94
10.19Bild eines Tonnenförmig verzeichnetem leicht perspektivisch verzerrtem Schachbretts(GRFIK AUSTAUSCHEN BILD IS KACKE)	95
10.20Algorithmisch detektierte Linie der dritten i-Reihe	95
10.21Bild eines Tonnenförmig verzeichnetem leicht perspektivisch verzerrtem Schachbretts	95
10.22Algorithmisch detektierte Linie der dritten i-Reihe	95
11.1 Weltkoordinatensystem (O, δ) mit $\delta = [\vec{d}_1, \vec{d}_1, \vec{d}_1, O]$ und Kamerakoordinatensysteme $\beta = [\vec{b}_1, \vec{b}_2, \vec{b}_3, C]$ und (C', β') mit $\beta' = [\vec{b}'_1, \vec{b}'_2, \vec{b}'_3, C']$	97
11.2 in blau ist die Abbildung des Quaders von Kamera eins und in rot die Abbildung des selben Quaders in Kamera zwei	101
11.3 Weltkoordinatensystem (O, δ) mit $\delta = [\vec{d}_1, \vec{d}_1, \vec{d}_1, O]$ und Kamerakoordinatensysteme $\beta = [\vec{b}_1, \vec{b}_2, \vec{b}_3, C]$ und (C', β') mit $\beta' = [\vec{b}'_1, \vec{b}'_2, \vec{b}'_3, C']$	104

Tabellenverzeichnis

7.1	Auflösungen Canon EOS 6D	65
7.2	Vgl [31]	65

Literaturverzeichnis

- [1] Zhengyou Zhang Gang Xu. *Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach*. Springer-Science and Business Media, 1996.
- [2] Lutz Pries. *Computer Vision, Einführung in die Verarbeitung und Analyse digitaler Bilder*. Springer-Verlag Berlin Heidelberg, 2014.
- [3] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in computer vision*. Cambridge, 2004, Second Edition.
- [4] Ferid Bajrmovic. *Self- Calibration of Multi- Camera Systems*. Logos Verlag Berlin GmbH, 2010.
- [5] Tomas Pajdla. *Elements of Geometry for Computer Vision*. "<http://people.ciirc.cvut.cz/pajdla/>", 2013, überarbeitet am 27.2.2017.
- [6] MathWorks. Mathworks documentation, stereo camera calibration app. "<https://de.mathworks.com/help/vision/stereo-camera-calibration.html>".
- [7] Ramalingam Tardif S.Gasparini J.Barreto R.Sturm, S. Camera models and fundamental concepts used in geometric computer vision. Mitsubishi Electric Research Laboratories, 2011.
- [8] Christian Heipke. *Photogrammetrie und Fernerkundung*. 2017 Springer, 1. Auflage.
- [9] Jianzhong Lu. *Ein dreidimensionales Bildverarbeitungssystem für die Automatisierung visueller Prüfvorgänge*.
- [10] Rolf Martin Ekbert Hering. *Photonik, Grundlagen, Technologien und Anwendung*. Springer-Verlag Berlin Heidelberg, 2006.
- [11] Zhengyou Zhang. *Epipolar Geometry*, pages 247–258. Springer US, Boston, MA, 2014.
- [12] Dipl.-Ing. Martin Roser. *Modellbasierte und positionsgenaue Erkennung von Regentropfen in Bildfolgen zur Verbesserung von viedeoisierten Fahrerassistenzfunktionen*. 1986, 1994 Springer Basel AG, KIT Scientific Publishing.
- [13] Jeanne Peiffer and Amy Dahan-Dalmedico. *Wege und Irrwege - Eine Geschichte der Mathematik*. 1986, 1994 Springer Basel AG, Editions du Seuil, Springer Basel AG, aus dem französischen von Klaus Volkert.
- [14] Norbert Köckler Hans Rudolf Schwarz. *Numerische Mathematik*. 2011, Springer Verlag, 8. Auflage.
- [15] Daniel Scholz. *Numerik interaktiv, Grundlagen verstehen, Modelle erforschen und Verfahren anwenden mit taramath*. 2016, Springer Verlag.
- [16] Zhengyou Zhang. Determining the epipolar geometry and its uncertainty: A review. Received July 16, 1996; Accepted February 13, 1997.
- [17] Richard I. Hartley. In defence of the 8-point algorithm. GE-Corporate Research and Development, Schenectady, NY, 12309.
- [18] Cui Guodong Yu Ming Zhang Mandun, Qi Lichao. A triangulation method in 3d reconstruction from image sequences. College of Computer Science and Software, Hebei University of Technology Tianjin, China, 2009, Second International Conference on Intelligent Networks and Intelligent Systems.

- [19] MathWorks. Mathworks documentation, rectify stereo images. "<https://de.mathworks.com/help/vision/ref/rectifystereoimages.html>".
- [20] Charles Loop and Zhengyou Zhang. Computing rectifying homographies for stereo vision. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Vol.1, pages 125–131, June 23-25, 1999. Fort Collins, Colorado, USA, 1999 Errors corrected on June 6, 2001.
- [21] Carlos VILLAGRÁ ARNEDOr Antonio Javier GALLEGÓ SÁNCHEZ, Rafael MOLINA CARMONA. *Scene reconstruction and geometrical rectification from stereo images*. Januar 2005,uploaded by Antonio Javier Gallego Sánchez on 21 May 2014, ResearchGate.
- [22] Luca Irsara Andrea Fusiello. Euclidean epipolar rectification of uncalibrated images. Eurac researc, IT.
- [23] Richard I. Hartley. Euclidean reconstruction from uncalibrated views. G.E. CRD, Schenectady, NY, 12301.
- [24] Andrea Fusiello. Euclidean epipolar rectification. "http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/FUSIELLO2/rectif_cv01.html".
- [25] Dongqing Li, editor. *Encyclopedia of Microfluidics and Nanofluidics*, pages 999–999. Springer US, Boston, MA, 2008.
- [26] S. Margulies. Fitting experimental data using the method of least squares. Department of Physics, University of Illinois at Chicago Circle, Chicago, Illinois 60680, 1967.
- [27] William T. Vetterling Brian P. Flannery William H. Press, Saul A. Teukolsky. *Numerical Recipes in Fortran 77 The Art Of Scientific Computing*. Copyright Numerical Recipes Software 1986, 1992, 1997 All Rights Reserved., Reprinted with corrections 1997, Volume 1 of Fortran Numerical Recipes.
- [28] James Demmel Dinesh Manocha. Algorithms for interesting parametric and algebraic curves 1: Simple intersections. ACM Transactions of Graphics:73–100, 1994.
- [29] MathWorks. Mathworks documentation,estimate camera parameters. "<https://de.mathworks.com/help/vision/ref/estimatecameraparameters.html>".
- [30] MathWorks. Mathworks documentation, disparity. "<https://de.mathworks.com/help/vision/ref/disparity.html>".
- [31] Canon. Eos 6d, eos 6d (wg), eos 6d (n), instruction manual, pages 188. "http://gdlp01.c-wss.com/gds/6/0300009626/01/EOS_6D_Instruction_Manual_DE.pdf".
- [32] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In Aleš Leonardis, Horst Bischof, and Axel Pinz, editors, *Computer Vision – ECCV 2006*, pages 404–417, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.