



LEHRSTUHL FÜR INFORMATIK VII  
ROBOTIK UND TELEMATIK

*Bachelorarbeit*

# **Robuste Kamerakalibrierung für Unterwasser 3D-Rekonstruktion**

Joschka van der Lucht

April 2016

Gutachter: Prof. Dr. Andreas Nüchter  
Betreuer: Michael Bleier



## Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit der robusten Kamerakalibrierung für Unterwasser 3D-Rekonstruktion. Hierzu wurde ein 3D-Kalibrieraufbau und eine entsprechende Software entwickelt. Anstelle eines Schachbrettmusters wurden AprilTags verwendet, welche den Vorteil besitzen, dass eine Kalibrierung auch dann durchgeführt werden kann, wenn nicht die komplette Struktur im Bild ist. Als Modell für die Abbildungsfunktion der Kamera wurde ein Lochkameramodell mit radialer und tangentialer Verzeichnung ausgewählt. Die Ergebnisse der 3D-Struktur wurden mit den Kalibrieraufbauten auf Grund eines DIN A3 und DIN A4 Musters verglichen. In Experimenten konnte gezeigt werden, dass selbst mit einer relativ geringen Anzahl von Bildern mit der 3D-Struktur bereits eine stabile Kalibrierung erreicht werden kann, wohingegen dies mit einem planaren Muster nicht möglich ist. Die intrinsischen Parameter wurden durch eine Untersuchung des Rückprojektionsfehlers und Analyse von korrigierten Bildern verifiziert. Dazu wurden die Aufnahmen mit Hilfe der geschätzten Kameraparameter entzerrt. Um eine möglichst exakte Validierung der extrinsischen Parameter durchzuführen, wurde die Kamera an einem KUKA Manipulator befestigt und eine planare Trajektorie abgefahren. Bei einem Vergleich zwischen der Referenztrajektorie des Roboters und den aus den Kamerabildern geschätzten Posen ist eine sehr hohe Genauigkeit auf der x- und y-Achse zu erkennen, lediglich beim Abstand zwischen Kamera und Kalibrierstruktur, der z-Achse, ist ein geringerer Fehler zu beobachten.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Zielsetzung und wissenschaftlicher Beitrag . . . . .	2
1.3	Aufbau der Arbeit . . . . .	3
<b>2</b>	<b>Unterwasser Kamerakalibrierung</b>	<b>5</b>
2.1	State of the Art . . . . .	5
2.2	Kameramodelle . . . . .	8
2.2.1	Ideales Lochkameramodell . . . . .	9
2.2.2	Lochkameramodell mit radialer und tangentialer Verzeichnung . . . . .	11
2.2.3	Physikalisches Modell der Lichtbrechung . . . . .	12
2.3	Kalibrierung . . . . .	14
<b>3</b>	<b>Versuchsaufbau und Rahmenbedingungen</b>	<b>19</b>
3.1	Kalibriereraufbau . . . . .	19
3.2	Kalibrierung und Dedektierung der Tags . . . . .	22
3.2.1	Installation und Verwendung . . . . .	22
3.2.2	Funktionsweise . . . . .	25
<b>4</b>	<b>Experimente</b>	<b>27</b>
4.1	Analyse der Kalibrierstabilität . . . . .	28
4.2	Vergleich zwischen planaren und 3D-Kalibriermuster . . . . .	33
4.3	Validierung der Intrinsik . . . . .	36
4.4	Validierung der Extrinsik . . . . .	40
<b>5</b>	<b>Zusammenfassung und Ausblick</b>	<b>43</b>
<b>A</b>	<b>Zusätzliche Ergebnisse</b>	<b>45</b>
	<b>Literaturverzeichnis</b>	<b>55</b>



# Kapitel 1

## Einleitung

### 1.1 Motivation

In Europa wurden über die letzten Jahrhunderte die meisten, leicht zu erreichenden, Mineralvorkommen ausgeschöpft. Jedoch zeigen Schätzungen, dass es in Tiefen von 500-1000 Metern noch europäische Bodenschätze im Wert von circa 100 Milliarden Euro gibt [3]. Da sich diese jedoch meist in dicht besiedelten Gebieten, ehemaligen Bergwerken, als minderwertige Einlagerungen oder in kleineren Einlagen befinden, erweist sich größerer, konventioneller Bergbau hier als nicht rentabel. Aus diesem Grund müssen, um eine nachhaltige Versorgung mit Rohstoffen zu gewährleisten, neue Methoden entwickelt werden. ¡VAMOS! (Viable Alternative Mine Operating System) [7] ist ein Projekt, im Rahmen des Horizont 2020 Programms der Europäischen Union, welches sich mit der Entwicklung und dem Bau eines Unterwasserbergbauroboters beschäftigt, der in vier europäischen Minen getestet wird. Einen Entwurf dieses Systems zeigt Abbildung 1.1

Um dem Bediener ein möglichst intuitives Umfeld zu schaffen, wird ihm eine 3D-Karte des Untergrunds zur Verfügung gestellt werden. Hierfür ist es notwendig, den Untergrund mit Kameras und Lasern zu erfassen und daraus 3D-Informationen zu gewinnen. Dies ist jedoch nur möglich, wenn das Kamerasystem vorher stabil kalibriert wurde. Da ansonsten das Bildmaterial aufgrund von Verzerrungen durch die Linse, Lichtbrechung im Wasser und durch das Gehäuse der Kamera, zu große Fehler besitzt, die eine zuverlässige Rekonstruktion von 3D Tiefendaten unmöglich macht.

Damit eine solche Kalibrierung durchgeführt werden kann, ist es notwendig, die Kameraparameter genau zu schätzen. Im speziellen Fall der Anwendung unter Wasser, müssen hierbei auch die Brechungseffekte des Wassers berücksichtigt werden. Insbesondere ist hierbei der Einfluss des Kameragehäuses zu beachten, denn hier kommt es zu Lichtbrechungen zwischen Medien mit unterschiedlichen Brechungsindizes. Das Licht wird einmal an der Grenzfläche zwischen dem Wasser außen und Glas und ein weiteres Mal an der Grenzfläche zwischen Glas und der Luft im Gehäuse gebrochen.

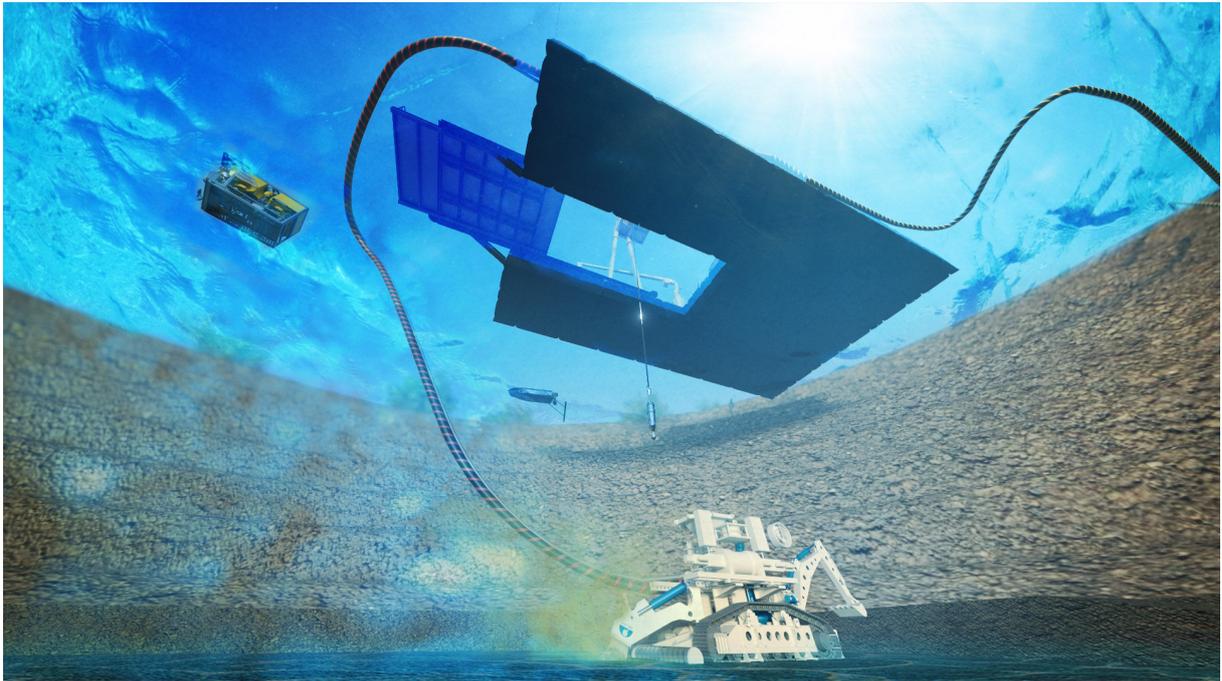


Abbildung 1.1: ¡VAMOS! Unterwasserbergbausystem (Rendering von DAMEN Dredging Equipment).

## 1.2 Zielsetzung und wissenschaftlicher Beitrag

Da es bereits Erkenntnisse darüber gibt, dass ein 3D-Kalibriermuster ein deutlich stabileres Ergebnis beim Kalibrieren von Kameras liefert als ein 2D-Muster, wie beispielsweise das herkömmliche Schachbrettmuster, wird im Rahmen dieser Arbeit eine solche 3D-Struktur zum Kalibrieren erstellt. Ziel der Arbeit ist es eine 3D-Struktur und eine Software zu erstellen mit der es möglich ist, eine Kamera in einem Unterwasserszenario möglichst schnell und einfach, jedoch auch genau, zu kalibrieren. Hierzu soll eine Kalibriertoolbox erstellt werden, die automatisiert alle Kalibrierpunkte erkennt und die intrinsischen und extrinsischen Parameter der Kamera bestimmt.

Die Ergebnisse sollen dann experimentell validiert und mit einer 2D-Struktur verglichen werden. Hierzu sollen in einem Testaufbau Bilder erstellt und ausgewertet werden und diese Ergebnisse im Anschluss mit den Erkenntnissen aus der Literatur [11, 12] verglichen werden. Hierbei soll im Speziellen die Stabilität der Kalibrierung analysiert und die Genauigkeit der intrinsischen und extrinsischen Parameter überprüft werden.

Im Laufe der Arbeit bestätigte sich, dass mit einer 3D-Struktur eine stabilere Schätzung der Parameter möglich ist. Durch den Einsatz von Kalibrierpunkten mit eindeutiger ID konnte die Kalibrieroutine vollständig automatisiert werden und Robustheit gegen beispielsweise teilweise Verdeckung erreicht werden.

## 1.3 Aufbau der Arbeit

Die vorliegende Arbeit beschäftigt sich mit der stabilen Kamerakalibrierung für Unterwasser 3D-Rekonstruktionen. Zu Beginn, in Kapitel 2, wird der aktuelle Forschungsstand anhand einiger Ergebnisse von Telem und Filin [19], Harvey und Shortis [11], Schewe et al. [16] und Menna et al. [14] aufgezeigt und deren unterschiedliche Methoden zur Kalibrierung kurz erläutert.

Darauf folgend werden die theoretischen Grundlagen zusammengefasst, welche für den weiteren Verlauf der Arbeit von Bestand sind. Hierzu wird das Lochkameramodell und ein physikalisches Modell, welches die Lichtbrechung am Gehäuse berücksichtigt, genauer erklärt. Dies wird im Folgenden dazu verwendet, um die von Zhang [21] entwickelte Methode zur Bestimmung der Kameraparameter mathematisch darzulegen.

Im dritten Kapitel wird der Aufbau des verwendeten Kalibrierungsmusters und die, im Rahmen der Arbeit entstandene, Software und deren Arbeitsweise kurz erläutert.

Kapitel 4 der Arbeit beschäftigt sich mit der Auswertung der durchgeführten Versuche. Hierzu wird zunächst der Rückprojektionsfehler zwischen Aufnahmen Unter- und Überwasser verglichen. Da dies jedoch kein gutes Maß für die Stabilität einer Kalibrierung ist, wird daraufhin anhand der geschätzten Brennweite über eine unterschiedliche Anzahl von Bildern bestimmt, ab wie vielen Bildern von einer stabilen Kalibrierung gesprochen werden kann. Anschließend werden mit der daraus gewonnenen Anzahl an Bildern Vergleiche zwischen der 3D-Struktur, einem DINA3 und einem DINA4 Muster durchgeführt. Dabei ist zu beobachten, dass eine Kalibrierung mit der 3D-Struktur deutlich stabilere Ergebnisse liefert, wohingegen die Schätzungen mit einem DINA3 oder DINA4 Muster großen Schwankungen unterliegen. Zuletzt wird hier noch eine feste Trajektorie mit Hilfe eines KUKA Manipulators abgefahren, um so die Genauigkeit der bestimmten extrinsischen Parameter zu verifizieren.



## Kapitel 2

# Unterwasser Kamerakalibrierung

Bei der Kamerakalibrierung unter Wasser gilt es einige zusätzlichen Herausforderungen zu bewältigen. Zunächst muss ein geeignetes Gewässer mit gutem Zugang gefunden werden, um dort Testaufnahmen angefertigt zu können. Hierbei werden jedoch Arbeitsbedingungen und auch die Ergebnisse vom Anteil der Schwebepartikel im Wasser und der damit unter Umständen beeinträchtigten Sicht beeinflusst. Ein weiterer Faktor, der die Sichtweite stark beeinträchtigt, ist die Lichtabsorption. Je tiefer sich die Kamera unter der Wasseroberfläche befindet, desto mehr Licht wird vom Wasser absorbiert und es wird gegebenenfalls notwendig, die Szene mit zusätzlichen Lichtquellen auszuleuchten. Dabei gilt es stets zu beachten, dass sämtliche Ausrüstung wasserdicht und dennoch verwendbar, verpackt werden muss. Dies führt im Fall der Kamera zu einem weiteren Problem, welches hier im Folgenden noch einmal genauer betrachtet wird, die Lichtbrechung am Übergang zwischen Wasser und dem Gehäuse der Kamera, sowie zwischen dem Gehäuse der Kamera und der darin eingeschlossenen Luft. Hierbei gibt es unterschiedliche Varianten, *Flat Port* (flaches Glas) oder *Dome Port* (gewölbtes Glas), welche sich unterschiedlich auf die Verzerrung auswirken und je nach verwendeter Kamera und verwendetem Gehäuse stark variieren. Hierbei ist auch die Lage der Kamera im Gehäuse entscheidend und ein Ein- und Ausbau kann zu einer Veränderung des optischen Pfads führen. Das heißt, die Lage der Wassergrenzfläche relativ zum Bildsensor verändert sich. Dies macht häufiges Nachkalibrieren der Kamera notwendig. Ein weiterer Punkt sind chromatische Abbildungsfehler, die aufgrund der verschieden starken Brechung von Lichtanteilen mit unterschiedlicher Wellenlänge auftreten.

### 2.1 State of the Art

Im Bereich der Unterwasser Kamerakalibrierung gibt es unterschiedliche Ansätze. Die drei am häufigsten anzutreffenden Methoden sind Absorption durch den physikalischen Standardparametersatz [11, 14, 16], geometrische Korrektur [12] und perspektivische Verzerrung [19]. Die Brechungseffekte können auch teilweise optisch korrigiert werden, indem die Kamera in ein Gehäuse mit einem so genannten *Dome Port*, ein Gehäuse mit einer halbkugelförmigen Glasscheibe, verpackt wird.

Wird die Kamera exakt im Zentrum der Halbkugel angebracht, so werden senkrecht zur Kugeloberfläche eintreffende Lichtstrahlen auf dem Weg zur Kamera nicht gebrochen (Annahme einer dünnen Glaskuppel). Damit scheint das Sichtfeld gleich zwischen über und unter Wasser. Doch auch diese Methode ist aus mehreren Gründen für die Anwendung in der 3D-Rekonstruktion nicht ohne Kalibrierung nutzbar. Zunächst ist es meist nicht möglich die Kamera exakt im Mittelpunkt der Halbkugel zu platzieren und es kommt somit zu Fehlern. Ebenso kann für eine Anwendung in größeren Tiefen keine solch dünne Glasscheibe verwendet werden, dass die Lichtbrechung an ihr vernachlässigt werden könnte. Zuletzt beeinflusst der *Dome Port* den Fokus und hat somit ähnliche Auswirkungen wie eine Linse. Somit ist ein unkalibrierter *Dome Port* für die hier gestellten Anforderungen unpraktikabel und wird im Folgenden nicht weiter betrachtet.

Die Absorptionsmethode wird beispielsweise von Harvey und Shortis [11], Schewe et al. [16] und Menna et al. [14] verwendet. Vorteil dieser Methode ist, dass Methoden und Algorithmen ohne Veränderung verwendet werden können, die sonst auch Überwasser benutzt werden. Dies ist jedoch zugleich auch ein kleiner Nachteil, da so nicht alle Auswirkungen des Wassers genau modelliert werden, sondern versucht wird, dies auch über die Parameter des Lochkameramodells abzufangen. Grundlegend wird davon ausgegangen, dass sich der Großteil der Abbildungsfehler durch eine radiale Verzeichnung um die optische Achse darstellen und korrigieren lassen. Die wenigen Fehler, die von dieser Korrektur nicht betroffen sind, werden durch die restlichen Kameraparameter absorbiert. Die zweite Methode, die geometrische Korrektur, wird von Li et al. [12] verwendet. Hierbei kann theoretisch, für ein bestimmtes Gehäuse, vor der Kamera, der korrekte Ausgleich berechnet werden. Die physikalischen Effekte werden hierbei durch ein 3D Raytracing Modell beschrieben. Dabei wird explizit die Doppelbrechung an Wasser-Glas und Glas-Luft modelliert. Dies ist jedoch sehr aufwendig, komplex und nur für jeweils diesen Aufbau gültig. Li et al. verwenden beispielsweise ein Gehäuse mit einer konkaven Linsenabdeckung. Die Kalibrierung findet in zwei Schritten statt. Zuerst wird nur die Kamera über Wasser kalibriert und somit die genauen Parameter der Kamera bestimmt. Anschließend wird eine Kalibrierung unter Wasser durchgeführt und somit die Parameter des Gehäuses und die zwei Brechungsindizes geschätzt. Die Brechungsindizes von Luft und Wasser können hierfür auch als bekannt angenommen werden.

Bei der, von Telem und Filin [19] verwendeten, perspektivischen Verzerrung ist es auch wieder möglich unter der Berücksichtigung aller Faktoren ein genaues Modell zur Korrektur aufzustellen. Der Grundgedanke dieses Ansatzes ist es, die gemessenen Bildpunkte so zu korrigieren, dass ein perspektivisches Projektionsmodell anwendbar ist. Jedoch ist auch dies wieder nur für ein ganz bestimmtes Gehäuse gültig und es werden einige physikalische Bedingungen vernachlässigt, die weiterhin durch die Standardparameter bei der Kalibrierung aufgenommen werden müssen. Telem und Filin haben sich speziell mit einem Gehäuse mit einer flachen Linsenabdeckung beschäftigt. Bei den von ihnen bestimmten Parametern handelt es sich um die physikalischen Standardparameter der Kamera, die Brechungsindizes von Glas und Wasser, dem Abstand zwischen dem Projektionszentrum der Kamera und der Glasfläche des Gehäuses, der Dicke dieser Fläche und dem Winkel zwischen optischer Achse und dem Gehäuse.

Ein wichtiger Punkt bei der Kalibrierung, unabhängig des verwendeten Modells und der Methode, ist das Objekt, welches hierfür verwendet wird. Häufig finden hier 2D-Schachbrettmuster Anwendung. Fast ebenso kritisch ist die Größe des Objektes. Es ist nach Möglichkeit das gesamte

Sichtfeld abzudecken oder zumindest sollten mehrere Aufnahmen im gesamten Sichtfeld gemacht werden. Ebenso muss die gesamte für die spätere 3D-Rekonstruktion relevante Tiefe des Raumes ausgenutzt werden, um eine stabile Kalibrierung zu erhalten.

Mark Shortis stellte bei seinen Arbeiten fest, dass die beste Möglichkeit, die Genauigkeit der Kalibrierung zu validieren und zu vergleichen, ist, ein Objekt bekannter Größe im Bild zu vermessen und die Abweichung zur realen Größe zu bestimmen [17].

Technik	Validierung	Relativer Fehler
Shortis und Harvey [11], Lochkameramodell	Längenmessungen von Stäben im gesamten Sichtfeld	0,2 – 0,7%
Petrell et al. 1997, geometrische Korrektur	Messungen von Lachs	1,5%
Harvey et al. 2003, Lochkameramodell	Messungen von Tunfisch	0,2%
Telem und Filin [19], perspektivische Verzerrung	Rekonstruktion von Linien, Messung von Referenzplatten	0,4%
Menna et al. [14], Lochkameramodell	Ähnlichkeitstransformation zwischen über und unter Wasser Netzwerken	0,3%

**Tabelle 2.1:** Relativer Fehler für verschiedene Methoden aus der Literatur [18].

An denen in Tabelle 2.1 aufgeführten Ergebnissen ist zu erkennen, dass sich alle drei Verfahren zur Kalibrierung eignen und selbst das relativ einfache Modell der Lochkamera dazu in der Lage ist, über die verschiedenen Parameter, die Lichtbrechung des Wassers am Gehäuse auszugleichen. Somit ist es nicht unbedingt notwendig ein komplexes physikalisches Modell zu erstellen und zu kalibrieren. Mit allen Verfahren kann ein relativer Fehler von maximal 1,5% erreicht werden. In dieser Arbeit wird der Ansatz die Absorptionsmethode verfolgt, da dessen Parameter im Vergleich zu den anderen Methoden einfacher zu schätzen sind.

In einem Versuch verglich Mark Shortis die Ergebnisse der Kalibrierung mit einem 3D-Würfel und einem DINA3 sowie einem DINA4 Schachbrettmuster. In Abbildung 2.1 sind die Fehler bei unterschiedlichen Entfernungen aufgetragen. Hier ist sehr gut zu sehen, dass der Fehler, unabhängig von der Entfernung, bei der Kalibrierung mit der 3D-Struktur unter 1% liegt. Mit dem DINA3 Schachbrettmuster steigt der Fehler für Entfernungen über 4m sehr schnell an. Die Ergebnisse mit einem DINA4 Schachbrettmuster erweisen sich als noch schlechter. Bereits bei 2m Entfernung ist der Fehler  $> 5\%$  und steigt bei 8m auf rund 20% an. Ebenfalls sind die Schwankungen der Fehler, bei gleichen Abständen, deutlich zu erkennen. Während dieser beim DINA4 Muster teils um 10% schwankt, ist dies bei dem, nur doppelt so großen, DINA3 Muster bei nur noch etwa 1 – 2%. Am genauesten ist auch hier wieder der 3D-Würfel, mit einer Abweichung von  $< 0,5\%$ .

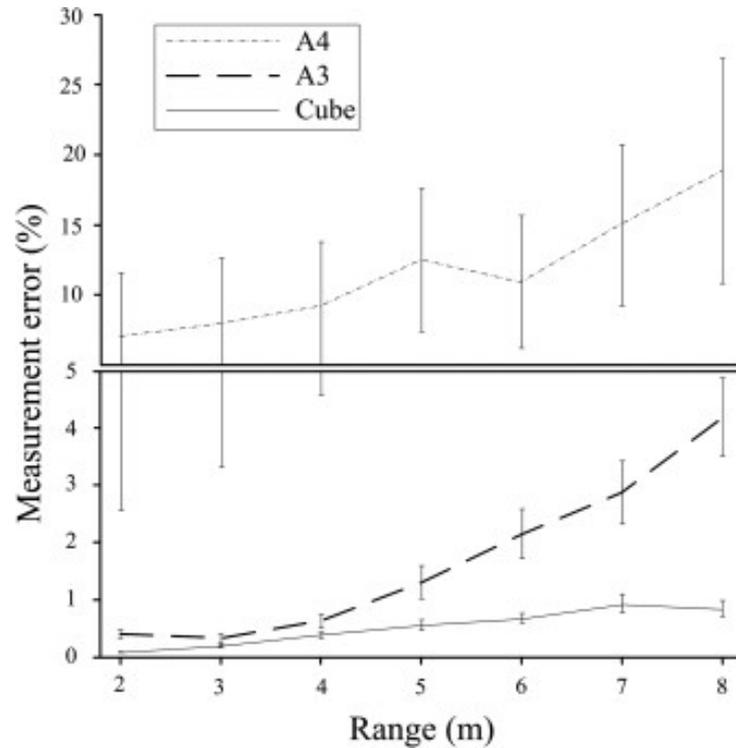


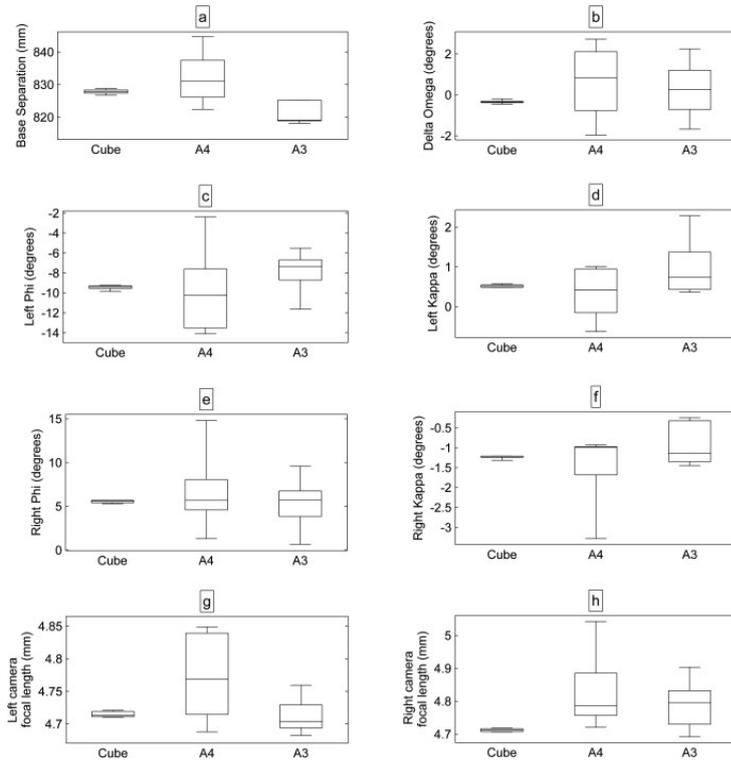
Abbildung 2.1: Messfehler in Abhängigkeit der Entfernung [9].

In der Grafik 2.2 ist ebenso zu erkennen, dass die unterschiedlichen Parameter, bei Verwendung des DIN A4 Schachbretts, sehr stark variieren. Während dieses Phänomen mit dem DIN A3 Brett schon deutlich nachlässt, sind sie jedoch mit dem Würfel nahezu eliminiert.

Es liegt somit nahe, dass eine Kalibrierung mit einer 3D-Struktur ein deutlich genaueres und stabileres Ergebnis liefert als ein konventionelles Schachbrettmuster.

## 2.2 Kameramodelle

Eine Kamera kann als eine Art Werkzeug zur Abbildung der 3D Welt auf ein 2D Foto betrachten. Hierbei treten jedoch häufig unerwünschte Effekte auf, welche das Bild von der Realität abweichen lassen. Diese können zum Beispiel eine Verzerrung durch die Lichtbrechung an der Linse sein oder auch durch Lichtbeugung am Verschluss entstehen. Um ein Foto nun für, wie in unserem Fall, die Rekonstruktion von 3D Informationen verwenden zu können, müssen diese Fehler, so weit wie möglich, beseitigt werden. Für diesen Vorgang werden jedoch einige Parameter der Kamera benötigt, wie beispielsweise die Brennweite. Da diese nun bestimmt oder abgeschätzt werden muss, wird hierfür ein geeignetes Kameramodell benötigt, welches die reale Abbildungsfunktion möglichst genau beschreibt und gleichzeitig möglichst einfach auf verschiedenen Kameras angewendet werden kann.



**Abbildung 2.2:** Schwankungen der einzelnen Kalibrierungsparameter mit den unterschiedlichen Kalibrieremustern [9].

### 2.2.1 Ideales Lochkameramodell

Die projektive Abbildung einer Kamera lässt sich in erster Näherung durch das Lochkameramodell bezeichnen. In der Abbildung 2.3 ist ein solches Modell zu sehen. Hierbei stellt  $C$  den Kameramittelpunkt und  $P$  den Bildmittelpunkt dar. Betrachtet man nun die Ebene auf der  $Z$ -Achse mit dem Abstand der Brennweite  $f$  zu  $C$ , auch Bildebene genannt, wird ein Punkt mit den Weltkoordinaten  $\mathbf{X} = (X, Y, Z)^T$  auf die Stelle der Bildebene  $\mathbf{x} = (x, y)$  abgebildet, an der die Gerade vom Kameramittelpunkt  $C$  die Bildebene schneidet. Der Punkt  $(X, Y, Z)^T$  wird somit auf  $(\frac{fX}{Z}, \frac{fY}{Z}, f)^T$  auf der Bildebene abgebildet. Lässt man nun die letzte Koordinate der Bildebene außen vor, so ergibt sich eine Abbildung von  $\mathbb{R}^3 \mapsto \mathbb{R}^2$ .

$$(X, Y, Z)^T \mapsto \left(\frac{fX}{Z}, \frac{fY}{Z}\right)^T \quad (2.1)$$

Der Mittelpunkt der Projektion wird Kameramittelpunkt genannt und die senkrechte Gerade von diesem zur Bildebene Hauptachse oder optische Achse (*principal axis*). Der Punkt an dem die Hauptachse die Bildebene schneidet wird als Bildmittelpunkt bezeichnet. Werden die Welt- und Bildkoordinaten nun als homogene Vektoren angesehen, so kann die Projektion aus (2.1) als

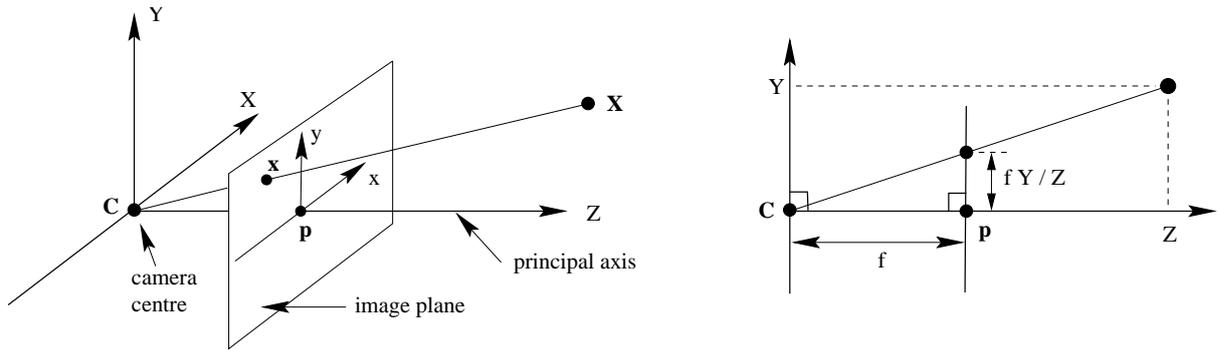


Abbildung 2.3: Lochkameramodell [10, S. 154].

Abbildung mit einer Matrixmultiplikation dargestellt werden:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & 0 \\ f & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}. \quad (2.2)$$

Die Gleichung (2.2) setzt nun aber voraus, dass der Bildmittelpunkt im Ursprung der Bildebene liegt, da dies jedoch in der Praxis nicht der Fall ist, muss daher die Abbildung folgendermaßen angepasst werden

$$(X, Y, Z)^T \mapsto \left( \frac{fX}{Z + p_x}, \frac{fY}{Z + p_y} \right)^T, \quad (2.3)$$

mit  $(p_x, p_y)^T$  als Koordinaten für den Bildmittelpunkt. Werden dies nun auch wieder als homogene Koordinaten betrachtet, ergibt sich folgender Zusammenhang:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & p_x & 0 \\ f & p_y & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (2.4)$$

Daraus ergibt sich nun die so genannte Kameramatrix  $A$ , welche die intrinsischen Parameter enthält,

$$K = \begin{bmatrix} f & p_x & 0 \\ f & p_y & 0 \\ 1 & 0 & 0 \end{bmatrix}. \quad (2.5)$$

Diese wird später für die Kalibrierungen benötigt.

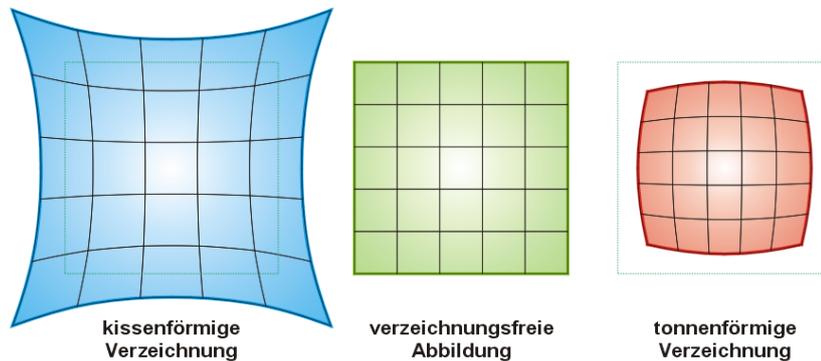
### 2.2.2 Lochkameramodell mit radialer und tangentialer Verzeichnung

Da Verzeichnungen einen großen Einfluss auf die Fehler im Bild haben, wird das ideale Lochkameramodell um radiale und tangentiale Verzeichnungsparameter erweitert.

Bei der radialen Verzeichnung handelt es sich um einen Effekt, der entsteht, wenn parallel eintreffende Lichtstrahlen nicht im selben Punkt konvergieren. Unter Wasser wird dieser Effekt auch durch eine Brechung des Wassers beeinflusst. Hierbei wird das eintreffende Licht rotationssymmetrisch um die optische Achse verzerrt. Hierbei gibt es zwei unterschiedliche Varianten, siehe auch Abbildung 2.4. Bei der kissenförmigen Verzeichnung werden Punkte, weiter vom optischen Zentrum entfernt projiziert. Werden die Punkte in Richtung des Bildmittelpunktes verschoben, wird von einer tonnenförmigen Verzeichnung gesprochen. Folgendes vereinfachtes Modell wird zur Beschreibung der radialen Verzeichnung verwendet:

$$\begin{aligned} x_{\text{corrected}} &= x(1 + k_1r^2 + k_2r^4 + k_3r^6) \\ y_{\text{corrected}} &= y(1 + k_1r^2 + k_2r^4 + k_3r^6) \quad , \end{aligned} \quad (2.6)$$

wobei  $r$  der Abstand zum Bildmittelpunkt ist.



**Abbildung 2.4:** Radiale Verzeichnung [6].

Bei der tangentialen Verzeichnung werden die Punkte nicht radial um die optische Achse verzerrt, sondern entlang einer Tangente um das optische Zentrum. Hierbei kann sich, entgegen zur radialen Verzeichnung, der Richtungsvektor verändern. Für die meisten Kamerasysteme spielt dieser Fehler jedoch eine eher untergeordnete Rolle, da er meist sehr klein ist und somit kaum ins Gewicht fällt. Es lässt sich durch Anwendung folgender Gleichung korrigieren:

$$\begin{aligned} x_{\text{corrected}} &= x[2p_1xy + p_2(r^2 + 2x^2)] \\ y_{\text{corrected}} &= y[p_1(r^2 + 2y^2) + 2p_2xy] \quad . \end{aligned} \quad (2.7)$$

Die exakte Bestimmung dieser Parameter wird im Abschnitt 2.3 Kalibrierung erläutert.

### 2.2.3 Physikalisches Modell der Lichtbrechung

Eine weitere Möglichkeit ist ein physikalisches Modell zu erstellen, das die Lichtbrechung zwischen den unterschiedlichen Medien berücksichtigt. Hier wird das von Treibitz et al. [20] verwendete Modell gezeigt, welches sich speziell auf ein Gehäuse mit flachem und vernachlässigbar dünnem Glas bezieht. Da die späteren Experimente in dieser Arbeit ebenfalls mit einem solchen Gehäuse durchgeführt werden, wurde dieses ausgewählt.

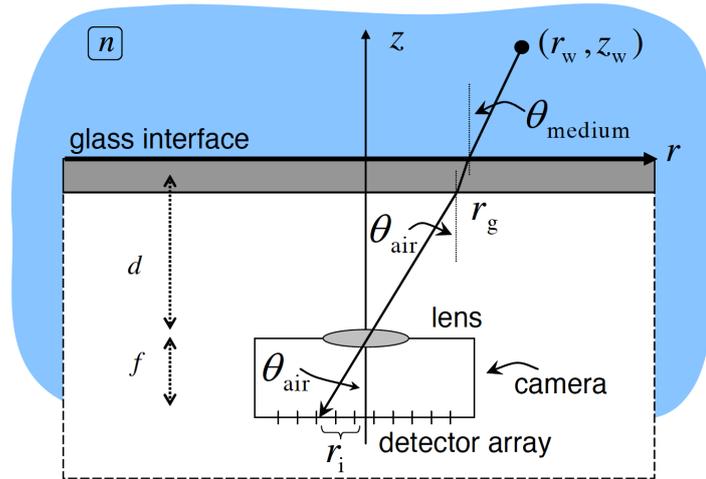


Abbildung 2.5: Brechung eines Lichtstrahls an der Grenzfläche Wasser-Glas und Glas-Luft [20].

In Abbildung 2.5 sind die geometrischen Zusammenhänge für die Brechung eines Lichtstrahls an der Glasfläche abgebildet. Hierbei beschreibt  $z$  die optische Achse,  $(r_w, z_w)$  die Objektkoordinaten und  $\theta$  den jeweiligen Einfallswinkel des Lichts im Medium, relativ zur  $z$ -Achse. So ergibt sich nach dem Snelliusschen Brechungsgesetz [8], mit  $n$  als Brechungsindex des entsprechenden Mediums relativ zur Luft, folgende Gleichung:

$$n \sin \theta_{\text{medium}} = n_{\text{glass}} \sin \theta_{\text{glass}} = \sin \theta_{\text{air}} \quad . \quad (2.8)$$

Damit hat die Glasfläche keinen Einfluss auf den Brechungswinkel  $\theta_{\text{air}}$  für einen gegebenen Einfallswinkel  $\theta_{\text{medium}}$ . Die Verschiebung des Austrittspunktes ist bei dünnem Glas sehr klein. Somit lassen sich die Auswirkungen des Glases vernachlässigen und es gilt lediglich die Effekte des anderen Mediums, in diesem Fall des Wassers, zu korrigieren.

Ein Lichtstrahl, eines Objektpunktes  $(r_w, z_w)$ , wird auf der Bildebene (in Abb. 2.5 *detector array*) abgebildet, dieser trifft an dem Punkt  $r_g$  auf die Glasfläche auf. Dieser Abstand  $r_g$  von der  $z$ -Achse kann mit dem fermatschen Prinzip berechnet werden. Danach ist der Weg des Lichts, zwischen zwei Punkten, jener, auf dem es die wenigste Zeit benötigt. Bezüglich  $n$  und der damit geringeren Geschwindigkeit im Medium Wasser, ergibt sich für die optische Pfadlänge

$$L = n\sqrt{(r_w - r_g)^2 + z_w^2} + \sqrt{r_g^2 + d^2} \quad , \quad (2.9)$$

wobei  $d$  den Abstand zwischen der Glasfläche und dem Linsenmittelpunkt darstellt. Somit erfüllt  $r_g$  folgende Bedingung

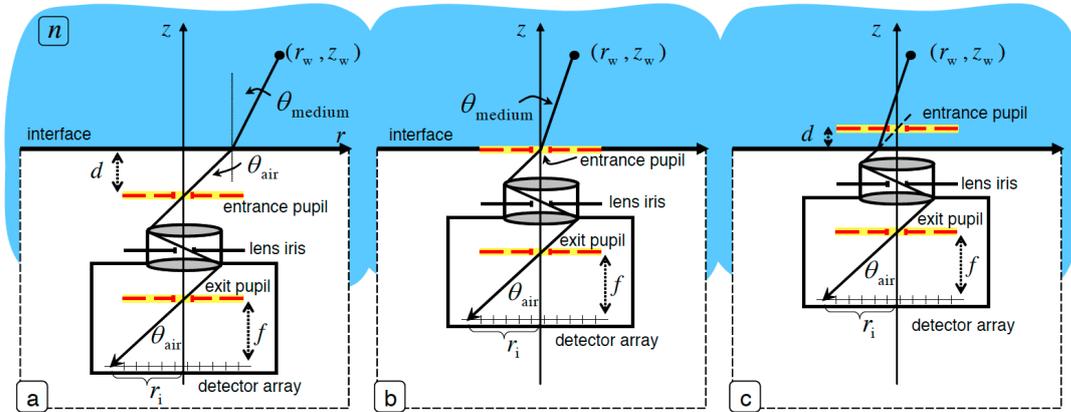
$$0 = \frac{\partial L}{\partial r_g} = n \frac{r_g - r_w}{\sqrt{(r_w - r_g)^2 + z_w^2}} + \frac{r_g}{\sqrt{r_g^2 + d^2}} \quad , \quad (2.10)$$

wobei zu beachten ist, dass das Maximum für  $L = \infty$  ist und somit der endliche Extremwert von  $L$  die minimale Pfadlänge beschreibt. Des Weiteren gilt für den Lichtstrahl, nachdem er durch das Glas in der Luft angekommen ist und auf die Linse trifft:

$$r_g = \frac{dr_i}{f} \quad . \quad (2.11)$$

Dabei beschreibt  $f$ , wie auch bei der Lochkamera, die Brennweite. Mit den Gleichungen (2.10) und (2.11) lässt sich nun eine weitere Gleichung aufstellen, welche die Abbildung eines Punktes  $(r_w, z_w)$  auf einen Punkt auf der Bildebene mit den Bildkoordinaten  $r_i$  beschreibt

$$\left(r_w - \frac{d}{f}r_i\right)^2 \left[\left(\frac{fn}{r_i}\right)^2 + (n^2 - 1)\right] = z_w^2 \quad . \quad (2.12)$$



**Abbildung 2.6:** Zeigt die drei unterschiedlichen Möglichkeiten, wo die Eintrittsstelle des Lichtstrahls liegen kann. In (a) befindet sie sich in der Luft, (b) zeigt sie auf der Grenzfläche und (c) im Wasser [20].

Das in Abbildung 2.5 dargestellte Modell wird nun in Abbildung 2.6(a) erweitert, da sich bei der Linse nicht mehr um eine flache Linse handelt, sondern um eine Verbundlinse. Diese wird hier als Blackbox mit zwei Ebenen behandelt, einer Eintrittspupille und einer Austrittspupille. Der Hauptstrahl ist ein Lichtstrahl, der stets die Mitte der Blendenöffnung (Iris) passiert. Dieser

kreuzt beim Eintritt und beim Austritt in gleicher Richtung und gleichem Winkel die Hauptachse  $z$ . Aus diesem Grund wird hier für den Abstand  $d$  für die Distanz zwischen der Grenzfläche und der Eintrittspupille gemessen. Somit gilt auch hier weiterhin die Gleichung (2.12).

Befindet sich die Eintrittspupille direkt auf der Grenzfläche, wie in Abbildung 2.6b, gilt  $d = 0$  und die Richtung des Strahls wird durch die Brechung am Glas beeinflusst. Für sehr kleine Winkel  $\theta_{\text{air}}$  zeigt die Gleichung (2.8), dass  $\theta_{\text{medium}} \approx \theta_{\text{air}}/n$  gilt. Da  $n > 1$  ist, sind die Winkel im Medium kleiner und es scheint so, als ob die Kamera eine größere Brennweite hat:

$$f_{\text{medium}}^{\text{effective}} = nf \quad (2.13)$$

Somit gilt für eine perspektivische Kamera, dass das gesamte System perspektivisch ist und daher folgendes gilt

$$r_i^{\text{persp}} = r_w \frac{f_{\text{medium}}^{\text{effective}}}{z_w} \quad (2.14)$$

Die lineare Approximation aus den Gleichungen (2.13) und (2.14) funktioniert nicht mehr, wenn  $\theta_{\text{air}}$  zu groß wird. Dann kann die nichtlineare Relation der Strahlen, mit der Gleichung (2.8), als radiale Verzeichnung behandelt werden.

Hier nun eine kompakte Lösung, für eine durch Brechung verursachte Verzeichnung. Mit Hilfe der Gleichung (2.12) und  $d = 0$  ergibt sich eine direkte Beziehung zwischen  $r_i$  und  $r_w$

$$r_i = fn[(z_w/r_w)^2 - (n^2 - 1)]^{-\frac{1}{2}} \quad (2.15)$$

Gemeinsam mit der Gleichung (2.14) kann nun eine Funktion zur Korrektur der Verzeichnung erstellt werden

$$r_i^{\text{persp}} = f_{\text{medium}}^{\text{effective}} [n^2 - 1 + (f_{\text{medium}}^{\text{effective}}/r_i)^2]^{-\frac{1}{2}} \quad (2.16)$$

Dies funktioniert jedoch nur für den Spezialfall  $d = 0$ . Für alle anderen  $d$  ist die Berechnung deutlich aufwendiger.

Die letzte Möglichkeit ist in Abbildung 2.6c zu erkennen. Hier befindet sich die Eintrittspupille außerhalb der Grenzfläche, nämlich im Medium, hier Wasser. In diesem Fall lässt sich die Gleichung (2.12) mit einem negativen  $d$  verwenden.

## 2.3 Kalibrierung

Eines der populärsten Verfahren zur Kalibrierung von Kameras wurde von Zhengyou Zhang entwickelt [21]. Da dies in OpenCV implementiert ist und somit später in dieser Arbeit verwendet wird, werden hier die theoretischen Grundlagen dieses Verfahrens dargestellt.

Diese Methode benutzt als Grundlage das Modell der Lochkamera mit der Kameramatrix

$$A = \begin{bmatrix} f_x & \gamma & u_0 \\ & f_y & v_0 \\ & & 1 \end{bmatrix}, \quad (2.17)$$

wobei  $\gamma$  die Verzerrung der beiden Bildachsen darstellt und der Gleichung

$$s\tilde{m} = A[R \ t]\tilde{M} \quad , \quad (2.18)$$

für die Abbildung eines 3D Punktes  $\tilde{M}$  in Weltkoordinaten auf einen 2D Punkt  $\tilde{m}$  mit einem beliebigen Skalierungsfaktor  $s$  und  $(R, t)$  als die extrinsischen Parameter der Kamera. Im Folgenden wird  $\tilde{m} = [u, v, 1]^T$  stets als Notation für die homogenen Koordinaten von  $m = [u, v]^T$  verwendet. Nehmen wir nun an, dass die Objektebene auf  $Z = 0$  des Weltkoordinatensystems liegt, so ergibt sich aus (2.18) mit  $r_i$  als  $i$ -te Spalte der Rotationsmatrix  $R$

$$\begin{aligned} s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= A[r_1 \ r_2 \ r_3 \ t] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} \quad . \\ &= A[r_1 \ r_2 \ t] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \end{aligned} \quad (2.19)$$

Da  $Z = 0$  ist, gilt hier nun  $M = [X, Y]^T$  und daher  $\tilde{M} = [X, Y, 1]^T$ . Damit wird ein Objektpunkt  $M$  nun durch die Homographie  $H$  auf den Punkt  $m$  abgebildet

$$s\tilde{m} = H\tilde{M} \quad \text{mit} \quad H = A[r_1 \ r_2 \ t] \quad . \quad (2.20)$$

Existiert nun ein Bild der Objektebene kann daraus die Homographie  $H = [h_1 \ h_2 \ h_3]$  bestimmt werden. Mit der Gleichung (2.20) folgt daraus

$$[h_1 \ h_2 \ h_3] = \lambda A[r_1 \ r_2 \ t]$$

mit  $\lambda$  als beliebigen Skalierungsfaktor. Wird nun die Eigenschaft der Rotationsmatrix  $R$  verwendet, dass  $r_1$  und  $r_2$  orthonormal sind, ergibt sich daraus

$$h_1^T A^{-T} A^{-1} h_2 = 0 \quad (2.21)$$

$$h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2 \quad . \quad (2.22)$$

Dies sind zwei grundlegende Beschränkungen für die, durch die intrinsischen Parameter, gegebene Homographie. Da die Homographie acht Freiheitsgrade besitzt und wir sechs extrinsische Parameter haben (drei für die Rotation und drei für die Translation), können daraus nur zwei Einschränkungen für die intrinsischen Parameter bestimmt werden.

Eine erste analytische Lösung für das Problem der Kalibrierung ist folgende. Nehmen wir an, dass

$$\begin{aligned}
B &= A^{-T} A^{-1} \equiv \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{f_x^2} & -\frac{\gamma}{f_x^2 f_y} & \frac{v_0 \gamma - u_0 f_y}{f_x^2 f_y} \\ -\frac{\gamma}{f_x^2 f_y} & \frac{\gamma^2}{f_x^2 f_y^2} + \frac{1}{f_y^2} & -\frac{\gamma(v_0 \gamma - u_0 f_y)}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} \\ \frac{v_0 \gamma - u_0 f_y}{f_x^2 f_y} & -\frac{\gamma(v_0 \gamma - u_0 f_y)}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} & \frac{(v_0 \gamma - u_0 f_y)^2}{f_x^2 f_y^2} + \frac{v_0^2}{f_y^2} + 1 \end{bmatrix} .
\end{aligned} \tag{2.23}$$

Zu Beachten ist hierbei, dass  $B$  symmetrisch ist und durch den 6D Vektor

$$b = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T , \tag{2.24}$$

definiert ist. Wird nun der  $i$ -ten Spaltenvektor von  $H$ ,  $h_i = [h_{i1}, h_{i2}, h_{i3}]^T$ , genommen, so ergibt sich

$$h_i^T B h_j = v_{ij}^T , \tag{2.25}$$

wobei

$$v_{ij} = [h_{i1} h_{j1}, h_{i1} h_{j2} + h_{i2} h_{j1}, h_{i2} h_{j2}, h_{i3} h_{j1} + h_{i1} h_{j3}, h_{i3} h_{j2} + h_{i2} h_{j3}, h_{i3} h_{j3}]^T$$

gilt. Mit Hilfe der beiden Gleichungen (2.21), (2.22) und dem Vektor  $b$  aus (2.24) kann dies, unter Beachtung der gegebenen Homographie, in folgende Form umgeschrieben werden

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{12})^T \end{bmatrix} b = 0 . \tag{2.26}$$

Betrachtet man nun  $n$  Bilder der Objektebene, indem  $n$  mal die Gleichung (2.26) in einer Matrix zusammenfasst wird, so entsteht daraus eine Gleichung mit einer  $2n \times 6$  Matrix  $V$

$$Vb = 0 . \tag{2.27}$$

Gilt  $n = 1$ , so können lediglich zwei der intrinsischen Parameter bestimmt werden, die Restlichen müssen vorher bekannt sein oder passend geschätzt werden. Um eine Lösung für  $n = 2$  zu bestimmen, kann zusätzlich gefordert werden, dass  $\gamma = 0$  gelten muss. Für alle  $n \geq 3$  kann, bis auf einen Skalierungsfaktor, eine eindeutige Lösung berechnet werden. Diese ist der Eigenvektor von  $V^T V$  mit dem kleinsten Eigenwert. Damit lässt sich nun die Matrix  $B$  aufstellen und daraus alle für  $A$  benötigten Werte berechnen.

Mit (2.23) kann

$$\begin{aligned}
v_0 &= (B_{12} B_{13} - B_{11} B_{23}) / (B_{11} B_{22} - B_{12}^2) \\
\lambda &= B_{33} - [B_{13}^2 + v_0 (B_{12} B_{13} - B_{11} B_{23})] / B_{11} \\
f_x &= \sqrt{\lambda / B_{11}} \\
f_y &= \sqrt{\lambda B_{11} / (B_{11} B_{22} - B_{12}^2)} \\
\gamma &= B_{12} f_x^2 f_y / \lambda \\
u_0 &= \gamma v_0 / f_y - B_{13} f_x^2 / \lambda
\end{aligned} \tag{2.28}$$

gewonnen werden und daraus lassen sich nun, mit Hilfe vom (2.20), die extrinsischen Parameter bestimmen

$$\begin{aligned} r_1 &= \lambda A^{-1} h_1 \\ r_2 &= \lambda A^{-1} h_2 \\ r_3 &= r_1 \times r_2 \\ t &= \lambda A^{-1} h_3 \end{aligned} \quad , \quad (2.29)$$

dabei gilt  $\lambda = 1/\|A^{-1}h_1\| = 1/\|A^{-1}h_2\|$ .

Eine andere Möglichkeit, dieses Problem zu lösen, ist eine nichtlineare Optimierungstechnik auf Basis der Maximum-Likelihood-Methode zu verwenden. Da die analytische Lösung sich der Minimierung des algebraischen Abstands bedient, dies aber keinerlei physikalisch Bedeutung besitzt, lässt sich dies durch Minimierung des Rückprojektionsfehlers verbessern.

Hat man nun  $n$  Bilder der Objektebene mit  $m$  Punkten, von denen einige durch Rauschen beschädigt sein können, muss hierzu folgende Funktion minimiert werden:

$$\sum_{i=1}^n \sum_{j=1}^m \|m_{ij} - \hat{m}(A, R_i, t_i, M_j)\|^2 \quad . \quad (2.30)$$

Hierbei stellt  $\hat{m}(A, R_i, t_i, M_j)$  die Abbildung des Punktes  $M_j$  im Bild  $i$  (vgl. (2.20)) dar.  $R$  bezeichnet die Rotation, als einen Vektor mit drei Parametern  $r$ . Dieser ist parallel zur Rotationsachse und der Winkel wird durch dessen Länge repräsentiert. Da es sich bei der Minimierung um ein nichtlineares Problem handelt, wird dies mit dem Levenberg-Marquardt-Algorithmus gelöst. Für diesen ist es notwendig eine initiale Schätzung für  $A$  und  $\{R_i, t_i | i = 1..n\}$  zu bestimmen. Dies kann zum Beispiel über das vorher beschriebene analytische Verfahren geschehen.

Bisher wurde die Verzeichnung durch die Linse der Kamera völlig vernachlässigt, da dies jedoch ein Faktor ist, der für gewöhnlich nicht zu vernachlässigen ist, gilt es nun, diesen noch in das bereits beschriebene System zu integrieren.

Nehmen wir hierzu  $(u, v)$  als die optimalen, nicht verzeichneten Bildkoordinaten und  $(\check{u}, \check{v})$  die entsprechenden Bildkoordinaten mit Verzeichnung. Nun können, mit  $(x, y)$  als nicht verzeichnete und  $(\check{x}, \check{y})$  als verzeichnete normalisierte Bildkoordinaten, sowie  $k_1$  und  $k_2$  als Koeffizienten der Verzeichnung, die Gleichungen

$$\begin{aligned} \check{x} &= x + x[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \\ \check{y} &= y + y[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \end{aligned}$$

beschrieben werden. Wird nun weiter angenommen, dass der Mittelpunkt der Verzeichnung auf dem Bildmittelpunkt liegt und  $\gamma = 0$  gilt, so können mit  $\check{u} = u_0 + f_x \check{x} + \gamma \check{y}$  und mit  $\check{v} = v_0 + f_y \check{y}$  folgende Gleichungen aufgestellt werden

$$\begin{aligned} \check{u} &= u + (u - u_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \\ \check{v} &= v + (v - v_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \end{aligned} \quad . \quad (2.31)$$

Ist zu erwarten, dass die Verzerrung nur sehr gering ist, kann diese auch ignoriert werden und eine Kalibrierung nur über eine Schätzung der anderen fünf intrinsischen Parameter durchgeführt werden. Eine weitere Möglichkeit ist es, erst alle anderen Parameter zu schätzen und daraus  $k_1$  und  $k_2$  zu bestimmen, um damit die optimalen Koordinaten  $(u, v)$  zu bestimmen. Mit Hilfe der Gleichungen (2.31) können nun zwei neue Gleichungen für jeden Punkt jedes Bildes aufgestellt werden

$$\begin{bmatrix} (u - u_0)(x^2 + y^2) & (u - u_0)(x^2 + y^2)^2 \\ (v - v_0)(x^2 + y^2) & (v - v_0)(x^2 + y^2)^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} \check{u} - u \\ \check{v} - v \end{bmatrix} .$$

Werden nun wieder alle  $m$  Punkte der  $n$  Bilder genommen, kann aus den daraus entstehenden Gleichungen eine Matrix der Form  $Dk = d$  mit  $k = [k_1, k_2]^T$  erstellt werden. Eine lineare Lösung hierfür ist gegeben durch

$$k = (D^T D)^{-1} D^T d . \quad (2.32)$$

Wurde  $k_1$  und  $k_2$  so bestimmt, können mit der Gleichung (2.30), in dem  $\hat{m}(A, R_i, t_i, M_j)$  durch (2.31) ersetzt wird, die übrigen Werte berechnet werden.

Da dies jedoch eine sehr hohe Laufzeit hat, werden die Parameter der Verzerrung der Gleichung (2.30) hinzugefügt. Dadurch wird folgende Gleichung erhalten, welche es zu minimieren gilt:

$$\sum_{i=1}^n \sum_{j=1}^m \|m_{ij} - \hat{m}(A, k_1, k_2, R_i, t_i, M_j)\|^2 . \quad (2.33)$$

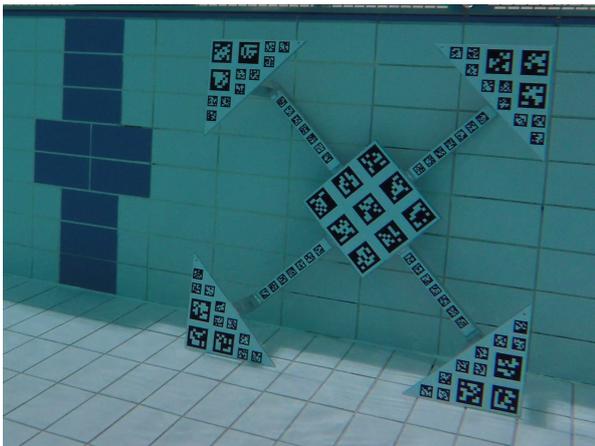
Auch hierbei ist  $\hat{m}(A, k_1, k_2, R_i, t_i, M_j)$  die Abbildung des Punktes  $M_j$  im Bild  $i$  mit der Verzerrung aus (2.31). Wie in (2.30) handelt es sich hierbei um ein nichtlineares Problem und daher wird auch hier, zur Lösung der Gleichung, der Levenberg-Marquardt-Algorithmus eingesetzt. Die benötigte initiale Schätzung von  $k_1$ ,  $k_2$ ,  $A$  und  $\{R_i, t_i | i = 1..n\}$  geschieht genau wie bereits vorher beschrieben.

Da die von Zhang entwickelte Methode nur für planare Kalibriermuster geeignet ist, muss für eine 3D-Struktur eine initiale Schätzung der Parameter mit Hilfe einer, in einer Ebene liegenden, Teilmenge der Kalibrierpunkte vorgenommen werden. Mit dieser Schätzung kann dann die Gleichung (2.33) mit den gesamten 3D-Punkten der Struktur verwendet werden, um die initiale Schätzung zu optimieren.

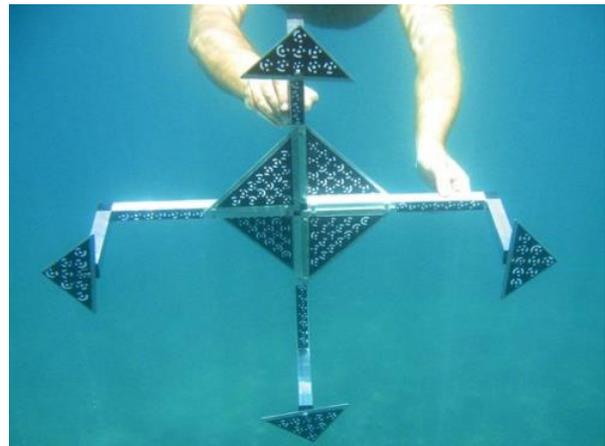
## Kapitel 3

# Versuchsaufbau und Rahmenbedingungen

### 3.1 Kalibrier Aufbau



(a) Im Rahmen der Arbeit erstelltes Kalibriermuster



(b) Kalibriermuster von Fabio Menna [13]

**Abbildung 3.1**

Da die Literatur darauf hinweist, dass 3D-Strukturen ein stabileres Ergebnis liefern, wurde ein eigenes konstruiert und gebaut. Hierbei wurde beachtet, dass es für den Einsatz unter Wasser geeignet ist. Dies bedeutet, es musste sowohl wasserfest, als auch formstabil und handhabbar sein. Angelehnt an eine Struktur von Fabio Menna [13], siehe Abbildung 3.1(b), entstand so die in Abbildung 3.1(a) und 3.2 zu sehende Struktur, welche später aus Aluminium gefertigt wurde.

Hinsichtlich der Markierungen auf dem Gestell, fiel die Wahl auf die von Edwin Olson ent-

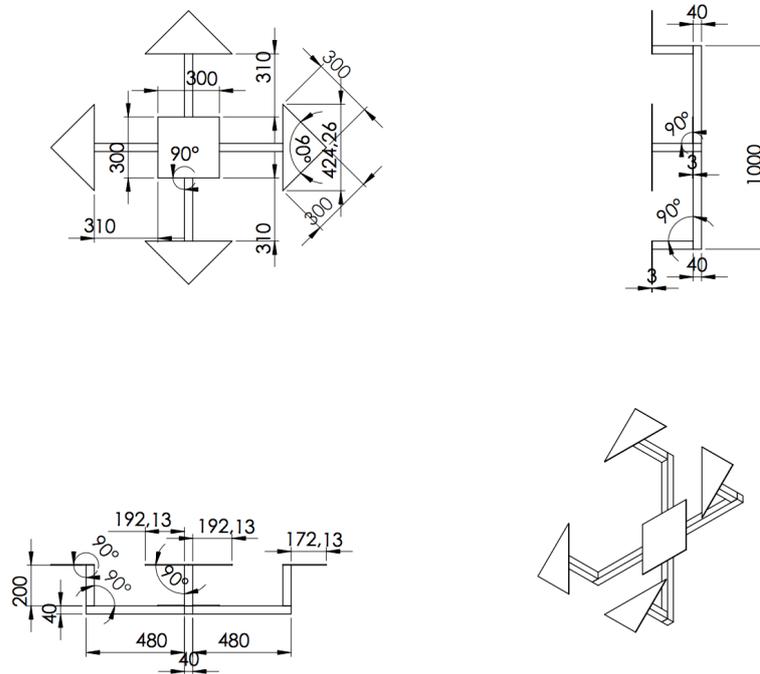


Abbildung 3.2: Konstruktionszeichnung des Kallibrieraufbaus.

wickelten AprilTags [15] der Tagfamilie *36h11*. Der Vorteil gegenüber einem herkömmlichen Schachbrettmuster ist die Möglichkeit, jeder Koordinate eine eigene und eindeutige Id zuzuweisen. Wodurch es möglich ist, jede Bildkoordinate zweifelsfrei einer Objektkoordinate zuzuordnen, auch wenn die Struktur nur teilweise im Bild ist oder einige Tags nicht erkannt werden. Um später eine Kalibrierung durchführen zu können, ist es notwendig, vorher die Objektkoordinate der Eckpunkte der Tags zu bestimmen. Hierfür wurden alle Ecken der Tags mit einem lokalen Positionierungssystem, relativ zur linken oberen Ecke des Tags mit der *ID0*, eingemessen (siehe Abbildung 3.4).

Anschließend wurde die Druckvorlage mit einer Auflösung von  $0,1 \text{ mm/Pixel}$  exportiert und die Koordinaten der Eckpunkte der Tags mit Hilfe des Tagdetektors bestimmt. Diese wurden dann mit den eingemessenen Punkten verglichen, um so alle vier Eckpunkte der Tags in der jeweiligen Ebene zu erhalten und etwaige Messfehler zu korrigieren. Die ermittelten Abweichungen zwischen den gemessenen Punkten und den aus der Druckvorlage bestimmten lag im Durchschnitt bei  $0,367 \text{ mm}$ . Dieser Fehler ist durch ungenaues händisches positionieren der Messnadel zu begründen. Um dies Messpunkte zu Verifizieren wurde ein 3D-Scan der Kalibrierstruktur, siehe Abbildung 3.3, mit einer Genauigkeit von  $2 \text{ mm}$  angefertigt.

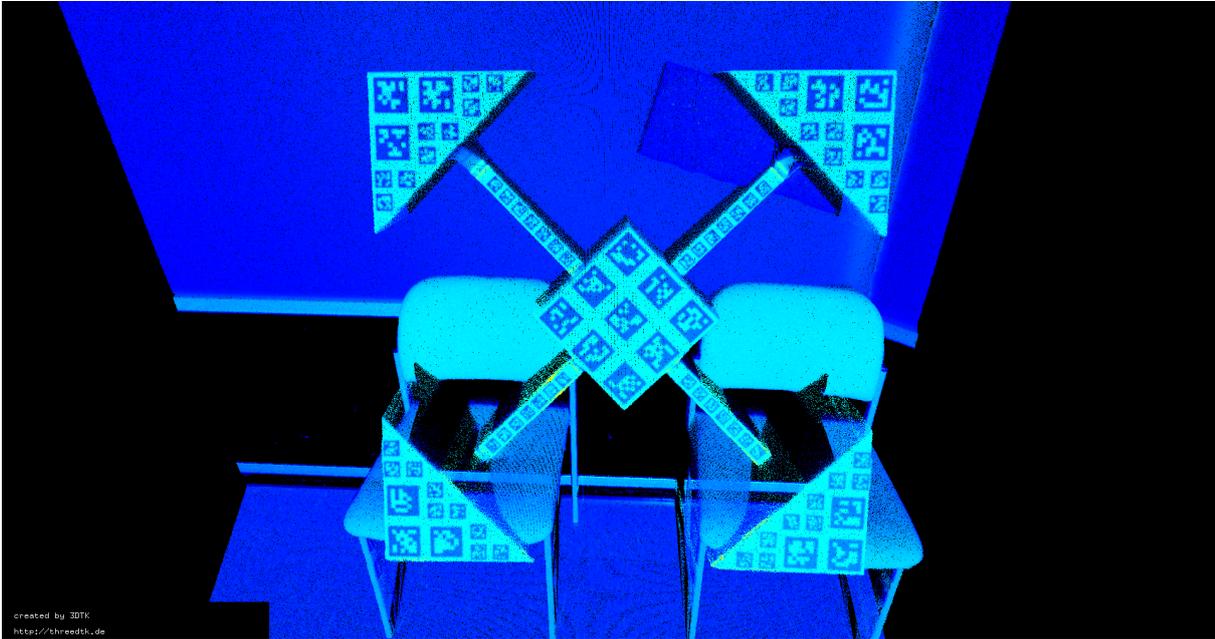


Abbildung 3.3: 3D Scan der Kalibrierstruktur.



Abbildung 3.4: Vermessung der Tags.

Insgesamt wurden auf der 3D-Struktur 85 Tags angebracht. Jeweils zwölf auf den Ecken, sieben auf jedem der Stege und neun in der Mitte. Um es zu ermöglichen die Punkte noch aus größeren Entfernungen erkennen zu können, sind alle möglichst groß und die freien Flächen mit kleineren, für kürzere Distanzen, aufgefüllt. Somit werden im Idealfall 340 Punkte erkannt.

## 3.2 Kalibrierung und Dedektierung der Tags

Zur Bestimmung der Kameraparameter wird eine eigens geschriebene Software verwendet, welche sich im Anhang befindet und deren Verwendung und Funktionsweise hier beschrieben wird.

### 3.2.1 Installation und Verwendung

#### Installation

Vorausgesetzt wird eine funktionsfähige Installation von OpenCV [4] in der Version 3.1 oder höher und boost [2]. Im Anschluss kann der Quelltext in Linux oder MacOS über das im Hauptverzeichnis befindliche CMake Skript kompiliert werden. Nun befindet sich dort eine ausführbare Datei mit dem Namen *CameraCalibrationTool*. Diese kann jetzt mit den im Folgenden beschriebenen Parametern ausgeführt werden.

#### Verwendung

Wird das Programm mit dem Parameter *-h* oder *--help* gestartet, erscheint eine Ausgabe (siehe Abbildung 3.5) aller gültigen Parameter mit ihren Standardwerten. Um eine Kalibrierung durchzuführen, genügt es zu Anfang dem Programm über den Parameter *-P* den Pfad der zu verwendenden Bilder mitzuteilen. Aus diesem Ordner werden alle Dateien mit der Dateieindung *.JPG* eingelesen. Liegen die Bilder in einem anderen Format vor, muss ebenfalls der Parameter *-T* verwendet werden und der Software das entsprechende Format mitgeteilt werden. Wird ein planares Muster verwendet muss mit dem Parameter *--patterntype* eines der akzeptierten Formate (*APRIL\_A3*, *APRIL\_A4*) angegeben werden. Wird eine 3D-Struktur verwendet ist es notwendig, die vorher bestimmten Objektpunkte als Datei im in Abbildung 3.6 dargestelltem Format über den Parameter *-C*, zu übergeben.

```
#APRIL_3D
#ID
X Y Z
X Y Z
X Y Z
X Y Z
X Y Z
#ID
X Y Z
...
```

**Abbildung 3.6:** Fomatierung der Datei mit den Objektpunkten der Tags.

Ist das Programm gestartet, beginnt es damit die AprilTags in den übergebenen Bildern zu suchen und anschließend eine Kalibrierung durchzuführen. Das Ergebnis wird in Form einer XML-Datei im angegebenen Pfad der Bilder gespeichert. Informationen über den Status des

```

Allowed options:
  -h [ --help ]                produce help message
  --patterntype arg            set Patterntype, default
  APRIL_3D,                    allowed APRIL_3D, APRIL_A4,
  APRIL_A3                     APRIL_A3
  -t [ --threads ] arg (=1)    set threads count for
  AprilTag                     detection
  --blur arg (=0.800000012)    set blur level for AprilTag
  -m [ --initial-cameramatrix ] arg (=10) set the number of Pictures
  to use for                   the initial camera matrix
  --debug arg (=0)             set AprilTag to debug mode
  -e [ --refine-edges ] arg (=1) set AprilTag refine-edges
  -d [ --refine-decodes ] arg (=1) set AprilTag refine-decodes
  -p [ --refine-pose ] arg (=1) set AprilTag refine-pose
  --tagFamily arg (=tag36h11)  set AprilTag TagFamily
  -P [ --path ] arg            set path to pictures
  -C [ --path-pattern ] arg    set path to file with
  pattern                      coordinates
  -T [ --picturetype ] arg (=JPG) set image extension
  -O [ --output-filename ] arg (=calibration23_4_9_49.xml) set name for outputfile
  -c [ --picture-count ] arg (=1) set number of pictures to
  read, default                -1 for read all pictures in
  directory                    directory

```

Abbildung 3.5: Output für -h.

Programms können der Konsole entnommen werden. Dort werden Informationen über das aktuell verarbeitete Bild, die Anzahl der im Bild gefundenen Tags und der Gesamtfortschritt angezeigt. Der XML-Datei sind unter anderem folgende Werte zu entnehmen:

**Kameramatrix** Unter dem Punkt *Camera\_Matrix* findet sich die von OpenCV bestimmte Kameramatrix [5]:

$$A = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.1)$$

Diese enthält die Faktoren  $f_x$  und  $f_y$ , welche die Brennweite der Kamera darstellen und  $u_0, v_0$ , welche den Bildmittelpunkt in Pixelkoordinaten darstellen.

**Verzerrungskoeffizienten** Der Vektor der Verzerrungskoeffizienten  $d$  findet sich unter *Distortion\_Coefficients*:

$$d = \begin{bmatrix} k_1 \\ k_2 \\ p_1 \\ p_2 \\ k_3 \end{bmatrix} \quad (3.2)$$

Dabei gilt für die radiale Verzerrung [5]:

$$\begin{aligned} x_{\text{corrected}} &= x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \\ y_{\text{corrected}} &= y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \end{aligned} \quad (3.3)$$

Somit gilt für einen Punkt  $(x, y)$  aus dem ursprünglichen Bild, dass der korrigierte Punkt durch  $(x_{\text{corrected}}, y_{\text{corrected}})$  dargestellt wird.

Ebenso gilt für die tangentielle Verzerrung:

$$\begin{aligned} x_{\text{corrected}} &= x[2p_1 xy + p_2(r^2 + 2x^2)] \\ y_{\text{corrected}} &= y[p_1(r^2 + 2y^2) + 2p_2 xy] \end{aligned} \quad (3.4)$$

auch hierbei gilt wieder, dass der Punkt  $(x_{\text{corrected}}, y_{\text{corrected}})$  dem korrigierten Punkt  $(x, y)$  aus dem ursprünglichen Bild entspricht.

**Rückprojektionsfehler** Der Rückprojektionsfehler steht in der *XML*-Datei unter dem Punkt *Avg\_Reprojection\_Error* und wird über den Fehler jedes, zur Kalibrierung herangezogenen, Fotos folgendermaßen aufsummiert

$$e = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}} \quad (3.5)$$

wobei  $e$  dem gesamten Rückprojektionsfehler,  $e_1$  dem Rückprojektionsfehler der einzelnen Bilder und  $n$  der Anzahl der verwendeten Bilder entspricht. Hierbei beschreibt  $e_i$  den euklidischen Abstand zwischen den detektierten Punkten im Bild und den zurück projizierten Objektpunkten. Die Rückprojektionsfehler der einzelnen Fotos sind in der Datei unter dem Abschnitt *Per\_View\_Reprojection\_Errors* zu finden.

Wird das Programm zusätzlich mit dem Parameter *-visualize* gestartet, wird für jedes Bild je ein weiteres Bild angelegt, in dem die erkannten Eckpunkte und die rückprojizierten Punkte eingezeichnet wurden. Ebenfalls wird eine Datei mit dem Namen *points.png* angelegt, welche die Verteilung des Kalibrieremusters im Sichtfeld der Kamera aufzeigt.

### 3.2.2 Funktionsweise

Der Aufbau der Software lässt sich grob in drei Stufen unterteilen: dem Einlesen und Vorbereiten der Bilder und der Objektkoordinaten des Kalibrierungsmusters, dem Suchen und Zuordnen der Tagkoordinaten in den Bildern und zuletzt die Berechnung der Kamerawerte. Um die Tags in den Fotos zu erkennen und die Pixelkoordinaten zu bestimmen, wurde das *April Robotics Toolkit* [1] eingebunden und verwendet. Nach der Detektierung der Tags, dargestellt in Abbildung 3.7, werden diese Bildkoordinaten in einer Datei im Bilderverzeichnis gespeichert, um bei einer erneuten Ausführung diese anstelle der Bilder einzulesen und somit die Laufzeit zu reduzieren.

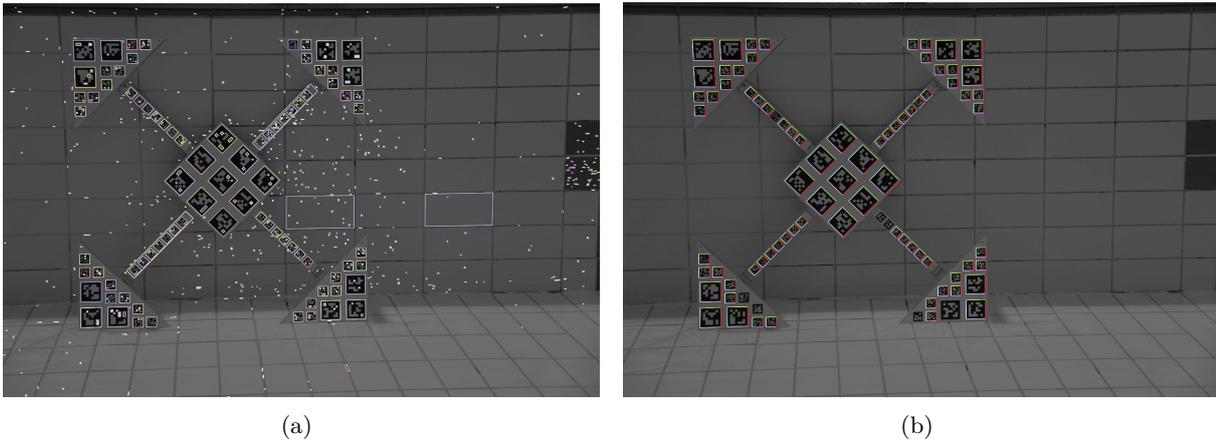


Abbildung 3.7: Erkannte Quads (a) und detektierte Tags (b).

Zu den gefundenen Tags werden im Anschluss die Objektkoordinaten gesucht und zuletzt die Kameraparameter berechnet. Diese werden als *XML*-Datei gespeichert. Wurde hierbei eine 3D-Struktur verwendet, wird im ersten Schritt der Kalibrierung eine initiale Schätzung über die neun planaren Punkte in der Mitte der Struktur vorgenommen, da die von Zhang entwickelte Routine, siehe Abschnitt 2.3, nur für 2D-Koordinaten berechnet werden kann. Anschließend können diese initialen Werte für die Kameramatrix nun für die Optimierung mit dem Levenberg-Marquardt-Algorithmus mit allen 3D-Punkten verwendet werden.



## Kapitel 4

# Experimente

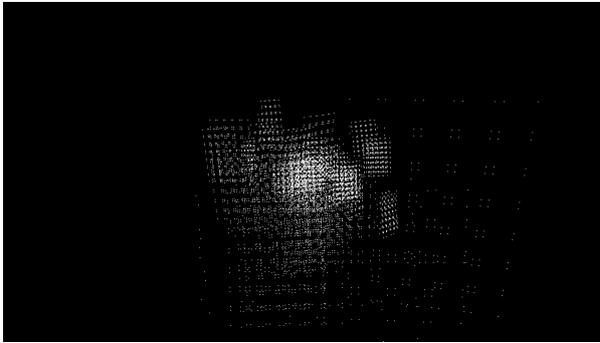
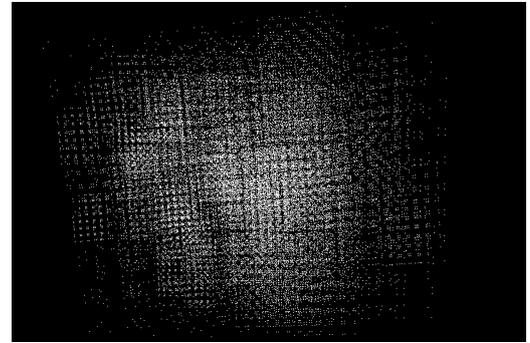
Zunächst gilt es herauszufinden wie viele Bilder für eine stabile Kalibrierung verwendet werden müssen. Hierbei ist zu beachten, dass die Qualität der Ergebnisse stark von den verwendeten Bildern abhängt. Es ist wichtig, dass mit den verwendeten Bildern möglichst das gesamte Sichtfeld und die gesamte Tiefe des Raums abgedeckt wird. Anschließend werden mit der so bestimmten Anzahl an Bildern weitere Versuche durchgeführt um die Stabilität der Kalibrierungen mit den unterschiedlichen Kalibriermustern zu vergleichen. Daraufhin wird die Genauigkeit der bestimmten Parameter überprüft. Hierzu werden mit den Parametern korrigierte Bilder analysiert. Für die Überprüfung der extrinsischen Parameter wird eine genaue Trajektorie mit einem KUKA Manipulator abgefahren und die aus den Bildern errechnete Trajektorie mit der Referenztrajektorie des Manipulators verglichen.

Für jedes der verwendeten Kalibriermuster, die 3D-Struktur, ein planares DINA3 und ein planares DINA4 Muster, wurden je zwei Bildersätze erstellt. Jedes dieser Sets umfasst mindestens 60 Fotos. Das erste Set wurde in Luft aufgenommen, das zweite besteht aus Unterwasseraufnahmen, die in einem Schwimmbad angefertigt wurden. Die *GoPro* befand sich bei beiden Aufnahmen in einem Gehäuse mit einem *Flat Port*. Bei der zweiten Kamera, einer *Sony Alpha6000*, wurde Überwasser auf ein zusätzliches Gehäuse verzichtet. Für die Aufnahmen unter Wasser wurde sie ebenfalls in einem Gehäuse mit *Flat Port* verstaut. Alle Versuche wurden jeweils mit beiden Kameras durchgeführt, da die Aussagen der Ergebnisse jedoch äquivalent sind, werden hier nur die Bilder der *GoPro* behandelt. Die Diagramme mit den Auswertungen der *Sony Alpha6000* finden sich im Anhang (siehe A). Für alle durchgeführten Experimente wurde die *GoPro Hero 4* im 4k-Video-Modus mit einer Auflösung von  $3840 \times 2160$  Pixel verwendet. Mit der *Sony Alpha6000* wurden mit einem 35 mm Objektiv Einzelbilder in einer Auflösung von  $3000 \times 2000$  Pixel erstellt. Zur Darstellung werden sogenannte Boxplots verwendet, hier stellt das Rechteck den Bereich in dem 15 – 85% der Ergebnisse liegen dar. Die rote Linie markiert den Median und die Extremwerte werden durch die gestrichelte Linie aufgezeigt.

## 4.1 Analyse der Kalibrierstabilität

Für die nachfolgenden Versuche wurde je ein Set mit ca. 60 Bildern benutzt, aus dem dann zufällig die Anzahl der benötigten Bilder gezogen wurde. Es wird ein Bild aus dem Set gezogen, für dies eine Kalibrierung durchgeführt, im Anschluss ein weiteres Bild hinzugefügt und wieder eine Kalibrierung durchgeführt. Dies wird so lange wiederholt bis mit 40 Bildern kalibriert wurde. Dieser Vorgang wird 20 mal durchgeführt.

Bei Betrachtung der Grafiken in der Abbildung 4.2 fällt sofort auf, dass der Fehler für die planaren Muster, Abbildungen 4.2(c)-(f), größer ist und deutlich mehr schwankt als mit der 3D-Struktur 4.2(a),(b), jedoch bei durchgehend klein ist.

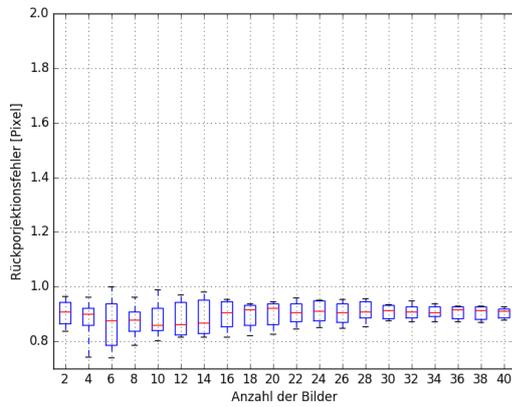
(a) *GoPro*(b) *Sony Alpha6000*

**Abbildung 4.1:** Verteilung der detektierten Objektpunkte des DIN A3 Musters im Sichtfeld der Kamera bei den Testaufnahmen unter Wasser.

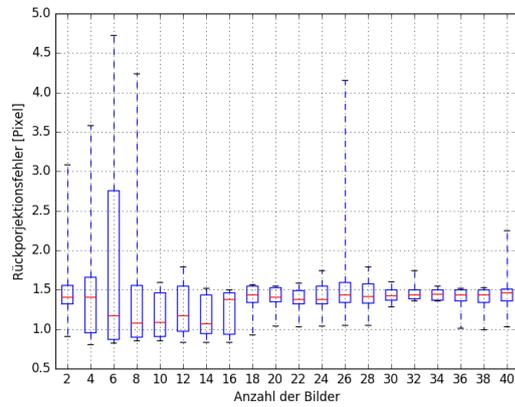
Dass der Fehler bei den Bildern der *Sony Alpha6000* im Anhang (Abbildungen A.1, A.2 und A.3) durchgehend geringer ist, als bei den Bildern der *GoPro*, liegt zum einen an der geringeren Verzerrung der *Sony Alpha6000* und zum anderen auch daran, dass es durch das fehlende Display der *GoPro* sehr viel schwerer ist gute Aufnahmen zu machen.

Auf Abbildung 4.1(a) ist beispielsweise die Verteilung des DIN A3 Musters auf den Aufnahmen mit der *GoPro* unter Wasser zu erkennen. Im Vergleich mit der in Abbildung 4.1(b) aufgezeigten Verteilung in den Bildern der *Sony Alpha6000* ist zu sehen, dass dort das Muster etwas besser auf das komplette Sichtfeld verteilt ist.

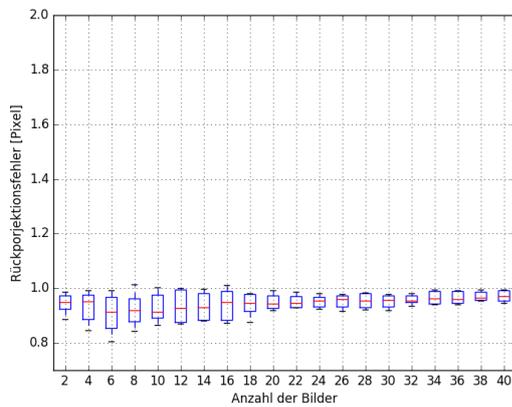
Der Rückprojektionsfehler alleine ist jedoch kein besonders gutes Maß für die Genauigkeit der Kalibrierung, da dieser lediglich anzeigt, wie gut die berechneten Werte für genau diese Bildersets waren. Was daher sehr viel interessanter und wichtiger für eine stabile Kalibrierung ist, ist zu betrachten, wie stark die bestimmten Kameraparameter schwanken und ab welcher Anzahl von Bildern dies möglichst gering wird. Hierzu werden im nächsten Schritt die Schätzungen der jeweiligen Brennweite über mehrere Kalibrierungen betrachtet.



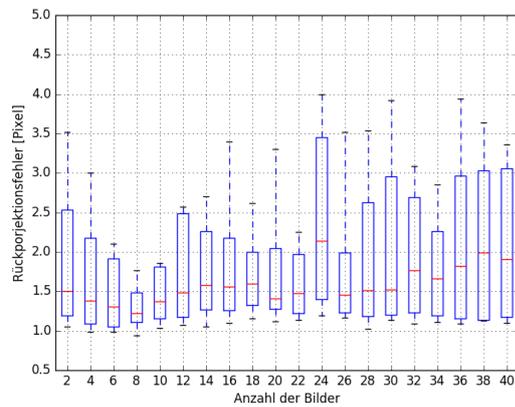
(a) 3D-Struktur unter Wasser



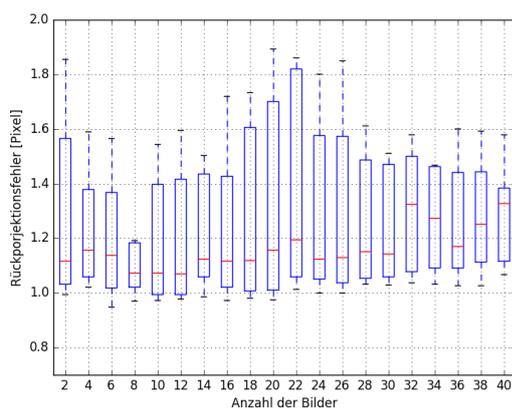
(b) 3D-Struktur in Luft



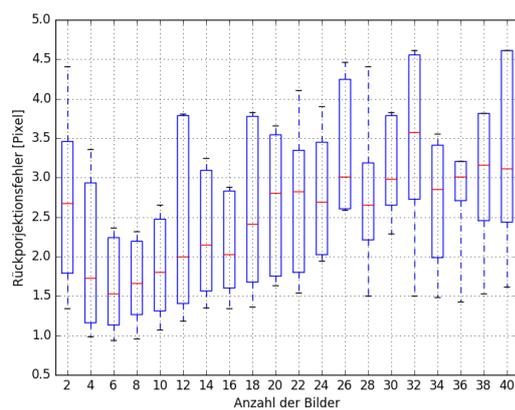
(c) DINA3 Muster unter Wasser



(d) DINA3 Muster in Luft

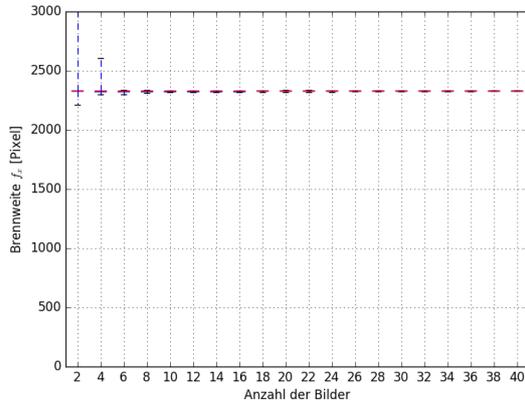


(e) DINA4 Muster unter Wasser

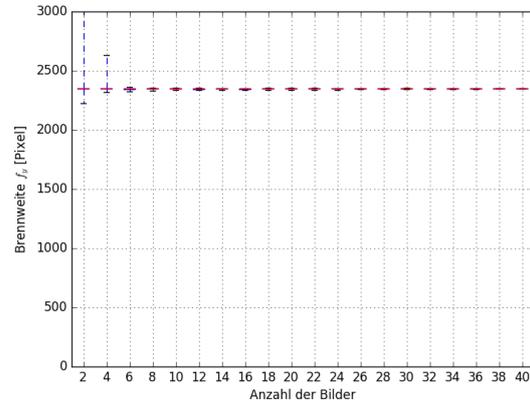


(f) DINA4 Muster in Luft

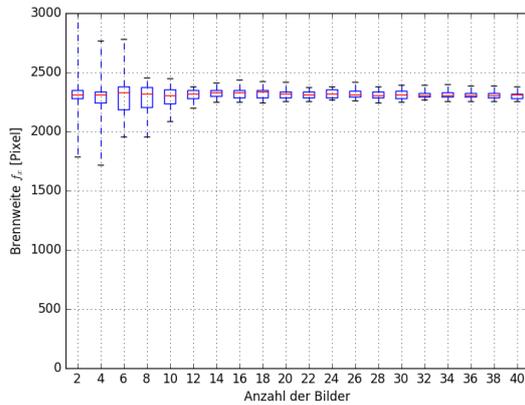
**Abbildung 4.2:** Veränderung des Rückprojektionsfehlers für eine unterschiedliche Anzahl an, mit einer GoPro aufgenommen, Bilder.



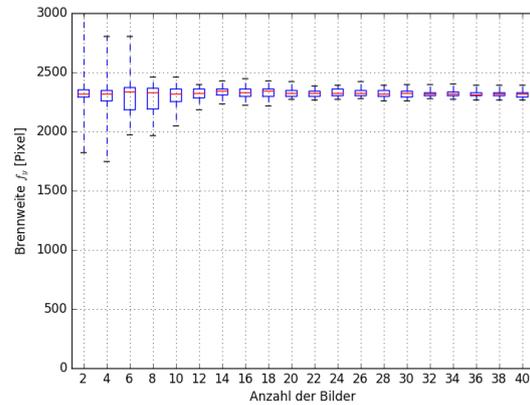
(a) unter Wasser 3D-Struktur



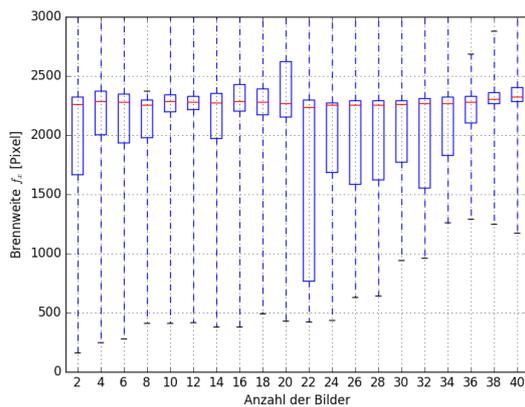
(b) unter Wasser 3D-Struktur



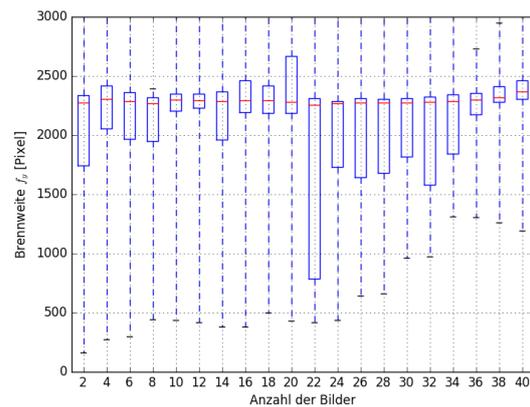
(c) unter Wasser DINA3 Muster



(d) unter Wasser DINA3 Muster

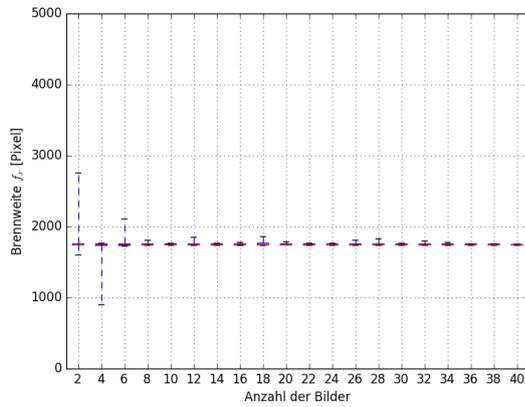


(e) unter Wasser DINA4 Muster

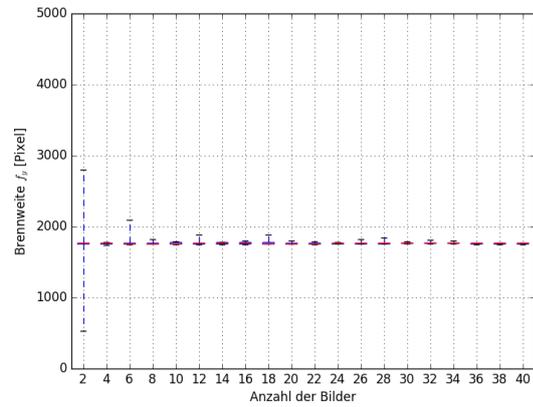


(f) unter Wasser DINA4 Muster

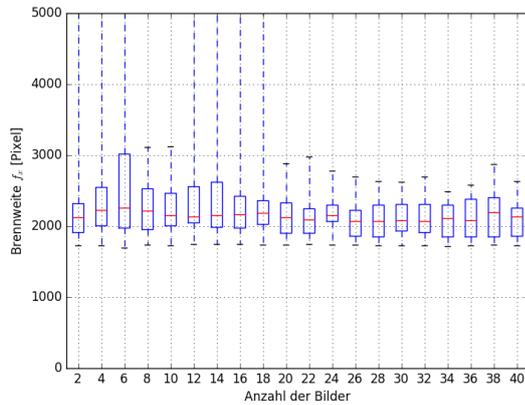
**Abbildung 4.3:** Veränderung der bestimmten Brennweite für eine unterschiedliche Anzahl an, mit einer GoPro unter Wasser aufgenommenen, Bildern.



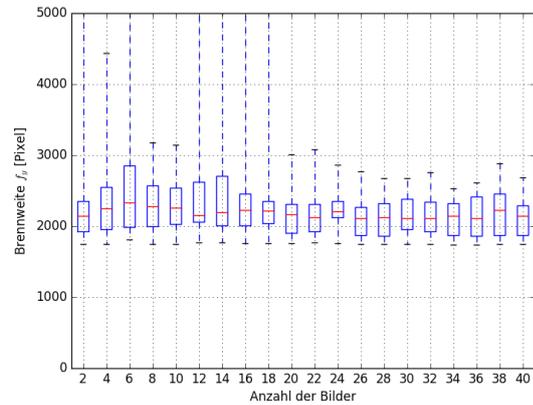
(a) 3D-Struktur in Luft



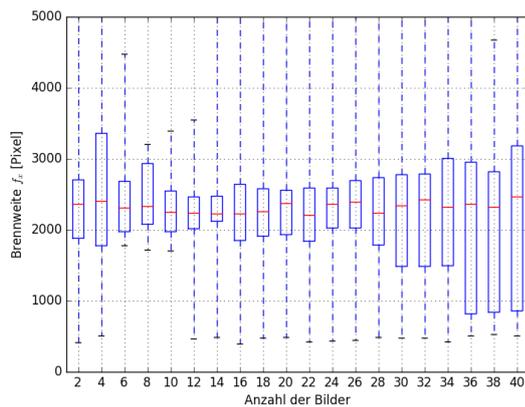
(b) 3D-Struktur in Luft



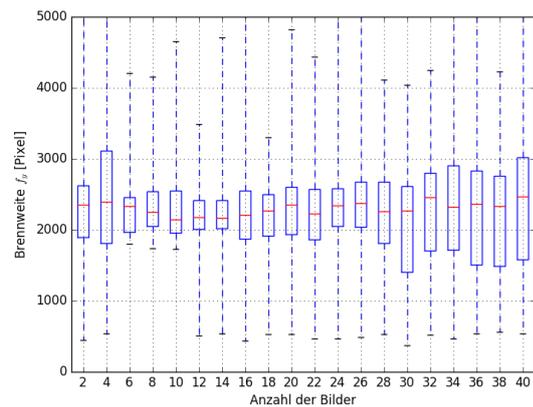
(c) DINA3 Muster in Luft



(d) DINA3 Muster in Luft

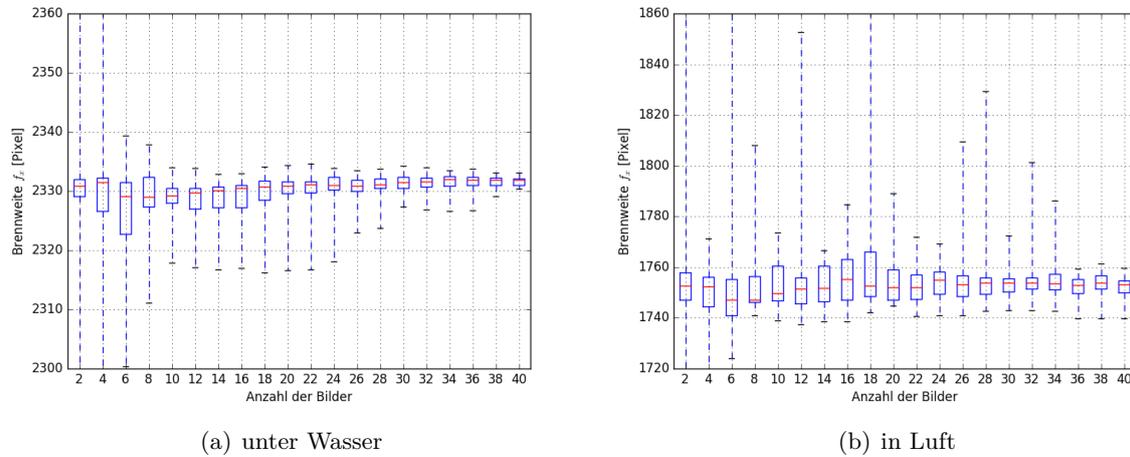


(e) DINA4 Muster in Luft



(f) DINA4 Muster in Luft

**Abbildung 4.4:** Veränderung der bestimmten Brennweite für eine unterschiedliche Anzahl an, mit einer GoPro in Luft aufgenommenen, Bilder.



**Abbildung 4.5:** Ausschnitt aus den Diagrammen 4.3(a) und 4.4(a), hier jeweils nur  $f_x$ .

Wird nun der Boxplot 4.5(a) für die Kalibrierung der *GoPro* unter Wasser unter Verwendung der 3D-Struktur betrachtet, ist zu erkennen, dass selbst bei einer Kalibrierung mit weniger als zehn Bildern 70% alle Ergebnisse in einem Bereich von 20 Pixeln liegt, was bei einer ermittelten Brennweite für Aufnahmen unter Wasser von ca. 2324 Pixeln einer Abweichung von  $\pm 1\%$  entspricht. Eine klar erkennbare Konvergenz findet ab 20 Bildern statt.

Hier besitzt ein großer Teil der Werte nur eine Abweichung von ca.  $\pm 2,5$  Pixel vom Median. Auch die Extremwerte befinden sich in einem Bereich von kleiner 20 Pixel und somit kann von einer hinreichend stabilen Kalibrierung ausgegangen werden. Daher werden für die folgenden Versuche stets 30 Bilder verwendet. Somit ist sichergestellt, dass mit der 3D-Struktur eine stabile Kalibrierung möglich ist.

## 4.2 Vergleich zwischen planaren und 3D-Kalibriermuster

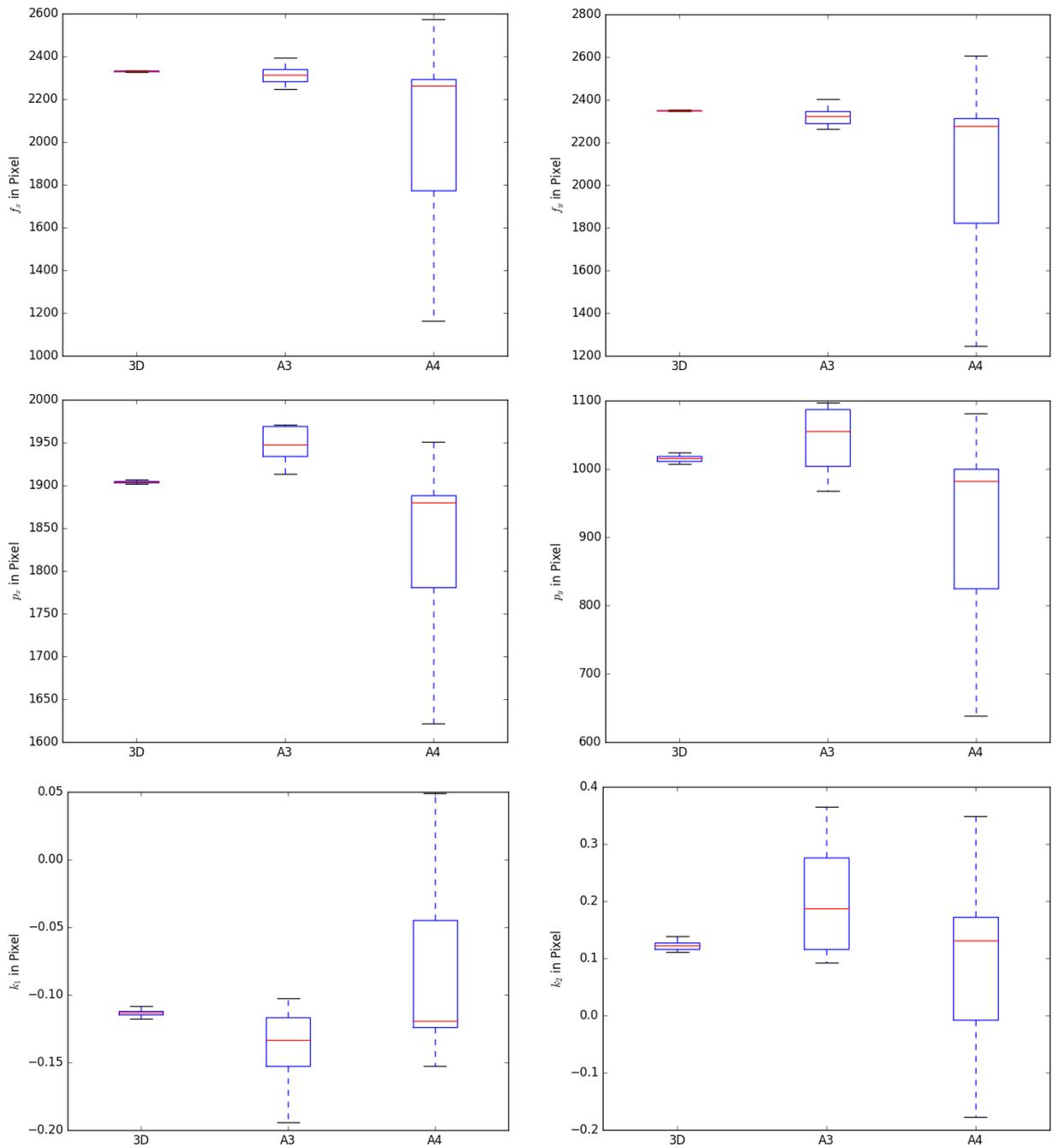
Im Folgenden werden nun für alle drei Kalibriermuster Kalibrierungen mit je 30 Bildern 20 Mal durchgeführt. Die somit bestimmten Parameter für die 3D-Struktur, das DINA3 und DINA4 Muster werden nun für je unter Wasser und in Luft gegenübergestellt. Des Weiteren werden diese mit Ergebnissen aus der Literatur verglichen.

In Abbildung 4.6 sind die Ergebnisse der Kalibrierung mit den unterschiedlichen Kalibriermustern unter Wasser direkt gegenübergestellt. Es sind jeweils die Parameter  $f_x$  und  $f_y$  für die Brennweite,  $p_x$  und  $p_y$  für den Bildmittelpunkt, sowie die Parameter  $k_1$  und  $k_2$  für die radiale Verzerrung aufgezeichnet.

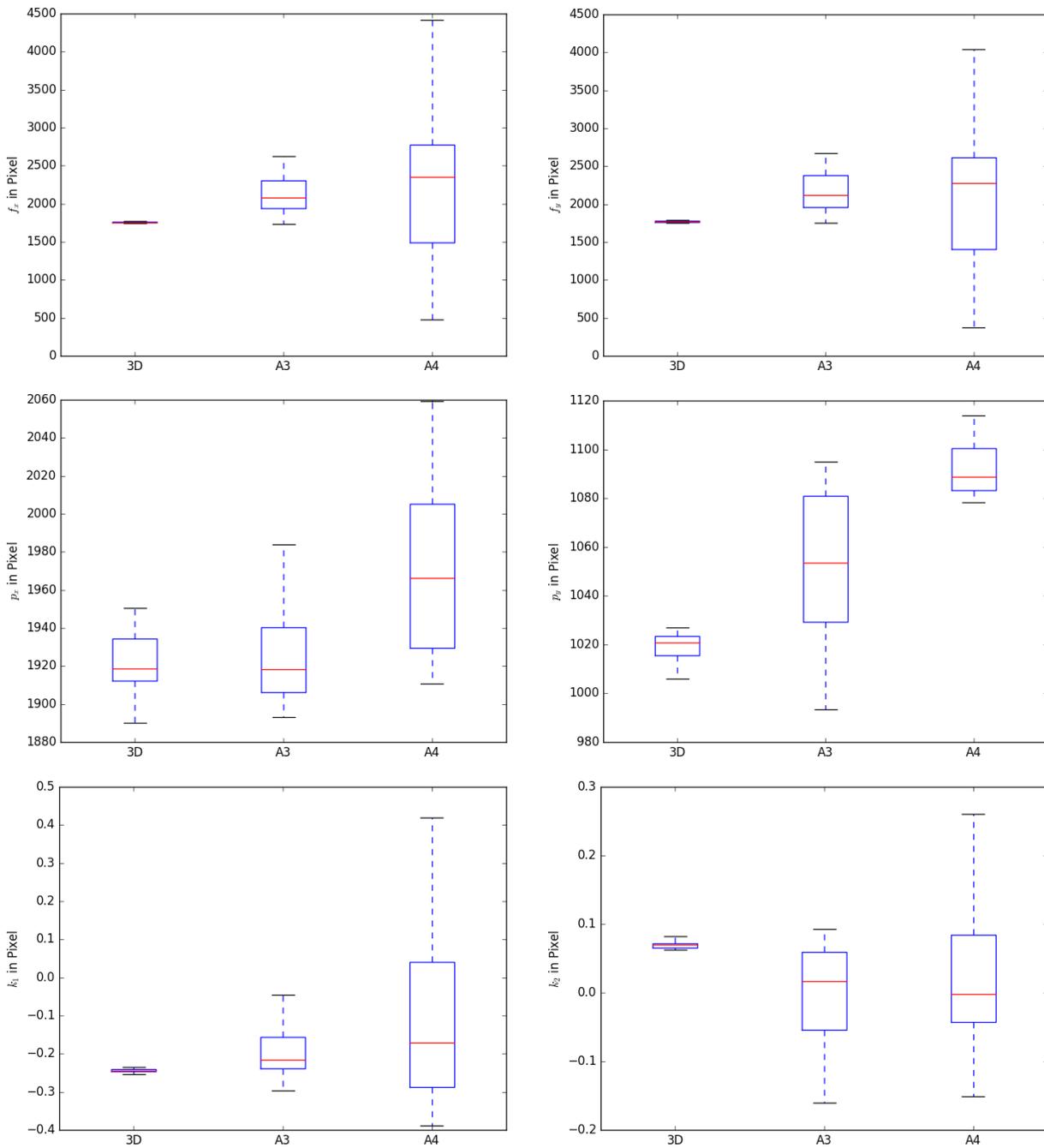
Bei einem Vergleich der einzelnen Strukturen gegeneinander, fällt sofort auf, dass die Abweichungen und Extremwerte unter Verwendung der 3D-Struktur bei allen Parametern sehr gering sind und daher von einer stabilen Kalibrierung ausgegangen werden kann. Hierbei ist zu erkennen, dass die Ergebnisse des planaren DINA3-Musters nicht die gleiche Stabilität besitzen, wie die der 3D-Struktur, jedoch deutlich näher zusammen liegen als die des DINA4-Musters. Die Parameter unter Verwendung des DINA4-Musters schwanken um solch große Werte, dass hier nicht von einer stabilen Kalibrierung ausgegangen werden kann. Die hierbei mit 30 Bildern ermittelten Werte sind für eine Kalibrierung unbrauchbar. Jedoch reicht auch die Genauigkeit der Schätzungen mit dem DINA3-Muster nicht aus, um eine stabile Kalibrierung durchführen zu können.

Die Abbildung 4.7 zeigt die gleiche Gegenüberstellung mit Überwasser angefertigten Aufnahmen. Auch hier liegen die Ergebnisse für die unterschiedlichen Parameter bei einer Kalibrierung mit der 3D-Struktur dicht beieinander. Die, im Vergleich zu unter Wasser, größeren Ausschläge bei den beiden Parametern  $p_1$  und  $p_2$  sind auf die wesentlich höher Verzerrung der Kamera in Luft zurückzuführen. Die Verhältnisse zwischen den einzelnen Mustern sind hier ähnlich den Ergebnissen unter Wasser.

Bei einem Vergleich mit den Ergebnissen von Shortis in der Abbildung 2.2 ist ein übereinstimmendes Ergebnis zu erkennen. Zu beachten ist hierbei, dass ein jeweils unterschiedliches Maß verwendet wurde. Während Shortis seine Angaben in mm angibt, wurde in den hier gezeigten Grafiken stets Pixel verwendet. Wichtig ist also jeweils nur die Verhältnisse der Ergebnisse zueinander zu betrachten und nicht die absoluten Werte.



**Abbildung 4.6:** Darstellung der Parameter der 3D-Struktur eines, mit einer *GoPro* unter Wasser aufgenommenen, planaren DINA3 und DINA4 Musters.



**Abbildung 4.7:** Darstellung der Parameter der 3D-Struktur eines, mit einer *GoPro* in Luft aufgenommenen, planaren DIN A3 und DIN A4 Musters.

### 4.3 Validierung der Intrinsik

Um die intrinsischen Parameter zu überprüfen, wird ein Bild korrigiert und anschließend, mit Hilfe von einigen eingezeichneten Linien, visuell der Qualität der Kalibrierung beurteilt. Hierfür wurden die Parameter der Kamera mit 30 Aufnahmen der 3D-Struktur, des DINA3 und des DINA4 Musters bestimmt. Diese wurden dann auf ein, nicht zur Kalibrierung genutztem, Bild angewendet. Daraufhin wurden einige Hilfslinien (in Rot) im Bild eingezeichnet. In Abbildung 4.8 sind Überwasseraufnahmen und in Abbildung 4.9 Unterwasseraufnahmen zu erkennen.

Bei Betrachtung des originalen Bildes 4.8(a) in Luft ist sehr deutlich eine radiale Verzeichnung zu erkennen. Die Ecken der 3D-Struktur bilden hier kein Quadrat, die Kanten sind deutlich nach außen gewölbt. In der sichtlich ungenauen Korrektur, unter Verwendung eines DINA4 Musters, des Bildes 4.8(d) spiegelt sich die Aussage des vorherigen Abschnitts wieder. Es ist eine leichte Verbesserung im Vergleich zum originalen Bild 4.8(a) zu erkennen, jedoch ist der Fehler noch deutlich sichtbar vorhanden.

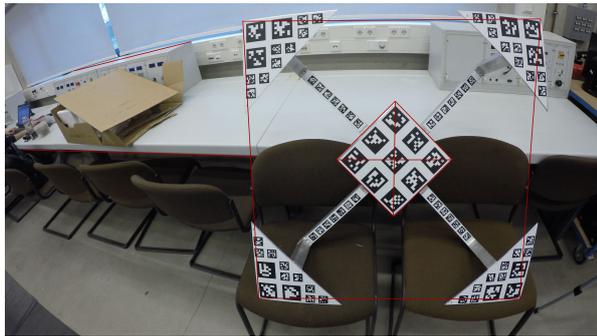
Wird ein DINA3 Muster zur Kalibrierung verwendet, sieht die Korrektur 4.8(c) etwas besser aus als das originale Bild. Das gedachte Viereck über der 3D-Struktur kommt hier schon näher an ein Quadrat als ohne diese Korrektur, ein perfektes Quadrat ist es jedoch noch nicht.

In Abbildung 4.8(b) ist das Ergebnis der Korrektur mit den Parametern aus der Schätzung mit der 3D-Struktur zu sehen. Hier ist zu erkennen, dass die Kanten der 3D-Struktur eine gerade Linie und somit ein Quadrat bilden. Auch das Quadrat in der Mitte der Struktur besitzt nun keine Verformungen mehr. Daraus lässt sich schließen, dass die Parameter der Kamera korrekt geschätzt wurden.

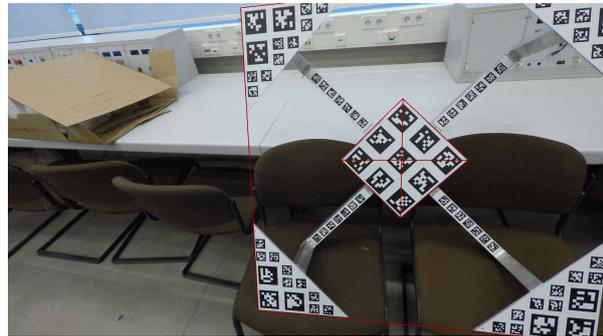
Die Abbildungen 4.9 zeigen eine Unterwasseraufnahme und 4.10 jeweils einen Ausschnitt aus diesen Bildern, auf die ebenfalls Korrekturen angewendet wurden. Hier treten selbst bei dem originalen Bild nur sehr geringe Fehler auf. Ein großer Teil der radialen Verzeichnung wird durch die Lichtbrechung des Wassers im Vergleich zu Luft kompensiert. Dennoch ist gerade in den Ecken des Bildes ein leichter Fehler zu erkennen. Wie auch schon Überwasser sind diese Fehler bei einer Kalibrierung mit der 3D-Struktur auf ein nicht erkennbares Maß reduziert worden und Abbildung 4.10(b) könnte somit für eine 3D-Rekonstruktion verwendet werden.

In Abbildung 4.10(c) ist im Bereich der Ecken eine Fehlkorrektur zu erkennen. Dies liegt an ungenau geschätzten Parametern der Kamera. Für das DINA3 Muster sind auch hier die verwendeten 30 Bilder sichtbar zu wenig für eine stabile Kalibrierung.

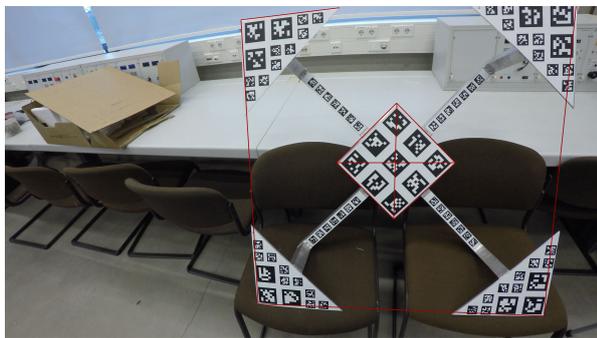
Wie an Abbildung 4.10(d) an der rechten unteren Ecke zu erkennen ist, wurde auch hier nicht korrekt kalibriert. Unter Verwendung des DINA4 Musters wurde der Bildfehler etwas zu stark korrigiert, so dass eine Verzeichnung nun entgegengesetzt der im originalen Bild zu beobachten ist.



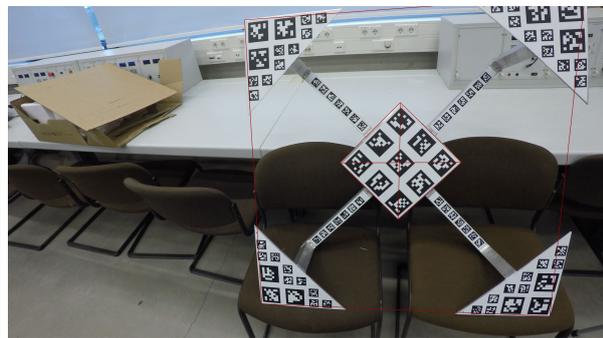
(a) originales Bild



(b) korrigiert mit Kameraparametern der 3D-Struktur

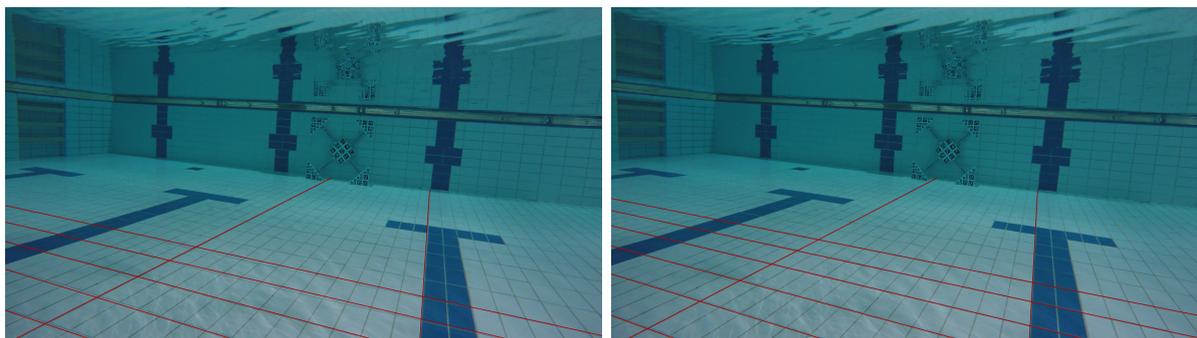


(c) korrigiert mit Kameraparametern des DINA3 Mus-



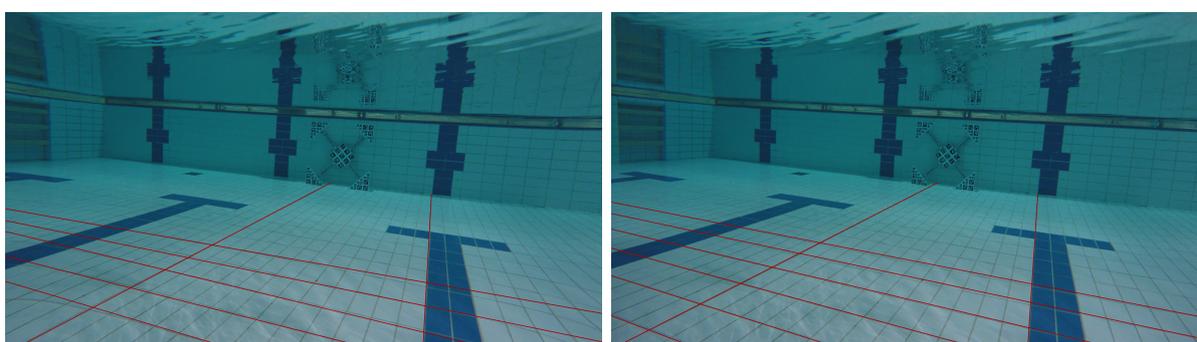
(d) korrigiert mit Kameraparametern des DINA4 Mus-

**Abbildung 4.8:** Ein unkorrigiertes Bild und je ein korrigiertes Bild der *GoPro* in Luft mit nachträglich eingezeichneten Linien, um die Qualität der Verzerrungskorrektur hervorzuheben.



(a) originales Bild

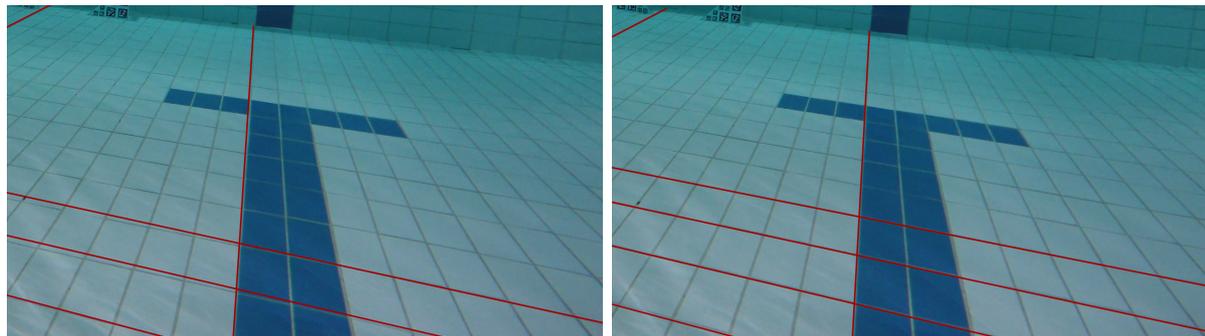
(b) korrigiert mit Kameraparametern der 3D-Struktur



(c) korrigiert mit Kameraparametern des DINA3 Musters

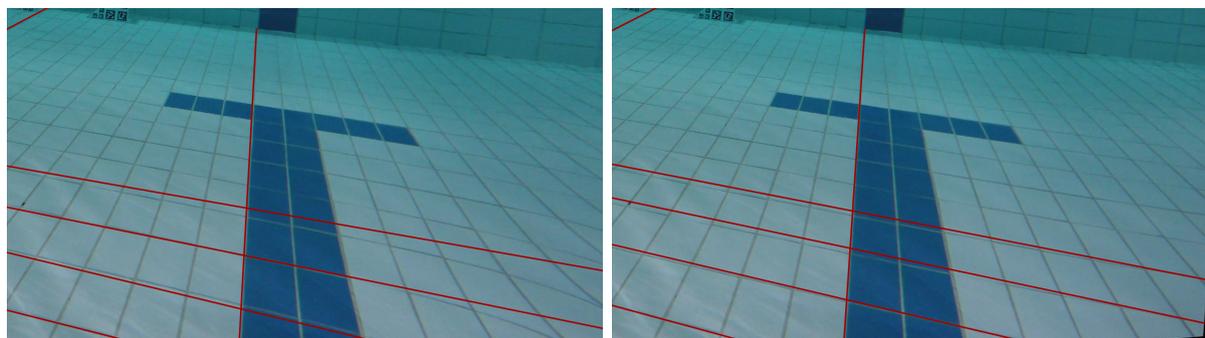
(d) korrigiert mit Kameraparametern des DINA4 Musters

**Abbildung 4.9:** Ein korrigiertes und ein unkorrigiertes Bild der *GoPro* unter Wasser mit nachträglich eingezeichneten Linien, um die Qualität der Verzerrungskorrektur hervorzuheben.



(a) originales Bild

(b) korrigiert mit Kameraparametern der 3D-Struktur

(c) korrigiert mit Kameraparametern des DINA3 Mus-  
ters(d) korrigiert mit Kameraparametern des DINA4 Mus-  
ters**Abbildung 4.10:** Vergrößerte Ausschnitte aus den Abbildungen in 4.9.

## 4.4 Validierung der Extrinsik

Um eine Validierung der extrinsischen Parameter vorzunehmen, wurde die Kamera auf einer Ebene in festem Abstand zum Kalibriermuster bewegt. Daraufhin wurde die Entfernung zum Muster geschätzt und mit der realen Entfernung verglichen. Hierzu wurde eine Reihe von Fotos der 3D-Struktur gemacht und damit die Kameraparameter bestimmt. Anschließend wurde die Kamera an einem KUKA Manipulator befestigt, siehe Abbildung 4.11. Zuvor wurde die Lage der 3D-Struktur zu den Manipulator eingemessen. Somit war es nun möglich Kamera in einer Ebene mit konstantem Abstand zum Kalibriermuster auf eine Genauigkeit von 0,1 mm zu bewegen. Die Trajektorien von Kamera und Roboterarm wurden zuerst durch Minimierung der Fehlerquadrate aneinander ausgerichtet, so dass der Punkt  $(0, 0, 0)$  dem ersten Punkt der Robotertrajektorie entspricht.

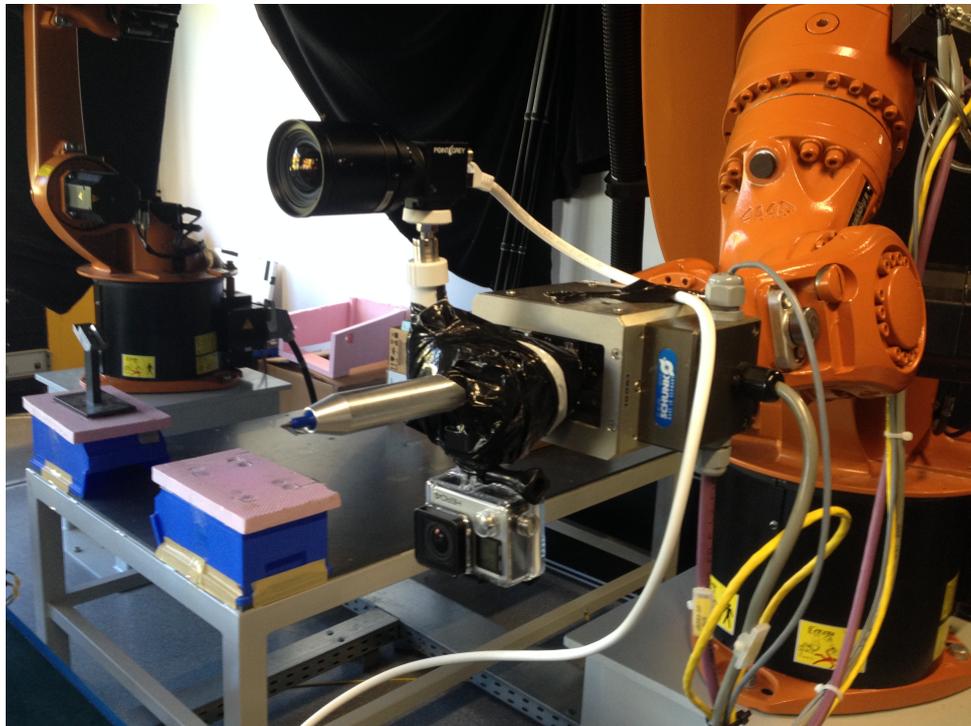
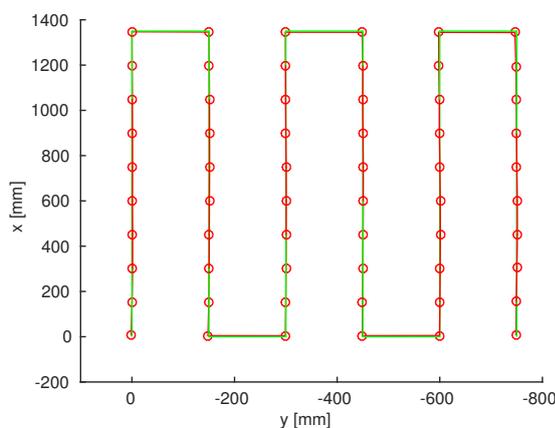


Abbildung 4.11: Kamera am KUKA Manipulator.

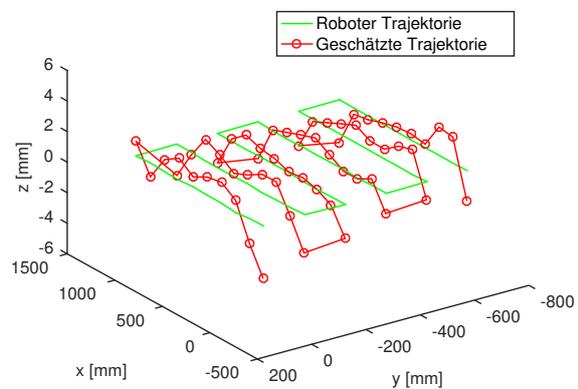
Daraufhin wurde die Kamera mit den zuvor bestimmten Werten kalibriert und aus den Aufnahmen die Entfernung zwischen Kamera und 3D-Struktur bestimmt. Vergleicht man dies nun mit der bekannten Trajektorie des Manipulators, siehe Abbildung 4.12(a)(b), ergibt sich daraus ein RMS Fehler der Kameraposition von 3,7 mm bei einer Distanz von ca. 750 mm. Somit lässt sich erkennen, dass auch die extrinsischen Parameter der Kamera sehr exakt bestimmt wurden. Außerdem lässt sich auch an dieser Grafik wieder die deutliche Verzerrung der *GoPro* zu den Rändern hin erkennen. Sind in der Mitte die ideale und die berechnete Trajektorie deckungsgleich,

so sieht man an den Außenbereichen eine deutliche Abweichung.

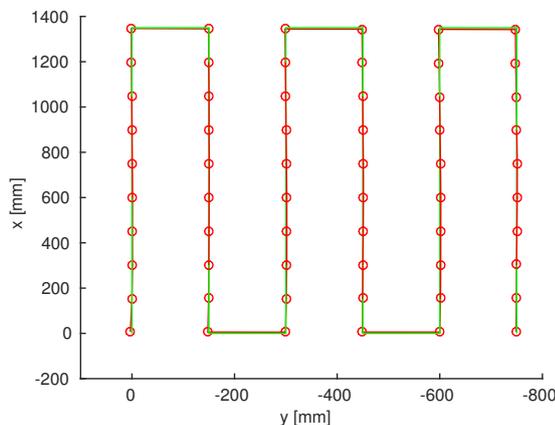
Für den Vergleich wurde der selbe Versuch mit einem Schachbrett der Größe 1,25 m auf 0,6 m durchgeführt. In der Grafik 4.12(c) ist die Abweichung auf der x-Achse und der y-Achse dargestellt. Auch hier ist keine signifikante Abweichung von der idealen Trajektorie zu erkennen. Bei Betrachtung der Abbildung 4.12(d) ist ein ähnliches Verhalten wie mit der 3D-Struktur 4.12(b) zu beobachten. An den Rändern ist ein ebenfalls ein negativer Fehler zu sehen, während in der Mitte, halbkreisförmig, ein positiver Fehler zu erkennen ist. Im Gesamten tritt hier ein RMS Fehler von 4,5 mm auf. Somit sind die Ergebnisse der 3D-Struktur und des großen Schachbretts in einer vergleichbaren Größenordnung.



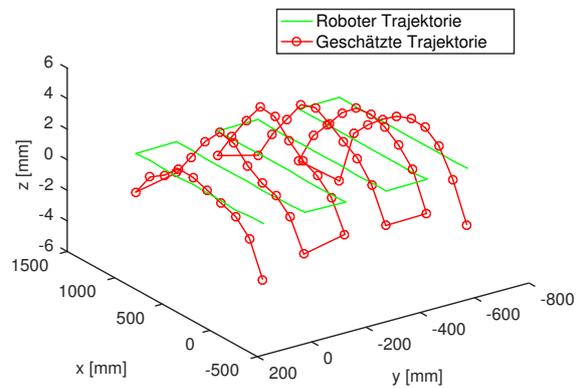
(a) Draufsicht auf die Trajektorie mit 3D-Struktur



(b) Trajektorie mit 3D-Struktur



(c) Draufsicht auf die Trajektorie mit Schachbrettmuster



(d) Trajektorie mit Schachbrettmuster

**Abbildung 4.12:** Roboter (in Grün) und geschätzte (in Rot) Trajektorie für die 3D-Struktur und ein Schachbrettmuster.



## Kapitel 5

# Zusammenfassung und Ausblick

Die Zielsetzung der Arbeit war die Erstellung und Validierung einer 3D-Struktur zur Kalibrierung von Kameras unter Wasser. Diese Struktur wurde erfolgreich erstellt und in einigen Versuchen mit einem planaren DIN A3 und DIN A4 Muster verglichen. Hierbei kristallisierte sich klar heraus, dass selbst mit einer geringen Anzahl von ca. 30 Bildern mit der 3D-Struktur eine stabile Kalibrierung erreicht werden kann. Wird hingegen ein planares Muster verwendet, ist es nicht möglich mit dieser Anzahl zuverlässig die Parameter der Kamera zu bestimmen. Die zu Beginn aufgegriffene Aussage, dass 3D-Strukturen besser seien als planare Muster, konnte hier bestätigt werden. Somit stimmen die Ergebnisse der hier angestellten Untersuchungen mit den Aussagen und Ergebnissen von Mark Shortis [11, 17] überein.

Im Speziellen wurde hier die Schätzung der Brennweite verglichen und die Konvergenz und Abweichungen dieser als Maß für die Stabilität verwendet, da sich herausstellte, dass der Rückprojektionsfehler kein geeignetes Maß für diesen Vergleich darstellt. Daraufhin wurden die geschätzten Parameter dazu verwendet um einige Bilder zu korrigieren und dieses Ergebnis zu beurteilen. Hierbei zeigte sich, dass die mit 30 Bildern geschätzten Parameter stabil sind. Die optischen Fehler in den Bildern konnten erfolgreich korrigiert werden.

Da für den speziellen Fall der 3D-Rekonstruktion jedoch nicht nur die Genauigkeit der intrinsischen Parameter, sondern auch die der extrinsischen Parameter wichtig ist, wurde auch dies in einem Versuch validiert. Für diesen Versuch wurde die Kamera an einem KUKA Manipulator befestigt und eine vorher bestimmte Trajektorie mit einer Genauigkeit von 0,1 mm abgefahren. Im Anschluss wurde aus den aufgenommenen Bildern mit Hilfe der bestimmten extrinsischen Parameter die Trajektorie berechnet und mit der Referenztrajektorie des Manipulators verglichen. Hierbei zeigte sich, dass auch diese Werte mit einer hohen Genauigkeit bestimmt werden konnte. Der RMS Fehler betrug 3,7 mm bei einer Distanz von 750 mm, somit besteht eine Toleranz von weniger als 0,5%. Der hauptsächliche Fehler befindet sich hierbei am Rand des Sichtfeldes, da die verwendete *GoPro* hier sehr starke Verzerrungen aufweist. Zu vermuten ist nun, dass dieser Fehler bei Aufnahmen unter Wasser mit der *GoPro* noch geringer ausfällt. Da die Verzerrung der Kamera unter Wasser deutlich kleiner ist und somit die Genauigkeit der Kalibrierung, wie in den Versuchen gezeigt wurde, deutlich besser ist. Hierzu kann in nachfolgender Arbeit ein

weiteres Experiment durchgeführt werden, in dem die Trajektorie der Kamera beispielsweise über ein 3D-Positioniersystem aufgezeichnet wird. Die verwendete 3D-Struktur und das benutzte Lochkameramodell erwiesen sich im Laufe der Arbeit als geeignete Methode zur Unterwasser Kamerakalibrierung. Die Genauigkeit der Kalibrierung war stets besser, als mit einem planaren DINA3 oder DINA4 Muster. In einem direkten Vergleich der Extrinsik, in Kapitel 4.4, mit einem Schachbrettmuster der Größe 1,25 m auf 0,6 m konnten zwar keine signifikant besseren Ergebnisse erzielt werden, dennoch sind die Abweichungen in der z-Achse sehr klein.

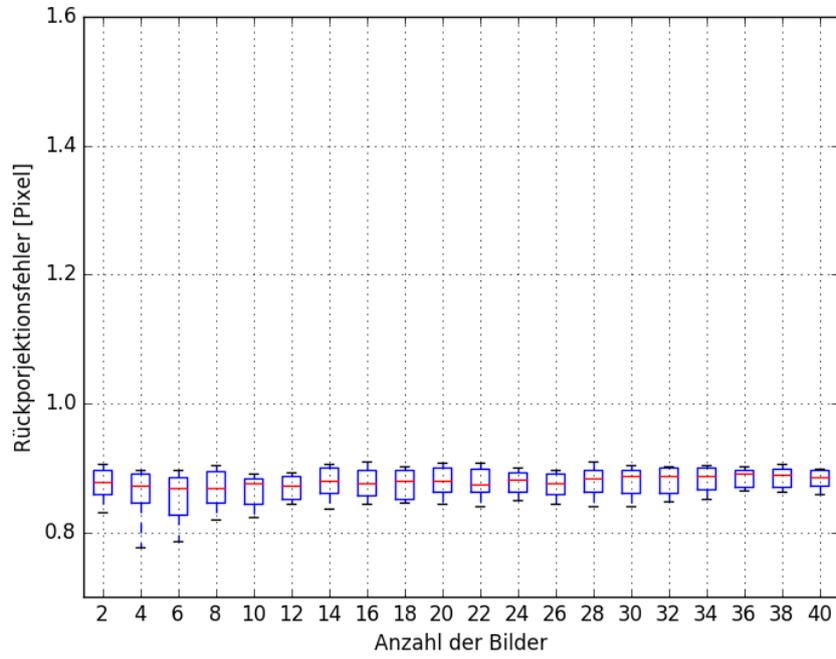
Bei allen Versuchen in dieser Arbeit hat sich ebenfalls die Verwendung von Tags mit eindeutiger ID anstelle eines herkömmlichen Schachmusters als großer Vorteil erwiesen. Die Detektion und Zuordnung der Tags funktioniert deutlich zuverlässiger und lässt sich einfacher automatisieren als dies bei einem Schachbrettmuster möglich ist. Ebenfalls hat sich die Entscheidung Freiräume auf der 3D-Struktur mit kleineren Tags aufzufüllen als praktisch erwiesen. Dadurch war es möglich die Struktur einerseits auf große Distanz zu erkennen und andererseits bei Nahaufnahmen die Anzahl der Kalibrierpunkte zu erhöhen.

Weiter könnte nun der Kalibrierprozess und das Kameramodel näher betrachtet werden. Hierbei ist es beispielsweise interessant sich den Levenberg-Marquardt-Algorithmus näher anzuschauen und zu untersuchen, ob dieser wirklich zu einem globalen Optimum konvergiert. Das hier verwendete Lochkameramodell hat sich zwar als gut herausgestellt, jedoch wäre es ebenfalls möglich, hier noch weitere Kameramodelle genauer zu untersuchen. Hierbei gilt es festzustellen, ob es ein Modell gibt, welches eine genauere Modellierung der Abbildungsfunktion bei ähnlicher Einfachheit bei den zu schätzenden Parametern erreicht. Auch wäre es möglich die Konvergenz der Kalibrierungsparameter auf einer größeren Anzahl an Testaufnahmen weiter zu analysieren und dabei den Einfluss verschiedener Faktoren, wie beispielsweise Trübheit oder der Salzgehalt des Wassers, zu betrachten. Ebenfalls könnten weitere Versuche zur Genauigkeit bezüglich 3D-Rekonstruktionen vorgenommen werden. Hierzu könnte eine 3D-Rekonstruktion einer bekannten Struktur vorgenommen werden und im Anschluss die Abweichung zwischen den Maßen im Model und den realen Werten bestimmt werden.

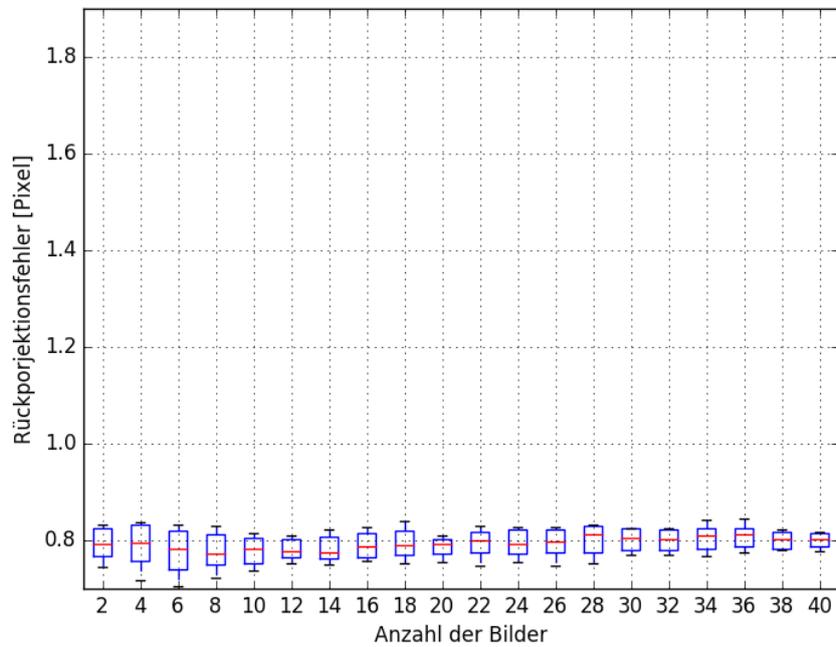
# Anhang A

## Zusätzliche Ergebnisse

In diesem Abschnitt befinden sich die Grafiken mit den Ergebnissen der *Sony Alpha6000* zu den in Kapitel 4 diskutierten Versuchen. Die Abbildungen A.1, A.2 und A.3 zeigen die Veränderung des Rückprojektionsfehlers in Abhängigkeit der Anzahl der Bilder. Die geschätzten Brennweiten in Luft für jeweils die 3D-Struktur, das DIN A3 und DIN A4 Muster finden sich in den Abbildungen A.9, A.5 und A.7. Die äquivalenten Grafiken für die Aufnahmen unter Wasser befinden sich in den Abbildungen A.8, A.4 und A.6.

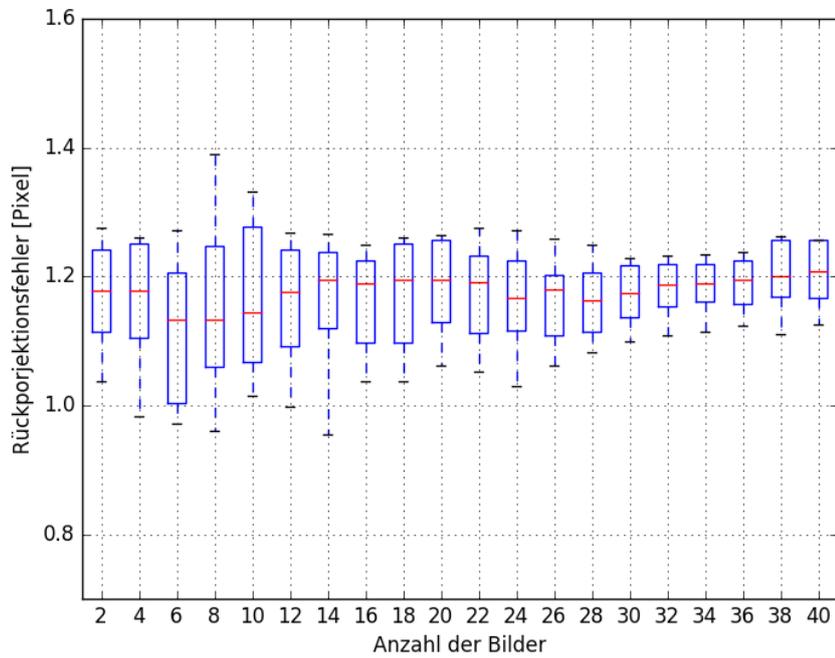


(a) unter Wasser

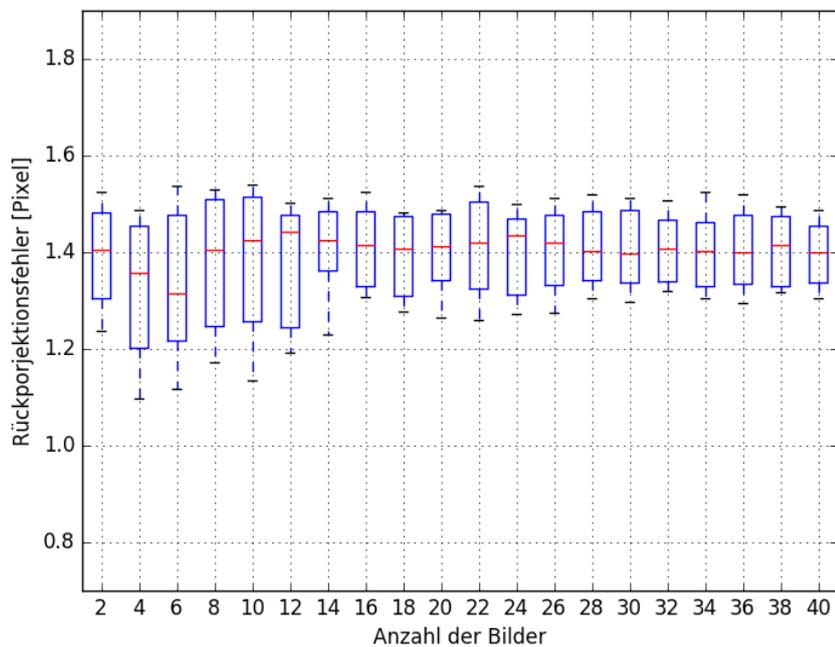


(b) in Luft

**Abbildung A.1:** Veränderung über Rückprojektionsfehler für eine unterschiedliche Anzahl an Bildern mit der 3D-Struktur, aufgenommen mit einer Sony Alpha6000.

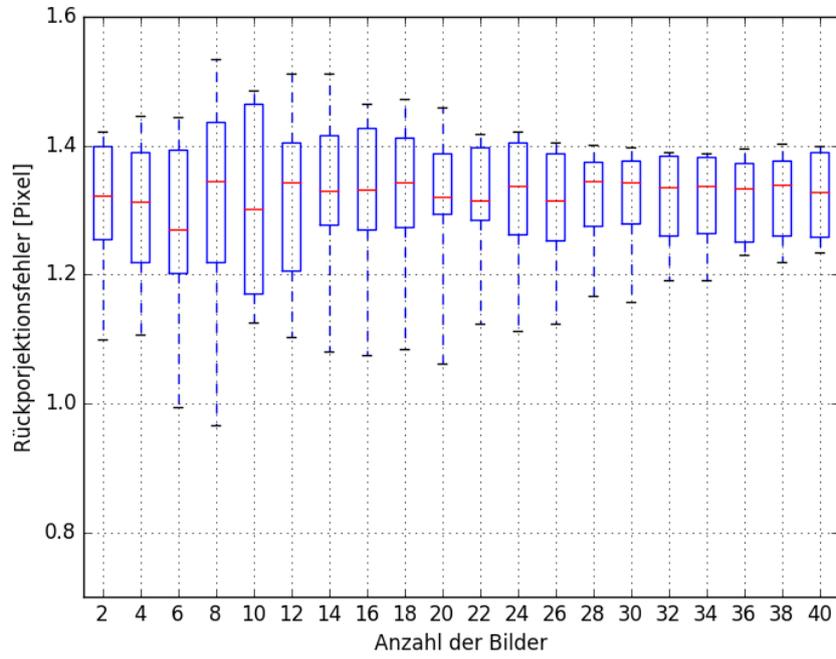


(a) unter Wasser

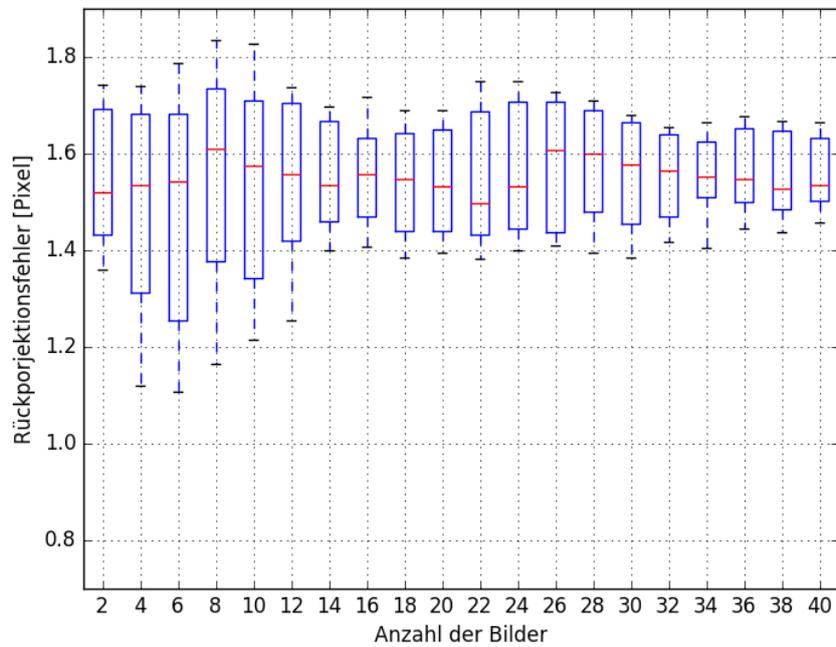


(b) in Luft

**Abbildung A.2:** Veränderung über Rückprojektionsfehler für eine unterschiedliche Anzahl an Bildern mit einem planaren DIN A3 Muster, aufgenommen mit einer Sony Alpha6000.

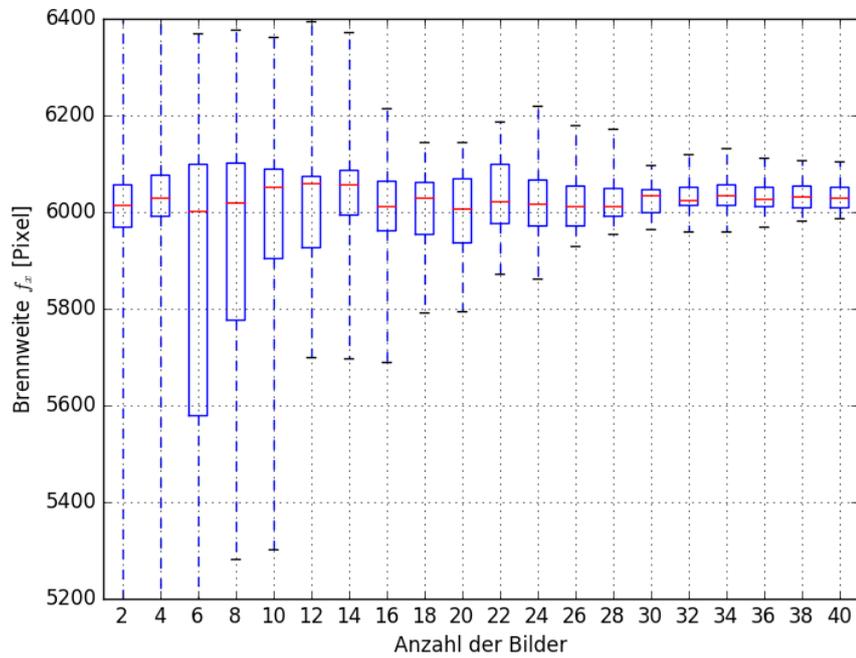
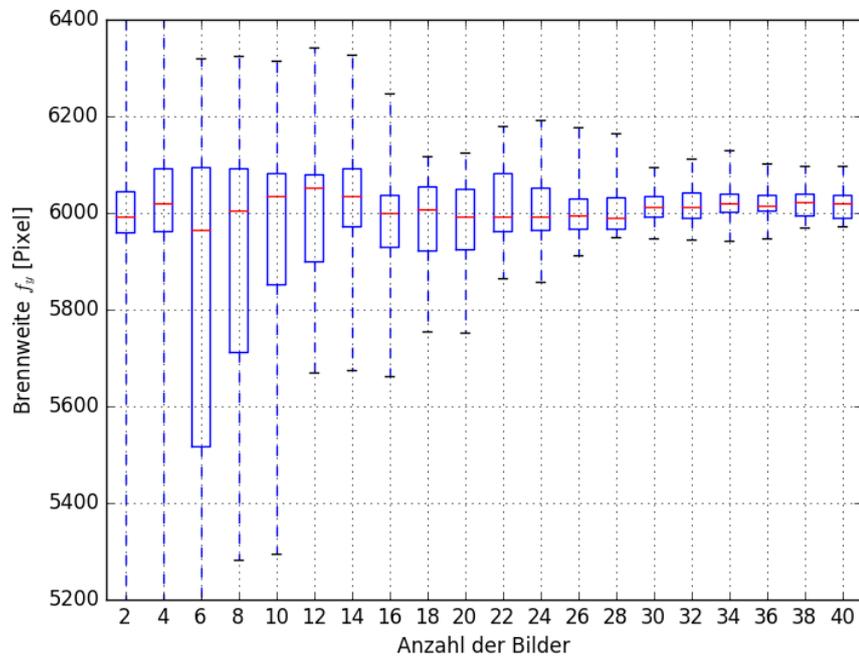


(a) unter Wasser

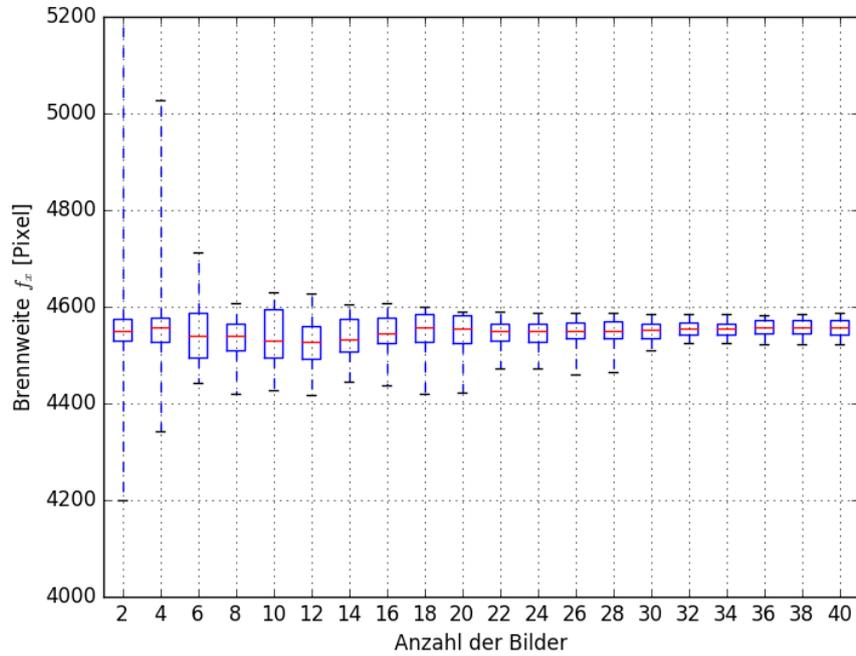
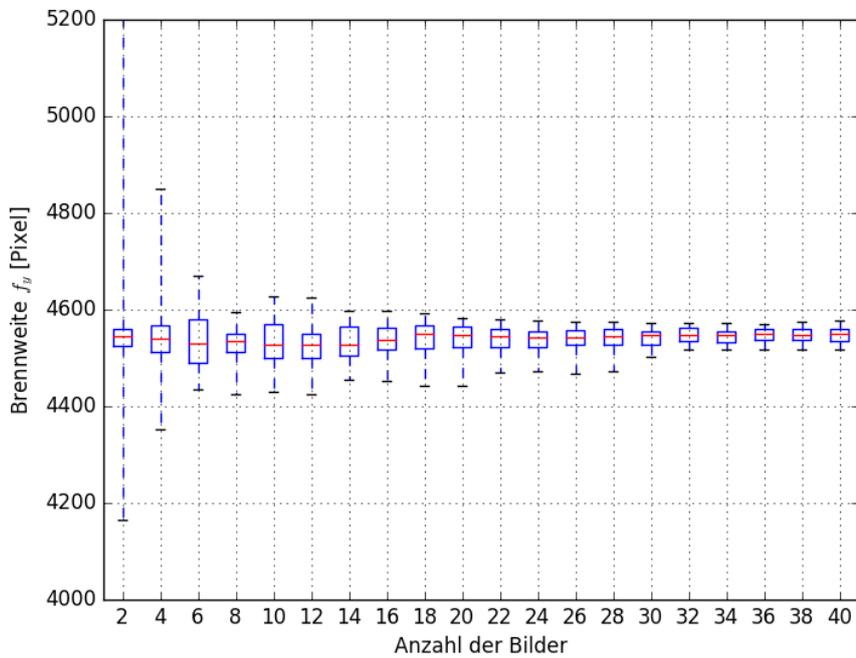


(b) in Luft

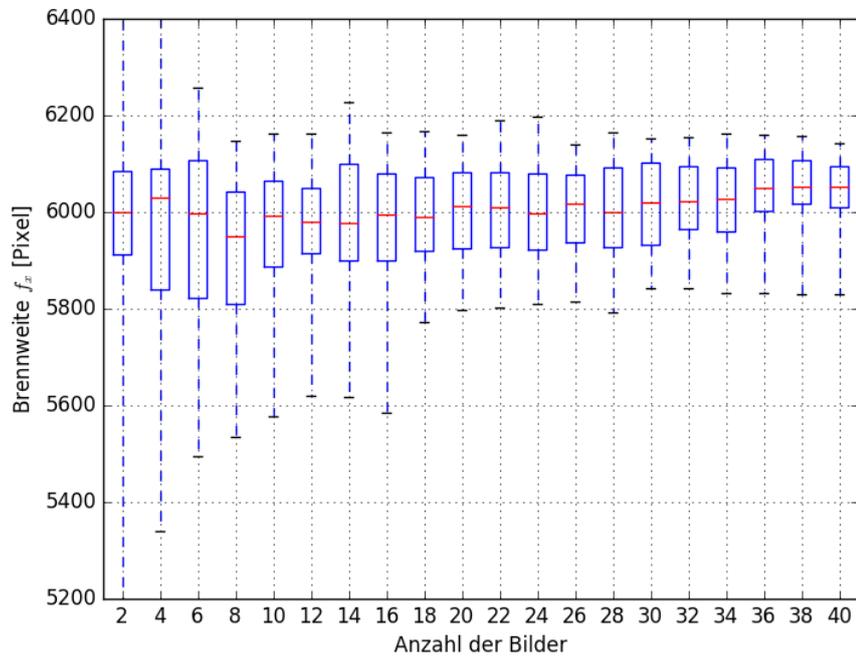
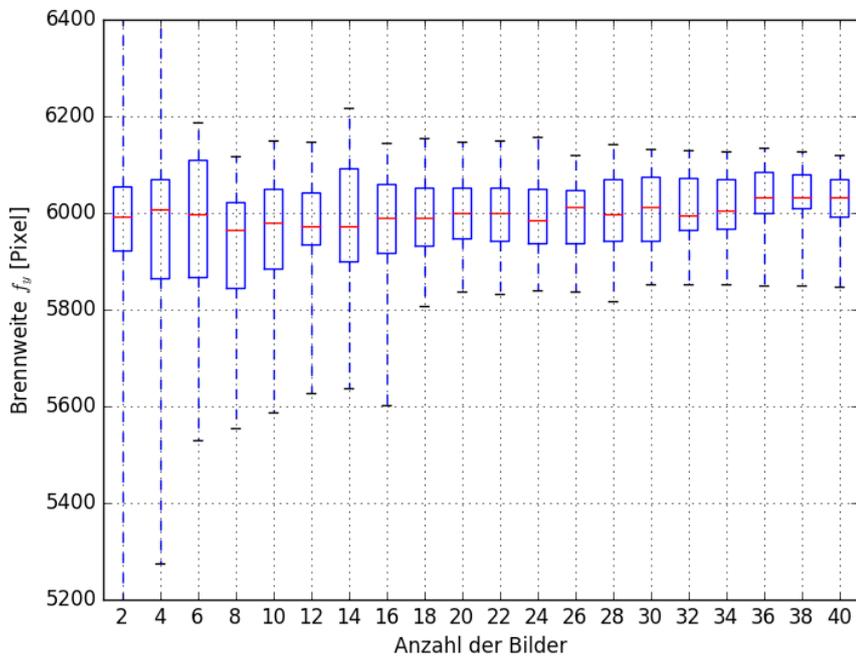
**Abbildung A.3:** Veränderung über Rückprojektionsfehler für eine unterschiedliche Anzahl an Bildern mit einem planaren DIN A4 Muster, aufgenommen mit einer Sony Alpha6000.

(a) bestimmte Brennweite  $f_x$  in Pixel(b) bestimmte Brennweite  $f_y$  in Pixel

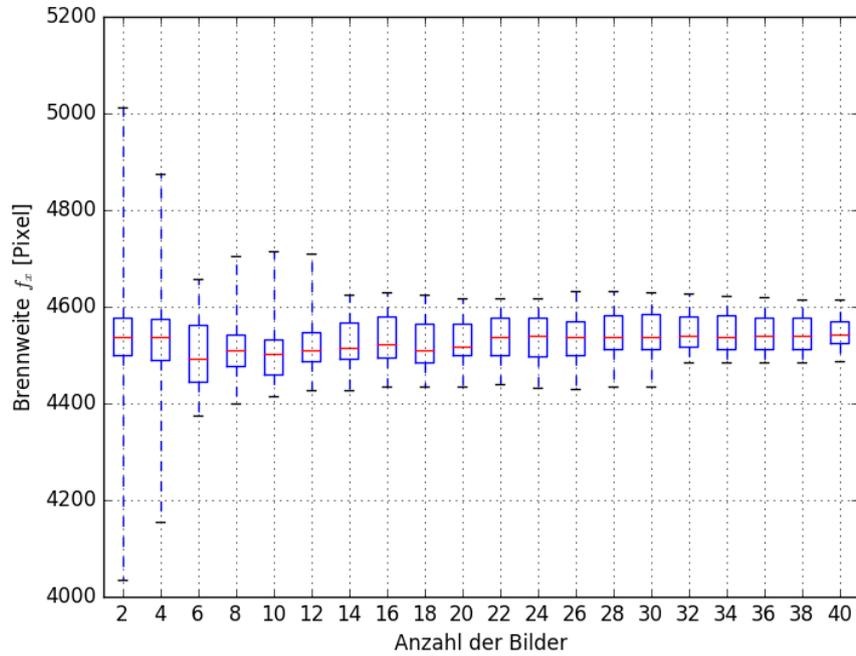
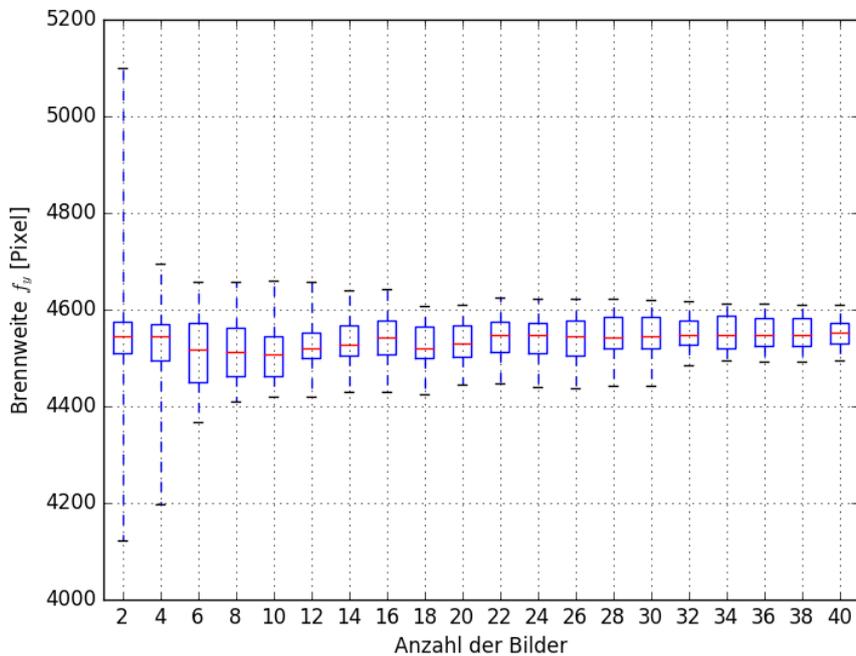
**Abbildung A.4:** Veränderung der bestimmten Brennweite für eine unterschiedliche Anzahl an Bildern, aufgenommen mit einer Sony Alpha6000 und einem planaren DIN A3 Muster unter Wasser.

(a) bestimmte Brennweite  $f_x$  in Pixel(b) bestimmte Brennweite  $f_y$  in Pixel

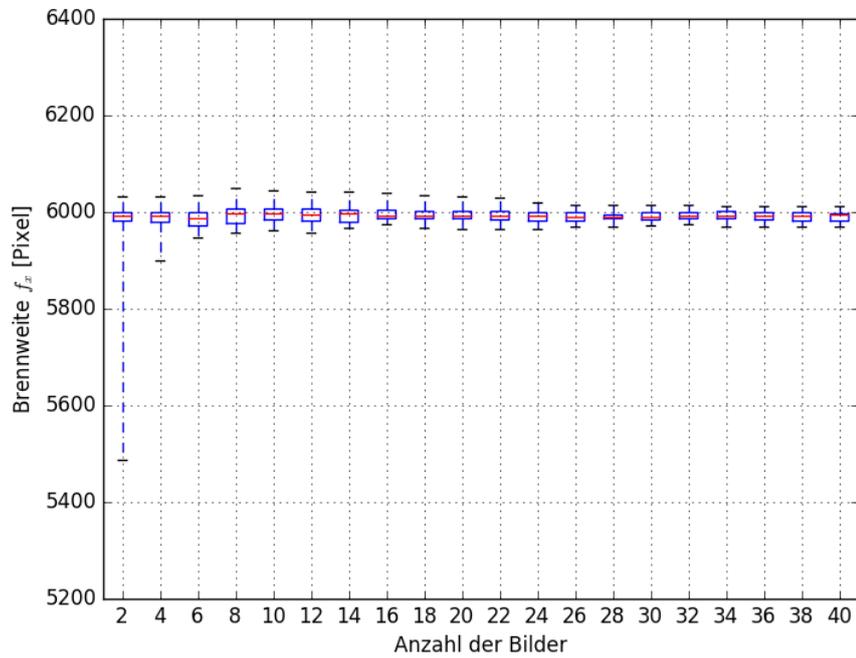
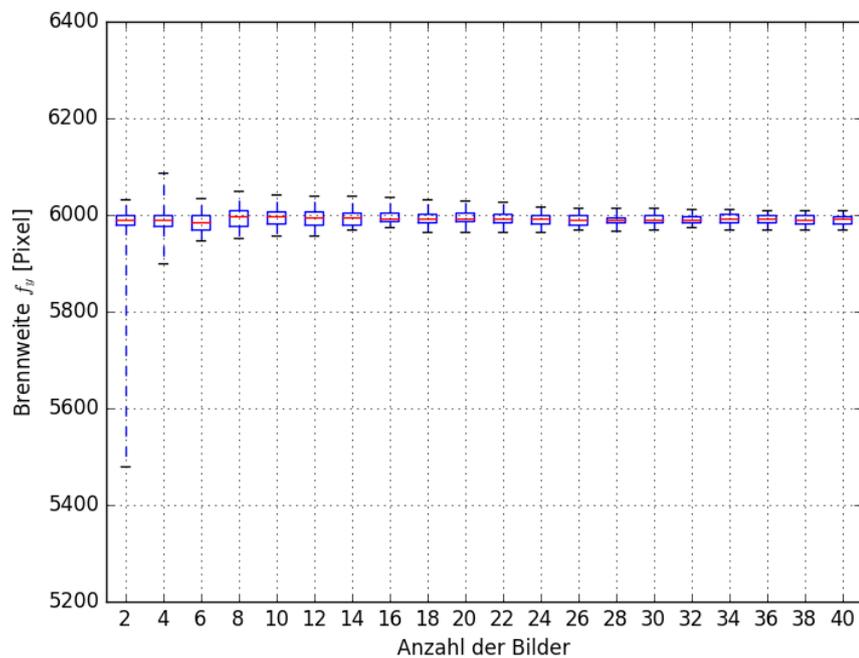
**Abbildung A.5:** Veränderung der bestimmten Brennweite für eine unterschiedliche Anzahl an Bildern, aufgenommen mit einer Sony Alpha6000 und einem planaren DIN A3 Muster.

(a) bestimmte Brennweite  $f_x$  in Pixel(b) bestimmte Brennweite  $f_y$  in Pixel

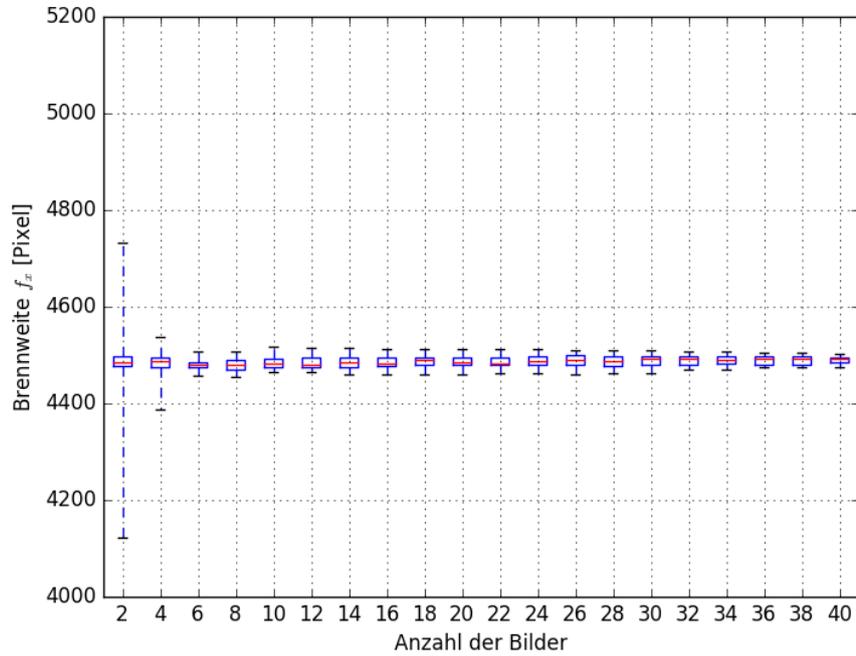
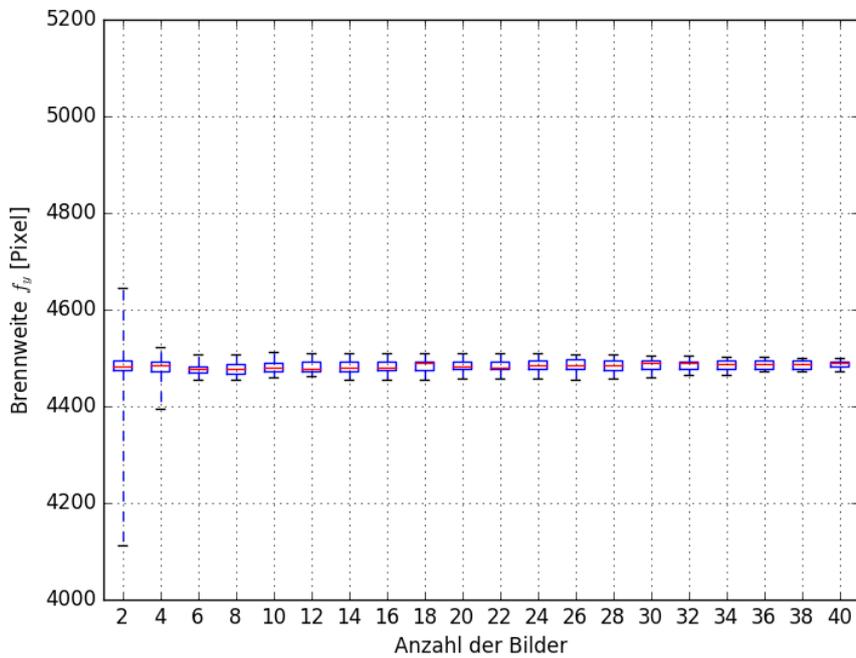
**Abbildung A.6:** Veränderung der bestimmten Brennweite für eine unterschiedliche Anzahl an Bildern, aufgenommen mit einer Sony Alpha6000 und einem planaren DINA4 Muster unter Wasser.

(a) bestimmte Brennweite  $f_x$  in Pixel(b) bestimmte Brennweite  $f_y$  in Pixel

**Abbildung A.7:** Veränderung der bestimmten Brennweite für eine unterschiedliche Anzahl an Bildern, aufgenommen mit einer Sony Alpha6000 und einem planaren DINA4 Muster.

(a) bestimmte Brennweite  $f_x$  in Pixel(b) bestimmte Brennweite  $f_y$  in Pixel

**Abbildung A.8:** Veränderung der bestimmten Brennweite für eine unterschiedliche Anzahl an Bildern, aufgenommen mit einer Sony Alpha6000 und der 3D-Struktur unter Wasser.

(a) bestimmte Brennweite  $f_x$  in Pixel(b) bestimmte Brennweite  $f_y$  in Pixel

**Abbildung A.9:** Veränderung der bestimmten Brennweite für eine unterschiedliche Anzahl an Bildern, aufgenommen mit einer Sony Alpha6000 und der 3D-Struktur.

# Literaturverzeichnis

- [1] *April Robotics Toolkit*. [https://april.eecs.umich.edu/wiki/index.php/Main\\_Page](https://april.eecs.umich.edu/wiki/index.php/Main_Page). – Zugriff am 25. Februar 2016
- [2] *boost*. <http://www.boost.org>. – Zugriff am 24. Februar 2016
- [3] *Europäische Kommission*. [http://europa.eu/rapid/press-release\\_IP-12-196\\_de.htm?locale=de](http://europa.eu/rapid/press-release_IP-12-196_de.htm?locale=de). – Zugriff am 04. April 2016
- [4] *OpenCV*. <http://opencv.org>. – Zugriff am 24. Februar 2016
- [5] *OpenCV Documentation*. [http://docs.opencv.org/2.4/doc/tutorials/calib3d/camera\\_calibration/camera\\_calibration.html](http://docs.opencv.org/2.4/doc/tutorials/calib3d/camera_calibration/camera_calibration.html). – Zugriff am 24. Februar 2016
- [6] *Radiale Verzeichnung*. <https://commons.wikimedia.org/w/index.php?curid=6708994>. – Zugriff am 21. April 2016
- [7] *VAMOS-Project*. <http://vamos-project.eu>. – Zugriff am 15. März 2016
- [8] BORN, M.; WOLF, E.: *Principles of Optics*. 7th expanded edition. Cambridge University Press. – 985 S. – ISBN 9780521642224
- [9] BOUTROS, Nader; SHORTIS, Mark R.; HARVEY, Euan S.: A comparison of calibration methods and system configurations of underwater stereo-video systems for applications in marine ecology. In: *Limnology and Oceanography: Methods* 13 (2015), Nr. 5, S. 224–236. – ISSN 1541–5856
- [10] HARTLEY, R. I.; ZISSERMAN, A.: *Multiple View Geometry in Computer Vision*. Second. Cambridge University Press, 2004
- [11] HARVEY, Euan; SHORTIS, Mark: A system for stereo-video measurement of sub-tidal organisms. In: *Marine Technology Society Journal* 29 (1995), Nr. 4, S. 10–22

- 
- [12] LI, Rongxing; LI, Haihao; ZOU, Weihong; SMITH, Robert G.; CURRAN, Terry u. a.: Quantitative photogrammetric analysis of digital underwater video imagery. In: *IEEE Journal of Oceanic Engineering* 22 (1997), Nr. 2, S. 364–375
- [13] MENNA, Fabio: *Photogrammetric Surveying of the Costa Concordia*. Presentation at Large Volume Metrology Conference & Exhibition, 2013
- [14] MENNA, Fabio; NOCERINO, Erica; TROISI, Salvatore; REMONDINO, Fabio: A photogrammetric approach to survey floating and semi-submerged objects. In: *SPIE Optical Metrology 2013* International Society for Optics and Photonics, 2013
- [15] OLSON, Edwin: AprilTag: A robust and flexible visual fiducial system. In: *Proceedings of the 2011 IEEE International Conference on Robotics and Automation (ICRA)* (2011), S. 3400–3407
- [16] SCHEWE, H: *Improvement of Fishfarm Pen Design Using Computational Structural Modelling and Large-Scale Underwater Photogrammetry (CoSMoLUP)*. Bd. 31. International Archives of Photogrammetry and Remote Sensing, 1996
- [17] SHORTIS, Mark: Calibration Techniques for Accurate Measurements by Underwater Camera Systems. In: *Sensors* 15 (2015), Nr. 12, S. 30810–30826
- [18] SHORTIS, Mark: *A Review of Calibration, Validation and 3D Measurements for Underwater Stereo-video Systems*. Presentation at ISPRS / CIPA Workshop on Underwater 3D Recording & Modeling, 2015
- [19] TELEM, Gili; FILIN, Sagi: Photogrammetric modeling of underwater environments. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (2010), Nr. 5, S. 433–444
- [20] TREIBITZ, Tali; SCHECHNER, Yoav Y.; SINGH, Hanumant: Flat refractive geometry. In: *Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008. – ISSN 1063–6919
- [21] ZHANG, Z: A flexible new technique for camera calibration. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (2000), November, Nr. 11, S. 1330–1334

# Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Würzburg, April 2016