

3D-Rekonstruktion mit Kamera und Spiegel

Bachelor-Thesis
31. März 2014

JULIAN STIER



Prof. Dr. Tomas Sauer

Abstract

Schnellere Computer, bessere Optimierungsalgorithmen und über 20 Jahre Forschung im Bereich der Kamerakalibrierung ermöglichen heute immer mehr Anwendungsmöglichkeiten für Aufnahmen von Kameras. Dreidimensionale Rekonstruktionen finden immer mehr Verwendung in Bereichen der Wirtschaft (Automobilbranche, Automatisierungstechnik), Unterhaltung (Augmented Reality, Filmtechnik) oder Kriminologie. Viele dieser Anwendungen benötigen hohe Präzision und damit einhergehend teure Ausrüstung: hochauflösende Kameras, störungsfreie Stabilisatoren oder robuste Materialien.

In dieser Bachelor-Arbeit wurde untersucht, ob es mit nur einer Kamera und einem Spiegel möglich ist, eine 3D-Rekonstruktion vorzunehmen. Dies wurde erreicht, indem die Installation einen zweiten Blickwinkel über das Spiegelbild auf ein zu beobachtendes Objekt anbietet.

Inhaltsverzeichnis

1. Einleitung	5
1.1. Motivation	5
1.2. Zielsetzung	5
2. Szenarien - Möglichkeiten & Analyse	6
2.1. Ausgangssituation	7
2.2. Entfernung des Objekts	8
2.3. Winkel des Spiegels	8
2.4. Parameter im dreidimensionalen Szenenmodell	9
2.5. Fazit zum Zwischenergebnis	10
3. Einführung in die Kamerakalibrierung	11
3.1. Die Kameramodellierung	11
3.1.1. Das Lochkameramodell	11
3.1.2. Die Lochkameraabbildung	12
3.1.3. Das Lochkameramodell in homogenen Koordinaten	14
3.1.4. Verzeichnungen	14
3.2. Kamerakalibrierung	15
3.2.1. Bestimmung der Kollinearität	16
3.2.2. Startwert für die Kalibrierung	16
4. Stereorekonstruktion	17
4.1. Stereorekonstruktion durch korrespondierende Punkte	18
4.1.1. Startwert für die Stereorekonstruktion	18
5. Spiegelkamera	19
5.1. Ermittlung der Spiegelposition	19
5.1.1. Spiegelmaße	20
5.2. Die Spiegelungsmatrix	21
6. Evaluation	23
6.1. Szenenaufbau	24
6.2. Kalibrierung der Kamera	25
6.3. Berechnung der Transformationen	25
6.4. Triangulierung	27
6.5. Messergebnisse & Bewertung	28
6.6. Aufwand & Anwendbarkeit	31
7. Fazit & Ausblick	34
A. Literatur	35
B. Abbildungs-, Tabellen- und Listenverzeichnis	36
C. Eidesstattliche Erklärung	37

1. Einleitung

1.1. Motivation

Mit *High Precision Camera Calibration* [Han11] existiert eine ausführliche Grundlage zu metrischer Stereorekonstruktion aus hochauflösenden Bildern. Neben der Kameramodellierung und -kalibrierung beschreibt das Werk eine Methode zur präzisen Triangulierung von Weltpunkten aus zwei gegebenen Bildpunkten mit Hilfe einer linearen Startwertberechnung und einer nicht-linearen Fehlerfunktion.

Solche Modelle mit zwei oder mehreren Kameras nennen sich Stereo- oder Multikamerasysteme. Diese ermöglichen durch ihre relativen Positionen zueinander eine dreidimensionale Rekonstruktion einer beobachteten Szene. Autonome Systeme können dadurch beispielsweise ihre Lage im Raum bestimmen und sich während einer Bewegung anhand der gewonnenen Informationen selbst kontrollieren und ggf. neu ausrichten.

Neben bewegten Systemen gibt es aber auch Anwendungen für statische Installationen: beispielsweise die Überwachung von Räumlichkeiten oder Fabrikbändern in der Fertigung. In diesen Fällen sind Stereo- oder Multikamerasysteme häufig vergleichsweise teuer, da sie entweder Kameras in großen Stückzahlen benötigen oder der Nutzen einer metrischen 3D-Rekonstruktion unter Berücksichtigung der Kosten gering ausfällt.

Um dennoch eine metrische 3D-Rekonstruktion vornehmen zu können, erscheint die Idee eines Spiegels als Ersatz für eine zweite Kamera einleuchtend. Unter Verlust von Präzision könnten Kosten beinahe halbiert und damit neue Anwendungsfälle an Relevanz gewonnen werden. Solche Fälle sind beispielsweise Stereosystemanwendungen mit sehr eingeschränkten technischen Möglichkeiten oder mit einem Bedarf an sehr vielen Installationen - wie in der Fertigung oder bei der Überwachung von Objekten.

Der Spiegel böte in einem solchen Fall durch die Reflexion einen zweiten Blickwinkel. Über eine einzelne Kamera entstünden somit zwei Bilder derselben Szene. Die Kalibrierung einer zweiten Kamera entfiere und es wäre lediglich wie bei einem üblichen Stereokamerasystem die relative Position einer (virtuellen) Kamera über den Spiegel zu berechnen.

1.2. Zielsetzung

Im nachfolgenden Kapitel wird sich zeigen, dass ein Spiegel einen zweiten, nahen Blickwinkel auf eine Szene bieten kann: ermittelt man die Position des Spiegels und darüber die Spiegelung eines beobachteten Punktes, erhält man zwei korrespondierende Bildpunkte, die denselben Weltpunkt repräsentieren. Ein Stereokamerasystem ist gegeben.

Diese Arbeit untersucht, ob die theoretische Überlegung eines zweiten Blickwinkels über einen Spiegel zu einer 3D-Rekonstruktion führen kann, betrachtet verschiedene Szenarien und Vorgehensweisen und evaluiert die Güte der Ergebnisse.

Es wird sich im Verlauf der Arbeit herausstellen, dass eine metrische 3D-Rekonstruktion mit hinnehmbarer Varianz möglich ist und welche Schwierigkeiten dabei zu berücksichtigen sind.

2. Szenarien - Möglichkeiten & Analyse

Ein Spiegel und ein zu beobachtendes Objekt lassen sich - insbesondere im Modell - in Abhängigkeit zu einer (Loch-)Kamera vielfältig positionieren. Da viele Konfigurationen in der Realität unpraktikabel oder für ein brauchbares Ergebnis von vornherein auszuschließen sind, gibt dieses Kapitel einen Einblick über mögliche Szenarien einer Kamera-Spiegel-Installation.

Da in der Fachliteratur keine Maße zur Beschreibung des Beobachtungsraums für das Lochkameramodell etabliert sind, werden im Nachfolgenden die Vorgehensweise und Ergebnisse der empirischen Erfassung eines Kamera-Spiegel-Szenarios beschrieben.

Im Verlauf der Arbeit werden verschiedene *Geogebra*¹-Modelle im zwei- und dreidimensionalen² Koordinatensystem betrachtet, um den Zusammenhang der verschiedenen Positionierungsmöglichkeiten zu veranschaulichen. Grundsätzlich wurde o.B.d.A. nur eine Seite der Kamera für die Positionierung des Spiegels betrachtet, da sich das Modell symmetrisch verhält. Die Geogebra-Modelle betrachten eine einfache Lochkamera, die in Kapitel 3 zur Einführung in das Lochkameramodell näher erläutert wird. Damit wird die Kamera als Ursprungspunkt im Koordinatensystem betrachtet. Die x -Achse gibt die Blickrichtung, also die sog. Hauptachse der Kamera vor. Im Lochkameramodell wird die Hauptachse der Kamera als z -Achse festgelegt. Um aber in den Geogebra-Modellen bei den theoretischen Untersuchungen der zwei- und dreidimensionalen Modelle konsistent bleiben zu können, werden die Achsen anders gewählt:

Während das zweidimensionale Modell eine Projektion von oben auf eine Szene beschreibt, wird das dreidimensionale Modell um eine z -Achse in die Höhe erweitert und beschreibt damit einen Raum. Die x - und y -Achse beschreiben dann weiterhin die Tiefe und Breite. Entsprechend stellt im 2D-Modell die Gerade durch $x = 1$ die später eingeführte $\{z = 1\}$ -Ebene des Lochkameramodells dar. Die Hauptebene der Kamera wird im 2D-Modell von Geogebra mit $x = 0$ repräsentiert und fällt im 3D-Modell mit der yz -Ebene ($\{(x, y, z) \in \mathbb{R}^3 \mid x = 0\}$) überein, während sie im Lochkameramodell mit der xy -Ebene identifiziert wird.

In den Szenarien wurden folgende Parameter betrachtet:

Im zweidimensionalen Szenenmodell:

- Die **Spiegelposition** (in Abbildung 1 als S_1 gekennzeichnet) legt einen Punkt im Modell fest. Dieser ist vor und zu einer Seite versetzt vom Kamerazentrum und beschreibt mit Hilfe der Spiegelmaße (Breite und im dreidimensionalen Modell auch Höhe) die Spiegelfläche. Die Position wird durch den **Spiegelabstand** (orthogonal zur Blickrichtung) und die **Spiegelentfernung** (entlang der Blickrichtung) festgelegt.
- Die **Spiegelbreite** w_m gibt an, wie weit sich die Spiegelfläche von einem Positionierungspunkt aus in Blickrichtung unter Beachtung von Neigung und Rotation ausdehnt.
- Der **Spiegelwinkel** α_m legt die Rotation des Spiegels in Richtung der Blickfläche fest. Im Zweidimensionalen ist der Spiegelpunkt S_2 gegeben durch S_1 mit $S_2 = (x(S_1) + w_m \cdot \cos(\alpha_m), y(S_1) + w_m \cdot \sin(\alpha_m))$ ³.
- Die **Objektbreite** gibt eine feste Länge für die Ausdehnung des Objekts entlang der y -Achse vor. Durch sie wird die von der Kamera aus betrachtete linke und rechte Objekthecke

¹Geogebra ist ein freies Geometrie- und Algebrasystem. Zur Zeit befindet es sich in Version 4.4.12 (Feb. 2014). <http://geogebra.org>

²Dreidimensionale Modelle werden in Geogebra offiziell noch nicht unterstützt. Daher wurde neben der offiziellen Version die Beta-Version 4.9.117.0 eingesetzt.

³ $x(A)$ und $y(B)$ bezeichnen in Geogebra Funktionen, die die x - bzw. y -Koordinate des jeweiligen Punktes zurückliefern.

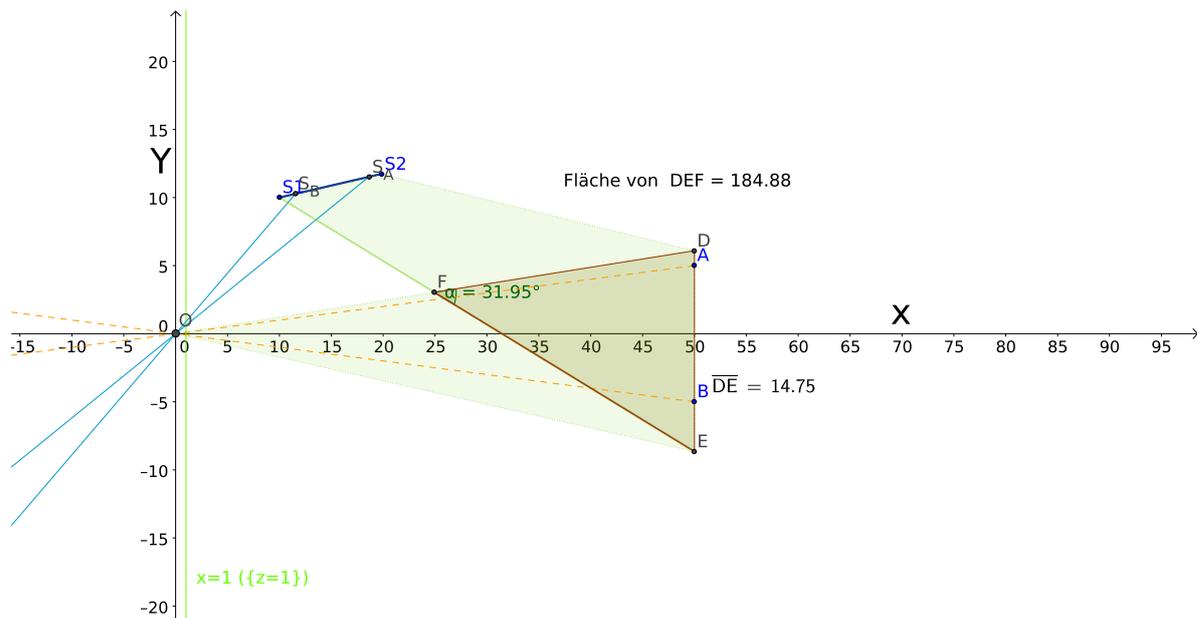
festgelegt. Über die Objektecken lassen sich vereinfacht die Strahlenverläufe über den Spiegel zur Lochkamera nachvollziehen.

- Genau wie der Spiegel lässt sich das Objekt über das **Objektzentrum** mit der **Objektverschiebung** entlang der y-Achse verschieben.
- Um das Verhalten bei unterschiedlicher Entfernung des Objekts zu untersuchen, lässt sich mit der **Objektentfernung** der Abstand des Objekts zur Lochkamera regeln.

Im dreidimensionalen Szenenmodell:

- Analog zur Spiegelbreite gibt die **Spiegelhöhe** im dreidimensionalen Szenenmodell die Länge des Spiegels entlang der z-Achse vor. Die Spiegelbreite definiert weiterhin eine Länge entlang der x- und y-Achse.
- Während man im 2D-Modell den Spiegel zur Kamera hin oder von der Kamera weg drehen kann, ist es im dreidimensionalen Raum auch möglich die **Spiegelneigung** zu betrachten.
- Ebenso wie der Spiegel ist auch das Objekt mit einer **Objekthöhe** parametrisierbar.

Abbildung 1: Projektion von oben auf eine Lochkamera (als Punkt), den Spiegel (als Strecke) und ein breites Objekt (als Strecke AB).



2.1. Ausgangssituation

Abbildung 1 dient den weiteren Erläuterungen als Ausgangssituation. Die Spiegelposition S_1 ist vom Kamerazentrum um zehn Einheiten zur linken Seite und nach vorne verschoben (in der Projektion jeweils durch Verschiebung entlang der y- und x-Achse dargestellt). Der Spiegel befindet sich in einem Winkel von zehn Grad zur optischen Achse und das betrachtete Objekt ist 50 Einheiten vom Kamerazentrum entfernt.

Während die Parameter für das Objekt den Abstand zweier Punkte A und B regeln, zeigt das Modell zugleich die maximale Breite (durch D und E begrenzt) eines Objekts auf der selben Distanz. Die Strecke \overline{DE} beträgt in der Ausgangssituation 14.75 Einheiten und beschreibt damit den gemeinsamen Sichtbereich (in der Breite) der Kamera und des Spiegels.

Die blau markierten Strahlen vom Spiegel ins Kamerazentrum (in Abbildung 1 als $\overline{OS_B}$ und $\overline{OS_A}$ gekennzeichnet) markieren die am Spiegel reflektierten Sichtstrahlen des beobachteten Objekts. Überschreitet das Objekt den gemeinsamen Sichtbereich, wird das Objekt aus Sicht der Kamera im Spiegel abgeschnitten. Der jeweilige Sichtstrahl wird nicht mehr über den Spiegel reflektiert (erkennbar z.B. in den Abbildungen 2 und 3).

Neben den reflektierten Sichtstrahlen des Objekts werden auch die Strahlenverläufe der Kamera und des Spiegels zum Objekt beobachtet. Aus dem Schnitt der beiden Flächen auf die gemeinsam beobachtbare Breite zu der vorgegebenen Objektentfernung und einer Geraden auf Höhe der Objektentfernung ergibt sich ein Dreieck, dessen Winkel α ein erstes Maß für die Weite der beobachtbaren Szene beschreibt.

Die Kamera muss über eine Brennweite verfügen, die es ermöglicht, sowohl den Spiegel als Ganzes, als auch das Objekt (zumindest zu gewünschten Teilen) zu erfassen. Dies führt bei praktischer Untersuchung zu weiteren Schwierigkeiten (z.B. mit Schärfentiefe und Belichtung), die in der Evaluation erörtert werden und nicht mehr Teil des Geogebra-Modells sind.

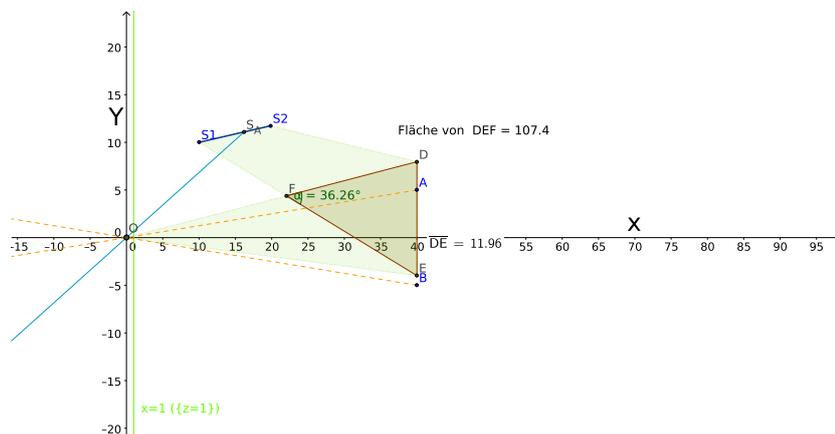


Abbildung 2: Objekt mit Abstand 40E bei einem Spiegelwinkel von 10°

2.2. Entfernung des Objekts

In Abbildung 2 wird der Abstand des Objekts auf 40 Einheiten verringert. Der Winkel α erhöht sich dabei von 31.95° auf 36.26° , während die Strecke \overline{DE} auf die Länge 11.96 Einheiten abnimmt.

Außerdem fällt auf, dass sich der beobachtbare Raum entlang der y -Achse verschoben hat. Daraus folgt, dass dynamische Objekte sehr viel schwieriger als mit üblichen Stereokamerasystemen zu beobachten sind, weil sie durch den deutlich eingeschränkten Sichtbereich zusätzlich zum kleineren Bildbereich auch leichter bei Bewegung aus diesem verschwinden können.

2.3. Winkel des Spiegels

Beobachtet man nun die reine Veränderung der Spiegelrotation (im Modell der Winkel des Spiegels zur Hauptachse), stellt man ebenfalls eine Veränderung des gemeinsamen Sichtbereichs fest.

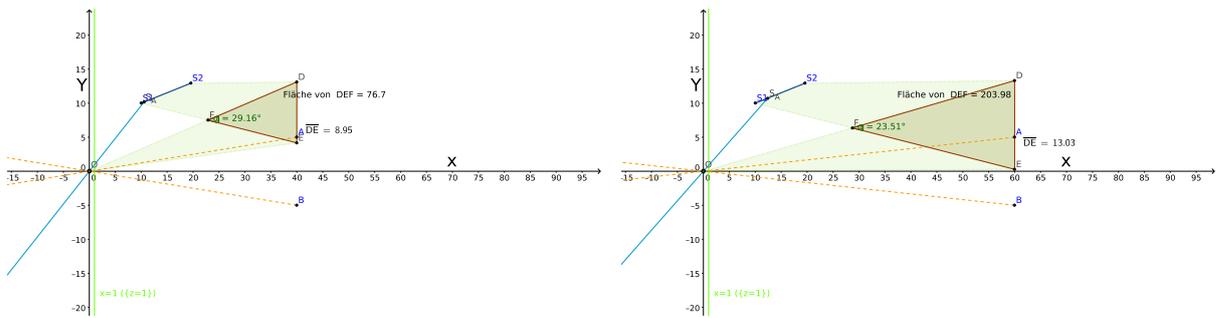


Abbildung 3: Veränderung des gemeinsamen Sichtbereichs bei gleichbleibendem Spiegelwinkel, aber veränderter Objektentfernung. Winkel 17° , Objektentfernung $40E$ und $60E$.

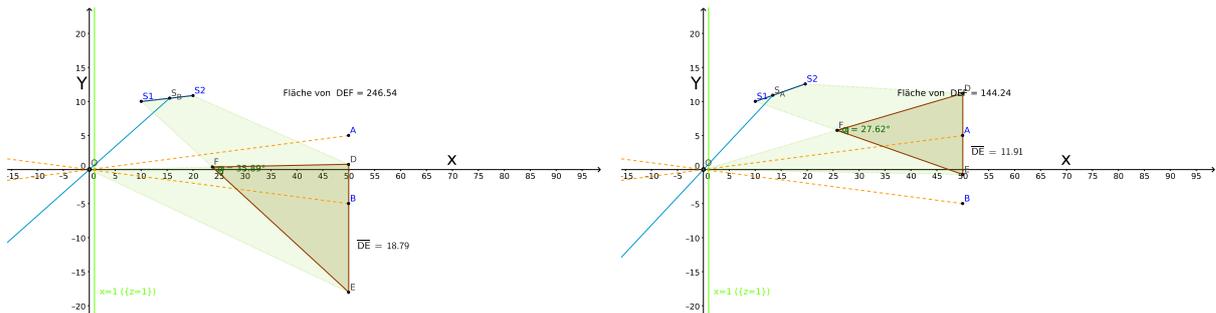


Abbildung 4: Veränderung des gemeinsamen Sichtbereichs bei verändertem Spiegelwinkel, aber gleichbleibender Objektentfernung. Objektentfernung $50E$, Spiegelwinkel 5° und 15° .

Wird der Spiegel von der Kamera weg gedreht, verringert sich der Winkel α des gemeinsamen Sichtbereiches. Umgekehrt vergrößert er sich zunächst, je weiter man ihn zur Kamera hin dreht. Dabei erreicht der Winkel bei starker Ausrichtung zur Kamera ein Extremum und nimmt danach wieder ab.

Es lässt sich erkennen, dass kleine Winkel im Bereich von 0° bis 30° von der Hauptachse der Kamera weg und zu ihr hin am geeignetsten für eine Installation sind.

2.4. Parameter im dreidimensionalen Szenenmodell

Erweitert man das Modell auf den dreidimensionalen Raum verändert sich die Aussagekraft nur geringfügig. Der gemeinsame Sichtbereich der Kamera wird neben der Breite des Spiegels zusätzlich durch die Höhe begrenzt - immer unter der Voraussetzung, dass der Spiegel vollständig im Bild der Kamera ist.

Der Spiegelwinkel im zweidimensionalen Modell entspricht nun der Rotation des Spiegels zu der Achse, die über die Spiegelposition parallel zur Hauptebene der Kamera definiert ist. Die Neigung erfolgt über die dazu orthogonale Achse ausgehend von der Spiegelposition. Während die Rotation den gemeinsamen Sichtbereich links oder rechts zur Kamera verschiebt, ändert die Neigung den Sichtbereich nach oben oder unten zur Hauptachse der Kamera.

Mit der Objektentfernung wird aus der Strecke für den gemeinsamen Sichtbereich im zweidimensionalen Modell eine Fläche, die sich durch Neigung, Rotation und Position des Spiegels beliebig positionieren lässt. Die Größe hängt ebenso wie im zweidimensionalen Modell wesentlich von Größe, Neigung und Rotation des Spiegels sowie der Entfernung des Objekts ab.

2.5. Fazit zum Zwischenergebnis

Es hat sich gezeigt, dass sich eine Kamera und ein Spiegel für eine 3D-Rekonstruktion trotz Einschränkungen vielfältig positionieren lassen. Zwar nimmt der Spiegel Beobachtungsraum der Kamera in Anspruch und ein Aufbau hängt stark von Abstand und Verschiebung des Objektes ab, dennoch lässt sich eine statische Beobachtung in der Regel sehr flexibel installieren, da bereits kleine Veränderungen in Neigung und Position des Spiegels starke Veränderungen des beobachtbaren Raums herbeiführen.

Entfernungen des beobachteten Objekts sind durch Anpassungen der Spiegelrotation und -neigung ausgleichbar. Da in Kapitel 5 festgestellt wird, dass die Berechnung einer sog. *Stereobasis* wesentlich von der Position des Spiegels abhängt, ist die Veränderung der relativen Spiegelposition zur Kamera aber mit hohem Aufwand (einer sog. *extrinsischen Kalibrierung*) verbunden.

Solange die Ermittlung der Spiegelposition nicht automatisiert vorgenommen werden kann, lässt sich der Schluss ziehen, dass dynamische Systeme ungeeignet für eine 3D-Rekonstruktion mit Kamera und Spiegel sind.

In der Evaluation in Kapitel 6 wird zudem eine weitere Einschränkung festgestellt, die durch dieses Modell nicht betrachtet wird: die Schärfentiefe beeinflusst zusätzlich die Installationsmöglichkeiten.

3. Einführung in die Kamerakalibrierung

Um aus einem Kamerabild Informationen zu gewinnen oder gar Rückschlüsse auf die reale Umgebung ziehen zu können, bedarf es dem Verständnis der Beziehung zwischen der Realität, der verwendeten Kamera und ihrer erzeugten Projektion auf der entstandenen Rastergrafik⁴.

Verschiedene **Kameramodelle** dienen dazu, diese Beziehungen mathematisch zu beschreiben. Dadurch wird es möglich, auftretende Probleme voneinander zu trennen und Lösungsansätze lokal zu untersuchen und berechenbar zu machen.

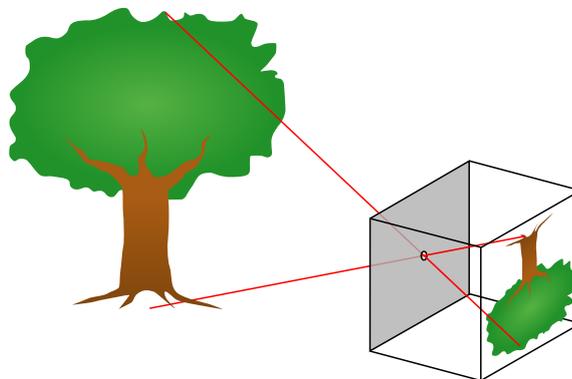
3.1. Die Kameramodellierung

3.1.1. Das Lochkameramodell

Eine **Lochkamera** ist eine sehr einfache Kamera, bestehend aus einer schwarzen, abgedunkelten Box und einem winzigen, verschließbaren Loch - eine sog. *camera obscura* ohne Linse. Durch das Loch führen die Lichtstrahlen eines Objektes nahezu direkt auf die Bildsensorebene⁵, ohne dass eine Sammellinse zur Bündelung benötigt wird. Das setzt voraus, dass das Loch möglichst klein ist, hat aber gleichzeitig den Effekt, dass die Lichtstärke gegenüber heute üblichen Kameras deutlich geringer ist.

Wird das Loch unendlich klein, wird die Abbildung unendlich scharf, aber die Lichtstärke gleichzeitig unendlich schwach. Man spricht dann von einer *idealen Lochkamera* und nutzt deren vereinfachte optische Eigenschaften zur Beschreibung des **Lochkameramodells**⁶.

Abbildung 5: Skizzierung einer Lochkamera. Quelle: de.wikipedia.org



Das Lochkameramodell betrachtet eine ideale Kamera mit dem unendlich kleinen Loch als Kamerazentrum. Wie in Abbildung 5 zu sehen ist, verlaufen alle Sichtstrahlen eines beobachteten Objekts durch dieses Kamerazentrum und projizieren so ein Bild auf die hintere Leinwand der Kamerabox. Diese hintere Wand wird als *Bildebene* (engl. *image plane*) bezeichnet.

Betrachtet man ein Koordinatensystem, das seinen Ursprung im Kamerazentrum hat und dessen xy -Ebene der Bildebene der Lochkamera gleicht, so verläuft die z -Achse entlang der optischen Achse. Diese bildet die Sichtrichtung der Lochkamera.

⁴Eine Rastergrafik (engl. *raster graphic*, *bitmap* oder *pixelmap*) ist eine Beschreibungsart für zweidimensionale Bilder auf dem Computer. Auf einem endlichen Feld wird Punkten ein Farbwert zugeordnet und so die Darstellung eines digitalen Farbbildes ermöglicht.

⁵Die Bildsensorebene ist ein Element zur Abbildung von Licht auf einer zweidimensionalen Ebene, das heute zumeist auf Halbleitertechnik basiert.

⁶Die vereinfachten Annahmen bzgl. Optik erster Ordnung und infinitesimal dünner Linse zur Modellierung des Lochkameramodells sind in [Han11] ausführlich beschrieben.

Während die tatsächliche Bildebene durch das Kamerazentrum - einer Punktspiegelung ähnlich - ein seitenverkehrtes, auf dem Kopf stehendes Bild einfängt, kann man im Modell ein ungespiegeltes Bild auf der vom Kamerazentrum aus gegenüberliegenden Seite beobachten, das rechnerisch den selben Ansprüchen genügt. Hierfür wird üblicherweise die Ebene $\{(x, y, z) \in \mathbb{R}^3 : z = 1\} =: \{z = 1\}$ definiert.

Zur Beschreibung des Lochkameramodells wird eine unendlich dünne Linse und die Optik 1. Ordnung⁷ angenommen. Es genügt drei Strahlenverläufe eines beobachteten Punktes zu betrachten; nämlich den Strahl parallel zur optischen Achse vor der Linse, den Mittelpunktstrahl durch das optische Zentrum und den Strahl parallel zur optischen Achse hinter der Linse. Alle drei Strahlen schneiden sich in der Fokalebene hinter der Linse in einem Punkt. Daher genügt es, nur den Mittelpunktstrahl zur Beschreibung des Lochkameramodells zu betrachten.

Im Nachfolgenden werden Abbildungen für einzelne Übergänge im Lochkameramodell beschrieben und schließlich zu einer Lochkameraabbildung zusammengefügt. Die Lochkameraabbildung beschreibt schließlich die Transformation eines Weltpunktes in einen Bildpunkt.

3.1.2. Die Lochkameraabbildung

Definition 1. Sei $P_z : \mathbb{R}^3 \setminus \{z = 0\} \rightarrow \{z = 1\}$ mit

$$P_z : (x, y, z) \mapsto \left(\frac{x}{z}, \frac{y}{z}, 1 \right)$$

Mit P_z ist eine Projektion von Punkten im dreidimensionalen Raum auf die $\{z = 1\}$ -Ebene gegeben. Durch die Projektion auf eine Ebene lässt sich die Abbildung nicht invertieren (die "Tiefeninformation" der z -Koordinate geht verloren) und das Urbild eines Bildpunktes (u, v) ist eine Gerade durch den Ursprung (das Kamerazentrum): $P_z^{-1}((u, v)) = \{\lambda (u, v, 1) \mid \lambda \in \mathbb{R}\}$ [Han11, 16].

In der *projektiven Geometrie*⁸ bilden alle Punkte auf dieser Geraden eine Äquivalenzklasse. Durch P_z wird im projektiven Sinne also ein Repräsentant seiner Äquivalenzklasse ausgewählt. Damit liegen alle Repräsentanten ihrer jeweiligen Äquivalenzklasse in $\{z = 1\}$.

Die wesentliche Skalierung und Verschiebung geschieht nun mit der Bildtransformation P , deren Parameter *intrinsische Kameraparameter*⁹ genannt werden.

Definition 2. Eine Abbildung $P : \{z = 1\} \rightarrow \mathbb{R}^2$ mit $A = \begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix}$, $\alpha, \beta \in \mathbb{R}_+^*$, $\gamma, u_0, v_0 \in \mathbb{R}$

$$P : (u, v, 1) \mapsto A \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$$

heißt Bildtransformation.

Die Verschiebung (u_0, v_0) orientiert sich am Hauptpunkt im Bildkoordinatensystem, der üblicherweise bei der Mitte der Rastergrafik liegt. Die Skalierungsparameter α und β werden mit Hilfe "der Brennweite f der Linse und den Ausmaßen $d_u \times d_v$ eines photosensitiven Elements

⁷Auch paraxiale Optik genannt. Durch den Abbruch der Taylorreihe für den Sinus nach dem ersten Glied, erhält man $\sin(x) \approx x$ für kleine Winkel [Ver06, S. 50f, S.98f].

⁸Eine übersichtliche Einführung in projektive Geometrie gibt [Kan93, S.15].

⁹Genau genommen werden auch die Parameter zur Beschreibung von P_z zu den intrinsischen Kameraparametern gezählt. Durch die Projektion ist dies aber von Punkt zu Punkt verschieden und ist in der Praxis lediglich die Division durch die z -Koordinate.

auf dem Sensor interpretiert als $\alpha = \frac{f}{d_u}$ und $\beta = \frac{f}{d_v}$ “ [Han11]. Der Parameter γ beschreibt eine zusätzlich mögliche Neigung des Sensors, falls dieser nicht senkrecht zur optischen Achse steht. Häufig wird ein hier entstehender Fehler durch tangentiale Verzeichnung (siehe Kapitel 3.1.4 zu Verzeichnungen) ausgeglichen.

Die Kamera wird im Modell üblicherweise in relativer Position zu einem bekannten Objekt beschrieben. Bei der Kamerakalibrierung werden beispielsweise eine Reihe von Kalibrierobjekten benötigt, um die intrinsischen Parameter zu bestimmen. In der Stereorekonstruktion wird die relative Position bezüglich einer weiteren Kamera benötigt, um darüber eine 3D-Rekonstruktion vorzunehmen. Zur Beschreibung dieser Transformation vom Kamerakoordinatensystem der betrachteten Kamera in ein Referenzkoordinatensystem werden sog. *extrinsische Parameter* definiert.

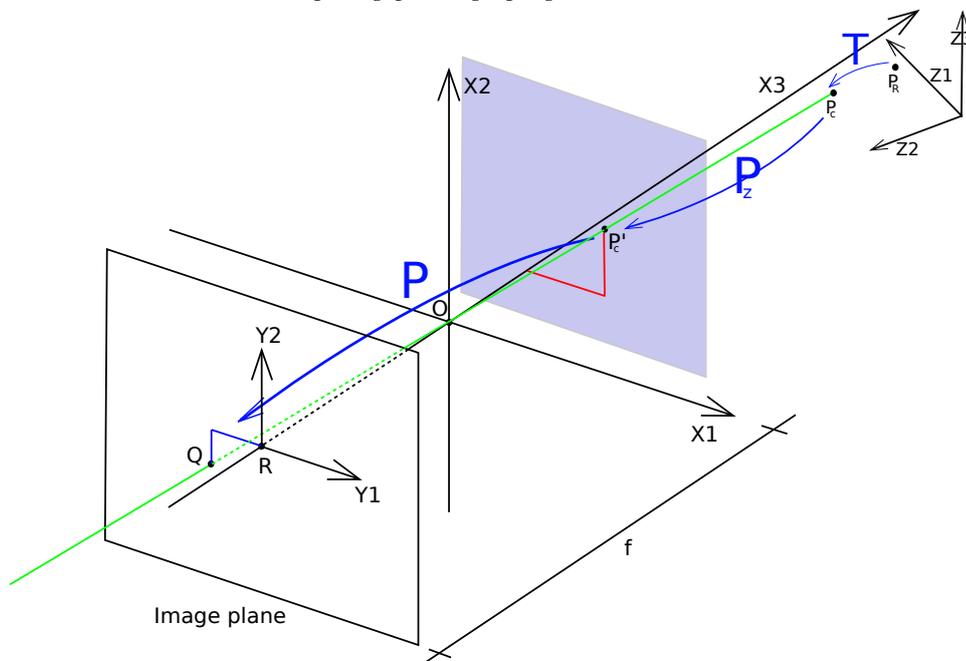
Definition 3. Sei R eine Rotation und $t \in \mathbb{R}^3$ ein Translationsvektor. Dann ist

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : p \mapsto Rp + t$$

die aus R und t zusammengesetzte (uneigentliche) Bewegung $(R \ t)$. Transformiert die Bewegung ein Referenzkoordinatensystem in das Kamerakoordinatensystem, werden die Koeffizienten von T *extrinsische Parameter* genannt.

Abbildung 6: Das Lochkameramodell

Über den Ursprung verläuft der Sichtstrahl des Weltpunktes P in die Bildebene auf Q . X_1, X_2 und X_3 legen das Kamerakoordinatensystem fest, Y_1 und Y_2 das Bildkoordinatensystem. Auf der gegenüberliegenden Seite des Bildes mit Abstand eins im Kamerakoordinatensystem lässt sich leicht die $\{z = 1\}$ -Ebene erkennen - und auch, dass das entstehende Bild dort nicht im Ursprungspunkt gespiegelt wird.



Die einzelnen Abbildungen lassen sich nun zu einer gemeinsamen Abbildung zusammenfügen:

Definition 4. Ist wie nach Definition 3 eine Bewegung T von einem Referenzkoordinatensystem in das Kamerakoordinatensystem gegeben und existiert eine Bildtransformation P , so heißt für $D := \mathbb{R}^3 \setminus T^{-1}(\{z = 0\})$

$$K : D \rightarrow \mathbb{R}^2, K = P \circ P_z \circ T$$

Kameraabbildung bzgl. des Lochkameramodells. Die Menge aller Kameraabbildungen wird als \mathbb{K} bezeichnet.

3.1.3. Das Lochkameramodell in homogenen Koordinaten

Für eine leichtere Lesbarkeit und geschickteren Umgang mit den Abbildungen hat es sich als nützlich erwiesen, das Lochkameramodell in homogenen Koordinaten [Kan93, S.15f] aus der *projektiven Geometrie* darzustellen.

Die projektive Geometrie kann als Erweiterung des \mathbb{R}^n oder \mathbb{C}^n betrachtet werden. Dabei ist durch $p, q \in \mathbb{C}^{n+1} p \sim q : \Leftrightarrow \exists \lambda \in \mathbb{R} \setminus \{0\} : p = \lambda q$ eine Äquivalenzrelation auf \mathbb{C}^{n+1} gegeben. Man spricht von der *projektiven Vervollständigung des \mathbb{C}^n* [Han11].

Für eine Äquivalenzklasse $[\lambda(x_1, \dots, x_n, x_{n+1})^t]$ mit $x_{n+1} \neq 0$ wird

$$\begin{pmatrix} \frac{x_1}{x_{n+1}} \\ \dots \\ \frac{x_n}{x_{n+1}} \\ 1 \end{pmatrix}$$

als ausgezeichneter Repräsentant ausgewählt (siehe auch Definition 1 zu P_z).

Die intrinsischen und extrinsischen Parameter lassen sich nun in homogenen Koordinaten darstellen und so zu einer projektiven Kamera zusammensetzen.

Sei $K \in \mathbb{K}$ eine Lochkameraabbildung mit $K = P \circ P_z \circ T$ und $T(p) = Rp + t$ sowie

$$P : \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}.$$

Dann seien

$$\begin{aligned} \tilde{T} &:= (Rt) \in \mathbb{R}^{3 \times 4} \\ \tilde{P} &:= \begin{pmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3 \times 3} \end{aligned}$$

und damit

$$\tilde{K} := \tilde{P}\tilde{T} \in \mathbb{R}^{3 \times 4}$$

die Lochkameraabbildung in homogenen Koordinaten.

3.1.4. Verzeichnungen

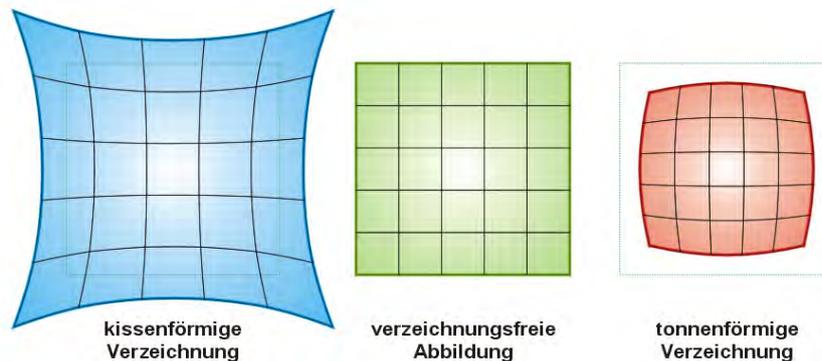
In der Realität hat man nicht die Idealbedingungen, die beim Lochkameramodell angenommen werden. Durch die Betrachtung einer dicken Linse und der Optik 3. Ordnung gelangt man zu

monochromatischen Aberrationen (d.h. optischen Abbildungsfehlern). Während von Seidel¹⁰ insgesamt fünf verschiedene monochromatische Aberrationen beschrieben wurden, reicht es üblicherweise aus, von diesen Aberrationen nur Verzeichnungen zu betrachten, da sie als einzige den Mittelpunktstrahl betreffen [Han11, 2.3.1].

Definition 5. Eine analytische Funktion $\delta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die einen Verzeichnungsfehler approximativ ausgleicht, heißt *Verzeichnungsfunktion*.

Abbildung 7: Verzeichnungen

Quelle: <http://de.wikipedia.org/wiki/Verzeichnung>



Während die Verzeichnungsfunktionen meist unendliche Reihen sind, werden in der Praxis nur endlich viele bzw. so wenig Parameter wie möglich genutzt (d.h. die Potenzreihe wird nach wenigen Entwicklungen abgebrochen).

Name	Potenzreihe
Radiale Verzeichnung $\delta_{rad}(u, v)_u, \delta_{rad}(u, v)_v$	$u \sum_{i=1}^{\infty} k_i (u^2 + v^2)^i$ $v \sum_{i=1}^{\infty} k_i (u^2 + v^2)^i$
Dezentrierungsverzeichnung $\delta_{dec}(u, v)_u, \delta_{dec}(u, v)_v$	$(p_1(3u^2 + v^2) + 2p_2uv) (1 + \sum_{i=3}^{\infty} p_i (u^2 + v^2)^{i-2})$ $(p_2(u^2 + 3v^2) + 2p_1uv) (1 + \sum_{i=3}^{\infty} p_i (u^2 + v^2)^{i-2})$
Prismaverzeichnung $\delta_{thp}(u, v)_u, \delta_{thp}(u, v)_v$	$s_1(u^2 + v^2) (1 + \sum_{i=3}^{\infty} s_i (u^2 + v^2)^{i-2})$ $s_2(u^2 + v^2) (1 + \sum_{i=3}^{\infty} s_i (u^2 + v^2)^{i-2})$

Tabelle 1: Beispiele für Verzeichnungsfunktionen

Die Verzeichnung wird nach der Projektion P_z in die $\{z = 1\}$ -Ebene und vor der Bildtransformation P angewandt. Mehrere Verzeichnungsfunktionen werden einfach aufsummiert. Um eine gegebene Verzeichnungsfunktion δ als Matrix darzustellen, wird $\tilde{\delta}$ definiert als $\tilde{\delta} : \{z = 1\} \rightarrow \{z = 1\}$, $\tilde{\delta} = \begin{pmatrix} \delta \\ 1 \end{pmatrix}$. Damit lässt sich die Kameraabbildung zur Kameraabbildung mit Verzeichnung erweitern: $K = P \circ P_z \circ \tilde{\delta} \circ T$.

3.2. Kamerakalibrierung

Die Kamerakalibrierung bezeichnet das Verfahren, konkrete Parameter für die Abbildungen des Lochkameramodells zu finden. Das findet in mehreren Schritten über Kalibrieremuster statt und

¹⁰Philipp Ludwig von Seidel (*1821, †1896)

beinhaltet mitunter mehrere nicht-lineare Optimierungsprobleme. Die nachfolgenden Erläuterungen orientieren sich an [Han11, 37ff] und geben einen groben Überblick über den Ablauf.

Zur Einzelkamerakalibrierung wird eine Fehlerfunktion mit mehreren¹¹ Targets minimiert. Dazu werden projektive Abbildungen für jedes Bild einer Kalibrierplatte benötigt. Diese werden als sog. Kollinearitäten bezeichnet (engl. *homography*). Aus diesen Abbildungen lässt sich ein überbestimmtes Gleichungssystem aufstellen, dessen nicht-triviale Lösung als Minimum linear gegeben ist [Han11, 3.2.7]. Daraus kann in der Regel die Projektionsmatrix als Startwert für die Fehlerfunktion berechnet werden.

Die Fehlerfunktion berechnet für jedes Target (Kalibrierplatte) den quadratischen¹² Abstand des jeweiligen Weltpunktes, der über die Kameraabbildung in das Bild projiziert wurde, und seiner Beobachtung im Bild.

Definition 6. Fehlerfunktion der Kalibrierung

Für m Bilder einer Kalibrierplatte $k \in \{1, \dots, m\}$ ergeben sich n_k Weltpunkte w_{k,c_k} mit ihrer Beobachtung i_{k,c_k} für $c_k \in \{1, \dots, n_k\}$. Durch

$$\Phi : (P, \delta, P_z, T_1, \dots, T_m) \mapsto \sum_{k=1}^m \sum_{c=1}^{n_k} \|i_{k,c} - P \circ \delta \circ P_z \circ T_c\|^2$$

ist die zu minimierende Fehlerfunktion gegeben.

3.2.1. Bestimmung der Kollinearität

Für die Kollinearität einer Kalibrierplatte mit Weltpunkten w_n und ihrer Beobachtung i_n in homogenen Koordinaten für $n \in I \subset \mathbb{N}, |I| \geq 4$ gilt:

$$i_n = H \cdot w_n$$

woraus sich das nicht-lineare Optimierungsproblem ergibt, in dem der Fehler

$$\sum_{n \in I} \|i_n - \tilde{P}_z(H \cdot w_n)\|$$

minimiert werden muss.

3.2.2. Startwert für die Kalibrierung

Für eine projektive Abbildung H aus Abschnitt 3.2.1 ergibt sich unter bestimmten Nebenbedingungen¹³ ein Gleichungssystem der Form $Ab = O$, wobei sich die Einträge von A als Linearkombination von H darstellen lassen.

Für n Kalibriermuster ergeben sich so $2n$ Gleichungen. “Eine nicht-triviale Lösung für $b_* = \operatorname{argmin} \|Ab\|$ ist dann durch den kleinsten Eigenvektor mit dem kleinsten Eigenwert von $A^t A$ gegeben.” [Han11, 3.2.7] Darüber lassen sich dann in der Regel linear die Parameter der Projektionsmatrix u_0, v_0, α, β und γ bestimmen.

¹¹Ein planares Target reicht nicht aus, was mit dem Strahlensatz leicht zu veranschaulichen ist.

¹²Durch das Quadrieren werden im Fehlerwert größere Abstände gegenüber geringeren Abständen stärker gewichtet. Außerdem entfällt durch das Quadrat das Radizieren in der euklidischen Norm.

¹³Siehe [Han11, 3.2.6]: Bedingungen der Projektionsmatrix \tilde{P} : $h_1^t \tilde{P}^{-t} \tilde{P}^{-1} h_2 = 0, h_1^t \tilde{P}^{-t} \tilde{P}^{-1} h_1 - h_2^t \tilde{P}^{-t} \tilde{P}^{-1} h_2$ für h_1, h_2, h_3 Spalten der projektiven Abbildung H .

4. Stereorekonstruktion

Da das Bild einer Kamera mit einem Spiegel bei bekannter relativer Position des Spiegels zur Kamera im Grunde wie zwei Bilder eines Stereokamerasystems verstanden werden kann, werden hier zunächst einige Grundbegriffe der 3D-Rekonstruktion eingeführt.

Ein Lochkamerastereosystem setzt sich aus zwei Lochkameraabbildungen zusammen und bildet damit eine sog. *Stereobasis*. Dabei betrachtet man eine der beiden Kameras und ihre relative Position zur anderen Kamera. Diese relative Position ist in Form der extrinsischen Parameter als Bewegung in das Koordinatensystem der ersten Kamera gegeben.

Definition 7. Seien (K, K') zwei Lochkameraabbildungen (mit Verzeichnung) und $T = (R \ t)$ die Transformation von der zweiten in die erste Kamera. Dann definieren

$$K = P(I \ \mathbf{O}) = (P \ \mathbf{O})$$

und $K' = P'(R \ t)$

ein *Lochkamerastereosystem*.

Definition 8. Korrespondierende Punkte

Korrespondierende Punkte sind ein Paar oder eine Menge von Bildpunkten, die allesamt Repräsentant des selben Weltpunkts sind.

Das Ziel der Stereorekonstruktion ist es, für zwei korrespondierende Punkte einen gemeinsamen Weltpunkt zu finden. Die Bestimmung einer ganzen Reihe von Weltpunkten ermöglicht die 3D-Rekonstruktion von beobachteten Objekten. Über die Qualität dieser Rekonstruktion lässt sich außerdem feststellen, wie gut das verwendete Verfahren bei der intrinsischen oder extrinsischen Kalibrierung funktioniert hat - metrisch beispielsweise dadurch, wie *nah* in Lage und Oberfläche das rekonstruierte Objekt dem realen Objekt kommt.

Die Schwierigkeit in der Stereorekonstruktion besteht nun darin,

1. für zwei Bilder korrespondierende Punkte zu finden und
2. für zwei korrespondierende Punkte den dazugehörigen Weltpunkt zu bestimmen.

Definition 9. Markante Punkte

Markante Punkte (engl. auch *features*) sind Bildpunkte mit *markanten* Extraktionseigenschaften. Diese Eigenschaften ermöglichen auch in verschiedenen Bildern mit unterschiedlichen Blickwinkeln der selben Szene einen hohen Wiedererkennungswert, so dass sie zur Bestimmung *korrespondierender Punkte* genutzt werden können.

Während in der Praxis vielerlei Algorithmen für das Finden von markanten Punkten existieren, beschränkt sich diese Arbeit auf die manuelle Extraktion von korrespondierenden Punkten. Das hat den Vorteil, dass die Punkte auf Subpixel-Genauigkeit¹⁴ ohne weiteren Programmieraufwand und mit mehr Flexibilität für eine Evaluation bestimmt werden können.

So können sowohl die qualitativen Unterschiede bei geringer und hoher Genauigkeit von korrespondierenden Punkten festgestellt, als auch die extrinsische Kalibrierung unabhängig davon sehr genau vorgenommen werden.

Definition 10. Koplanare Punkte

Koplanare Punkte sind Punkte, die allesamt in der selben Ebene liegen.

¹⁴Subpixel sind Fließkommawerte zwischen tatsächlichen Bildpunkten, die nur als Ganzzahl existieren. Weltpunkte liegen eher selten genau auf einem Bildpunkt, sondern meist leicht versetzt dazu.

Für die Qualität der rekonstruierten Weltpunkte wird ein bekanntes Objekt mit sehr günstigen Extraktionseigenschaften genutzt: eine Kalibrierplatte. Durch die Koplanarität der Punkte lässt sich später leicht veranschaulichen, wie die Weltpunkte im Raum liegen und sie sich zu ihrem Referenzobjekt verhalten.

4.1. Stereorekonstruktion durch korrespondierende Punkte

Für zwei korrespondierende Punkte $i_{p,1}$ und $i_{p,2}$ unter K und K' , gilt für einen Weltpunkt p :

$$i_{p,1} = K(p) \text{ und } i_{p,2} = K'(p)$$

Durch die Fehler in der Abbildung, der Diskretisierung (Pixel statt Weltpunkte) und der Extraktion markanter Punkte, weicht die tatsächlich ermittelte Position von der idealen Position in der Regel ab. Aus diesem Grund wird in der Praxis eine Fehlerfunktion minimiert, um den Weltpunkt zu bestimmen:

Definition 11. Seien (K, K') zwei Lochkameraabbildungen (mit Verzeichnung) in homogenen Koordinaten und $i_{p,1}$ und $i_{p,2}$ korrespondierende Punkte in Bildern zu K und K' . Dann ist der gesuchte Weltpunkt $p \in \mathbb{R}^3$, der durch $i_{p,1}$ und $i_{p,2}$ repräsentiert wird, gegeben durch

$$p := \underset{x}{\operatorname{argmin}} f(x)$$

für

$$f : p \mapsto \|K(p) - i_{p,1}\| + \|K'(p) - i_{p,2}\|$$

Dies ist wiederum ein nicht-lineares Optimierungsproblem, “für das ein guter Startwert ermittelt werden sollte” [Han11].

4.1.1. Startwert für die Stereorekonstruktion

Mit $i_{p,1}, i_{p,2}$ und p folgt aus der Lochkameragleichung¹⁵ für $\tilde{K} = (\tilde{k}_{i,j})_{1 \leq i,j \leq 3}$ und $\tilde{K}' = (\tilde{k}'_{i,j})_{1 \leq i,j \leq 3}$:

$$A = \begin{pmatrix} \tilde{k}_{1,1} - u_1 \tilde{k}_{3,1} & \tilde{k}_{1,2} - u_1 \tilde{k}_{3,2} & \tilde{k}_{1,3} - u_1 \tilde{k}_{3,3} \\ \tilde{k}_{2,1} - v_1 \tilde{k}_{3,1} & \tilde{k}_{2,2} - v_1 \tilde{k}_{3,2} & \tilde{k}_{2,3} - v_1 \tilde{k}_{3,3} \\ \tilde{k}'_{1,1} - u_2 \tilde{k}'_{3,1} & \tilde{k}'_{1,2} - u_2 \tilde{k}'_{3,2} & \tilde{k}'_{1,3} - u_2 \tilde{k}'_{3,3} \\ \tilde{k}'_{2,1} - v_2 \tilde{k}'_{3,1} & \tilde{k}'_{2,2} - v_2 \tilde{k}'_{3,2} & \tilde{k}'_{2,3} - v_2 \tilde{k}'_{3,3} \end{pmatrix}$$

$$d = \begin{pmatrix} u_1 \tilde{k}_{3,4} - \tilde{k}_{1,4} \\ v_1 \tilde{k}_{3,4} - \tilde{k}_{2,4} \\ u_2 \tilde{k}'_{3,4} - \tilde{k}'_{1,4} \\ v_2 \tilde{k}'_{3,4} - \tilde{k}'_{2,4} \end{pmatrix}$$

und dadurch

$$p := \underset{q \in \mathbb{R}^3}{\operatorname{argmin}} \|Aq - d\|^2$$

Die Lösung dieser Gleichung erhält man über $p = (A^t A)^{-1} A^t d$, falls $A^t A$ regulär ist.

¹⁵Siehe [Han11, Kap.3]: $\exists \lambda \in \mathbb{R} \setminus \{0\} : \tilde{K} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$.

5. Spiegelkamera

5.1. Ermittlung der Spiegelposition

Um die Position des Spiegels zu ermitteln gibt es mehrere Ansatzmöglichkeiten, von denen eine in der vorliegenden Arbeit ausführlich beschrieben und zur praktischen Umsetzung verwendet wird. Da der Blick aus dem Spiegel heraus als virtuelle Kamera in einem Stereosystem betrachtet wird, ist das Wissen um die Spiegelposition notwendig, um über deren Spiegelebene die Transformation von der realen in die virtuelle Kamera zu berechnen und damit eine Stereorekonstruktion vorzunehmen.

Es wurden zwei verschiedene Möglichkeiten zur Ermittlung der Spiegelposition betrachtet:

- 1.) Ermittlung über ein im Spiegelbild beobachtetes Kalibriermuster und
- 2.) Ermittlung über eine Spiegelmarkierung.

Mit der Spiegelmarkierung werden markante Punkte (siehe Definition 9) bezeichnet, die am Spiegel angebracht werden und nicht Teil des Spiegelbilds sind.

Der Vorteil eines Kalibrierusters ist die hohe Anzahl markanter Punkte die zu einer extrinsischen Kalibrierung genutzt werden können, wogegen eine manuelle Spiegelmarkierung außerhalb des Spiegels nicht genügend Platz für eine vergleichbare Anzahl markanter Punkte bietet. Für die manuelle Spiegelmarkierung ist außerdem eine Vermessung der Markierungen notwendig, da diese im Gegensatz zu einem Kalibriermuster nicht symmetrisch angeordnet sind.

Gegen ein beobachtetes Kalibriermuster spricht aber, dass die Größe der Platte eine weitere Einschränkung für den späteren Aufbau der Szene darstellt, weil sie derart gestaltet sein muss, dass die Platte zu großem Teil sichtbar ist, um ihren Vorteil auszunutzen. Darüber hinaus verzerrt der Spiegel die Punkte der beobachteten Kalibrierplatte, so dass die Extraktion der Punkte ungenauer gegenüber äußeren Markierungen ist.

Während sich in der Praxis unter Berücksichtigung der entsprechenden Nachteile gegebenenfalls ein über den Spiegel beobachtetes Kalibrierobjekt besser eignen mag, wurde für die Untersuchung eine äußere Markierung vorgezogen, da sie bei der Extraktion eine geringere Fehlerquelle darstellt. In der Evaluation wird sich jedoch herausstellen, dass sich zwar eine höhere Genauigkeit bei der Extraktion der markanten Punkte erreichen lässt, aber an dieser Stelle dagegen andere Probleme auftreten können.

Zur extrinsischen Kalibrierung werden mindestens drei Punkte benötigt. Dazu wird an jeder der vier Ecken des Spiegels eine Markierung mit Kreuzen¹⁶ angebracht.

Mit Hilfe dieser vier Punkte lässt sich ein Koordinatensystem für den Spiegel aufspannen. Dabei wird die linke obere Ecke des Spiegels als Ursprung und der Normalenvektor der Spiegelebene als z-Achse angenommen. Bei der Festlegung der Punkte im Spiegelkoordinatensystem ist zu beachten, dass die Markierung in der linken oberen Ecke nicht exakt mit dem Ursprungspunkt übereinstimmt. Durch Ausmessen der Abstände zwischen den Punkten ergeben sich so vier Weltpunkte (bezogen auf das Spiegelkoordinatensystem) und deren korrespondierenden Bildkoordinaten. Mit diesen Punktkorrespondenzen lassen sich extrinsische Parameter¹⁷ berechnen und die Spiegelposition relativ zum Kamerakoordinatensystem ist ermittelt.

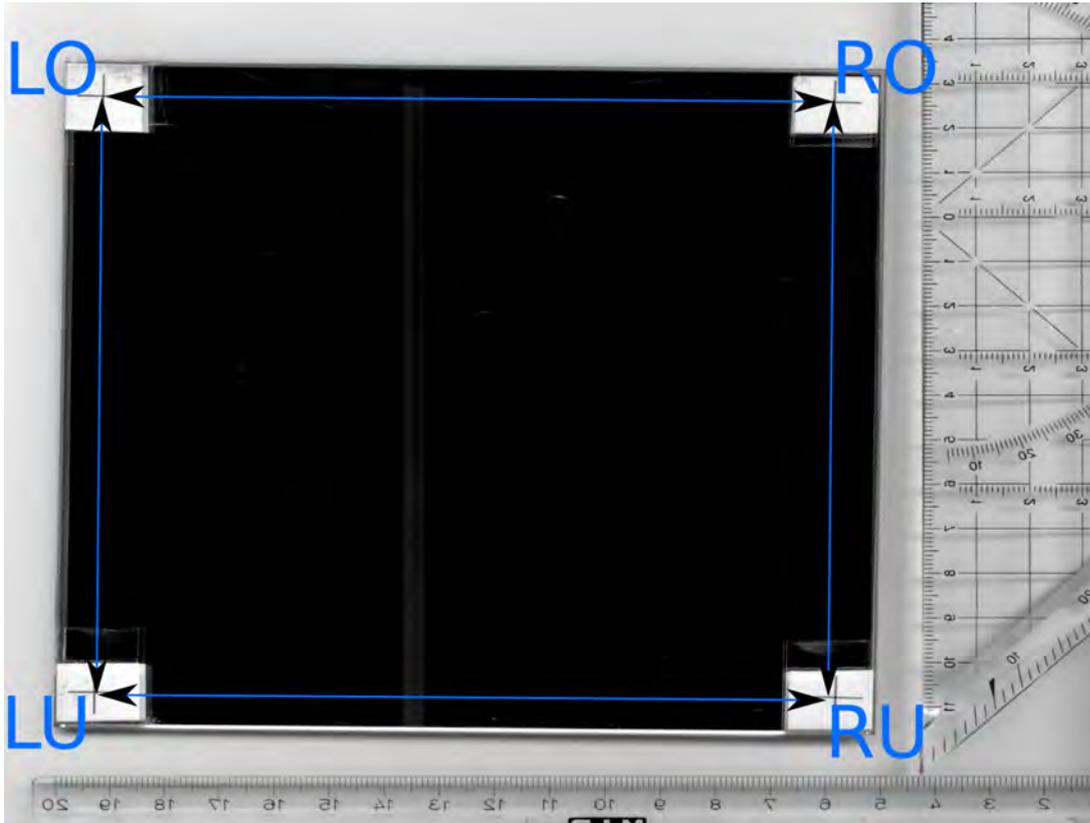
¹⁶Kreuze erweisen sich bei der Extraktion markanter Punkte als sehr zuverlässig. Selbst bei Verzeichnungen lässt sich der Mittelpunkt noch gut ermitteln und auf Subpixelgenauigkeit bestimmen. Siehe auch http://docs.opencv.org/modules/imgproc/doc/feature_detection.html#cornersubpix.

¹⁷Über die Spiegelposition in Form extrinsischer Parameter lassen sich die "Spiegelkoordinaten" in "Kamerakoordinaten" umrechnen; siehe Definition 3 zu extrinsischen Parametern.

5.1.1. Spiegelmaße

Die Genauigkeit der korrespondierenden Punkte zur Ermittlung der Spiegelposition tragen wesentlich zur Genauigkeit der Resultate bei. Da die markierten Ecken nicht notwendigerweise ein Rechteck bilden, wäre eine einfache Zuweisung von Koordinaten zu den jeweiligen markanten Punkten ohne genaue Maße unzulänglich (bspw. *links oben* := (0,0), *rechts oben* := (0,1), usw.). Aus diesem Grund muss der genaue Abstand der einzelnen Punkte bestimmt und darüber ein passendes Koordinatensystem gefunden werden.

Abbildung 8: Mit vier Kreuzen markierter Spiegel und beiliegende Lineale, aufgenommen mit einem Scanner.



Im Nachfolgenden bezeichnet M_{tl} die linke obere Ecke (“LO” in Abbildung 8), M_{tr} die rechte obere Ecke (“RO”), M_{bl} die linke untere Ecke (“LU”) und M_{br} die rechte untere Ecke (“RU”).

Für eine genaue Messung der Abstände wird der Spiegel in einem Scanner abgelichtet. Um Ungenauigkeiten der Technik vorzubeugen, werden Lineale entlang der Kanten des Spiegels gelegt. So können zusätzlich zu einer Vermessung des Spiegels mit einem Lineal (außerhalb des Bildes) die Pixelabstände im Scanbild in metrische Werte umgerechnet werden. Die Werte stimmen bei den gemessenen Spiegeln auf wenige Zehntelmillimeter überein.

Im Scanbild lassen sich außerdem Verschiebungen der Pixel auf $\frac{1}{100}$ mm feststellen. Die Strecken $M_{tl}M_{bl}$ und $M_{bl}M_{br}$ liegen beispielsweise nicht orthogonal zueinander. Mit dem Satz des Pythagoras ergibt sich eine Verschiebung von M_{tl} zu M_{bl} von $d_{M_{tl}.x \leftrightarrow M_{bl}.x} = 2$ mm entlang der x-Achse. M_{tr} ist gegenüber M_{tl} um $d_{M_{tr}.y \leftrightarrow M_{bl}.y} = 1,5$ mm entlang der y-Achse verschoben. Damit lässt sich M_{tl} auf einem Koordinatensystem mit Genauigkeit von $\frac{1}{10}$ mm mit $M_{tl} = (20, 0)$ identifizieren.

Analog wird mit den übrigen drei Punkten vorgegangen. Für Abbildung 8 ergeben sich damit Koordinaten wie exemplarische in Tabelle 3.

Strecke	Abstand
$M_{tl} \leftrightarrow M_{tr}$	133,0mm
$M_{tl} \leftrightarrow M_{bl}$	133,8mm
$M_{tr} \leftrightarrow M_{br}$	133,8mm
$M_{bl} \leftrightarrow M_{br}$	135,0mm

Tabelle 2: Abstände der markierten Punkte in den Spiegelecken.

Ecke	Koordinate
M_{tl}	(20, 0)
M_{tr}	(1330, 15)
M_{bl}	(0, 1338)
M_{br}	(1350, 1353)

Tabelle 3: Koordinaten des Spiegelkoordinatensystems.

5.2. Die Spiegelungsmatrix

Während zuvor die extrinsischen Parameter die Spiegelposition festgelegt haben, muss nun ermittelt werden, wie ein Punkt an der davon festgelegten Ebene gespiegelt wird. Eine Spiegelung¹⁸ an einer beliebigen Ebene im Raum lässt sich über mehrere Schritte als affine Transformation in einer 4×4 -Matrix darstellen.

Die einzelnen Schritte für eine Ebene mit Normalenvektor n und beliebigem Punkt p auf der Ebene sind nach [Kov12] wie folgt:

- i) Verschiebe den Punkt p der Spiegelebene in den Ursprung des Koordinatensystems.
- ii) Rotiere den Normalenvektor n der Spiegelebene in den Ursprung, so dass er mit der z -Achse übereinstimmt und die Spiegelebene nun in der xy -Ebene liegt.
- iii) Spiegle das Objekt mit einer Householdertransformation¹⁹ an der xy -Ebene.
- iv) Führe die inversen Transformationen durch.

Führt man die einzelnen Transformationen in homogenen Koordinaten hintereinander aus, ergibt sich:

Definition 12. Seien eine Ebene E in \mathbb{R}^3 mit ihrem normierten Normalenvektor $n \in E, n = (n_x, n_y, n_z)^t$ und ein Punkt $o \in E, o \neq n, o = (o_x, o_y, o_z)^t$ gegeben. Dann ist $d = -n_x o_x - n_y o_y - n_z o_z$ der Abstand der Ebene zum Ursprung und mit

$$M = \begin{pmatrix} 1 - 2n_x^2 & -2 \cdot n_x \cdot n_y & -2 \cdot n_x \cdot n_z & -2 \cdot n_x \cdot d \\ -2 \cdot n_x \cdot n_y & 1 - 2 \cdot n_y^2 & -2 \cdot n_y \cdot n_z & -2 \cdot n_y \cdot d \\ -2 \cdot n_x \cdot n_y & -2 \cdot n_y \cdot n_z & 1 - 2 \cdot n_z^2 & -2 \cdot n_z \cdot d \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

ist die **Spiegelungsmatrix** definiert, die einen Punkt in \mathbb{R}^3 an der gegebenen Ebene E spiegelt.

Für die Transformation $T_{B'}^B$ von der virtuellen Kamera mit Basis B' ins Kamerakoordinatensystem mit Basis $B = I$ gilt $T_{B'}^B = (B')^{-1}$, weil B die Standardbasis ist. Da die Spiegelmatrix die

¹⁸Ausführliche Erläuterungen zu den Strahlenverläufen bei einer Reflexion an einem ebenen Spiegel nach dem Reflexionsgesetz finden sich unter [Ver06, S. 14].

¹⁹Mit einer Householdertransformation lässt sich eine Spiegelung an einer Ebene durch den Ursprung berechnen. Die Householder-Matrix ist orthogonal und daher invertierbar.

affine Transformation in das Kamerakoordinatensystem der virtuellen Kamera beschreibt, gilt $B' = M$ und damit $T_{B'}^B = M^{-1}$.

Da der Normalenvektor n durch die Spiegelposition und die vier verwendeten markanten Punkte des Spiegels gegeben sind, ist die Spiegelungsmatrix M durch die extrinsische Kalibrierung festgelegt.

Definition 13. Sei $K = P \circ T$ eine Kameraabbildung, wobei T eine Transformation aus einer Spiegelebene in die xy -Ebene des Kamerakoordinatensystems beschreibt. Für den Normalenvektor n_m und einen Punkt p_m der Spiegelebene mit $n_m = (0, 0, 1)^t$ sind dann $n_c = T(n_m)$ der in das Kamerakoordinatensystem transformierte Normalenvektor und $p_c = T(p_m)$ ein in das Kamerakoordinatensystem transformierter Punkt der Spiegelebene. Sei M die Spiegelungsmatrix für n_c und p_c , die eine Spiegelung im Kamerakoordinatensystem an der gegebenen Spiegelebene definiert. Dann ist

$$K' = P(I \quad \mathbf{0})$$

die eigentliche Kamera und

$$K'' = P \circ M^{-1}$$

die virtuelle Kameraabbildung des Spiegelbilds. Mit (K', K'') ist dann ein Lochkamerastereosystem gegeben.

Mit der Spiegelungsmatrix ist ein Lochkamerastereosystem wie in Kapitel 4 gegeben. Die extrinsische Kalibrierung zur Ermittlung von T dient ausschließlich zur Bestimmung der Spiegelungsmatrix und der darüber definierten *virtuellen* Kamera K'' . Mit K' und K'' ist über die Stereorekonstruktion aus korrespondierenden Punkten eine Methode zur 3D-Rekonstruktion über eine Kamera und einen Spiegel gegeben.

6. Evaluation

In diesem Kapitel wird exemplarisch erläutert wie Werte für die einzelnen Teilabschnitte berechnet werden. Die Teilabschnitte umfassen die intrinsische Kalibrierung einer Kamera, die Ermittlung der Spiegelposition (sog. *extrinsische Kalibrierung*), die Berechnung der Spiegelungsmatrix und anschließend die Triangulierung einer Menge von Punktkorrespondenzen. Als Ergebnis erhält man für jede Punktkorrespondenz einen ermittelten Weltpunkt.

Für die Punktkorrespondenzen wird ein Schachbrettmuster gewählt, da die Ermittlung von Punktkorrespondenzen unbekannter Objekte in verschiedenen Bildern bei der Stereorekonstruktion ein nicht-triviales Problem darstellt. Darüber hinaus sind Punkte auf einem Schachbrett koplanar und ermöglichen so für die triangulierten Weltpunkte eine leichtere Untersuchbarkeit bezüglich der Abweichung von ihrem Sollwert.

Anschließend an den exemplarischen Ablauf folgt eine Reihe von Messergebnissen und deren Bewertung, sowie eine Erläuterung zum Aufwand und der Anwendbarkeit der Methode. Die Messergebnisse werden mit einem eigenen Programm berechnet und mit Hilfe von *GnuPlot*²⁰ visualisiert.

Programm zur Berechnung der Messergebnisse

Zur Berechnung werden für jede 3D-Rekonstruktion

1. die intrinsischen Parameter der kalibrierten Kamera eingelesen,
2. die Bildpunkte der Markierungen auf dem Spiegel und die zugeordneten 3D-Punkte eingelesen,
3. die Koordinaten der Punktkorrespondenzen auf dem Spiegel eingelesen,
4. eine extrinsische Kalibrierung über die Markierungen auf dem Spiegel durchgeführt,
5. der Normalenvektor und ein Punkt der Spiegelebene in der xy -Ebene des ermittelten Referenzkoordinatensystems in das Kamerakoordinatensystem transformiert,
6. die Spiegelungsmatrix mit der Normalen im Kamerakoordinatensystem berechnet,
7. die Inverse der Spiegelung als Transformation der virtuellen Kamera in die reale Kamera berechnet,
8. ein Weltpunkt für jede Punktkorrespondenz über einen nicht-linearen Optimierer und eine Fehlerfunktion nach Kapitel 4.1 ermittelt,
9. eine Ebenenschätzung für alle ermittelten Weltpunkte durchgeführt,
10. sowie die Abstände zwischen den Weltpunkten selbst und den Weltpunkten und der abgeschätzten Ebene berechnet.

Das Konsolenprogramm besteht aus rund 1500 Zeilen Quelltext und baut weitestgehend auf Bildverarbeitungsmethoden der *libforwiss*²¹ auf. Listing 1 zeigt die verschiedenen Parameter. Durch die Angabe eines vollständigen "records", bestehend aus Dateien mit intrinsischen Kameraparametern, Spiegelmarkierungen und Punktkorrespondenzen, lässt sich schnell eine größere Menge an Szenen verarbeiten und strukturiert abspeichern.

²⁰Siehe <http://www.gnuplot.info>

²¹Eine proprietäre Bibliothek des *FORIWSS*-Institut. Aus ihr werden vorwiegend Methoden zur Bildverarbeitung genutzt. Siehe auch <http://forwiss.de>.

Listing 1: Liste der Argumente für *PlaneReconstruction*

```

1 Allowed options:
2 -R [ --record ] arg          record directory path. if
   it contains a 'camera.cm4', a 'mirror' and a '
   correspondences' file, it is enough to kickstart the
   program
3 -P [ --camera-file ] arg     camera file
4 -I [ --image ] arg          image file
5 -M [ --mirror-file ] arg     mirror file
6 -C [ --correspondences-file ] arg correspondences file
7 -A [ --mirror-adjustment ] arg (=0) mirror adjustment factor
8 -D [ --measurement-file ] arg measurement file
9 -o [ --optimizer ] arg (=2) optimizer (2 for NM, 1
   for BFGS)

```

6.1. Szenenaufbau

Es werden ein handelsüblicher Spiegel und ein Oberflächenspiegel eingesetzt. Der Spiegel wird jeweils zur rechten Seite der Kamera positioniert, da sich der Aufbau bis auf äußere Faktoren (Lichtverhältnisse) symmetrisch zur Hauptachse der Kamera verhält. Als beobachtetes Objekt wird ein Schachbrettmuster gewählt. Abbildung 9 zeigt einen exemplarischen Aufbau einer Szene.



Abbildung 9: Aufnahme einer exemplarischen Szene: Zur Kamera rechts versetzt ist ein Spiegel so aufgebaut, dass er gerade noch ganz im Bild ist. Zur Linken wird ein Kalibrieremuster beobachtet. Das Objekt lässt sich durch Schrauben mikrometerweise verschieben.

6.2. Kalibrierung der Kamera

Die Kamerakalibrierung wird mit Hilfe von *fw_monocal*, einer proprietären Software des FORWISS-Institut²², durchgeführt. Dazu werden die Kolinearitäten von zehn bis zwanzig verschiedenen Aufnahmen eines Kalibriertargets und anschließend Startwerte für die intrinsischen Kameraparameter nach [Zha98] bestimmt. Nach der Ermittlung der intrinsischen Kameraparameter werden die Parameter für die radiale Verzeichnung berechnet (ebenso nach [Zha98]).

Listing 2: Ausgabe der Verzeichnungsparameter und der Projektions- und Reprojektionsmatrix.

```

1 radial distortion: (-1.579393e-01, 2.018388e-01)
2 ReProjection:
3 [ [2.848076e-04, 8.996877e-08, -5.648936e-01]
4   [0.000000e+00, 2.849397e-04, -3.883718e-01]
5   [0.000000e+00, 0.000000e+00, 1.000000e+00] ]
6 undistortion: 1.596833e-01, -1.685195e-01)
7
8 [ [3.511143e+03, -1.108632e+00, 1.982991e+03]
9   [0.000000e+00, 3.509515e+03, 1.362997e+03]
10  [0.000000e+00, 0.000000e+00, 1.000000e+00] ]

```

Listing 3: Statistische Werte der Fehlerwerte der Projektion mit Verzeichnung

```

1 Error w. dist. :
2 Projective: Samples: 4779.000000
3   Mean: 2.793961e-01 Var: 6.042687e-02
4   Min: 3.381474e-03 Max: 1.163305e+00
5 Euclidean: Samples: 4779.000000
6   Mean: 1.090279e-04 Var: 7.454357e-09
7   Min: 1.441176e-06 Max: 4.822589e-04
8 repro: Samples: 4779.000000
9   Mean: 1.108760e-04 Var: 7.755985e-09
10  Min: 1.472229e-06 Max: 4.829925e-04

```

6.3. Berechnung der Transformationen

Mit Hilfe der markierten Kreuze auf dem Spiegel und der Spiegelkoordinaten aus Kapitel 5.1.1 kann man vier Weltpunkte (denen unmittelbar eine metrische Einheit vorgegeben ist) und deren Bildpunkte zur extrinsischen Kalibrierung verwenden.

Listing 4 zeigt den Funktionsaufruf in der *OpenCV*-Bibliothek und mit *libforwiss*. In der *OpenCV* arbeitet man mit dem *PnP-Algorithmus* (engl. *Perspective-n-Point Camera Pose Estimation*) und übergibt die Weltpunkte des Spiegels, seine entsprechenden Bildpunkte, die intrinsischen Parameter und die Verzeichnung.

Die *libforwiss* arbeitet direkt mit einem Calibration-Objekt, in das die kalibrierten Kameraparameter geladen werden. Danach wird die extrinsische Kalibrierung mit einer Kalibrierplatte ausgeführt, die über Punktkorrespondenzen des Spiegels initialisiert wurde.

Listing 4: Extrinsische Kalibrierung in der *OpenCV* und mit *libforwiss*

```

1 solvePnP(settings.mirrorCoords, settings.mirrorImagePoints,
2 settings.Intrinsic, settings.Distortion,

```

²²Siehe <http://forwiss.de>

```

3  rotation_vector , translation_vector , false , 0);
4
5  Calibration calibration ;
6  calibration.load_cm4(cameraPath.c_str());
7  plate.init_from_correspondences(correspondences);
8  calibration.extrinsic_calib(plate);
9  tcamera_c camera = calibration.get_camera();
10 transformation3d extrinsic
11  = camera.get_camera_to_world_transformation();

```

In beiden Fällen erhält man die relative Position des Spiegels über die extrinsischen Parameter. In der *OpenCV* sind sie per Translationsvektor *translation_vector* und Rotationsvektor *rotation_vector* gegeben, in der *libforwiss* durch eine 4×3 -Matrix im *transformation3d*-Objekt *extrinsic*.

Listing 5: *Extrinsische Parameter der Kamera im Verhältnis zu den Markierungen auf dem Spiegel.*

```

1  Rotation:
2  [ [0.226939, -0.000229236, 0.973909]
3    [-0.0460712, 0.998878, 0.0109706]
4    [-0.972819, -0.0473588, 0.226674] ]
5  Translation:
6  (100.363, -69.3089, 462.336)

```

Durch die vier markierten Punkte des Spiegels existieren bereits vier verschiedene Punkte, die zur Berechnung der Normalen und damit zur Spiegelungsmatrix genutzt werden können. Für w_1, w_2, w_3, w_4 ergibt sich die Normale als Kreuzprodukt zweier Vektoren auf der Spiegelebene: Sei $u, v, normal \in \mathbb{R}^3$:

$$\begin{aligned}
 u &= w_2 - w_1 \\
 v &= w_3 - w_1 \\
 normal &= u \times v
 \end{aligned}$$

Der Punkt w_4 und der Normalenvektor *normal* können dann in die Spiegelungsmatrix aus Kapitel 5.2 eingesetzt werden und man erhält die Reflexionsmatrix im Kamerakoordinatensystem. Ein Beispiel dafür ist Listing 6.

Listing 6: *Berechnung der Reflexion an der Ebene, bestimmt durch die extrinsischen Parameter*

```

1  origin o = (100.363, -69.3089, 462.336)
2  normal n = (17427.9, 125.877, 4495.26)
3  euclideanNorm e = 17998.8
4  Distance d = -212.165
5  Normal/euclideanNorm n = (0.968284, 0.00699365, 0.249754)
6  Reflection: Rotation:
7  [ [-0.875148, -0.0135437, -0.483666]
8    [-0.0135437, 0.999902, -0.00349339]
9    [-0.0135437, -0.00349339, 0.875246] ]
10 Translation:
11 (410.872, 2.96762, 105.978)
12
13 det( Reflection ) = -0.814495

```

Als 4×4 -Matrix lässt sich die Spiegelung invertieren und man erhält die Transformation der virtuellen Kamera ins Kamerakoordinatensystem (in homogenen Koordinaten). Da der letzte

Zeilenvektor $(0,0,0,1)$ entspricht, ist sie als Matrix $(R \ t)$ mit spiegelndem und translatorischen Anteil wie in Listing 7 gegeben. Die Ausgabe des Objekts spricht in der *libforwiss* zwar von einer Rotation, aber die Determinante ist $\det(\text{virtualToCamera}) = -1.22775$ und es handelt sich in der Tat um eine Reflexion.

Nun sind durch die Kamera mit der intrinsischen Kalibrierung und der Transformation von der virtuellen Kamera ins Kamerakoordinatensystem alle Voraussetzungen für eine Stereorekonstruktion gegeben.

Listing 7: Relative Position der virtuellen Kamera

```

1 Rotation:
2 [ [-1.10129, -0.0262355, -0.542078]
3   [-0.02367, 0.999704, -0.00610623]
4   [-0.0263597, 0.00491665, 1.10159] ]
5 Translation:
6 (398.853, 4.49287, -74.7468)

```

6.4. Triangulierung

Zur Triangulierung werden nun eine Menge korrespondierender Bildpunkte der aufgenommenen Szene nach einer Startwertberechnung mit der Fehlerfunktion aus Kapitel 4.1 an einen nicht-linearen Optimierer übergeben. Dazu werden die numerischen Optimierer *BFGS*²³ und der Simplex-Optimierer *Nelder-Mead* eingesetzt. Hierbei werden ausschließlich Implementie-

Bildpunkt (u_1, v_1)	Spiegelbild (u_2, v_2)	Bildpunkt (u_1, v_1)	Spiegelbild (u_2, v_2)
406 653	3304 854	496 664	3233 853
403 696	3305 889	495 707	3232 889
401 739	3304 924	493 750	3233 924
400 783	3304 960	491 792	3233 959
397 826	3304 996	489 836	3233 995
396 870	3304 1032	487 879	3232 1031
394 914	3304 1068	485 922	3233 1066
392 957	3304 1103	484 965	3232 1102
390 1002	3305 1138	482 1009	3232 1138
389 1045	3304 1175	480 1052	3232 1174
⋮	⋮	⋮	⋮

Tabelle 4: Korrespondierende Punkte im Bild

rungen der *libforwiss* verwendet. Listing 8 zeigt zwei exemplarische Funktionsaufrufe für eine Punktkorrespondenz mit vorher ermitteltem Startwert.

Der ermittelte Weltpunkt wird in *result* geschrieben. Der Parameter *errorFunc* übermittelt jeweils den Zeiger auf die Fehlerfunktion, die die nicht-lineare Funktion aus Definition 11 für den jeweiligen Optimierer implementiert. Über *data* wird ein Objekt mit den korrespondierenden Punkten und den beiden Kamera-Objekten an den Optimierungsalgorithmus und damit für jeden Schritt an die Fehlerfunktion weitergereicht. Die Optimierer führen jeweils eine maximale Anzahl an Schritten durch und lassen sich über Statusfunktionen in ihrer Veränderung beobachten.

²³Broyden–Fletcher–Goldfarb–Shanno-Verfahren

Durch die Projektion ins Bild wird für beide Kameras in jedem Schritt die Abweichung zum jeweiligen Bildpunkt als Fehler aufsummiert.

Listing 8: Aufrufe der Optimierungsverfahren zur Minimierung der Fehlerfunktion.

```

1 find_local_min_point_bfgs(&result , errorFunc , &data , 0 , 0 ,
2   &startValue , bfgs_epsilon , 10000 , GOLDSTEIN_ORMIJO , statusFunc ,
3   NULL );
4
5 find_local_min_point_nelder_mead(&result , &step_size , errorFunc ,
6   &data , nm_exactness , 5000 , statusFunc_NelderMead , NULL ,
7   &errorFin );

```

Tabelle 5 zeigt die triangulierten Weltpunkte für die Punktkorrespondenzen aus Tabelle 4.

Triangulierter Weltpunkt (x,y,z)	Triangulierter Weltpunkt (x,y,z)
-433.988 -174.549 938.474	-401.636 -169.532 922.655
-434.836 -163.715 938.909	-401.859 -158.731 922.979
-435.718 -152.985 939.978	-402.146 -148.003 922.787
-435.649 -141.72 939.651	-402.742 -137.516 923.253
-436.885 -130.85 940.815	-403.271 -126.603 923.615
-436.881 -119.615 940.559	-404.144 -115.924 924.697
-437.499 -108.486 940.988	-404.487 -105.214 924.54
-438.216 -97.6499 941.557	-404.825 -94.4244 924.974
-438.534 -86.4425 941.343	-405.476 -83.5222 925.447
-438.948 -75.3036 941.845	-406.1 -72.7423 925.887
⋮	⋮

Tabelle 5: Triangulierte Weltpunkte (Einheit durch das vorgegebene Koordinatensystem in mm)

6.5. Messergebnisse & Bewertung

Für die Messergebnisse werden verschiedene Szenen mit unterschiedlichen Kameras und unterschiedlichen Einstellungen untersucht. Kameraeinstellungen, die eine neue Kalibrierung voraussetzen, werden für eine Reihe von Aufnahmen nur einmalig vorgenommen. Es werden die Kameras *Canon EOS 1000D*[Can14a] und die *Canon EOS 7D*[Can14b] eingesetzt. Durch den statischen Aufbau können lange Belichtungszeiten von 0,5-2,5s und kleine Blendenöffnungen²⁴ genutzt werden, um eine hohe Schärfentiefe zu erreichen.

Neben der Entfernung des Schachbretts und der Spiegelposition wird auch die millimetergenaue Verschiebung des Schachbretts entlang einer metrischen Laborschraube aufgenommen.

Die Extraktion der Punkte auf dem Schachbrettmuster wird sowohl manuell (über *Gimp*²⁵) als auch automatisch über die Schachbrettmustererkennung von *libforwiss* durchgeführt. Die Zuordnung markanter Punkte zu korrespondierenden Punkten geschieht manuell, da hierzu keine Verfahren zur Verfügung stehen.

²⁴Bei kleinerer Blendenöffnung besteht ein Kompromiss zwischen Schärfentiefe und Beugungsunschärfe. Man spricht dann von einer *förderlichen Blende*.

²⁵The GNU Image Manipulation Program auf <http://gimp.org>

Probleme in der Praxis

Während in den theoretischen Überlegungen in Kapitel 2 deutlich mehr Möglichkeiten einer Installation in Betracht kommen, hängt ein Aufbau in der Praxis jedoch wesentlich vom Winkel des Spiegels und der Schärfentiefe der Kamera ab.

Für die Bestimmung der Spiegelposition müssen die markierten Punkte am Spiegel möglichst präzise aus dem Bild zu extrahieren sein. Dies ist nur möglich, wenn der Winkel des Spiegels zur Hauptachse nicht zu gering ist. Je weiter der Spiegel zur Hauptachse hin rotiert wird, desto weiter auseinander liegen auch die gesuchten Bildpunkte. Ist der Abstand der Bildpunkte im Verhältnis zur Spiegelbreite zu gering, stellt ein einziger Pixelfehler gegenüber dem Idealwert bereits eine große Verschiebung in der Realität dar. Da die 3D-Rekonstruktion wesentlich von der Ermittlung der Spiegelposition abhängt, führt das wiederum zu schlechten oder gar falschen Messergebnissen.

Gleichzeitig ist die Extraktion der markierten Punkte auf dem Spiegel auch abhängig von der Bildschärfe der Markierungen. Je weiter sich ein Objekt aus der Schärfentiefe bewegt, desto größer wird der Unschärfekreis im Bild. Der Mittelpunkt der zur Markierung verwendeten Kreuze ist dann nicht mehr genau zu ermitteln, was ebenfalls zu Fehlern in der Bestimmung der Spiegelposition führt. Daraus folgt, dass der Aufbau einer Szene wesentlich von der Schärfentiefe der Kamera abhängt, da sowohl das Objekt als auch der Spiegel im Schärfentiefebereich der Kamera liegen müssen.

In der Untersuchung konnte durch eine längere Belichtungszeit und kleinere Blendenöffnungen eine höhere Schärfentiefe erreicht werden als es in der Praxis oft möglich ist. Dies ist in der Anwendung beim Aufbau einer Szene zusätzlich zu beachten.



Abbildung 10: Blick auf eine Ecke eines handelsüblichen Spiegels: die Spiegelung tritt deutlich hinter der Oberfläche auf.

Neben den Fehlern bei der extrinsischen Kalibrierung können auch Fehler bei der Spiegelung selbst auftreten. Ein Blick auf die 3D-Rekonstruktion der Ebene und das Histogramm zu den Abständen der Punkte zu ihrer gemeinsamen Ebene lässt eine Verzeichnung des Spiegels vermuten. Außerdem reflektieren handelsübliche Spiegel erst deutlich hinter der Oberfläche, wie man in Abbildung 10 erkennen kann.

Der Verschiebungsfehler des Spiegels lässt sich durch eine Verschiebung der Reflexionsebene entlang des Normalenvektors in negativer Richtung der z -Achse des Spiegelkoordinatensystems ausgleichen.

Tabelle 6 veranschaulicht deutlich, wie sehr die tatsächliche Reflexionsebene die Messergebnisse beeinflusst. Für die Werte ohne Spiegelverschiebung lässt sich die Ebene zu $x \cdot n = d$ mit $n = (0.40852, -0.00587489, 0.91273)^t$ und $d = -822.285$ abschätzen, während sich für die Werte mit einer Verschiebung der Spiegelungsebene entlang des Normalenvektors des Spiegels um 35mm die Parameter $n = (0.408496, -0.00591259, 0.912741)^t$ und $d = -679.678$ ergeben. Der Abstand der Ebene zum Ursprung des Kamerakoordinatensystems verändert sich also um 142.607 mm.

Da eine kleine Verschiebung der Spiegelebene entlang ihres Normalenvektors bereits eine große Veränderung in der Entfernung der abgeschätzten Ebene des Schachbretts ergibt, lässt sich leicht ermitteln, in welcher Ebene der jeweils eingesetzte Spiegel tatsächlich reflektiert. Denn nur dann stimmen die erwarteten Messwerte, die Entfernung der abgeschätzten Schachbrettebene und die Entfernung des Spiegels bis auf geringe Ungenauigkeiten überein.

Ohne Spiegelverschiebung	Mit Spiegelverschiebung $a = -35$
-525.043 -211.172 1135.39	-433.988 -174.549 938.474
-526.134 -198.083 1136.01	-434.836 -163.715 938.909
-527.087 -185.064 1137.16	-435.718 -152.985 939.978
-527.11 -171.479 1136.9	-435.649 -141.72 939.651
-528.56 -158.306 1138.23	-436.885 -130.85 940.815
-528.577 -144.717 1137.96	-436.881 -119.615 940.559
-529.333 -131.255 1138.5	-437.499 -108.486 940.988
-530.146 -118.133 1139.1	-438.216 -97.6499 941.557
-530.56 -104.578 1138.89	-438.534 -86.4425 941.343
-531.053 -91.1006 1139.47	-438.948 -75.3036 941.845
⋮	⋮

Tabelle 6: Vergleich der triangulierten Punkte bei Verschiebung der Spiegelebene um 35mm.

Messungen

Die rekonstruierten 3D-Punkte werden mit Hilfe von *GnuPlot* visualisiert und lassen sich so in einem virtuellen Raum untersuchen. Dadurch lässt sich die Struktur der Ebene als Ganzes einschätzen und mit weiteren Messwerten überprüfen. Abbildung 11 zeigt eine Ansicht auf die Punkte eines rekonstruierten Schachbretts. Die Punkte zeigen eine auffällige Verengung zum unteren Ende des Schachbretts. Die Kamera für diese 3D-Rekonstruktion ist dabei ohne tangentielle Verzeichnung kalibriert.

Es werden unterschiedliche Messungen für die Erkennung von Fehlern vorgenommen. Zunächst werden die Abstände der Punkte auf dem Schachbrett zueinander gemessen. Hierbei hat sich ein Durchschnitt von 10.3288 mm bei einer Varianz von 0.197985 mm ergeben. Das Schachbrettmuster hat in der Realität einen Kastenabstand von einem Zentimeter. Die Messung wird daher als adäquat beurteilt, obwohl 3D-Rekonstruktionen unter optimalen Bedingungen eine zehnmal so hohe Genauigkeit erreichen können. Wesentliche Fehler werden auf die Extraktion der markanten Punkte für die extrinsische Kalibrierung und die Punktkorrespondenzen zurückgeführt.

Da die Abstände auf dem Schachbrett einen Fehler in der Ebene repräsentieren, wird zusätz-

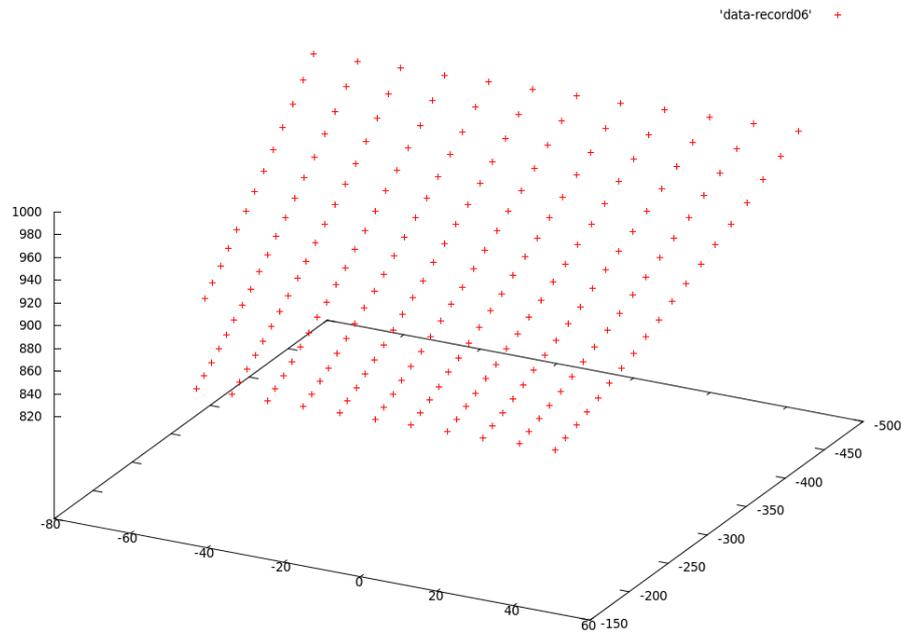


Abbildung 11: Ansicht der rekonstruierten Punkte des Schachbretts in GnuPlot. Es ist eine auffällige Verengung nach unten hin zu beobachten. Der durchschnittliche Abstand der Punkte auf der Ebene zum nächsten Nachbarn (Sollwert ein Centimeter) ist zufriedenstellend.

lich eine Ebenenschätzung durchgeführt. Dies geschieht unter der Annahme, dass die Punkte im Schnitt in der gesuchten Ebene liegen. Damit lassen sich die Abstände jeden Punktes zu der abgeschätzten Schachbrettebene berechnen. In Abbildung 12 ist das zugehörige Histogramm abgebildet. Über das Histogramm ist die erwartete Normalverteilung mit dem Erwartungswert und der Varianz aus den Daten gelegt. Dadurch lässt sich erkennen, dass der Mittelwert der Abstände in den negativen Bereich verschoben ist, während der Durchschnitt in der Nähe von null liegt.

Vergleicht man diesen Umstand mit Abbildung 13, ist eine Wölbung in der Mitte der Punktwolke zu erkennen. Abbildung 13 zeigt eine seitliche Ansicht derselben Ebene wie in Abbildung 11. Die Wölbung lässt Verzeichnungen des Spiegels vermuten.

6.6. Aufwand & Anwendbarkeit

Die Messergebnisse zeigen, dass sich eine 3D-Rekonstruktion mit Kamera und Spiegel durchführen lässt. In der Praxis haben sich gegenüber den Überlegungen in Kapitel 2 weitere Einschränkungen bei der Installation feststellen lassen. Ein Aufbau mit Kamera und Spiegel für statische Anwendungen ist unter diesen Einschränkungen aber problemlos möglich.

Wesentliche Problemfelder der 3D-Rekonstruktion sind die extrinsische Kalibrierung zur Ermittlung der Spiegelposition und die Ermittlung von korrespondierenden Punkten im Bild und Spiegelbild.

Für die Ermittlung der Spiegelposition sind in einer praktischen Anwendung mehrere Szenarien möglich:

1. Analog zur Untersuchung in dieser Arbeit lässt sich eine fixe Spiegelposition und manuelle extrinsische Kalibrierung sehr gut durchführen. Diese Umsetzung leidet aber unter einer möglicherweise notwendigen Flexibilität in der Anwendung.

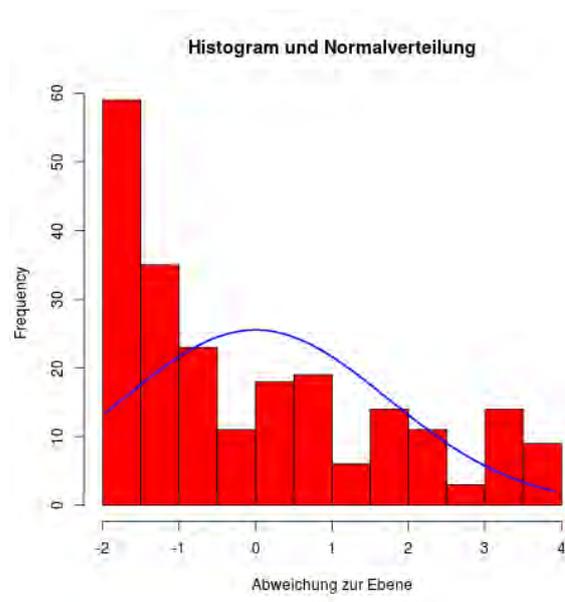


Abbildung 12: Histogramm der Abstände jedes Punktes d_f zur approximated Ebene und die unter Erwartungswert und Varianz erwartete Normalverteilung mit $E(d_f) = 1.766667 \cdot 10^{-7}$, $Var(d_f) = 0.7993224$

2. Kamera und Spiegel können über eine Montage zueinander fixiert werden. Damit ist eine einmalige extrinsische Kalibrierung notwendig. Unter möglichen Fehlern durch Störungen der Montage leidet gegebenenfalls die Qualität der Ergebnisse.
3. Eine technische Automatisierung der Bewegung der Kamera oder des Spiegels könnte eine einmalige extrinsische Kalibrierung anhand der Bewegung neu berechnen. Die Qualität der Ergebnisse hinge dann wesentlich mit der Genauigkeit der Bewegung zusammen.
4. Eine automatische Spiegelerkennung wäre durch verschiedene Ansätze realisierbar und würde damit eine beliebige (dynamische) Spiegelposition ermöglichen. Dazu könnten weitere Markierungen des Spiegels (beispielsweise farbige, leicht zu extrahierende Streifen) beitragen. Die Qualität solcher Verfahren müsste eingehend untersucht werden.

Neben der Ermittlung der Spiegelposition bliebe weiterhin das Problem der Ermittlung korrespondierender Punkte offen. Da Verfahren zur Ermittlung korrespondierender Punkte in getrennten Bildern existieren, beschränkt sich dieses Problem wesentlich darauf, den Spiegelbildbereich vom eigentlichen Bildbereich zu trennen und das entsprechende Verfahren mit dem getrennten Bildbereich als Bild 1 einer Kamera 1 und den gespiegelten Bildbereich als Bild 2 einer Kamera 2 zu behandeln.

Der Aufwand einer 3D-Rekonstruktion mit Kamera und Spiegel ist bei abgeschlossenem Aufbau im Vergleich zu einer üblichen 3D-Rekonstruktion mit Hilfe zweier Kameras geringfügig unterschiedlich. Bei einem Stereoaufbau mit zwei Kameras muss die relative Position der zweiten Kamera ebenso bekannt sein wie die Position des Spiegels bei einem Kamera-Spiegel-Aufbau. In beiden Fällen kann man eine einmalige fixe Installation zueinander vornehmen. Bewegt man die fixe Installation als Ganzes, ist auch die Beobachtung eines dynamischen Objekts möglich - unter der Einschränkung eines kleineren gemeinsamen Sichtbereichs gegenüber einer üblichen Stereoinstallation.

Die extrinsische Kalibrierung über einen Spiegel ist aber deutlich aufwändiger als für ein übliches Stereokamerasystem. Zwei Kameras können mit deutlich geringerem Aufwand nebeneinander fixiert werden, so dass die extrinsische Kalibrierung überwiegend aus einer Verschie-

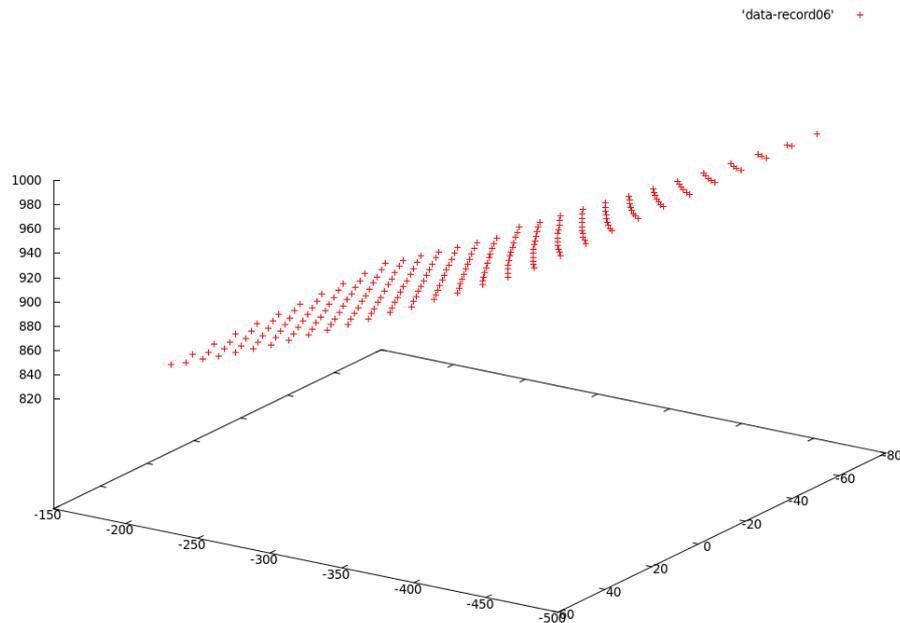


Abbildung 13: Seitlicher Blick auf die Punktwolke - im Idealfall würden die Punkte alle koplanar in einer Ebene liegen. Von der Seite ist aber eine leichte Beugung der Ebene zur Mitte hin zu erkennen, die auf eine Verzeichnung im Spiegel vermuten lassen könnte.

bung entlang einer Achse (idealerweise sogar einer Achse des Kamerakoordinatensystems) besteht. Dies wird häufig über ein gemeinsames Aufnahmeset von Kalibriertargets durchgeführt. Ein Spiegel muss zur Beobachtung jedoch vor und gegebenenfalls versetzt zur Hauptachse der Kamera positioniert werden. Daher muss eine Ermittlung der Spiegelposition erfolgen. Dies beinhaltet grundsätzlich eine Verschiebung und eine Rotation.

Im beschriebenen Beispiel dieser Evaluation handelt es sich aber wie bei einer Stereoinstallation mit zwei Kameras um eine einmalige Kalibrierung, weshalb der Aufwand für eine längerfristige Nutzung vergleichsweise gering ist.

Vorteilhaft gegenüber einer Stereorekonstruktion mit zwei Kameras ist die Tatsache, dass bei einer Kamera und einem Spiegel nur eine Kamerakalibrierung für intrinsische Parameter durchgeführt werden muss, da für den Spiegel mit einer *virtuellen Kamera* dieselben intrinsischen Kameraparameter angenommen werden.

Es zeigt sich also, dass die Anwendbarkeit für Beobachtungen nahe am Schärfentiefebereich an die Möglichkeiten einer üblichen Stereoinstallation heranreicht und sich der Aufwand wesentlich auf die Ermittlung einer Spiegelposition und das Finden von korrespondierenden Punkten beschränkt.

7. Fazit & Ausblick

Die vorliegende Arbeit legt dar, dass eine 3D-Rekonstruktion korrespondierender Punkte mit nur einer Kamera und einem Spiegel mit bekannten Maßen möglich ist. Außerdem wurde im Modell veranschaulicht, wie sich ein Spiegel relativ zur Kamera geeignet positionieren lässt, um das gewünschte Objekt zu beobachten. Dabei wurden Einschränkungen gegenüber einem üblichen Stereokamerasystem festgestellt. In der Praxis haben sich diese Einschränkungen durch die beschränkte Schärfentiefe nochmals erhöht.

Unter geeigneten Bedingungen lässt sich mit der aufgeführten Methode jedoch eine statische Installation vornehmen und eine zweite Kamera durch einen Spiegel ersetzen. Die Qualität der metrischen 3D-Rekonstruktion hängt dann maßgeblich von der Kalibrierung der Kamera und der extrinsischen Kalibrierung des Spiegels ab.

Werden der Spiegel und das beobachtete Objekt folglich geeignet positioniert, ist sowohl eine gute extrinsische Kalibrierung, als auch eine gute Extraktion markanter Punkte möglich. Dies führt zu Messergebnissen mit Millimeter-Genauigkeit und geringer Varianz. In diesem Fall ist die Verwendung eines Spiegels bspw. unter ökonomischen Gesichtspunkten bei der 3D-Rekonstruktion zu empfehlen.

In weiteren Untersuchungen könnten bessere Methoden zur Bestimmung der Spiegelposition evaluiert oder auch Untersuchungen zur Verbesserung der Präzision vorgenommen werden.

Wie in den Messergebnissen und schematischen 3D-Rekonstruktionen des beobachteten Schachbrettmusters zu erkennen ist, scheint auch der Spiegel Verzeichnungen zu verursachen. In der vorliegenden Arbeit wird eine Totalreflexion an einem ebenen Spiegel angenommen. Das Modell ließe sich aber analog zum Kameramodell unter Berücksichtigung von Verzeichnungen des Spiegels erweitern.

Neben der Gestaltung der Installation und der Präzision der Ergebnisse, ließe sich auch nach Methoden für die automatische Extraktion von markanten Punkten aus einem Bild und dessen Spiegelbild weiter forschen. Die Zuordnung von Punktkorrespondenzen auf einem extrahierten Schachbrettmuster im Bild und Spiegelbild musste für diese Arbeit getrennt vorgenommen und manuell zusammengeführt werden.

Im Sinne dieser weiteren Fragestellungen stellt die vorliegende Arbeit eine geeignete Grundlage für die Machbarkeit einer 3D-Rekonstruktion mit einem Spiegel dar.

A. Literatur

- [Can14a] Canon. Canon EOS 1000D - EOS DSLRs und kompakte Systemkameras, März 2014.
- [Can14b] Canon. Canon EOS 7D - EOS DSLRs und kompakte Systemkameras, März 2014.
- [Han11] Tobias Hanning. High Precision Camera Calibration. Vieweg+Teubner, 2011.
- [HZ04] Richard Hartley and Andrew Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2004.
- [IGC] Intel, Willow Garage, and OpenCV Community. OpenCV documentation <http://docs.opencv.org/>, zuletzt aufgerufen am 06. März 2014.
- [Kan93] Kenichi Kanatani. Geometric Computation for Machine Vision. Oxford Engineering Science Series, 1993.
- [Kov12] Emőd Kovács. Rotation about an arbitrary axis and reflection through an arbitrary plane. Annales Mathematicae et Informaticae, 40:175–186, 2012.
- [Ver06] R. K. Verma. Ray Optics. Discovery Publishing House, 2006.
- [Zha98] Zhengyou Zhang. A Flexible New Technique for Camera Calibration. Microsoft Research, 1998.

B. Abbildungs-, Tabellen- und Listenverzeichnis

Abbildungsverzeichnis

1.	2D-Szene	7
2.	2D-Szene mit geringerem Objektabstand	8
3.	Veränderung der Szene bei unterschiedlicher Objektentfernung	9
4.	Veränderung der Szene bei unterschiedlichem Spiegelwinkel	9
5.	Lochkamera	11
6.	Lochkameramodell	13
7.	Verzeichnung	15
8.	Scanbild eines Spiegels	20
9.	Aufnahme einer exemplarischen Szene	24
10.	Verschiebung der Spiegelebene	29
11.	Rekonstruiertes Schachbrett	31
12.	Histogramm zur Verteilung der Fehlerabstände zur approximierten Ebene	32
13.	Seitlicher Blick auf die Ebene des Schachbretts	33

Tabellenverzeichnis

1.	Beispiele für Verzeichnungsfunktionen	15
2.	Abstände der markierten Punkte in den Spiegelecken.	21
3.	Koordinaten des Spiegelkoordinatensystems.	21
4.	Korrespondierende Punkte im Bild	27
5.	Triangulierte Weltpunkte (Einheit durch das vorgegebene Koordinatensystem in mm)	28
6.	Vergleich der triangulierten Punkte bei Verschiebung der Spiegelebene um 35mm.	30

Listings

1.	Liste der Argumente für PlaneReconstruction	24
2.	Ausgabe der Verzeichnungsparameter und der Projektions- und Reprojektionsmatrix.	25
3.	Statistische Werte der Fehlerwerte der Projektion mit Verzeichnung	25
4.	Extrinsische Kalibrierung in der <i>OpenCV</i> und mit <i>libforwiss</i>	25
5.	Extrinsische Parameter der Kamera im Verhältnis zu den Markierungen auf dem Spiegel.	26
6.	Berechnung der Reflexion an der Ebene, bestimmt durch die extrinsischen Parameter	26
7.	Relative Position der virtuellen Kamera	27
8.	Aufrufe der Optimierungsverfahren zur Minimierung der Fehlerfunktion.	28

C. Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt und durch meine Unterschrift, dass ich die vorliegende Bachelorarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Ausführungen an Stellen der Arbeit, die dem Wortlaut oder Sinn anderer Werke entnommen wurden, sind als solche kenntlich gemacht und ihre Quellen vermerkt. Die Arbeit wurde weder in gleicher oder ähnlicher Form bereits einer anderen Prüfungsbehörde vorgelegt, noch anderswo veröffentlicht.

Passau, den 31. März 2014

Ort, Datum

Julian J. Stier

Julian J. Stier