



INSTITUT FÜR INFORMATIK VII
ROBOTIK UND TELEMATIK

Masterarbeit

**Evaluierung verschiedener Kameramodelle
zur horizontbasierten Färbung von
3D-Punktwolken mithilfe von Fish-Eye
Objektiven**

Michael Derleth

April 2021

Erstgutachter: Prof. Dr. Andreas Nüchter
Zweitgutachter: Prof. Dr. Klaus Schilling

Zusammenfassung

3D-Laserscanner finden immer öfter Einsatz zur Erstellung von Punktwolken oder 3D-Modellen. Ob im Bereich des autonomen Fahrens, der Dokumentation von Unfällen und Tatorten oder bei der Vermessung von Immobilien ist es immer von Vorteil den Messpunkten, die technologiebedingt keinerlei Farbinformationen besitzen, durch den Einsatz von einer oder mehreren Kameras Farben zuzuweisen. Eingefärbte Punktwolken bieten einen deutlich höheren Informationsgehalt, da sich Objekte dadurch klarer voneinander trennen lassen, die gescannte Umgebung sich intuitiv besser begreifen lässt und nicht nur ausschließlich die Tiefeninformationen vorliegen.

Hierfür bieten sich Fischaugenkameras an, da sie durch den großen Öffnungswinkel mit nur wenigen Bildern die komplette Umgebung abbilden können. Für diesen Vorteil müssen leider starke radiale Verzeichnungen in Kauf genommen werden, die aber mit speziellen Kamera- und Distortionmodellen korrigiert werden können. Diese Kameramodelle werden mit verschiedenen Kamerasystemen evaluiert und miteinander verglichen.

Um eine 3D-Punktwolke mithilfe einer Kamera einzufärben, sind deren genaue Position, Ausrichtung und einige Parameter, die deren Abbildungsverhalten beschreiben, nötig. Diese Daten zu erhalten ist in der Regel mit einem großen Aufwand verbunden.

Im Laufe dieser Arbeit wird eine Methode vorgestellt und getestet, die es ermöglichen soll, einen Laserscan mithilfe einer beliebigen Kamera, anhand von manuell oder automatisch gefundenen Korrespondenzen auf der Horizontlinie, einzufärben.

Danksagung

Ich möchte mich hiermit bei meinem Erstgutachter **Herrn Prof. Dr. Andreas Nüchter** bedanken, der mich bei unseren wöchentlichen Meetings mit Rat und Tat unterstützt hat. Ohne seine vorangegangenen Vorlesungen *Advanced Automation*, *3D-Pointcloud Processing* und *Photogrammetric Machine Vision* wäre diese Arbeit nicht möglich gewesen.

Mein Dank gilt auch **Frau Dr. Dorit Borrmann**, die mir den Umgang mit dem 3D-Laserscanner erklärt hat und mich bei allen dabei auftretenden Problemen unterstützt hat.

Herzlich bedanken möchte ich mich auch bei meinen **Eltern** für ihre Unterstützung während dieser Arbeit, ohne sie wäre dieses Studium nicht möglich gewesen.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Aufbau der Arbeit	2
2	Technische Grundlagen	3
2.1	Aufbau und Funktionsweise eines 3D-Laserscanners	3
2.2	Aufbau und Funktionsweise einer Kamera	5
2.2.1	Besonderheiten einer Fischaugenkamera	6
2.3	Kamerakalibrierung	7
2.3.1	Chromatische Aberration	7
2.3.2	Verzeichnungen	8
2.3.3	Kameramodelle	9
2.3.4	Funktionsweise der Kalibrierung	12
3	Evaluation der Kameramodelle	15
3.1	Versuchsaufbau	16
3.2	Verwendete Kamerasysteme	17
3.2.1	Fischaugenobjektiv Walimex pro 8mm F3,5 Fisheye II	18
3.2.2	Canon 18-55mm f/3.5-5.6 IS STM	18
3.2.3	Xiaomi Mi 9 SE	18
3.2.4	Xiaomi Mi 9 SE mit Fischaugenobjektiv	19
3.3	Evaluation der Kalibrierungen	20
3.3.1	Vergleich des Kalibrierfehlers mit Distortionparametern	20
3.3.2	Optischer Vergleich der Kalibrierungen mit Distortionparametern	24
3.3.3	Ergebnisse der Kalibrierung ohne Distortionparameter	32
3.3.4	Optischer Vergleich der Ergebnisse ohne Distortionparameter anhand des Walimex Pro Fisheye Objektivs	36
3.4	Laufzeitvergleich der Projektionsmodelle	38

4	Färben einer 3D-Punktwolke	41
4.1	Prinzip der Färbung	42
4.2	Evaluation der Färbungen mit den verschiedenen Kameramodellen	43
4.2.1	Verwendete Datensätze	43
4.2.2	Evaluierung mit dem Ground Datensatz	45
4.2.3	Evaluierung mit dem Horizon Datensatz	58
5	Automatische Korrespondenzenerkennung	61
5.1	Automatische Horizontextraktion	61
5.2	Automatische Extraktion der Keypoints	63
5.3	Anwendung des Feature Mappers	64
5.4	Evaluation der automatisch extrahierten Korrespondenzen	69
6	Fazit und Ausblick	71

Abbildungsverzeichnis

2.1	Riegl VZ-400 Laserscanner mit aufgesetzter Kamera.	3
2.2	Aufbau eines 3D-Laserscanners.	4
2.3	Schematischer Aufbau einer Kamera.	5
2.4	Schema einer Fischaugenoptik.	6
2.5	Chromatische Aberration.	8
2.6	Beispiele für Verzeichnungen.	8
2.7	Aufbau einer Lochkamera.	9
2.8	Perspektivische Projektion.	10
2.9	Auswirkung der Kalibrierung.	13
3.1	Beispiele für Kalibrierfotos.	16
3.2	Eingesetzte Kamerasysteme.	17
3.3	Beispielbild des Clip-On Objektivs.	19
3.4	Ergebnisse der Kalibrierungen für das Walimex Pro Fisheye Objektiv.	25
3.5	Abbildungsfehler des perspektivischen Modells.	26
3.6	Ergebnisse der Kalibrierungen für das Canon 18-55mm f/3.5-6.6 Objektiv.	27
3.7	Ergebnisse der Kalibrierungen für das Xiaomi Mi 9 SE Smartphone.	29
3.8	Ergebnisse der Kalibrierungen für das Xiaomi Mi 9 SE Smartphone mit aufgesetztem Fischaugenobjektiv.	31
3.9	Ergebnisse der Kalibrierungen für das Walimex Pro Fisheye Objektiv ohne Distortionparameter.	37
3.10	Laufzeitvergleich der verschiedenen Projektionsmodelle anhand des Walimex Pro Fisheye Objektivs mit 100 Kalibrierfotos.	38
3.11	Laufzeitvergleich der verschiedenen Projektionsmodelle anhand des Canon 18-55mm Objektivs mit 100 Kalibrierfotos.	39
4.1	Verschiedene Darstellungsvarianten einer 3D-Punktwolke.	41
4.2	Beispielbild eines für die Färbung verwendeten Bildes.	42
4.3	Reflectanceimage des Scans mit Beispielen für Punktkorrespondenzen.	44
4.4	Färbung mit dem perspektivischen Kameramodell mit Initial Guess.	46
4.5	Färbung mit dem perspektivischen Kameramodell ohne Initial Guess.	48
4.6	Färbung mit dem stereographischen Kameramodell mit Initial Guess.	49
4.7	Färbung mit dem stereographischen Kameramodell ohne Initial Guess.	50
4.8	Färbung mit dem equidistanten Kameramodell mit Initial Guess.	51

4.9	Färbung mit dem equidistanten Kameramodell ohne Initial Guess.	53
4.10	Färbung mit dem orthogonalen Kameramodell mit Initial Guess.	54
4.11	Färbung mit dem orthogonalen Kameramodell ohne Initial Guess.	55
4.12	Färbung mit dem equisolidangle Kameramodell mit Initial Guess.	56
4.13	Färbung mit dem equisolidangle Kameramodell ohne Initial Guess	57
4.14	Punktkorrespondenzen in Photo Nummer 5.	59
4.15	Beispiele für Punktkorrespondenzen des Horizon Datensatzes.	60
5.1	Screenshot der Horizontentfernung aus einem Panoramabild.	62
5.2	Screenshot der Horizontentfernung aus einem Foto.	63
5.3	Binärbild des Laserscans ungefiltert.	64
5.4	Binärbild des Laserscans mit Filter.	65
5.5	Parameter zur Horizontextraktion aus Fotos.	66
5.6	Screenshot gefundener Matches mittels Sift.	67
5.7	Parameter zur Erstellung eines Kantenbildes.	68
5.8	Screenshot gefundener Matches mittels Sift im Kantenbild.	69

Tabellenverzeichnis

2.1	Verschiedene Projektionsmodelle für Fischaugenkameras.	11
2.2	Eigenschaften und Bezeichnungen der Kameramodelle.	11
3.1	Reprojection Errors der perspektivischen Projektion.	21
3.2	Reprojection Errors der stereographischen Projektion.	21
3.3	Reprojection Errors der equidistanten Projektion.	22
3.4	Reprojection Errors der orthogonalen Projektion.	23
3.5	Reprojection Errors der equisolidangle Projektion.	23
3.6	Walimex Pro Fisheye Kalibrierwerte mit Distortionparameter.	24
3.7	Canon 18-55mm f/3.5-5.6 IS STM Kalibrierwerte mit Distortionparameter. . . .	26
3.8	Xiaomi Mi 9 SE Kalibrierwerte mit Distortionparameter.	28
3.9	Kalibrierwerte mit Distortionparameter für das Xiaomi Mi 9 SE Smartphone mit aufgesetztem Fischaugenobjektiv.	30
3.10	Reprojection Errors der perspektivischen Projektion ohne Distortionkoeffizienten.	32
3.11	Reprojection Errors der stereographischen Projektion ohne Distortionkoeffizienten.	33
3.12	Reprojection Errors der equidistanten Projektion ohne Distortionkoeffizienten. .	33
3.13	Reprojection Errors der orthogonalen Projektion ohne Distortionkoeffizienten. . .	34
3.14	Reprojection Errors der equisolidangle Projektion ohne Distortionkoeffizienten. .	35
3.15	Kalibrierwerte ohne Distortionparameter für das Walimex Pro Fisheye Objektiv.	36
4.1	Für die Färbung verwendete Datensätze.	43
4.2	Kalibrierfehler auf dem Ground Datensatz.	44
4.3	Kalibrierfehler auf dem Horizon Datensatz.	44
4.4	Intrinsik und Extrinsik - Perspektivisch mit Initial Guess.	47
4.5	Intrinsik und Extrinsik - Perspektivisch ohne Initial Guess.	48
4.6	Intrinsik und Extrinsik - Stereographisch mit Initial Guess.	49
4.7	Intrinsik und Extrinsik - Stereographisch ohne Initial Guess.	50
4.8	Intrinsik und Extrinsik - Equidistant mit Initial Guess.	52
4.9	Intrinsik und Extrinsik - Equidistant ohne Initial Guess.	53
4.10	Intrinsik und Extrinsik - Orthogonal mit Initial Guess.	54
4.11	Intrinsik und Extrinsik - Orthogonal ohne Initial Guess.	55
4.12	Intrinsik und Extrinsik - Equisolidangle mit Initial Guess.	56
4.13	Intrinsik und Extrinsik - Equisolidangle ohne Initial Guess.	57
4.14	Anzahl der Korrespondenzen.	58

Kapitel 1

Einleitung

3D-Modelle spielen in der heutigen Gesellschaft eine immer größere Rolle und finden auch nicht mehr ausschließlich in Forschung und Industrie Anwendung. Durch sie ist es inzwischen bereits möglich, Wohnungen und ganze Gebäude digital zu besichtigen. Ein farbiges Modell mit hoher Auflösung ermöglicht es hierbei, einen guten Eindruck des Objekts zu erhalten. Es lassen sich bereits darin beliebige Strecken messen, um beispielsweise herauszufinden, ob die eigenen Möbel in die neue Wohnung passen oder nicht.

Auch für das autonome Fahren spielen solche eingefärbte Punktwolken eine große Rolle. Denn anders als bei dem Einsatz einer bloßen Kamera lässt sich mit einem 3D-Scan unterscheiden, ob ein Mensch oder anderes Objekt nur beispielsweise auf einem LKW oder einer Plakatwand abgebildet, oder tatsächlich vorhanden ist. Außerdem lassen sich mithilfe von künstlicher Intelligenz oder anderen Objekterkennungsverfahren durch die zusätzlichen Farbwerte bessere Ergebnisse erzielen als ohne.

Die Verknüpfung der zweidimensionalen Farbwerte einer Kamera mit den dreidimensionalen eines 3D-Laserscans ist jedoch eine technische Herausforderung. Insbesondere bei Scannern, die nicht von Werk aus damit ausgestattet sind, können Kameras oft nur mit einem großen Versatz zu diesem montiert werden. Dies führt dazu, dass sich die Position und damit auch die Perspektive derselben, stark von dem Scanner unterscheidet. Außerdem können diese in der Regel nicht so fest montiert werden, dass eine einmalige Konfiguration ausreicht. Sie müssen daher bei jedem Einsatz neu kalibriert werden, da sich durch die erneute Montage die Position geringfügig verändern kann.

Für die Kalibrierung ist es nötig, korrespondierende Punkte zwischen der 3D-Punktwolke und den Kamerabildern zu kennen, also zu wissen, welcher Messpunkt des Scanners auf welchem Pixel des Kamerasensors abgebildet wird. Hiermit lassen sich sowohl Parameter finden, welche die Kamera selbst beschreiben, wie beispielsweise die Brennweiten, als auch diejenigen, welche die Position und Ausrichtung eben dieser relativ zum 3D-Scanner festlegen.

In der vorliegenden Arbeit werden unterschiedliche Projektionsmodelle getestet und verglichen, welche die Abbildungsfunktion für Fischaugenkameras beschreiben. Dieser Kameratyp hat durch

den großen Öffnungswinkel den Vorteil, dass bereits mit wenigen Bildern ein kompletter 360° Scan eingefärbt werden kann. Außerdem haben Fischaugenoptiken den Vorteil, dass sie Bereiche direkt vor ihnen auf einem größeren Bereich des Sensors abbilden als normale Objektive.

Nachteilig ist jedoch die starke radiale Verzerrung, welche in jedem Bild auftritt und nachträglich mithilfe von den in dieser Arbeit vorgestellten Kameramodellen und Distortionmodellen entfernt werden muss.

Um die oben genannten Korrespondenzen zwischen Scan und Kamerabildern zu erhalten, sind in der Regel entweder spezielle Objekte nötig, auf denen anschließend automatisch Punkte detektiert und extrahiert werden können, oder diese müssen manuell gesucht werden. In dieser Arbeit wird ein Verfahren vorgestellt und getestet, welches Punkte auf der Horizontlinie in Bildern und Scan miteinander verbindet und somit die gesuchten Korrespondenzen erhält.

1.1 Aufbau der Arbeit

Die vorliegende Arbeit ist in sechs Kapitel unterteilt:

Kapitel 1 ist die vorliegende Einleitung.

Kapitel 2 erläutert die technischen Grundlagen. Hierbei wird auf die Funktionsweise des verwendeten Laserscanners und der Kameras eingegangen. Außerdem wird der Kalibriervorgang von Kameras skizziert und die verschiedenen Projektionsmodelle vorgestellt.

Kapitel 3 beschreibt, wie verschiedene Kameras mit den unterschiedlichen Projektionsmodellen kalibriert wurden und stellt die Ergebnisse sowohl mit als auch ohne Distortionparametern vor. Hierbei werden der Reprojection Error, der optische Eindruck und die Laufzeit verglichen.

Kapitel 4 stellt die Kamerakalibrierung und Färbung einer 3D-Punktwolke mit den verschiedenen Kameramodellen vor. Dafür werden sowohl beliebige Korrespondenzen als auch welche, die sich ausschließlich auf der Horizontlinie befinden, verwendet.

Kapitel 5 stellt ein Programm und Verfahren vor automatisch die Horizontlinie in Scan und Bild zu erhalten und mit Featureextraktoren Korrespondenzen zu bestimmen.

Kapitel 6 fasst die Ergebnisse der Arbeit zusammen und erläutert Lösungsvorschläge für gefundene Problemstellungen.

Kapitel 2

Technische Grundlagen

2.1 Aufbau und Funktionsweise eines 3D-Laserscanners

3D-Laserscanner erfreuen sich immer größerer Beliebtheit. Dies lässt sich dadurch erklären, dass die Technik sowohl immer günstiger wird, als auch, dass es einen immer größeren Bedarf an 3D-Modellen von realen Umgebungen gibt. 3D-Laserscanner werden unter anderem für autonomes Fahren, von Maklern zur Erstellung von virtuellen Wohnungsbesichtigungen oder von der Polizei zur Aufklärung und Rekonstruktion von Verkehrsunfällen und Tatorten eingesetzt. Der Vorteil liegt darin, dass innerhalb von einigen Sekunden bis hin zu wenigen Minuten, je nach benötigter Auflösung, ein Bereich komplett gescannt und digitalisiert werden kann. Nachteilig sind in der Regel die Größe der Geräte, der hohe Anschaffungspreis und die Probleme beim Scannen von spiegelnden Oberflächen. 3D-Laserscanner funktionieren vom Prinzip her genauso



Abbildung 2.1: Riegl VZ-400 3D-Laserscanner mit aufgesetzter Kamera.

wie die günstigen Laser-Distanzmesser, die inzwischen aufgrund ihrer schnellen Anwendung und der hohen erreichbaren Genauigkeit oft den klassischen Zollstock ablösen. Ein Laserstrahl wird ausgesendet und anschließend die Zeit gemessen, die dieser benötigt, um ein anvisiertes Objekt zu erreichen, davon reflektiert zu werden und wieder am Messgerät einzutreffen. Daher lässt sich die Entfernung d mit folgender Formel berechnen:

$$d = \left(\frac{c \cdot t}{2}\right) \quad (2.1)$$

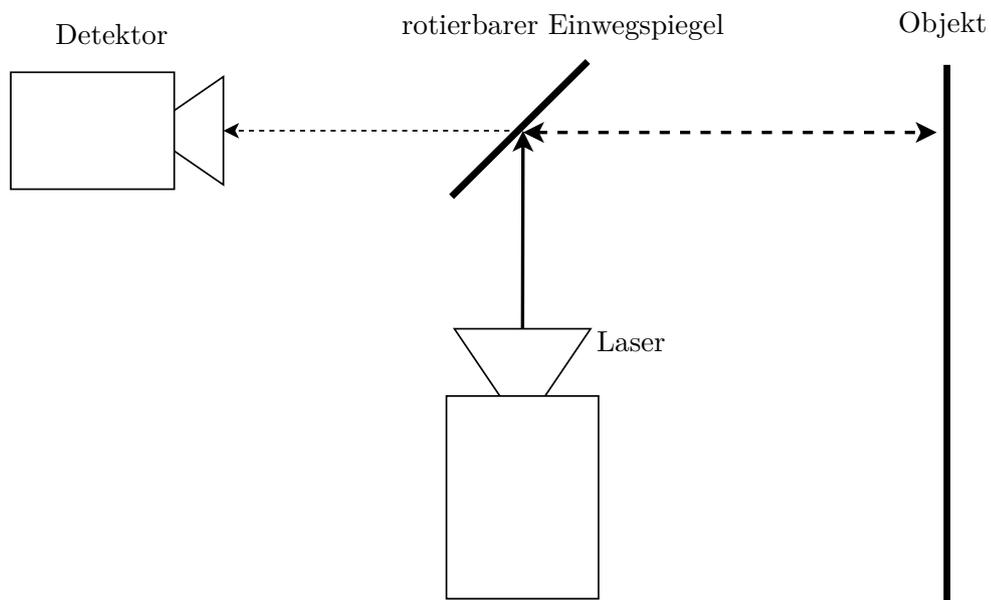


Abbildung 2.2: Aufbau eines Laserscanners. Adaptiert von [8].

Die gesuchte Distanz ergibt sich also, indem die gemessene Zeit t mit der Lichtgeschwindigkeit c multipliziert und anschließend halbiert wird. Um nun nicht nur die Entfernung, sondern auch die Position relativ zum Laserscanner zu erhalten, wird zusätzlich die Richtung benötigt, in die der Laserstrahl emittiert wurde. Diese stellt kein Problem dar, da der in dieser Arbeit verwendete Scanner, der in Abb. 2.2 zu sehen ist, den Laserstrahl mithilfe eines Spiegels um insgesamt bis zu 100° vertikal ablenkt und dessen Winkel kennt. Zusätzlich kann sich der Scanner horizontal um bis zu 360° drehen, wodurch die gesamte Umgebung mit Ausnahme von zwei kleinen Bereichen direkt unter- und oberhalb seiner Position gescannt werden kann. Damit sind alle Informationen für einen 3D-Punkt in Polarkoordinaten bekannt. Diese Polarkoordinaten werden anschließend in die gebräuchlicheren kartesischen Koordinaten umgerechnet. Durch die Verwendung von moduliertem Licht kann die Scanfrequenz deutlich erhöht und zusätzlich auch die Intensität des reflektierten Impulses erfasst werden.

2.2 Aufbau und Funktionsweise einer Kamera

Kameras sind aus dem heutigen Leben kaum noch wegzudenken. Beinahe jeder trägt inzwischen mindestens eine - oft sogar mehrere mit verschiedenen Objektiven - in seinem Handy integriert mit sich. Üblich sind dort nicht nur die in der Rückseite verbauten Kameras und die Selfiekamera, sondern inzwischen werden auch oft verschiedene Teleobjektive, Makrolinsen und teilweise sogar Wärmebildkameras eingesetzt. Alle heutigen Kameras besitzen im Prinzip den selben Aufbau, der in Grafik 2.3 schematisch und vereinfacht dargestellt ist.

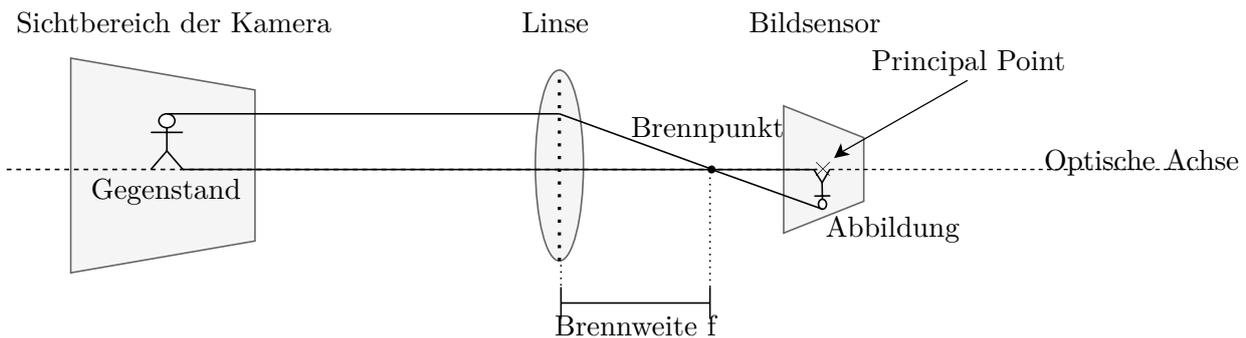


Abbildung 2.3: Aufbau einer Kamera.

Licht, das von einem Gegenstand reflektiert wird und in die Kamera trifft, wird durch eine, oder oftmals durch ein komplexes System von Linsen im Brennpunkt gebündelt und dahinter um 180° gedreht auf dem Bildsensor abgebildet. Der Bildsensor hat im Fall einer Farbkamera auf dem Bildsensor ein Gitternetz aus Pixeln. Für jeden Pixel wird der Rot-, Grün-, und Blau-Anteil des darauf treffenden Lichtes gemessen und gespeichert. Günstigere Kameras speichern das Bild automatisch in einem gebräuchlichen Format wie *JPEG*. Hierbei legt der Nutzer oder die Kamera automatisch Einstellungen zu Kontrast, Schärfe und Weißabgleich fest. Hochwertigere Kameras speichern oft die einzelnen Werte für die Farbkanäle jedes Pixels im so genannten *RAW*-Format. Dies ermöglicht eine bessere Weiterverarbeitung der Daten, benötigt aber auch deutlich mehr Speicherplatz.

2.2.1 Besonderheiten einer Fischaugenkamera

Fischaugenkameras besitzen Objektive mit einem sehr großen Öffnungswinkel. Der Vorteil liegt darin, dass - im Gegensatz zu normalen Objektiven - Sichtbereiche von 180° und mehr erreicht werden können. Anders als bei konventionellen Kameras wird hier der Bereich, der senkrecht zur optischen Achse steht, nicht proportional abgebildet. Dies liegt daran, dass diese Objektive eine starke radiale Verzerrung haben, die vom Bildmittelpunkt zum Bildäußeren zunimmt.

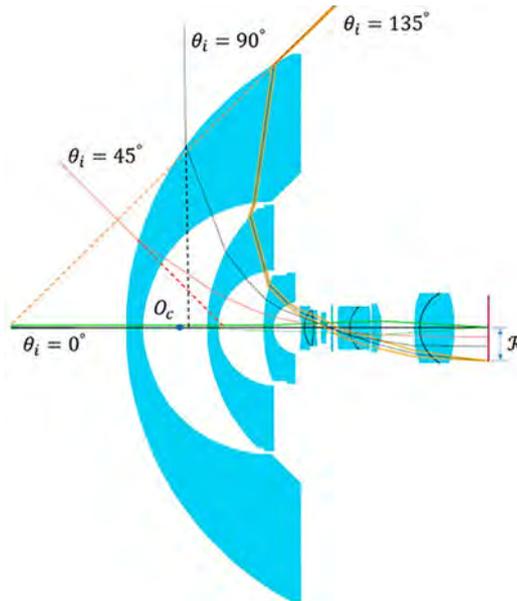


Abbildung 2.4: Aufbau einer Fischaugenoptik. Quelle: [11]

Wie in Grafik 2.4 zu sehen ist, werden einfallende Strahlen nahe der optischen Achse sehr wenig gebrochen, wohingegen Strahlen mit einem großen Einfallswinkel sehr stark abgelenkt werden und somit am Rand des Sensors auftreffen. Das bewirkt, dass der Bereich direkt vor der Kamera auf einem verhältnismäßig großen Bereich des Sensors abgelichtet wird. Dies ist für einige Anwendungsfälle ein entscheidender Vorteil gegenüber von Standardobjektiven. Lichtstrahlen mit einem geringen Einfallswinkel werden bei einem Fischaugenobjektiv also weniger abgelenkt und somit auf einem größeren Bereich des Sensors abgebildet. Strahlen, die mit einem großen Winkel einfallen, werden ebenfalls noch auf dem Sensor abgebildet, allerdings sehr weit außen und eng beieinander. Dies führt zu der typischen radialen Verzerrung, die vor allem am Bildrand stark auffällt.

2.3 Kamerakalibrierung

Die Kamerakalibrierung dient dazu sowohl die linsentypischen Verzerrungen als auch Abbildungsfehler, die fertigungsbedingt auftreten, zu entfernen und die intrinsischen und die extrinsischen Parameter einer Kamera zu bestimmen. Das folgende Kapitel beschreibt die Intention und den Vorgang der Kalibrierung stark verkürzt und vereinfacht. Details können in [10] nachgelesen werden. Zu den intrinsischen Parametern einer Kamera gehören die Brennweiten f_x , f_y und der principal Point (c_x, c_y) , der den Mittelpunkt des Sensors darstellt. Dieser sollte im Idealfall genau in der Bildmitte liegen. In der Regel weicht er aber durch Fertigungstoleranzen je nach Qualität der Kamera mehr oder weniger weit ab. Außerdem gehören zu den intrinsischen Parametern noch ein Scherungsfaktor α , der die Scherung zwischen x und y beschreibt, bei Digitalkameras jedoch in der Regel ungefähr 0 ist und daher oft ignoriert wird. Desweiteren gehören zu der Intrinsik sämtliche Verzerrungsparameter,

$$\mathbf{k} = [k_1, k_2, p_1, p_2, k_3, k_4, k_5, k_6]$$

welche je nach Kameramodell verwendet werden, um radiale und/oder tangentielle Verzerrungen zu modellieren. Die extrinsischen Parameter beschreiben die Pose, also die Ausrichtung und Position der Kamera im Weltkoordinatensystem. Folglich bestehen diese aus zwei Vektoren, der Translation

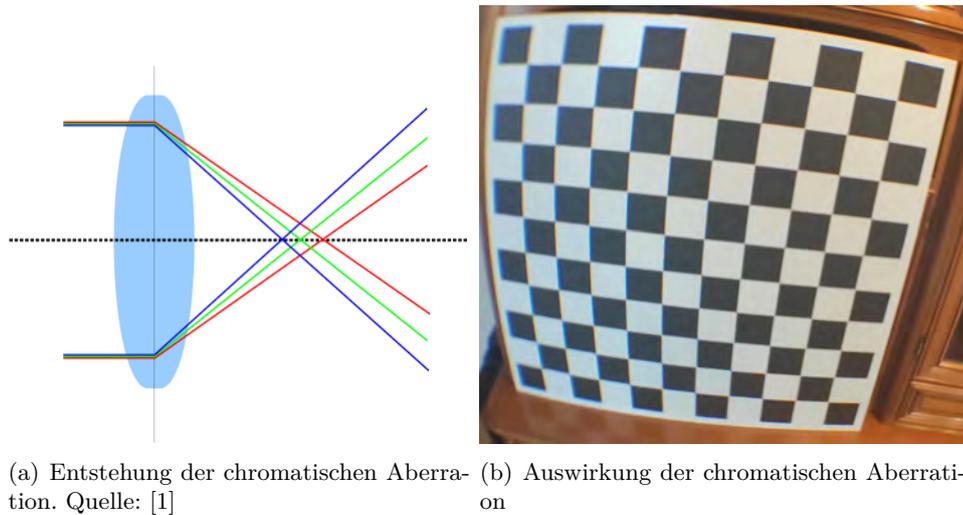
$$\mathbf{T} = [tx, ty, tz]$$

und der Rotation

$$\mathbf{R} = [rx, ry, rz].$$

2.3.1 Chromatische Aberration

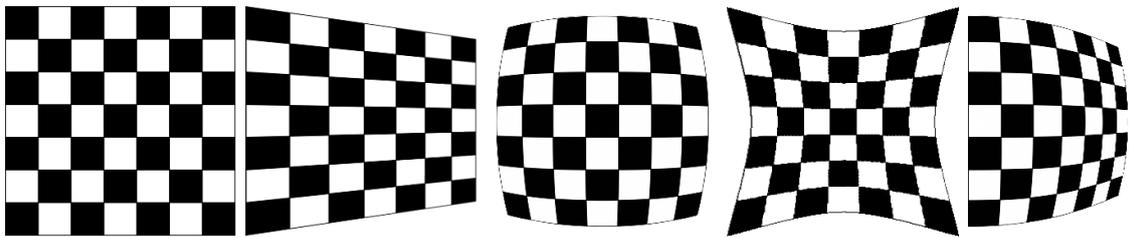
Die chromatische Aberration entsteht dadurch, dass die verschiedenen Wellenlängen von Licht durch eine Linse unterschiedlich stark gebrochen werden. Dieser Effekt ist in Abbildung 2.5(a) dargestellt und führt zu Farbverschiebungen, die besonders an Kanten und am Rand des Bildes auffallen. Bei hochwertigen Kameras tritt dieser Effekt kaum auf, da er durch den Einsatz einer Blende oder durch mehrere Linsen minimiert werden kann. Bei einem der in dieser Arbeit verwendeten Kamerasysteme (Xiaomi Mi 9SE mit aufgesetztem Fischaugenobjektiv) ist dieser Effekt, wie in Grafik 2.5(b) vor allem links im Bild zu sehen ist, jedoch stark ausgeprägt.



(a) Entstehung der chromatischen Aberration. Quelle: [1] (b) Auswirkung der chromatischen Aberration

Abbildung 2.5: Chromatische Aberration.

2.3.2 Verzeichnungen



(a) Bild ohne Verzeichnung (b) tangentielle Verzeichnung (c) radiale tonnenförmige Verzeichnung (d) radiale kissenförmige Verzeichnung (e) tangentielle und radiale Verzeichnung

Abbildung 2.6: Beispiele für Verzeichnungen.

Es gibt radiale und tangentielle Verzeichnungen. Typischerweise entstehen tangentielle Abbildungsfehler, wenn der Bildsensor nicht senkrecht zur optischen Achse steht. Dies ist besonders bei günstigen Webcams der Fall. Radiale hingegen sind von der Linse abhängig und gerade bei Fischaugenobjektiven nicht zu vermeiden. Natürlich können auch beide Fälle gemeinsam auftreten (s. Abb. 2.6). Radiale Verzeichnungen werden außerdem noch in tonnenförmige (s. Abb. 3.1(d)) und kissenförmige (s. Abb. 2.6(d)) unterteilt. Meistens werden zur Korrektur von Verzeichnungen fünf Parameter verwendet. Es ist aber auch möglich, komplexere Modelle zu verwenden. Üblicherweise werden Parameter für die radiale Verzeichnung k_i und für die tangentielle p_i genannt. In *OpenCV* werden damit die korrigierten Pixelwerte folgendermaßen berechnet (vgl. [2]):

$$x_{correctedTangential} = x + [2p_1xy + p_2(r^2 + 2x^2)]$$

$$y_{correctedTangential} = y + [p_1(r^2 + 2y^2) + 2p_2xy]$$

$$x_{correctedRadial} = x(1 + k_1r^2 + k_2r^4 + k_3^6 + k_4^8 + k_5^{10} + k_6^{12})$$

$$y_{correctedRadial} = y(1 + k_1r^2 + k_2r^4 + k_3^6 + k_4^8 + k_5^{10} + k_6^{12})$$

Wobei $r^2 = x^2 + y^2$ gilt. In *OpenCV* werden Pixel zuerst tangential und anschließend radial korrigiert.

2.3.3 Kameramodelle

In *OpenCV* und den meisten anderen verfügbaren Kamerabibliotheken wird der Einfachheit halber zur Kalibrierung das Lochkameramodell verwendet. Dieses ist leicht zu berechnen und mögliche Verzerrungen werden nachträglich hinein modelliert.

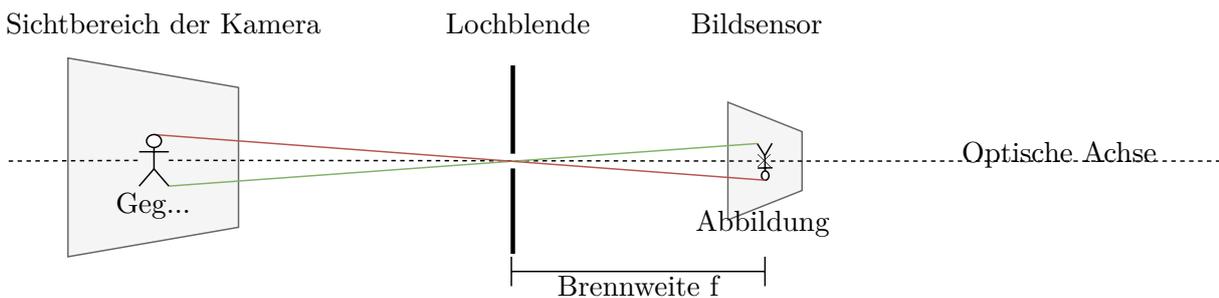


Abbildung 2.7: Aufbau einer Lochkamera.

Wie in Abb. 2.7 zu sehen ist, werden bei diesem Modell die Strahlen durch eine Lochblende auf den Bildsensor projiziert, wo sie das Objekt auf dem Kopf stehend abbilden.

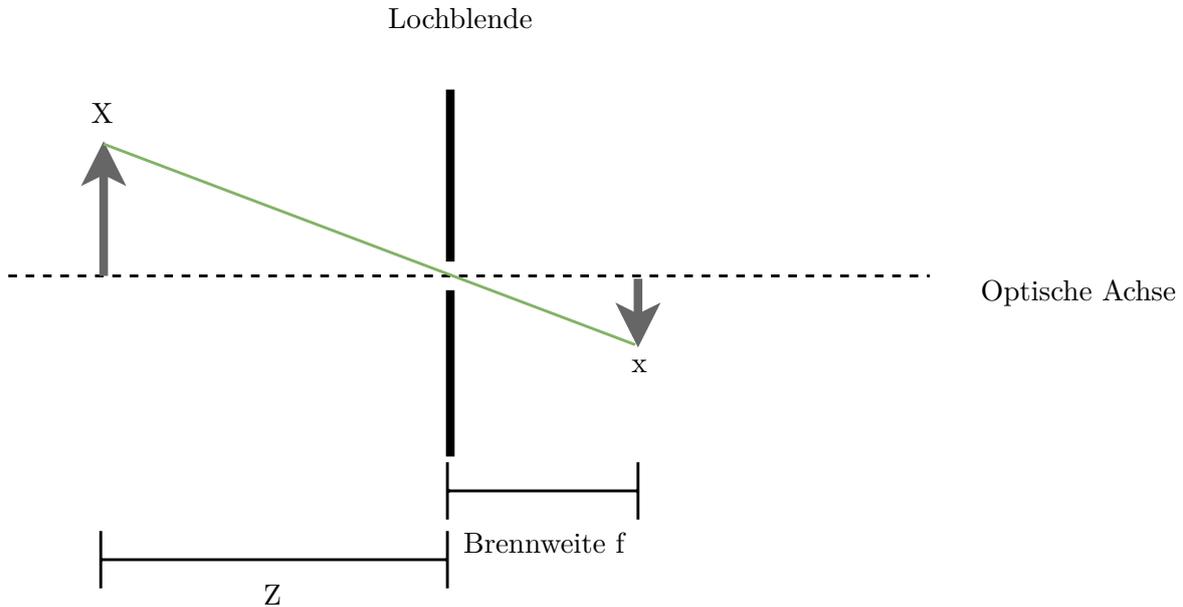


Abbildung 2.8: Perspektivische Projektion.

Von diesem Modell lässt sich mithilfe des Strahlensatzes wie in Abb. 2.8 einfach das perspektivische Modell ableiten, welches aus den X , Y und Z 3D-Koordinaten die x und y Koordinaten im Bild berechnet. Voraussetzung hierfür ist, dass die 3D-Koordinaten im Kamerakoordinatensystem vorliegen.

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

Analog funktioniert dies für y , so dass folgt:

$$\frac{Y}{y} = \frac{Z}{f}$$

Nach x bzw. y aufgelöst:

$$x = \frac{Xf}{Z}$$

$$y = \frac{Yf}{Z}$$

Zusätzlich werden die Bildkoordinaten normalisiert bzw. durch f geteilt. Daraus folgen die perspektivischen Abbildungsgleichungen $(X, Y, Z) \rightarrow (x^*, y^*)$:

$$x^* = \frac{X}{Z}; \quad y^* = \frac{Y}{Z} \quad (2.2)$$

Weitere Kameramodelle, die für Fischaugenkameras verwendet werden, sind die Folgenden:

Tabelle 2.1: Verschiedene Projektionsmodelle für Fischaugenkameras (vgl. [6]).

Projektionsmodell	$(X, Y, Z) \rightarrow (x^*, y^*)$	$(x^*, y^*) \rightarrow (X, Y, Z)$
Perspektivisch $r' = c \tan \phi$	$x^* = \frac{X}{Z}$ $y^* = \frac{Y}{Z}$	$X = x^*$ $Y = y^*$ $Z = 1$
Stereographisch $r' = c \tan \left[\frac{\phi}{2} \right]$	$x^* = \frac{X}{\sqrt{X^2 + Y^2 + Z^2} + Z}$ $y^* = \frac{Y}{\sqrt{X^2 + Y^2 + Z^2} + Z}$	$X = \frac{2x^*}{1 + x^{*2} + y^{*2}}$ $Y = \frac{2y^*}{1 + x^{*2} + y^{*2}}$ $Z = \frac{1 - (x^{*2} + y^{*2})}{1 + x^{*2} + y^{*2}}$
Equidistant $r' = c\phi$	$x^* = \frac{X}{\sqrt{X^2 + Y^2}} \arctan \left[\frac{\sqrt{X^2 + Y^2}}{Z} \right]$ $y^* = \frac{Y}{\sqrt{X^2 + Y^2}} \arctan \left[\frac{\sqrt{X^2 + Y^2}}{Z} \right]$	$X = \frac{x^*}{\sqrt{x^{*2} + y^{*2}}} \sin \left[\sqrt{x^{*2} + y^{*2}} \right]$ $Y = \frac{y^*}{\sqrt{x^{*2} + y^{*2}}} \sin \left[\sqrt{x^{*2} + y^{*2}} \right]$ $Z = \cos \left[\sqrt{x^{*2} + y^{*2}} \right]$
Orthogonal $r' = c \sin [\phi]$	$x^* = \frac{X}{\sqrt{X^2 + Y^2 + Z^2}} \quad (Z > 0)$ $y^* = \frac{Y}{\sqrt{X^2 + Y^2 + Z^2}}$	$X = x^*$ $Y = y^*$ $Z = \sqrt{1 - (x^{*2} + y^{*2})}$
Equisolidangle $r' = c \sin \left[\frac{\phi}{2} \right]$	$x^* = \frac{X}{\sqrt{2(X^2 + Y^2)}} \sqrt{1 - \frac{Z}{\sqrt{X^2 + Y^2 + Z^2}}}$ $y^* = \frac{Y}{\sqrt{2(X^2 + Y^2)}} \sqrt{1 - \frac{Z}{\sqrt{X^2 + Y^2 + Z^2}}}$	$X = 2x^* \sqrt{1 - (x^{*2} + y^{*2})}$ $Y = 2y^* \sqrt{1 - (x^{*2} + y^{*2})}$ $Z = \sqrt{1 - (x^{*2} + y^{*2})}$

Die Modelle, welche in Tabelle 2.1 zu sehen sind, wurden in der weit verbreiteten Open Source Programmbibliothek *OpenCV* [2] implementiert und in Kapitel 3.3 evaluiert. In Tabelle 2.2 sind

Tabelle 2.2: Eigenschaften und Bezeichnungen der Kameramodelle(vgl. [3, 10]).

Englische Bezeichnung	Deutsche Bezeichnung	Eigenschaften	Einschränkungen
Perspective	Perspektivisch, Gnomonisch	- Erhält Geraden	Gilt für $\phi < 180^\circ$
Stereographic	Stereographisch, Stereografisch	- Erhält Kreise	
Equidistant	Equidistant, Äquidistant	- Erlaubt Winkelmessungen in Kugelkoordinaten	
Orthogonal	Orthogonal, Orthographisch, Orthografisch		Gilt für $\phi \leq 180^\circ$
Equisolidangle	Equisolidangle, Flächentreu, Raumwinkelgleich	- Geeignet für Flächenvergleiche	

die besonderen Eigenschaften, Einschränkungen und weitere Bezeichnungen für die verschiedenen Modelle aufgelistet. In dieser Arbeit werden aus Übersichtlichkeitsgründen ausschließlich die englischen Bezeichnungen verwendet.

2.3.4 Funktionsweise der Kalibrierung

Durch die Kamerakalibrierung müssen mindestens 10, in der Regel aber deutlich mehr Parameter festgelegt werden. Dazu gehören jeweils die drei der Extrinsik, welche Rotation und Translation der Koordinaten im Weltkoordinatensystem in das der Kamera beschreiben und mindestens vier der Intrinsik. Zwei Parameter für die Brennweite in x und y Richtung und die Position des Principal Points. Wie oben beschrieben, werden außerdem für die Bildkorrektur Parameter gesucht, welche die Verzerrungen des Bildes korrigieren. Dies können theoretisch beliebig viele sein. Desweiteren kann der Scherungsfaktor α festgelegt werden. Im Folgenden wird die Kalibrierung verkürzt und vereinfacht dargestellt. Details dazu können in [10] nachgelesen werden.

Die 3D-Koordinaten werden mithilfe der Extrinsik in das Koordinatensystem der Kamera transformiert.

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} \mathbf{R} + \mathbf{T}$$

Anschließend können aus diesen je nach Projektionsmodell aus 2.1 die normalisierten Bildkoordinaten berechnet werden. Diese werden dann mithilfe der Distortionparameter korrigiert und in die tatsächlichen Pixelkoordinaten umgewandelt.

Um die Parameter zu erhalten, sind Kontrollpunkte nötig, von denen sowohl die 3D-Koordinaten als auch die Bildkoordinaten bekannt sind. Somit sind für diese Punkte die als richtig angenommenen Pixelkoordinaten p_i und auch das Ergebnis der Projektion p'_i bekannt. Mithilfe des Levenberg-Marquardt-Algorithmus wird dann der in *OpenCV* folgendermaßen berechnete

$$reprojectionError = \sqrt{\left(\frac{\sum_{i=1}^n d(p_i, p'_i)^2}{n}\right)}$$

minimiert. Wobei

$$d(p, q) = (p_x - q_x)^2 + (p_y - q_y)^2$$

gilt. Dieser beschreibt also die Entfernung zwischen den gemessenen Pixelkoordinaten und denen, die durch die Projektion erhalten wurden. Ist dieser Null, ist das Modell also (zumindest für die eingesetzten Kontrollpunkte) perfekt. Um Overfitting zu vermeiden ist es daher sinnvoll, sofern genügend Kontrollpunkte vorhanden sind, diese nach der Kalibrierung mit Punkten zu überprüfen, welche nicht für die Berechnung der Parameter verwendet wurden. Hierbei ist zu beachten, dass der Reprojection Error oftmals, wie auch in diesem Fall, mit dem Root Mean Square Error (RMSE) gleichgesetzt wird. Eine weitere gebräuchliche Variante den Reprojection Error zu berechnen ist die Summe der euklidischen Distanzen über alle Punktpaare zu bilden und anschließend durch deren Anzahl zu teilen. Da für die euklidische Distanz jedoch für jedes Punktpaar die Berechnung einer Wurzel nötig ist, wird dies offenbar in *OpenCV* aus Effizienzgründen vermieden.

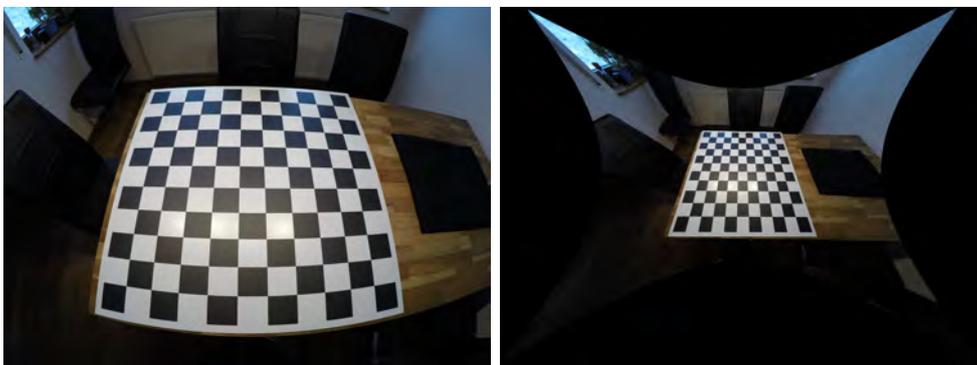
Erhalten von Kontrollpunkten

Kontrollpunkte für die Kalibrierung können auf unterschiedliche Arten erhalten werden. Dabei ist immer darauf zu achten, dass möglichst Korrespondenzen in allen Bereichen des Kamerasensors vorhanden sind und sich diese nicht auf der selben Ebene befinden. Um diese zu erhalten gibt es drei Möglichkeiten:

Mithilfe eines 3D-Scanners Ist ein 3D-Scanner vorhanden, können in jeder beliebigen Umgebung Kontrollpunkte erhalten werden, indem im Bild der Kamera markante Punkte und deren korrespondierender Messpunkt in der 3D-Punktwolke gesucht werden. Dies muss in der Regel manuell erfolgen und ist daher zeitaufwendig und fehlerbehaftet.

Mithilfe eines dreidimensionalen Kalibrierobjekts Ist ein dreidimensionales Kalibrierobjekt vorhanden, bei dem zum Beispiel durch QR-Codes oder April-Tags Punkte markiert sind, deren Position bekannt ist, können diese automatisch extrahiert werden.

Mithilfe von planaren Kalibrierobjekten Es ist auch möglich mit planaren Kalibrierobjekten eine Kamera zu kalibrieren. Hierfür muss lediglich ein planares Objekt, oft genügt schon ein Blatt Papier, mit einem bekannten Muster (Beispiel s. 2.9(a)) vorhanden sein. Dieses wird aus unterschiedlichen Winkeln fotografiert und die Kontrollpunkte werden automatisch extrahiert. Mehr Details dazu siehe [15]. Dies ist die wohl am häufigsten verwendete Variante, da sie ohne teure Hardware umsetzbar ist. Natürlich ist die Güte der Kalibrierung stark von der Planarität und der Qualität des Objekts abhängig. Ist das Objekt beispielsweise nicht perfekt planar, werden die Kontrollpunkte fehlerhaft extrahiert und somit erhöht sich der Kalibrierfehler.



(a) Originalbild

(b) Korrigiertes Bild

Abbildung 2.9: Auswirkung der Kalibrierung.

Wie in Abbildung 2.9 zu sehen, ist es mithilfe der Parameter, welche aus einer erfolgreichen Kalibrierung gewonnen wurden, möglich, die Effekte und Verzerrungen selbst für Fischaugenobjektive nachträglich zu entfernen. Dadurch werden beispielsweise die vorher stark radial verzerrten Kanten des Schachbrettmusters wieder geradlinig.

Kapitel 3

Evaluation der Kameramodelle

Zum Vergleich der verschiedenen Kameramodelle wurden in der weit verbreiteten Computer Vision Bibliothek *OpenCV* der Version 3.4.1 neben dem dort bereits verwendeten perspektivischen sämtliche Modelle aus Tabelle 2.1 implementiert. Hierfür wurde ein Enum *projectionMethod* mit Einträgen für die einzelnen Modelle eingeführt:

Listing 3.1: Enum für die Projektionsmodelle

```
1 //calib3d_c.h
2 enum projectionMethod {
3     PERSPECTIVE = 0,
4     STEREOGRAPHIC = 1,
5     EQUIDISTANT = 2,
6     ORTHOGONAL = 3,
7     EQUISOLIDANGLE = 4
8 };
```

Bei sämtlichen Funktionen, in denen das Projektionsmodell eine Rolle spielt, wurde ein zusätzlicher optionaler Parameter mit dem Default-Wert PERSPECTIVE für das oben genannte Enum (s. Listing 3.1) eingeführt. So lassen sich alle Funktionen weiterhin wie gewohnt nutzen, bei expliziter Angabe des Projektionsmodells wird jedoch dieses verwendet. Somit ist es zu jedem Code kompatibel, der für *OpenCV* 3.4.1 geschrieben wurde. Ein möglicher Aufruf für eine Kalibrierung lautet somit:

Listing 3.2: Aufruf der Kalibrierfunktion in OpenCV

```
1 cv::projectionMethod method = cv::projectionMethod::EQUIDISTANT;
2 cv::calibrateCamera(object_points, img_points, imgSize, K, D, rvecs,
   tvecs, flags, termCriteria, method);
```

3.1 Versuchsaufbau

Um die Projektionsmodelle zu vergleichen, wurden mit mehreren Kameras aus verschiedenen Winkeln und Entfernungen Fotos eines auf Aluminium gedruckten Schachbrettmusters angefertigt. Dieses Schachbrettmuster (s. Abb. 3.1) der Größe 90x90 cm ermöglicht es bei geringer Entfernung mit jedem einzelnen Bild einen großen Bereich des Kamerasensors abzudecken und ist annähernd perfekt planar.



Abbildung 3.1: Beispiele für Kalibrierfotos.

Beim Erstellen der Fotos wurde darauf geachtet, dass der gesamte Sensor abgedeckt wurde, so dass insbesondere die Ränder und Ecken mitkalibriert werden. Außerdem wurde sichergestellt, dass die Kanten der einzelnen Schachbrettfelder möglichst scharf sind und keine Reflektionen auftreten, die zu falsch extrahierten Koordinaten führen können. Anschließend wurden die Bilder in eine zufällige Reihenfolge gebracht und in 15 Gruppen eingeteilt. Zur Evaluierung wurden anschließend immer 14 der Splits zum Kalibrieren verwendet und das Ergebnis mit dem 15. überprüft. Somit ist es ausgeschlossen, dass das Modell nur auf den zur Kalibrierung verwendeten Bildern gut performt.

3.2 Verwendete Kamerasysteme

Um die Projektionsmodelle vergleichen zu können, wurden diese mit verschiedenen Kameras getestet (s. Abb. 3.2).



Abbildung 3.2: Eingesetzte Kamerasysteme.

Die Kamerasysteme unterscheiden sich stark voneinander und sollen so möglichst viel Auskunft über die verwendeten Projektionsmodelle liefern. Es wurden sowohl gewöhnliche, als auch Fischaugenobjektive mit ihnen getestet. Es wurden Kameras mit vergleichsweise großem APS-C Sensor und Kameras mit sehr kleinem Sensor (1/2 Zoll) wie im Smartphone eingesetzt. Auch große Objektive mit verstellbarer Mechanik und starre wie in den Action Cams wurden getestet. Für die Versuche wurden die Brennweiten, soweit es möglich ist, auf unendlich gestellt und der Autofokus deaktiviert. Die beiden Kameraobjektive (s. Abb. 3.2(a) und 3.2(b)) wurden in Verbindung mit einer Canon EOS 1000D Spiegelreflexkamera eingesetzt, welche Bilder mit einer Auflösung von 3888x2592 Pixeln liefert.

3.2.1 Fischaugenobjektiv Walimex pro 8mm F3,5 Fisheye II

Dieses Fischaugenobjektiv gehört zu den günstigsten seiner Art. Es verfügt über eine abnehmbare Streulichtblende und besitzt lediglich die Möglichkeit Blende und Brennweite manuell festzulegen. Der Öffnungswinkel beträgt laut Herstellerwebsite [12] 167 Grad.

3.2.2 Canon 18-55mm f/3.5-5.6 IS STM

Auch dieses Objektiv gehört zur Einstiegsklasse, besitzt aber bereits einen optischen Bildstabilisator und Autofokus. Beide Funktionen wurden aber für die Versuche deaktiviert. Der optische Zoom wurde zu Beginn der Versuche eingestellt und anschließend nicht mehr verändert. Dieser ist aber sehr leichtgängig, wodurch es nicht ausgeschlossen werden kann, dass er sich durch Vibrationen beim Fotografieren geringfügig verstellt. Im Gegensatz zu allen anderen verwendeten Systemen, mit Ausnahme der Smartphonekamera, handelt es sich hierbei nicht um eine Kamera mit Fischaugencharakteristik.

3.2.3 Xiaomi Mi 9 SE

Das verwendete Smartphone Xiaomi Mi 9 SE besitzt eine typische Handykamera aus dem Mittelklassebereich. In der für die Versuche eingesetzten Hauptkamera sitzt ein Sony 1/2" Sensor mit einer Auflösung von 48 Megapixeln [5]. Die Besonderheit liegt darin, dass immer 4 Pixel anschließend zu einem einzelnen zusammengefasst werden, wodurch das Bild schließlich mit nur 12 Megapixeln gespeichert wird. Dies geschieht durch KI und soll zu besseren und farbenfroheren Bildern führen. Die Kamera besitzt einen recht großen Öffnungswinkel von 123 Grad und gehört damit aber noch nicht zu den Fischaugenobjektiven. Die Bilder haben keine sichtbaren radialen Verzerrungen, wie sie bei Fischaugenobjektiven typisch sind.

3.2.4 Xiaomi Mi 9 SE mit Fischaugenobjektiv

Auf die oben beschriebene Smartphonekamera wurde ein Fischaugenobjektiv aufgesetzt. Dieses preislich sehr günstige und qualitativ minderwertige Objektiv führt zu einem deutlich größeren Öffnungswinkel, vermindert die Fotoqualität aber drastisch. Wie in Abb. 3.3 zu sehen ist, gibt es sehr ausgeprägte radiale Verzeichnungen. Lediglich in der Mitte ist das Bild scharf. Je näher am Rand, desto unklarer werden die Umrisse und desto größer wird die chromatische Aberration. Außerdem führt der Aufsatz dazu, dass das Bild im Gegensatz zu den anderen Systemen, kein Full-Frame oder Vollformatbild mehr ist.



Abbildung 3.3: Beispielbild mit aufgesetztem Fischaugenobjektiv.

GoPro Hero 4

Die GoPro Hero 4 besitzt laut [4] einen 1/2.3" Sensor und eine Auflösung von 12 Megapixeln. In der verwendeten Einstellung *Ultrawide* werden Fotos mit einem Öffnungswinkel von 170 Grad geschossen. Die Fotos sind damit so weitwinklig, dass sehr deutliche radiale Verzeichnungen entstehen.

Andoer BCF6412630187900JS

Die Andoer Action Kamera ist ein extrem günstiges Modell. Sie löst mit 12 Megapixeln auf und der Öffnungswinkel beträgt ebenfalls 170 Grad. Bei dieser Kamera können keine Einstellungen zum Öffnungswinkel vorgenommen werden und die Bildqualität ist deutlich schlechter als die der GoPro Hero 4. Aufgrund des Öffnungswinkels treten hier genauso wie bei der GoPro deutliche radiale Verzeichnungen auf.

3.3 Evaluation der Kalibrierungen

Um die Güte einer Kalibrierung bewerten zu können, müssen zuerst Kriterien festgelegt werden, die eine Aussage über deren Qualität treffen. Am weitesten verbreitet ist hierfür der Root-Mean-Square-Error (RMSE), der in diesem Zusammenhang oft auch als Reprojection Error bezeichnet wird. Dieser beschreibt für eine Menge von 3D-Messpunkten mit zugehörigen 2D-Abbildungspunkten, deren Position als richtig angenommen wird, die mittlere Abweichung, die entsteht, wenn die 3D-Messpunkte mithilfe des kalibrierten Modells in ein zweidimensionales System projiziert werden. *OpenCV* gibt diesen Error bei der Kalibrierung an, wobei zu beachten ist, dass dieser auf den Kalibrierbildern berechnet und somit bei der Kalibrierung minimiert wurde. Es handelt sich hierbei also um den Fehler auf dem Trainingsdatensatz, wodurch nicht sichergestellt ist, dass das Modell gut generalisiert. Zur Vermeidung des beschriebenen Problems wurde der RMSE in den Versuchen immer auf einem Teil des Datensatzes berechnet, der nicht zur Kalibrierung verwendet wurde. Damit der RMSE aussagekräftig ist, muss dieser auf Bildern berechnet werden, die möglichst den ganzen Sensor abdecken. Wird beispielsweise nur mit Bildern getestet, bei denen das Kalibrierobjekt in der Mitte des Sensors liegt, wo die Verzerrungen sehr gering sind, ist mit einem sehr guten Wert zu rechnen, auch wenn das Modell am Bildäußeren womöglich sehr schlecht funktioniert. Im Folgenden wurde der RMSE bei allen Durchläufen auf zufälligen 10% des Datensatzes berechnet. Hierdurch sollte eine ausreichende Aussagekraft des RMSEs gewährleistet sein.

Eine weitere Möglichkeit um sicherzustellen, dass eine Kalibrierung mit geringem RMSE gut ist, ist eine Sichtprüfung. Hierfür wird ein Foto mit den Kalibrierparametern entzerrt und überprüft, ob das Ergebnis zufriedenstellend ist. Dies ist der Fall, wenn Kanten und Linien beispielsweise von Wänden, Türen oder vom Kalibrierpattern nach der Korrektur wirklich geradlinig sind. Außerdem sollten natürlich keine offensichtlichen Verzerrungen oder Überkorrekturen auftreten.

3.3.1 Vergleich des Kalibrierfehlers mit Distortionparametern

Im Folgenden werden zuerst die RMSE-Ergebnisse der einzelnen Kameramodelle kurz betrachtet. Anschließend werden für die verschiedenen Kameras die Projektionsmodelle verglichen, indem die korrigierten Bilder genauer untersucht werden.

Perspektivisch

Die perspektivische Projektion (s. Tabelle 3.1), welche mit Abstand die am einfachsten zu berechnende ist, führt zu sehr guten Ergebnissen. Der RMSE liegt in vier von sechs Fällen unter eins und sowohl die Varianz als auch die Standardabweichung sind sehr gering. Das bedeutet, dass das perspektivische Modell auf allen Datensätzen ähnlich gut funktioniert. Eine Ausnahme hierbei ist das Walimex Fischaugenobjektiv, bei dem der RMSE eine Spanne von 0.38 bis 6.01 aufweist.

Tabelle 3.1: Reprojection Errors der perspektivischen Projektion.

Perspektivisch						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	2.77	0.99	4.06	0.67	2.87	1.61
2	1.40	0.84	3.21	0.67	2.71	0.73
3	6.01	0.65	3.35	0.60	2.85	1.23
4	2.87	0.50	4.45	0.87	2.52	1.68
5	0.51	0.74	3.57	0.58	4.84	1.55
6	2.20	0.57	4.73	0.80	2.74	0.55
7	3.46	0.87	4.31	0.50	2.48	1.35
8	0.38	1.27	2.81	0.92	3.01	0.72
9	5.52	1.68	4.15	0.66	2.70	0.79
10	0.56	0.54	4.38	0.71	2.89	0.93
11	0.48	1.75	1.78	0.46	2.61	0.85
12	2.20	0.52	2.86	2.11	2.71	0.72
13	4.92	0.72	4.74	0.77	2.66	0.55
14	5.01	0.43	5.53	0.49	2.97	0.39
15	3.86	0.48	3.81	0.71	2.80	0.84
Varianz	3.77	0.18	0.89	0.16	0.31	0.17
Standard- abweichung	1.94	0.42	0.94	0.39	0.56	0.41

Stereographisch

Das stereographische Modell (s. Tabelle 3.2) führt zu deutlich schlechteren Werten als das perspektivische. Bei allen Kameras liegt der Error deutlich über zwei. In der Regel wird von einer guten Kalibrierung gesprochen, wenn dieser unter 1.0 liegt. Auch Varianz und Standardabweichung sind sehr groß. Das Modell performt also auf den verschiedenen Datensätzen sehr unterschiedlich.

Tabelle 3.2: Reprojection Errors der stereographischen Projektion.

Stereographisch						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	56.71	13.95	13.94	20.53	6.06	31.08
2	4.86	37.66	20.01	14.47	8.50	19.81
3	11.99	8.26	13.45	16.56	8.26	26.46
4	11.71	8.97	26.51	16.52	6.20	17.66
5	1,120.70	8.62	15.30	17.72	11.70	19.60
6	8.34	22.61	12.30	21.71	2.76	19.83
7	6.58	15.85	24.77	15.75	6.83	18.23
8	12.20	5.79	3.62	16.65	8.32	28.50
9	8.40	12.53	11.93	14.57	9.15	29.13
10	4.28	16.89	15.86	15.58	12.45	21.76
11	12.66	12.59	5.53	8.85	9.52	17.09
12	4.95	4.79	6.99	16.78	2.92	21.76
13	4.12	12.88	25.69	17.25	9.82	19.17
14	8.97	7.47	29.68	15.19	6.69	9.07
15	12.79	8.15	12.48	18.03	7.24	19.26
Varianz	82,104.32	68.29	63.02	8.38	7.43	31.55
Standard- abweichung	286.54	8.26	7.94	2.89	2.73	5.62

Equidistant

Wie in Tabelle 3.3 ersichtlich ist, funktioniert das equidistante Modell zumindest bei zwei Kamerasystemen (Canon Objektiv und Smartphone) ähnlich gut wie das perspektivische. Hieraus lässt sich folgern, dass es Modelle gibt, die sich für manche Kameras gut eignen, jedoch nicht für alle. Die beiden Kamerasysteme, für die das equidistante Modell sehr gut performt, sind diejenigen ohne Fischaugencharakteristik. Es wäre also möglich, dass sich das Modell gut für Standardkameras eignet. Ohne Tests mit weiteren Systemen kann dies aber nicht sicher festgestellt werden.

Tabelle 3.3: Reprojection Errors der equidistanten Projektion.

Equidistant						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	5.26	10.47	7.95	0.73	5.31	1.63
2	5.13	1.58	4.84	0.67	3.37	0.75
3	6.33	7.45	7.49	0.61	3.83	1.24
4	4.39	7.99	5.10	0.89	10.96	1.69
5	2.38	4.47	5.77	0.62	6.12	1.58
6	5.13	8.24	9.32	0.82	5.19	0.57
7	5.22	6.31	5.63	0.59	3.77	1.38
8	6.00	8.44	4.00	0.93	3.17	0.77
9	5.91	6.09	9.76	0.73	3.85	0.84
10	5.38	5.16	6.08	0.78	10.85	0.95
11	3.75	6.62	10.07	0.51	2.95	0.86
12	6.30	9.49	12.11	2.20	3.13	0.77
13	6.62	5.95	5.10	0.84	2.96	0.60
14	5.05	7.31	6.21	0.50	4.52	0.44
15	6.61	6.97	6.82	0.72	5.47	0.87
Varianz	1.31	4.60	5.38	0.17	6.69	0.16
Standard- abweichung	0.14	2.14	2.32	0.41	2.59	0.41

Orthogonal

Bei dem orthogonalen Modell (s. Tabelle 3.4) fehlen leider einige Einträge. Das ist dadurch begründet, dass die Kalibrierung in manchen Fällen fehlschlug oder der Kalibriervorgang nach 12h abgebrochen wurde. Dies kann beim orthogonalen Modell passieren, wenn die Z -Koordinate ≤ 0 ist, was bei den Versuchen definitiv nicht der Fall war, da die Z -Achse im *OpenCV*-Koordinatensystem die Entfernung zum Kalibrierpattern ist. Ein Z -Wert kleiner gleich Null würde demnach bedeuten, dass das Pattern auf- oder hinter dem Projektionszentrum der Kamera liegt. Durch Fehler in der Kalibrierung und den daraus entstehenden Folgefehlern bei der Berechnung der Extrinsik kann es aber passieren, dass dies doch das Problem ist und somit zu fehlerhaften Werten führt. Der Wert *.NaN* (Not a Number) entsteht durch unzulässige Rechenoperationen wie das Teilen durch 0 oder beim Ziehen der Wurzel aus einer negativen Zahl. Durch unerwartete Fehler im Kalibrierungsprozess muss es zu solch einem Fall gekommen sein. Dass die Kalibrierung nach sehr langer Zeit nicht endet, kann wiederum aus zwei Gründen geschehen. Entweder es konvergiert bei der Optimierung extrem langsam oder auch gar nicht.

Tabelle 3.4: Reprojection Errors der orthogonalen Projektion.

Orthogonal						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	.Nan	9.55	7.54	0.70	6.39	1.63
2	5.09	.Nan	4.33	0.67	3.89	0.75
3	6.44	7.20	7.48	0.63	5.70	1.24
4	4.31	7.29	4.96	0.89	11.00	1.74
5	.Nan	3.84	5.52	0.59	6.91	1.61
6	5.28	.Nan	9.45	0.81	5.22	0.59
7	6.84	12.29	14.23	0.53	3.83	1.41
8			11.04	0.93	3.27	0.82
9			7.54	0.68	4.30	0.84
10			5.63	0.77	10.93	0.95
11			9.95	0.52	3.49	1.13
12			11.99	2.21	3.58	0.78
13			5.00	0.78	3.40	0.58
14			5.96	0.51	4.57	0.46
15			6.31	0.73	5.48	0.88
Varianz			8.60	0.17	6.24	0.17
Standard- abweichung			2.93	0.41	2.50	0.41

Equisolidangle

Auch bei dem equisolidangle Modell (s. Tabelle 3.5) trat mehrfach der Fall auf, dass kein Ergebnis erreicht wurde. Hier ist auch die Varianz und entsprechend die Standardabweichung recht groß. Tatsächlich sind die Ergebnisse für jede Kamera bei diesem Modell am schlechtesten. Es ist somit offenbar zumindest für keine der getesteten Kameras gut geeignet.

Tabelle 3.5: Reprojection Errors der equisolidangle Projektion.

Equisolidangle						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	.Nan	29.03	42.56	25.11	7.68	47.79
2	5.28	.Nan	55.59	22.86	16.67	44.80
3	13.44	31.54	18.71	25.29	16.19	44.57
4	13.79	17.14	36.79	20.35	6.14	32.45
5	.Nan	24.61	26.80	26.18	11.70	44.54
6	54.37	.Nan	18.68	22.51	4.02	41.04
7	5.36	66.23	88.62	21.31	9.24	44.01
8	22.27	36.93	4.49	20.39	10.59	72.50
9	8.46	36.42	36.70	18.67	14.85	39.31
10	38.84	35.59	28.81	25.63	13.01	53.46
11	25.68	73.59	6.89	19.33	9.45	22.39
12	10.67	32.67	8.65	20.49	8.00	47.68
13	20.12	68.28	41.41	26.26	13.68	39.57
14	12.72	36.21	71.54	10.70	8.13	21.60
15	21.49	26.11	32.52	27.90	11.33	29.10
Varianz			556.83	18.46	13.63	158.80
Standard- abweichung			23.60	4.30	3.69	12.60

3.3.2 Optischer Vergleich der Kalibrierungen mit Distortionparametern

Im Folgenden werden die Ergebnisse der Kalibrierung für einige der vorgestellten Kamerasysteme anhand der entzerrten Originalbilder untersucht. Hier ist leider nur ein subjektiver Vergleich möglich, da außer dem Reprojection Error keine Daten verglichen werden können. Die Parameter der Kalibrierung werden jeweils genannt, ohne an dieser Stelle darauf genauer einzugehen.

Kalibrierung des Walimex pro 8mm F3,5 Fisheye II Objektivs

Bei der Betrachtung der entzerrten Bilder (s. Abb. 3.8) zeigt sich bereits, dass es sehr schwierig ist, allein durch den optischen Eindruck die beste Kalibrierung zu erkennen. So sind sich beispielsweise das equidistante und das perspektivische Modell (Abb. 3.8 (b) und (d)), trotz des großen Unterschieds, bei dem RMSE sehr ähnlich. Insbesondere, wenn wie üblich, nur der größtmögliche rechteckige Bildbereich betrachtet wird, der (unter Berücksichtigung des originalen Bildformates) ohne schwarzen Rand möglich ist. Dennoch haben alle Projektionsmodelle ihre eigenen Charakteristika. Alle Modelle bis auf das orthogonale sind in der Lage die tonnenförmigen Verzeichnungen sehr gut zu entfernen. Das lässt sich beispielsweise an der Leinwand und den Türrahmen gut erkennen. Diese sind im Originalbild deutlich gekrümmt, in den korrigierten Bildern aber kaum bis gar nicht. Daraus ist ersichtlich, dass die Modelle mit den Distortion Parametern funktionieren. Während durch das perspektivische und das equidistante Modell die Randbereiche deutlich gestreckt werden, sichtbar am Fenster und dem Kachelofen, tun das stereographische und vor allem das orthogonale dies nicht. Das equisolidangle, also flächentreue Modell, hingegen erhält die Größenverhältnisse zwischen Fernseher, Wohnwand und Tür am besten. Dafür verliert es aber sehr viele Informationen, indem es am Bildrand große Bereiche, wie in diesem Fall den Kachelofen und die Balkontür, entfernt (Dies geschieht intern in *OpenCV*. Das Bild wurde nicht zugeschnitten).

Tabelle 3.6: Walimex Pro Fisheye Kalibrierwerte mit Distortionparameter.

	Perspective	Stereographic	Equidistant	Orthogonal	Equisolidangle
fx	1,473.55	3,393.32	1,495.04	2,726.66	5,061.59
fy	1,475.82	3,435.28	1,466.66	1,933.50	4,438.73
cx	1,966.60	1,945.35	1,933.86	2,584.21	1,958.48
cy	1,309.44	1,300.55	1,295.09	1,309.51	1,292.37
k1	2.33E-01	-5.34E-03	-8.38E-04	2.48E-03	-3.26E-02
k2	1.11E-03	-7.67E-02	1.16E-03	7.25E-02	-2.50E-01
p1	8.69E-06	-6.51E-03	-3.44E-04	-5.29E-02	-1.51E-03
p2	7.16E-06	-1.92E-04	3.10E-04	-1.47E-02	-2.27E-03
k3	2.72E-04	-9.47E-02	1.56E-03	2.01E-01	2.43E+00
k4	5.12E-01	5.67E-03	8.31E-04	-3.03E-03	3.11E-02
k5	1.72E-02	8.10E-02	-1.15E-03	-6.56E-02	2.32E-01
k6	8.25E-04	1.14E-01	-1.55E-03	-1.78E-01	-2.59E+00



(a) Originalbild



(b) Perspective (RMSE 0.38)



(c) Stereographisch (RMSE 4.12)*



(d) Equidistant (RMSE 2.38)



(e) Orthogonal (RMSE 4.31)



(f) Equisolidangle (RMSE 5.28)

Abbildung 3.4: Ergebnisse der Kalibrierungen für das Walimex Pro Fisheye Objektiv (Mit * gekennzeichnete Bilder wurden zugeschnitten).

Kalibrierung des Canon 18-55mm f/3.5-5.6 IS STM Objektivs

Bei diesem Objektiv ohne Fischaugeneffekt ist erkennbar, dass die Modelle (s. Abb 3.6) mit Ausnahme des equisolidangle und des stereographischen Modells, jeweils sehr niedrige RMSE haben und das Originalbild nur geringfügig ändern, indem sowohl oben als auch unten das Bild leicht in Richtung Bildmittelpunkt gekrümmt wird. Dies deutet darauf hin, dass dieses Objektiv eine, wenn auch mit bloßem Auge kaum erkennbare, tonnenförmige Verzeichnung besitzt. Das stereographische Modell überkompensiert diese Effekte, was dazu führt, dass vorher richtig abgebildete Geraden nun leicht gekrümmt sind. Obwohl der RMSE des equidistanten etwas größer ist als der des perspektivischen Modells, lässt sich rein optisch auch hier nicht erkennen, welches Modell besser ist. Tatsächlich unterscheiden sich die Ergebnisse nur minimal. Das perspektivische führt zu einem kreisförmigen Abbildungsfehler, welcher aber nur bei entsprechender Vergrößerung erkennbar ist und sich durch das gesamte Bild zieht. Dieser ist in Abb. 3.5 zu sehen. Daher kann hier das equidistante Modell dem perspektivischen vorgezogen werden, obwohl der RMSE deutlich höher ist.

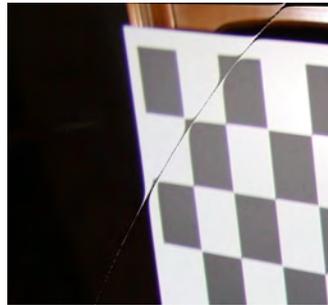


Abbildung 3.5: Abbildungsfehler des perspektivischen Modells.

Tabelle 3.7: Canon 18-55mm f/3.5-5.6 IS STM Kalibrierwerte mit Distortionparameter.

	Perspective	Stereographic	Equidistant	Orthogonal	Equisolidangle
fx	3,226.29	4,541.52	3,220.92	3,215.81	4,032.57
fy	3,223.37	4,794.11	3,219.91	3,214.52	4,007.26
cx	1,994.88	1,949.30	1,955.37	2,013.25	1,928.61
cy	1,311.62	1,302.83	1,297.73	1,342.66	1,324.35
k1	-1.65E+00	4.52E-02	7.14E-02	1.32E+01	7.93E-01
k2	-6.59E-01	-9.88E-01	8.35E-02	-6.88E+01	-3.93E+00
p1	-1.95E-05	9.89E-04	-9.07E-04	3.16E-03	1.81E-02
p2	-3.18E-03	-4.44E-02	-5.63E-03	-2.61E-03	7.45E-03
k3	-3.20E+00	4.68E+00	1.31E-01	4.82E+01	5.99E+00
k4	-1.46E+00	-4.35E-02	-6.87E-02	1.28E+01	-6.28E-01
k5	-1.15E+00	1.05E+00	-1.03E-01	-7.26E+01	2.62E+00
k6	-3.14E+00	-4.51E+00	-1.08E-01	6.49E+01	-1.50E+01

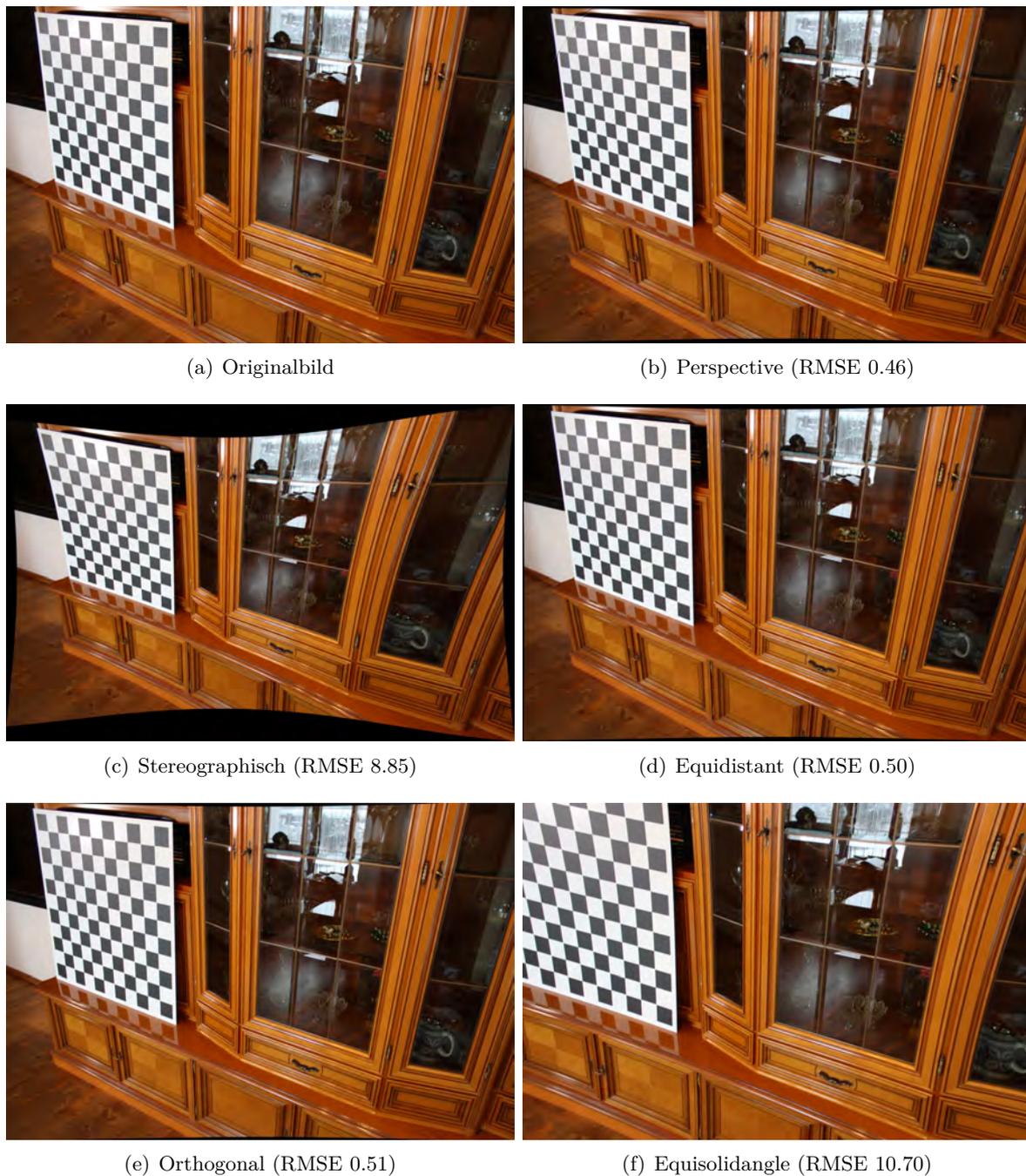


Abbildung 3.6: Ergebnisse der Kalibrierungen für das Canon 18-5mm f/3.5-6.6 Objektiv.

Kalibrierung des Xiaomi Mi 9 SE Smartphones

Vergleichbar mit dem Standard Canon Objektiv sind die Resultate mit dem Smartphone (s. Abb. 3.7). Die Ergebnisse des perspektivischen, equidistanten und orthogonalen Modells sind sich sehr ähnlich. Sowohl der RMSE als auch die entzerrten Bilder unterscheiden sich nur geringfügig. Einen Abbildungsfehler wie in Abb. 3.5 gibt es nicht. Auch hier überkompensiert das stereographische Modell die radialen Verzeichnungen.

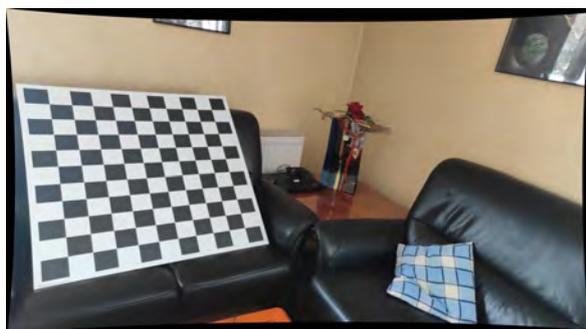
Tabelle 3.8: Xiaomi Mi 9 SE Kalibrierwerte mit Distortionparameter.

	Perspective	Stereographic	Equidistant	Orthogonal	Equisolidangle
fx	2,986.07	3,792.86	2,982.84	2,977.63	2,670.08
fy	2,985.63	3,857.70	2,983.38	2,978.55	2,726.25
cx	1,994.26	2,015.53	2,005.57	2,014.28	1,937.89
cy	1,123.72	1,140.47	1,124.07	1,123.75	1,103.00
k1	3.74E+00	4.90E-01	2.09E-01	2.98E-01	3.25E-01
k2	-4.60E+01	1.16E-01	4.69E-02	1.78E-01	1.02E+00
p1	-2.35E-05	2.92E-03	-3.11E-04	-6.22E-04	1.54E-02
p2	-2.99E-05	1.14E-02	1.75E-03	3.10E-03	-2.24E-03
k3	1.03E+02	-1.18E+00	-1.70E-01	-4.92E-02	-2.68E-01
k4	3.60E+00	-4.16E-01	-1.96E-01	-2.72E-01	-2.37E-01
k5	-4.53E+01	-5.78E-02	-5.35E-02	-1.64E-01	-6.55E-01
k6	1.02E+02	8.33E-01	2.69E-01	4.36E-01	5.94E-01

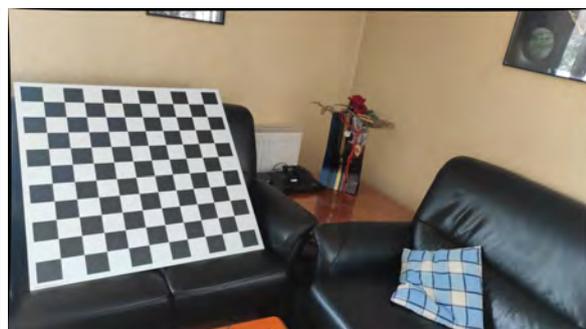


(a) Originalbild

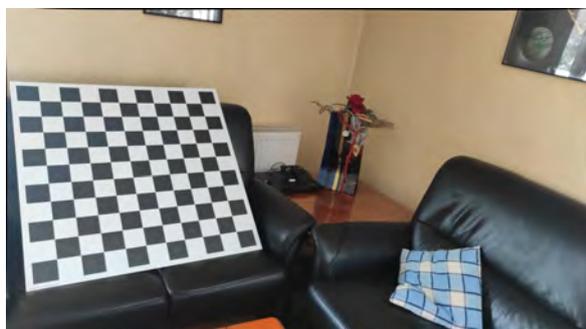
(b) Perspective (RMSE 0.39)



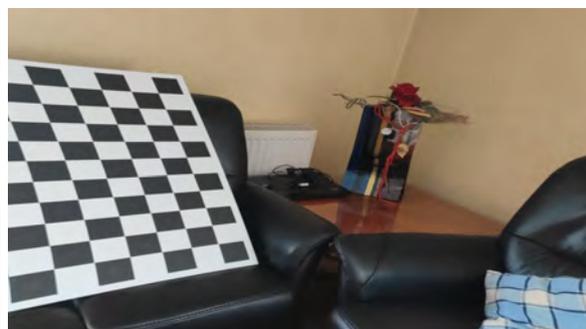
(c) Stereographisch (RMSE 9.07)



(d) Equidistant (RMSE 0.44)



(e) Orthogonal (RMSE 0.46)



(f) Equisolidangle (RMSE 21.60)

Abbildung 3.7: Ergebnisse der Kalibrierungen für das Xiaomi Mi 9 SE Smartphone.

Kalibrierung des Xiaomi Mi 9 SE mit Fischaugenobjektiv

Das auf das Smartphone aufgeklebte Fischaugenobjektiv unterscheidet sich in einigen Punkten von den zuvor betrachteten Systemen. Trotz der starken chromatischen Aberration und der Unschärfe, die von der Bildmitte zum Bildrand sehr stark zunimmt, erreichen alle Projektionsmodelle einen Error zwischen 2.48 und 4.02. Der immer relativ hohe Kalibrierfehler lässt sich damit erklären, dass das Kalibrierpattern am Bildrand nicht verwendet werden konnte, da durch die Unschärfe die Kanten zwischen den Schachbrettfeldern nicht mehr erkannt werden können. Die Extraktion der Kanten führt auch bei Bildern wie 3.8(a) bereits zu Fehlern, da beispielsweise die Kante im oberen linken Bildrand, welche auch den Ursprung des Koordinatensystems festlegt, fehlerhaft extrahiert wird. Dennoch sind alle Modelle in der Lage, die tatsächlich sehr großen radialen Verzeichnungen zumindest im Bereich der Bildmitte erstaunlich gut zu kompensieren. So werden die Kanten des Kalibrierobjektes, welche im Originalbild stark gerundet sind, bei allen Modellen beinahe perfekt gerade. Wird nur der mittlere Bildbereich betrachtet, welcher einigermaßen scharf ist, ist es auch hier schwierig zu entscheiden, welches Modell am besten geeignet ist.

Tabelle 3.9: Kalibrierwerte mit Distortionparameter für das Xiaomi Mi 9 SE Smartphone mit aufgesetztem Fischaugenobjektiv.

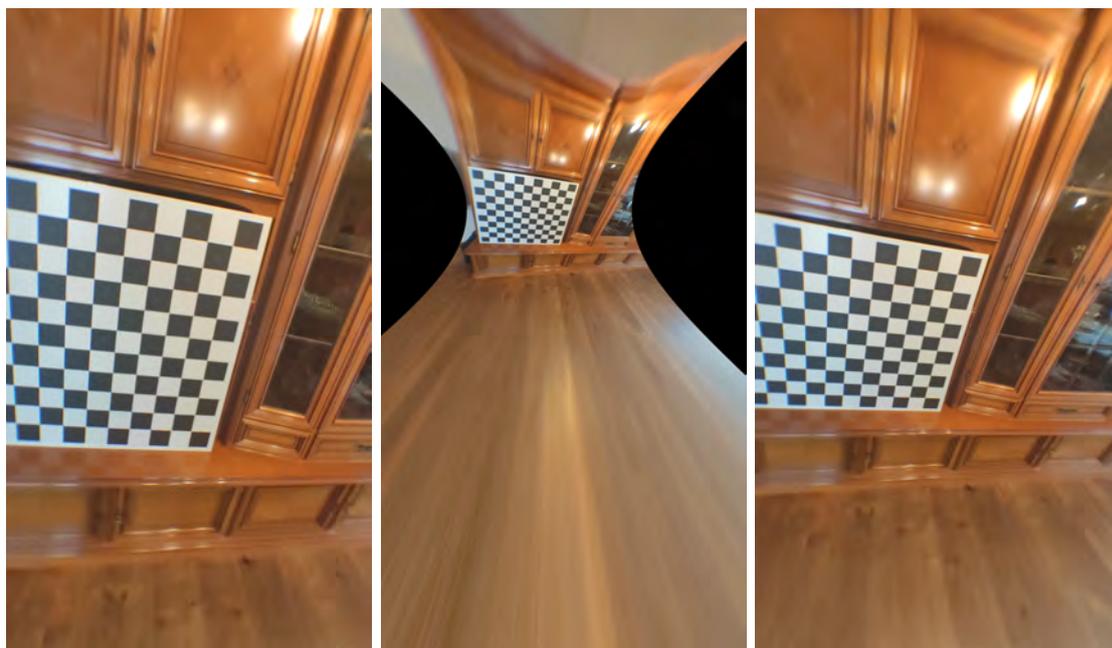
	Perspective	Stereographic	Equidistant	Orthogonal	Equisolidangle
fx	1,619.72	3,257.38	1,709.71	1,743.20	2,905.90
fy	1,618.12	3,257.78	1,704.04	1,730.68	2,865.30
cx	1,138.42	1,136.21	1,153.06	1,122.22	1,123.35
cy	1,973.65	1,996.31	2,040.61	2,007.97	1,997.70
k1	-5.69E+00	-2.04E-01	8.21E-02	4.75E-02	3.77E-02
k2	8.41E+00	-8.99E-01	1.94E+00	3.62E-02	-9.06E-02
p1	7.57E-04	-1.77E-03	-8.65E-03	-1.05E-02	7.22E-03
p2	3.30E-03	1.01E-02	4.27E-03	1.18E-02	1.07E-02
k3	8.21E-01	4.22E+00	1.06E+01	1.78E-02	-5.23E-01
k4	-5.33E+00	2.05E-01	2.50E-01	-4.18E-02	-3.73E-02
k5	6.30E+00	9.59E-01	1.25E+00	1.40E-02	9.60E-02
k6	4.10E+00	3.36E-01	1.26E+01	1.98E-01	5.80E-01



(a) Originalbild

(b) Perspective (RMSE 2.48)

(c) Stereographisch (RMSE 2.76)*



(d) Equidistant (RMSE 2.95)*

(e) Orthogonal (RMSE 3.27)

(f) Equisolidangle (RMSE 4.02)

Abbildung 3.8: Ergebnisse der Kalibrierungen für das Xiaomi Mi 9 SE Smartphone mit aufgesetztem Fischaugenobjektiv (Mit * gekennzeichnete Bilder wurden zugeschnitten).

3.3.3 Ergebnisse der Kalibrierung ohne Distortionparameter

Um zu untersuchen, ob das Distortionmodell mit manchen Modellen nicht gut funktioniert oder gar zu Problemen führt, wurde derselbe Versuch wie im obigen Kapitel nochmals durchgeführt. Hierbei wurden jedoch die Distortionparameter deaktiviert, so dass das Distortionmodell keine Anwendung findet. Hierdurch lässt sich herausfinden, welches Kameramodell für welche Kamera am besten geeignet ist.

Perspektivisch

Auf den ersten Blick ist ersichtlich, dass die Werte in Tabelle 3.10 offensichtlich schlechter sind als in Tabelle 3.1. Dies ist aber kaum verwunderlich, da die Korrekturterme genau für dieses Modell entwickelt wurden und nun fehlen. Auffällig ist, dass die Werte für die Kamerasysteme ohne Fisheye-Charakteristik (Canon Objektiv und Smartphone) besser sind als bei den anderen Modellen. Dies zeigt, dass das perspektivische Modell für solche Kameras am besten geeignet ist.

Tabelle 3.10: Reprojection Errors der perspektivischen Projektion ohne Distortionkoeffizienten.

Perspektivisch						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	6.83	13.70	15.57	3.28	18.19	2.19
2	6.28	10.72	15.96	2.48	19.59	1.42
3	6.94	18.81	18.40	3.08	18.05	1.71
4	6.17	11.42	10.38	3.40	22.04	2.11
5	6.59	10.38	13.00	2.62	16.05	1.94
6	6.90	14.41	16.58	2.33	17.45	1.05
7	6.09	11.82	14.20	2.76	15.18	1.80
8	7.45	12.48	15.88	2.67	15.26	1.37
9	6.74	9.47	16.20	3.31	18.39	1.49
10	7.20	9.60	12.35	2.99	18.30	1.37
11	7.37	11.63	14.49	2.38	17.89	1.24
12	6.09	14.59	15.07	3.15	21.40	1.35
13	7.73	14.50	16.67	3.00	18.33	1.50
14	5.77	12.53	12.03	3.18	15.44	1.17
15	6.80	12.93	15.21	3.17	15.83	1.49
Varianz	0.32	5.78	4.42	0.13	4.37	0.11
Standard- abweichung	0.57	2.40	2.10	0.35	2.09	0.34

Stereographisch

In den durchgeführten Versuchen war das stereographische Modell für keine Kamera die beste Wahl. Dies ist unerwartet, da auf der Website *fishyelens.de* [3] geschrieben steht, dass das Walimex Pro Fisheye Objektiv eine beinahe perfekt stereographische Abbildung erzeugt und somit dieses Modell am besten performen sollte. Doch es hat sich gezeigt, dass das equidistante Modell qualitativ hochwertigere Ergebnisse liefert. Dies ist bei allen 15 Durchläufen der Fall, so dass nicht von einem Zufall auszugehen ist. Auffällig ist hierbei der enorme Fehler in Split-Nr. 5 bei dem Walimex Pro Objektiv, ausgerechnet in jenem Split, bei dem das equidistante Modell das beste Ergebnis für dieses System liefert.

Tabelle 3.11: Reprojection Errors der stereographischen Projektion ohne Distortionkoeffizienten.

Stereographisch						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	83.37	12.94	22.54	26.06	15.36	31.21
2	9.02	38.62	26.95	15.71	18.18	28.41
3	12.82	8.78	18.67	20.01	10.62	29.93
4	11.74	15.87	27.18	17.36	14.22	23.80
5	1,291.57	11.43	17.28	18.03	17.94	20.19
6	9.26	25.94	21.00	21.73	12.46	20.28
7	7.09	18.83	24.65	19.84	15.96	25.02
8	13.21	7.61	10.67	19.00	16.11	29.49
9	10.60	14.22	15.93	16.38	14.13	26.36
10	6.93	19.40	17.28	18.71	17.15	28.88
11	14.83	12.63	12.42	8.58	16.76	17.79
12	4.44	6.93	11.92	21.15	15.72	26.23
13	5.55	12.89	25.74	19.97	20.89	22.27
14	10.04	9.21	32.28	18.38	12.53	22.04
15	21.17	8.90	16.07	20.93	15.02	27.09
Varianz	108,889.00	69.22	40.79	14.14	6.62	16.53
Standardabweichung	329.98	8.32	6.39	3.76	2.57	4.07

Equidistant

Das equidistante Modell überzeugt gleich bei zwei Kameras am meisten, sowohl für die GoPro Actioncam als auch für das Walimex Pro Fisheye Objektiv. Das Modell liefert für alle Kamerasysteme, wie bereits bei dem Versuch zuvor mit Distortion Parametern (s. Tabelle 3.3), vergleichsweise solide Werte und ist somit eine gute Wahl für Kameras, bei denen die Abbildungsart nicht bekannt ist.

Tabelle 3.12: Reprojection Errors der equidistanten Projektion ohne Distortionkoeffizienten.

Equidistant						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	5.66	10.92	8.17	9.25	13.43	12.59
2	6.28	10.44	5.86	5.76	14.52	12.36
3	7.08	8.12	9.16	7.03	9.36	11.65
4	6.07	8.15	5.02	8.91	12.41	10.95
5	3.64	5.02	6.39	5.90	16.35	10.97
6	5.33	8.94	9.16	6.14	12.80	9.28
7	6.35	6.04	5.83	6.95	10.52	12.20
8	6.86	11.47	7.47	6.15	6.28	13.78
9	8.11	6.68	8.57	6.46	11.13	12.61
10	5.73	6.14	6.47	6.90	13.86	13.34
11	5.43	6.85	10.58	6.85	10.82	10.37
12	8.52	14.44	12.79	6.35	12.73	12.86
13	7.11	6.83	5.85	7.41	11.06	12.89
14	6.77	9.98	6.63	7.26	9.77	9.05
15	7.76	7.26	8.04	7.71	10.85	11.29
Varianz	1.53	6.46	4.40	1.03	5.87	2.00
Standardabweichung	1.24	2.54	2.10	1.01	2.42	1.41

Orthogonal

Das orthogonale Modell zeigt leider wie bei dem vorherigen Experiment das Problem, dass es bei manchen Datensätzen nicht innerhalb einer annehmbaren Zeit konvergiert. Es liefert auch hier wieder Werte, die im Vergleich mit den anderen Modellen, weder sehr gut noch sehr schlecht sind. Es ist aber für keines der getesteten Systeme die erste Wahl.

Tabelle 3.13: Reprojection Errors der orthogonalen Projektion ohne Distortionkoeffizienten.

Orthogonal						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	.Nan	15.14	13.43	17.44	10.29	23.83
2	6.51	.Nan	12.85	10.63	9.10	20.03
3	7.23	17.71	13.53	13.14	5.87	18.85
4	6.69	13.63	10.68	16.09	13.45	17.47
5	.Nan	12.80	10.81	11.18	16.56	18.44
6	9.88	.Nan	15.03	12.39	9.31	15.04
7	6.05	20.86	17.79	12.92	8.26	21.29
8	8.18	18.32	7.94	11.81	6.49	25.54
9	8.08	15.62	10.82	11.68	5.91	21.97
10	9.11	13.69	12.27	12.61	16.01	21.76
11	7.74	19.05	10.53	12.11	9.59	16.58
12	9.83	18.41	12.10	11.24	9.37	21.82
13	7.81	22.19	14.40	13.45	7.48	20.97
14	6.97	15.46	17.05	14.20	7.88	15.11
15	9.12	14.00	9.49	14.26	7.93	17.95
Varianz		8.95	7.39	3.51	11.03	9.44
Standard- abweichung		2.99	2.72	1.87	3.32	3.07

Equisolidangle

Dieses Modell funktioniert für zwei der getesteten Kameras am besten, nämlich für das Smartphone mit aufgesetztem Fisheye Objektiv und die Andoer Actioncam. Auch hier lieferte das Modell leider nicht immer ein Ergebnis und für manche Systeme, wie die GoPro Kamera oder diejenigen ohne Fisheye Charakteristik, sind die Werte ungenügend.

Tabelle 3.14: Reprojection Errors der equisolidangle Projektion ohne Distortionkoeffizienten.

Equisolidangle						
Split-Nr.	Walimex Pro Fisheye Objektiv	Gopro	Andoer	Canon Objektiv	Clip-On Objektiv	Smartphone
1	.Nan	38.45	39.82	38.46	8.24	57.29
2	21.94	.Nan	89.25	26.99	17.83	52.52
3	14.96	56.90	24.03	45.32	18.36	50.81
4	21.49	41.51	46.16	38.52	8.83	47.89
5	.Nan	61.48	31.44	25.79	17.56	54.05
6	90.07	.Nan	23.74	37.98	4.84	58.00
7	5.61	104.45	122.83	33.47	9.48	59.27
8	38.89	65.86	4.71	37.13	19.84	93.80
9	19.79	47.81	39.44	28.05	25.11	56.23
10	55.40	65.87	50.87	31.78	15.07	71.88
11	50.45	97.93	7.50	25.80	14.54	33.66
12	16.36	50.85	9.80	29.75	12.87	53.24
13	21.19	102.79	63.86	29.36	21.85	50.66
14	11.47	59.70	95.43	47.73	9.05	51.40
15	27.17	47.58	53.20	40.01	12.09	48.81
Varianz			1,169.24	48.58	32.75	172.43
Standard- abweichung			34.19	6.97	5.72	13.13

3.3.4 Optischer Vergleich der Ergebnisse ohne Distortionparameter anhand des Walimex Pro Fisheye Objektivs

Bei Betrachtung der Ergebnisse in Abbildung 3.9 fällt auf, dass sich jeweils drei Bilder stark ähneln. So sieht das entzerrte Foto des perspektivischen und des equisolidangle Modells fast genauso aus, wie das Originalbild. Beispielsweise ist bei Betrachtung des linken Türrahmens ersichtlich, dass dieser bei dem equisolidangle deutlich weniger gekrümmt ist als bei dem perspektivischen oder dem Originalbild. Das Originalbild ist tatsächlich von dem Ergebnis des perspektivischen Modells kaum zu unterscheiden. Die drei anderen Modelle sind sich ebenfalls sehr ähnlich. Bei allen dreien werden die radialen Verzerrungen bereits etwas korrigiert, aber nicht vollständig behoben. Auch dies lässt sich gut an der eingerollten Leinwand und den Türrahmen erkennen. Hier zeigt sich wieder, dass ein teilweise deutlich größerer RMSE optisch kaum zu unterscheiden ist. Denn bei diesen drei Modellen ist es nur durch die Betrachtung nicht eindeutig, welches am besten ist.

Tabelle 3.15: Kalibrierwerte ohne Distortionparameter für das Walimex Pro Fisheye Objektiv.

	Perspective	Stereographic	Equidistant	Orthogonal	Equisolidangle
fx	3,489.07	4,053.15	1,502.33	3,405.43	5,046.35
fy	2,725.83	3,875.00	1,486.56	3,128.58	4,430.12
cx	1,546.95	1,938.30	1,943.86	1,953.93	1,957.08
cy	2,054.38	1,301.60	1,295.94	1,308.53	1,293.04



(a) Originalbild



(b) Perspective (RMSE 5.77)



(c) Stereographisch (RMSE 4.44)



(d) Equidistant (RMSE 3.64)



(e) Orthogonal (RMSE 6.51)



(f) Equisolidangle (RMSE 5.61)

Abbildung 3.9: Ergebnisse der Kalibrierungen für das Walimex Pro Fisheye Objektiv ohne Distortionsparameter.

3.4 Laufzeitvergleich der Projektionsmodelle

Die Projektionsmodelle aus 2.1 haben offensichtlich sehr unterschiedliche Komplexitäten. Im Gegensatz zu dem perspektivischen Modell, welches lediglich die Division benötigt, ist bei den anderen Modellen das Ziehen von Wurzeln und teilweise das Berechnen von trigonometrischen Funktionen nötig. Dies führt zu einem höheren Rechenaufwand und damit zu einer längeren Laufzeit.

Um dies zu untersuchen wurde mit 100 der bereits oben verwendeten Kalibrierfotos des Walimex Pro Fisheye und des Canon 18-55mm Objektivs kalibriert und dabei die Zeit gemessen. Um ein möglichst vergleichbares Ergebnis vorweisen zu können, wurde das Experiment 15 mal wiederholt und der Median der gemessenen Zeitspannen berechnet. Dies ist erforderlich, da obwohl alle anderen Programme beendet wurden, der Computer im Hintergrund verschiedene Aufgaben erledigt und somit Verfälschungen auftreten, da die CPU möglicherweise die Berechnung kurzzeitig für andere Aufgaben unterbricht. Dies lässt sich leider nicht verhindern. Durch die mehrmalige Wiederholung des Experiments und da alle Modelle auf dem selben System berechnet wurden, sind die Ergebnisse dennoch durchaus aussagekräftig und vergleichbar. Der eingesetzte PC verwendet als Betriebssystem Ubuntu 18.04.5 LTS, hat einen AMD Ryzen 7 3700x 8-Kern Prozessor mit 16 Threads, welcher auf 4,2 GHz übertaktet wurde. Außerdem besitzt er 32Gb Arbeitsspeicher. Durch die vielen Kerne sollte es unwahrscheinlich sein, dass die CPU-Zeit für andere Berechnungen unterbrochen wird. Ausgeschlossen werden kann es aber nicht.

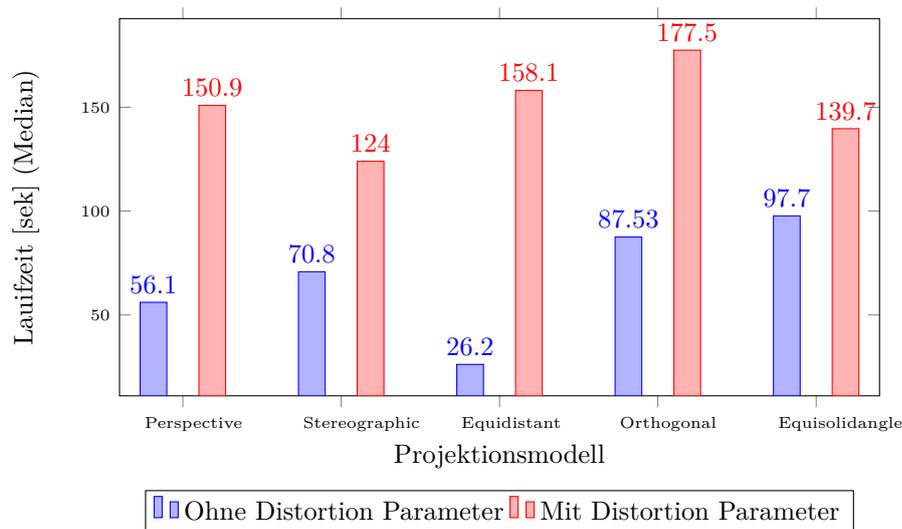


Abbildung 3.10: Laufzeitvergleich der verschiedenen Projektionsmodelle anhand des Walimex Pro Fisheye Objektivs mit 100 Kalibrierfotos.

Wie in Grafik 3.10 ersichtlich ist, erhöht der Einsatz von Distortionparametern die Laufzeit erheblich. Anders als erwartet, benötigt das perspektivische Modell aufgrund der mathematischen Einfachheit weder mit, noch ohne Distortionparametern die geringste Rechenzeit. Auffällig ist

auch, dass das equidistante Modell ohne Distortionparametern mit Abstand am schnellsten ein Ergebnis liefert, mit Distortionparametern jedoch fast am meisten Zeit benötigt. Durch diesen Versuch lässt sich also zeigen, dass die Komplexität des Modells nicht unbedingt mit der Laufzeit korreliert und, dass die Hinzunahme der Distortionparameter immer zu einer schlechteren Laufzeit führt. Dies lässt sich dadurch erklären, dass mehr Unbekannte optimiert werden müssen.

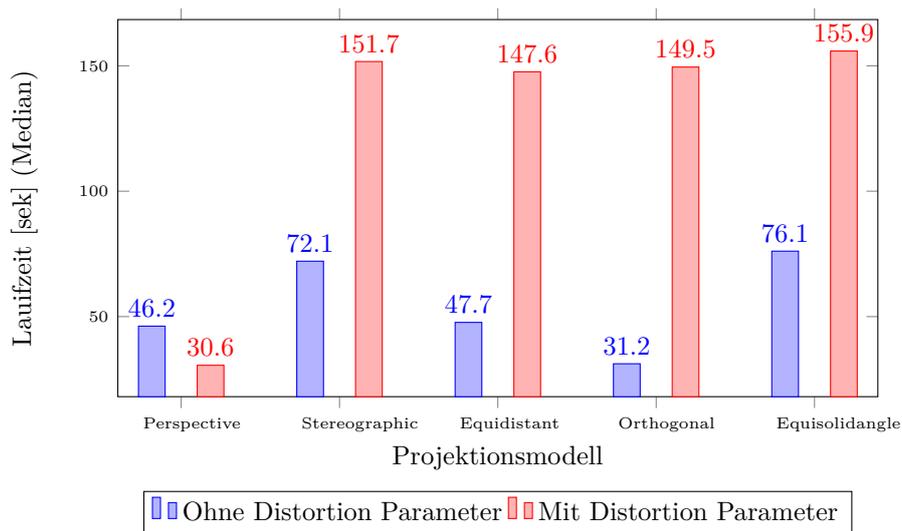


Abbildung 3.11: Laufzeitvergleich der verschiedenen Projektionsmodelle anhand des Canon 18-55mm Objektivs mit 100 Kalibrierfotos.

Wird die Laufzeit anhand des Canon 18-55 mm Objektivs (s. Abb. 3.11) betrachtet, so fällt sofort auf, dass hier das perspektivische Modell mit Distortionparametern schneller konvergiert als das perspektivische ohne diese. Desweiteren wird deutlich, dass das equidistante nicht mehr das Schnellste ist, sondern das orthogonale Modell ohne Distortionkoeffizienten einen deutlichen Vorsprung hat. Aus dem Experiment lässt sich auch ableiten, dass die Laufzeit nicht direkt mit der Güte der Kalibrierung korreliert. Denn auch wenn bei dem Canon Objektiv das perspektivische Modell sowohl mit Abstand die geringste Rechenzeit benötigt als auch das beste Ergebnis liefert, so ist dies bei dem Walimex Pro Fisheye Objektiv nicht der Fall. Denn bei diesem benötigt das stereographische Modell am wenigsten Zeit, das perspektivische führt jedoch zu einem geringeren Reprojection Error und einer besseren Korrektur der radialen Verzerrungen.

Kapitel 4

Färben einer 3D-Punktwolke

Wie in Kapitel 2.1 beschrieben, liefert ein 3D-Laserscanner Messpunkte im Raum. Diese Punkte bestehen aber lediglich aus Koordinaten und in manchen Fällen aus dem Reflektionswert, der angibt wie viel Licht reflektiert wurde und somit vom Material und dem Winkel des Objekts zum Scanner abhängt.



(a) Färbung nach Entfernung



(b) Färbung nach Reflectance Wert



(c) Eingefärbt mithilfe einer Kamera

Abbildung 4.1: Verschiedene Darstellungsvarianten einer 3D-Punktwolke.

Es gibt drei Varianten 3D-Punktwolken darzustellen. Die Erste (Färbung nach Entfernung s. Abb.4.1(a)) ist mit jedem 3D-Scanner möglich, da lediglich die Koordinaten der Messpunkte verwendet werden. Hierbei werden die Punkte, analog zu einer Kamera, auf eine zweidimensionale Fläche projiziert. Anders als bei der Kamera werden nun die Pixel aber nicht abhängig von der Farbe des abgebildeten Objekts, sondern abhängig von der Entfernung eingefärbt. Somit lässt sich ein Tiefeneffekt generieren und es ist möglich Entfernungen im Bild abzuschätzen. Sind die Reflectance-Werte vorhanden, lässt sich das Bild statt mit den Entfernungen zu arbeiten mit diesen einfärben (s. Abb. 4.1(b)). Hierdurch erhält der Betrachter deutlich mehr Informationen über die Struktur und einen besseren Eindruck von der abgebildeten Szene. Im Gegensatz zur ersten Variante lässt sich von dem Bild schwer oder gar nicht auf die Entfernung schließen, was beispielsweise an den Bäumen im Hintergrund deutlich wird. Die dritte Variante (s. Abb. 4.1(c)) hingegen verwendet Bilder einer Kamera (hier 4.2), um den Messpunkten ihre tatsächlichen Farbwerte zu geben. Das Ergebnis ermöglicht es, einen guten Eindruck der abgebildeten Szenerie zu erhalten und zusätzlich sind Menschen instinktiv in der Lage, Entfernungen zu schätzen, wenn sich bekannte Referenzobjekte darin befinden. Somit lässt sich beispielsweise die Entfernung zu dem Kleintransporter schätzen, da ungefähr bekannt ist, wie groß ein Solcher in der Realität ist.



Abbildung 4.2: Für die Färbung verwendetes Bild (Canon 1000D mit Walimex Pro Fisheye Objektiv).

4.1 Prinzip der Färbung

Um eine 3D-Punktwolke zu färben, sind also Bilder einer Kamera nötig, von der sowohl die Intrinsik als auch die Extrinsik bekannt sind. Natürlich spielt die Ausrichtung der Kamera relativ zum Laserscanner eine sehr große Rolle, denn geringe Abweichungen des Winkels führen bereits dazu, dass alle Farbwerte verschoben werden. Dies fällt bereits bei sehr geringen Fehlern insbe-

sondere an Objektkanten auf. Um die Punktwolke zu färben, wird diese mittels der Extrinsik in das Koordinatensystem der Kamera transformiert. Anschließend können die 3D-Koordinaten mit dem Kameramodell und der Intrinsik in 2D abgebildet werden. Diese 2D-Koordinaten entsprechen dann den Pixelkoordinaten des Bildes, dessen Farbwerte an der entsprechenden Stelle schließlich den 3D-Koordinaten zugeordnet und wieder in das Koordinatensystem des Laserscanners zurück transformiert werden. Um eine 3D-Punktwolke farblich darzustellen, gibt es mehrere Möglichkeiten. Es werden, wie oben beschrieben, die Intrinsik und Extrinsik der Kamera benötigt. Die Extrinsik muss immer durch Punktkorrespondenzen mit dem Laserscanner erhalten werden. Die Intrinsik kann ebenfalls durch diese Korrespondenzen bestimmt werden. Es ist aber auch möglich, diese durch eine vorherige Kalibrierung beispielsweise mittels Checkerboard, wie in Kapitel 3.1 beschrieben, zu erhalten. Dies hat den Vorteil, dass mit wenig Aufwand viele Punktkorrespondenzen genutzt werden können. Der Nachteil ist, dass die Kalibrierung nicht zeitgleich mit den für die Färbung verwendeten Fotos geschieht, wodurch die Parameter sich durch thermische Unterschiede oder auch durch vibrationsbedingte Änderungen der Optik (Verstellung der Brennweite oder ähnliches) verändern können. Werden Punktkorrespondenzen aus dem Laserscan verwendet, können diese Probleme größtenteils umgangen werden. Dafür müssen diese aber händisch gesucht werden, was sowohl zeitaufwendig als auch fehlerbehaftet ist.

4.2 Evaluation der Färbungen mit den verschiedenen Kameramodellen

Um die verschiedenen Kameramodelle bei der Färbung von 3D-Punktwolken zu vergleichen, wurde ein 3D-Scan zwischen dem Rechenzentrum und dem Institut für Informatik der Universität Würzburg aufgenommen. Diese Position wurde gewählt, da der Scan sowohl einige markante Punkte auf der Horizontlinie aufweist, was für die folgenden Versuche interessant ist, als auch, weil der Boden dort gepflastert ist und somit viele Punktpaare zwischen Fotos und Scan möglich sind.

4.2.1 Verwendete Datensätze

Für die Färbung wurden zwei Datensätze angefertigt. Einer besitzt lediglich Korrespondenzen auf der Horizontlinie, der andere überwiegend am Boden.

Tabelle 4.1: Für die Färbung verwendete Datensätze.

	Anzahl Korrespondenzen
Horizon	87
Ground	255

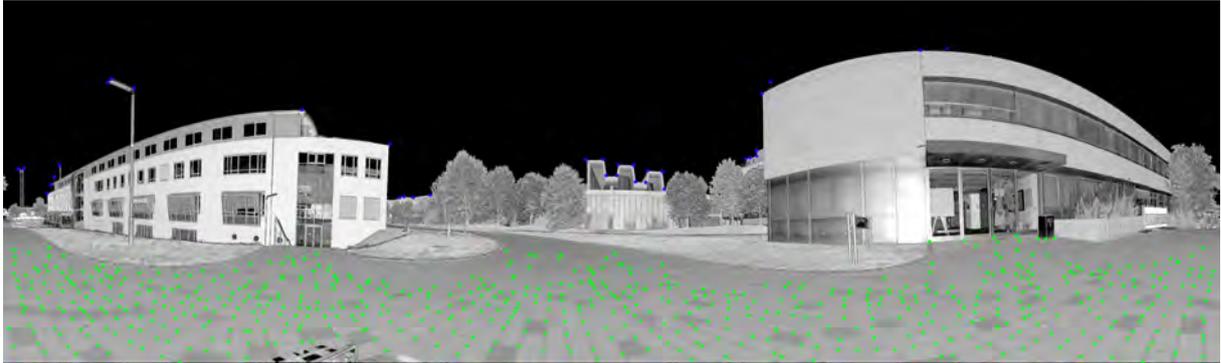


Abbildung 4.3: Reflectanceimage des Scans mit Beispielen für Punktkorrespondenzen.
Blau: Punkte des Horizon Datensatzes *Grün:* Punkte des Ground Datensatzes

Die Datensätze haben verschieden viele Punktpaare, was daran liegt, dass es auf der Horizontlinie sehr viel weniger markante Punkte gibt als beispielsweise auf dem Boden, wo jeder Pflasterstein mehrere Punkte ermöglicht. Beispiele hierfür sind in Grafik 4.3 zu sehen. Um die Punktwolke zu färben, wurde jeweils mit den fünf Kameramodellen und den zwei Datensätzen eine Kalibrierung durchgeführt. Zusätzlich wurde einmal mit einem Initial Guess aus der jeweils besten Kalibrierung aus Kapitel 3.3.1 und einmal ohne kalibriert. Die Kalibrierfehler sind in den Tabellen 4.2 und 4.3 zu sehen. Auf den ersten Blick ist bereits ersichtlich, dass die Fehler in allen

Tabelle 4.2: Kalibrierfehler auf dem Ground Datensatz.

Ground					
	Perspective	Stereographic	Equidistant	Orthogonal	Equisolidangle
Mit Initial Guess	3.7	8.3	4.2	29.1	6.1
Ohne Initial Guess	3.7	447.3	16.2	708.0	623.3

Tabelle 4.3: Kalibrierfehler auf dem Horizon Datensatz.

Horizon					
	Perspective	Stereographic	Equidistant	Orthogonal	Equisolidangle
Mit Initial Guess	139.9	524.1	160.1	390.0	405.0
Ohne Initial Guess	64.5	489.7	300.5	.NaN	.NaN

Fällen deutlich über denen der Schachbrettkalibrierung liegen. Dies lässt sich durch die manuelle und fehlerbehaftete Korrespondenzsuche erklären. Bei der Schachbrettkalibrierung werden beispielsweise die Pixel sogar mit Subpixelgenauigkeit bestimmt, wohingegen bei der manuellen Auswahl und dem Verwenden natürlicher Objekte fast immer mehrere Pixel zur Auswahl stehen. So ist die Kante eines Pflastersteins keineswegs eindeutig, da dessen Kanten abgerundet und auch farblich nicht klar gekennzeichnet sind wie bei dem Schachbrett. Bei der Auswahl des korrespondierenden 3D-Punktes besteht das selbe Problem. Daher ist hier auch ein schlechteres

Ergebnis zu erwarten.

Dies könnte optimiert werden, indem zum Beispiel mehrere Schachbrettmuster in der Szene verteilt werden, deren Punkte sowohl im Bild als auch im Reflectance-Image des 3D-Scans mit geringem Fehler ausgewählt werden können. Desweiteren wurden sehr viel weniger Korrespondenzen verwendet. Bei der Schachbrettkalibrierung wurden für jeden Split 140 Fotos eingesetzt, die in jedem Foto 100 Korrespondenzen ermöglichen. Daher wurde jeweils mit 14.000 Korrespondenzen gerechnet. Ein weiterer Grund für die größeren Fehler ist die Sensorabdeckung. Bei beiden, aber ganz besonders bei dem Horizon Datensatz, wurde nur ein kleiner Bereich des Kamerasensors abgedeckt, da die Horizontlinie naturgemäß nur im oberen Bereich liegt.

4.2.2 Evaluierung mit dem Ground Datensatz

Im Folgenden werden die Ergebnisse der Färbung erläutert. Neben den Endergebnissen wird auch auf Parameter der Kalibrierung und deren Auswirkung auf das Ergebnis eingegangen. Hierzu nochmals die Bedeutung der Parameter:

$f = (f_x, f_y)$ Brennweite in x- und y-Richtung

$c = (c_x, c_y)$ x- und y-Koordinate des Principal Points

$\mathbf{R} = (r_x, r_y, r_z)$ Rotation relativ zum 3D-Laserscanner

$\mathbf{T} = (t_x, t_y, t_z)$ Translation relativ zum 3D-Laserscanner

Die Brennweiten hängen hierbei sehr von dem jeweiligen Projektionsmodell ab. Dennoch sollten sich bei der verwendeten Kamera die Brennweiten für x- und y-Richtung nicht stark unterscheiden. Der Principal Point, welcher den Mittelpunkt der Projektion darstellt, liegt bei einer idealen Kamera genau im Sensormittelpunkt. Bei der Canon Kamera und der eingesetzten Auflösung wäre dies also bei $c = (1296, 1944)$. Ebenso wie c sind die extrinsischen Parameter nicht vom Projektionsmodell abhängig. Sie beschreiben schließlich die Rotation \mathbf{R} und Translation \mathbf{T} der Kamera relativ zum 3D-Laserscanner (für das erste Photo) und sind somit für alle Kameramodelle gleich.

Perspective mit Initial Guess



(a) Position A



(b) Position B



(c) Position C

Abbildung 4.4: Färbung mit dem perspektivischen Kameramodell mit Initial Guess (Reprojection Error: 3.7).

Bei der Kalibrierung mit dem perspektivischen Modell und einem Initial Guess wurde ein gutes Ergebnis erzielt. Auf den ersten Blick ist aus der Ausgangsposition (4.4(a)) nur in großer Entfernung, beispielsweise bei den Schornsteinen hinten mittig im Bild, eine Farbverschiebung zu erkennen. Dass der Baum einen bläulichen Rand hat, ist selbst bei einer sehr guten Kalibrierung kaum zu vermeiden, da die Fotos und der Scan nicht gleichzeitig entstehen und somit bereits leichter Wind hier zu Fehlern führt. Bei Abb. 4.4(b) ist außer am Blitzableiter oben links am Rechenzentrumsgebäude kaum ein Fehler zu sehen. Obwohl hier mehrere Bilder zusammengefügt wurden, sind kaum Übergänge und Fehler beim Stitching festzustellen. Am besten sind Fehler bei einer solchen Färbung an Objekten wie der Straßenlaterne in 4.4(c) zu erkennen. Durch den schwarzen Hintergrund im oberen Bereich und den Entfernungsunterschied zwischen Laterne und dem Institut für Informatik im Hintergrund fallen hier bereits sehr kleine Fehler in Rotation und

Translation deutlich auf. Auch wenn die gesuchten Parameter nicht bekannt sind, ist ersichtlich, dass die gefundenen Werte (s. Tabelle 4.4) durchaus plausibel sind. So zeigt der Translationsvektor, dass die Kamera circa 22,8cm über dem Laserscanner montiert ist. Außerdem ist sie nur geringfügig in x-Richtung und circa -10,8cm in z-Richtung vom Koordinatenursprung verschoben. Der Principal-Point weicht von dem geschätzten Wert ab, ist aber dennoch realistisch. Die schwarzen Stellen in den Fenstern stammen nicht aus Fehlern bei der Färbung. Dort existieren keine Messpunkte, da der Laserstrahl vom Glas reflektiert wurde. Daran lässt sich gut erkennen, dass die Farbverschiebung nach oben hin zunimmt, was aus Fehlern des Rotationsvektors folgt.

Tabelle 4.4: Intrinsik und Extrinsik - Perspektivisch mit Initial Guess.

Intrinsik				Extrinsik					
fx	fy	cx	cy	rx	ry	rz	tx	ty	tz
1,481.90	1,456.48	1,332.87	2,011.51	77.78°	-57.37°	56.95°	0.42	22.76	-10.83

Perspective ohne Initial Guess



(a) Position A



(b) Position B



(c) Position C

Abbildung 4.5: Färbung mit dem perspektivischen Kameramodell ohne Initial Guess (Reprojection Error: 3.7).

Das Ergebnis der Färbung ohne Initial Guess unterscheidet sich nur geringfügig von dem mit. Sowohl die Färbung als auch die einzelnen Parameter sind sich sehr ähnlich. Selbst der Reprojection Error unterscheidet sich nur um 0.005. Dadurch zeigt sich, dass das perspektivische Modell auch ohne einen Initial Guess ein gutes Ergebnis liefern kann.

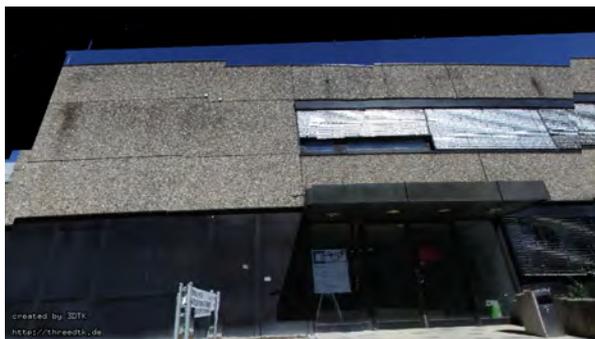
Tabelle 4.5: Intrinsik und Extrinsik - Perspektivisch ohne Initial Guess.

Intrinsik				Extrinsik					
fx	fy	cx	cy	rx	ry	rz	tx	ty	tz
1,478.13	1,457.33	1,332.45	2,012.26	77.79°	-57.36°	56.94°	0.43	22.43	-10.59

Stereographisch mit Initial Guess



(a) Position A



(b) Position B



(c) Position C

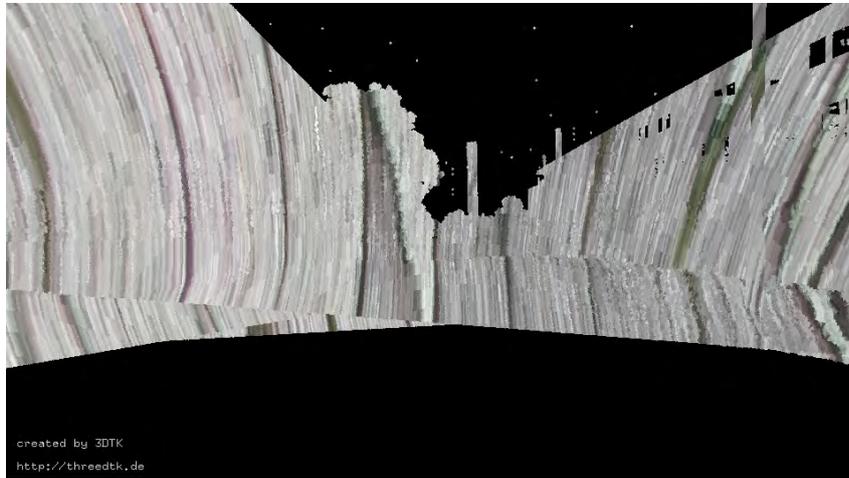
Abbildung 4.6: Färbung mit dem stereographischen Kameramodell mit Initial Guess (Reprojection Error: 8.3).

Anders als bei dem perspektivischen ist bei der Färbung mit dem stereographischen Modell mit initial Guess die Verschiebung gravierend. Wird nur die unmittelbare Umgebung des Scanners in Bild 4.6(a) betrachtet, fällt diese beispielsweise am Boden kaum auf. Auch die Bildübergänge sind dort fast nicht erkennbar. Bereiche oberhalb des Bodens sind aber offensichtlich stark verschoben. Bei Betrachtung der Parameter (s. Tabelle 4.6) wird ersichtlich, dass die Rotation mit der des perspektivischen Modells fast übereinstimmt. Die Translation weicht aber um mehrere Zentimeter ab und ist somit hauptverantwortlich für diese Fehler.

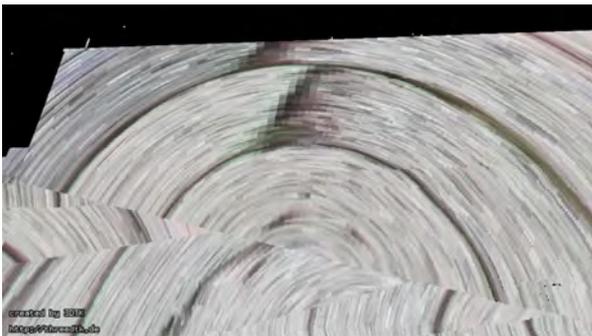
Tabelle 4.6: Intrinsik und Extrinsik - Stereographisch mit Initial Guess.

Intrinsik				Extrinsik					
fx	fy	cx	cy	rx	ry	rz	tx	ty	tz
3,097.61	3,147.80	1,301.20	1,963.37	77.45°	-56.48°	56.24°	1.55	25.32	-2.31

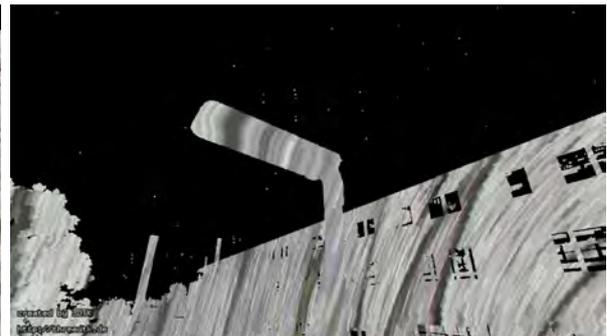
Stereographisch ohne Initial Guess



(a) Position A



(b) Position B



(c) Position C

Abbildung 4.7: Färbung mit dem stereographischen Kameramodell ohne Initial Guess (Reprojection Error: 447.3).

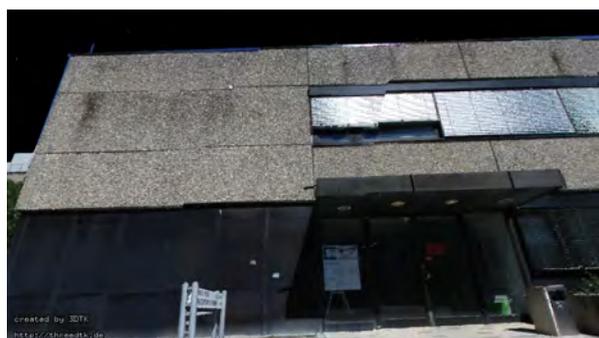
Bei Verwendung des stereographischen Modells ohne Initial Guess werden alle Parameter sehr fehlerbehaftet ermittelt. Bereits an den unterschiedlichen Brennweiten gemäß Tabelle 4.7 ist erkennbar, dass diese Kalibrierung keine brauchbaren Ergebnisse liefern kann. Die daraus generierte Färbung ist demnach ebenso unbrauchbar.

Tabelle 4.7: Intrinsik und Extrinsik - Stereographisch ohne Initial Guess.

Intrinsik				Extrinsik					
f_x	f_y	c_x	c_y	r_x	r_y	r_z	t_x	t_y	t_z
1,451.84	-73.23	1,244.43	2,908.84	10.63°	-10.83°	59.07°	-13.64	67.98	-58.96

Equidistant mit Initial Guess

(a) Position A



(b) Position B



(c) Position C

Abbildung 4.8: Färbung mit dem equidistanten Kameramodell mit Initial Guess (Reprojection Error: 4.2).

Die Färbung mit dem equidistanten Modell und Initial Guess ähnelt denen mit dem perspektivischen Modell. Die Parameter aus Tabelle 4.12 gleichen denen des perspektivischen mit initial Guess in Tabelle 4.4 ebenfalls sehr. Die Translation weicht leicht ab und zeigt, dass die Position der Kamera näher am Laserscanner geschätzt wird. Die Verschiebung an der Laterne in Abb. 4.8(c) fällt geringer aus als bei den Färbungen mit dem perspektivischen Modell (Abb. 4.4(c) und 4.5(c)). Offenbar wird mit dem equidistanten Modell hier eine bessere Extrinsic berechnet. So werden sogar die sehr weit entfernten Schornsteine teilweise richtig eingefärbt. Das Stitching ist hingegen schlechter als bei dem perspektivischen Modell und besonders am Rechenzentrumsgebäude in Abbildung 4.8(b) deutlich sichtbar. Dies ist eine Folge des höheren Reprojection Errors.

Tabelle 4.8: Intrinsik und Extrinsik - Equidistant mit Initial Guess.

Intrinsik				Extrinsik					
fx	fy	cx	cy	rx	ry	rz	tx	ty	tz
1,476.32	1,480.21	1,300.42	1,961.10	76.36°	-55.23°	56.86°	0.87	20.81	-8.82

Equidistant ohne Initial Guess



(a) Position A



(b) Position B



(c) Position C

Abbildung 4.9: Färbung mit dem equidistanten Kameramodell ohne Initial Guess (Reprojection Error: 16.2).

Ohne Initial Guess liefert das equidistante Modell hingegen deutlich schlechtere Ergebnisse. Auch hier ist die Rotation (s. Tabelle 4.13) wieder recht nahe an den Werten der Kalibrierung mit Initial Guess (s. Tabelle 4.12). Die Translation und die Brennweiten hingegen weichen stark ab. Dies führt zu großen Fehlern beim Stitching, Färbem (s. Abbildung 4.9) und einem hohen Reprojection Error.

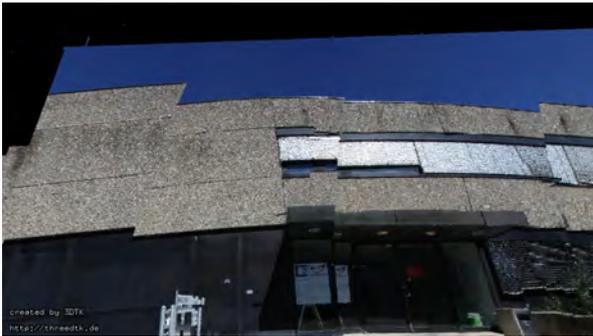
Tabelle 4.9: Intrinsik und Extrinsik - Equidistant ohne Initial Guess.

Intrinsik				Extrinsik					
fx	fy	cx	cy	rx	ry	rz	tx	ty	tz
1,330.51	1,160.37	1,303.47	1,971.11	74.9°	-52.43°	57.93°	1.03	32.02	-38.31

Orthogonal mit Initial Guess



(a) Position A



(b) Position B



(c) Position C

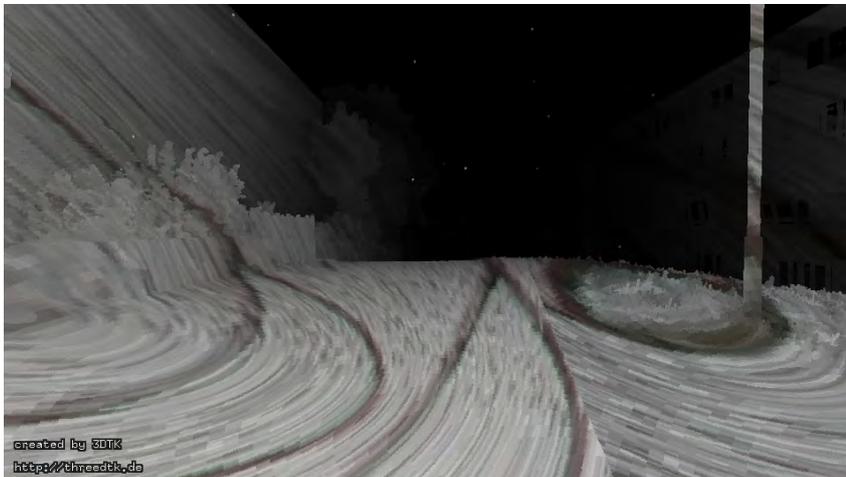
Abbildung 4.10: Färbung mit dem orthogonalen Kameramodell mit Initial Guess (Reprojection Error: 29.1).

Auch bei der orthogonalen Kalibrierung mit Initial Guess sind der Principal Point und die Rotation (s. Tabelle 4.10) plausibel. Die restlichen Parameter, insbesondere die Translation, liegen weit neben den erwarteten Werten. Somit ist auch diese Färbung (s. Abbildung 4.10) stark verschoben und fehlerhaft.

Tabelle 4.10: Intrinsik und Extrinsik - Orthogonal mit Initial Guess.

Intrinsik				Extrinsik					
fx	fy	cx	cy	rx	ry	rz	tx	ty	tz
1,896.46	1,985.70	1,270.56	1,883.17	78.16°	-57.34°	55.05°	4.28	36.93	23.84

Orthogonal ohne Initial Guess



(a) Position A



(b) Position B



(c) Position C

Abbildung 4.11: Färbung mit dem orthogonalen Kameramodell ohne Initial Guess (Reprojection Error: 708.0).

Bereits am Reprojection Error von 708.0 ist ersichtlich, dass die Parameter dieser Kalibrierung (s. Abb. 4.11) nicht zu einem guten Ergebnis führen können. Hier sind sowohl alle Werte der Intrinsik als auch der Extrinsik (s. Tabelle 4.11) weit von denen entfernt, die sich bei den anderen Kalibrierungen bewährt haben (beispielsweise Tabelle 4.4). Die Färbung ist dementsprechend wie bereits bei dem stereographischen Modell ohne Initial Guess unbrauchbar.

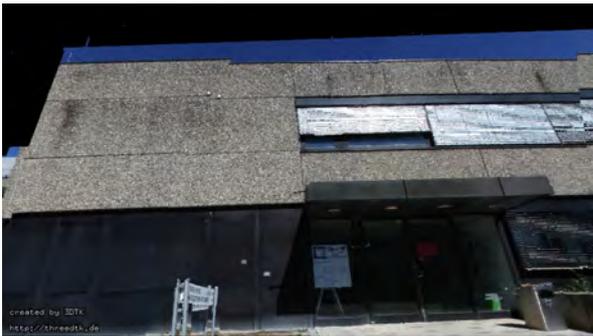
Tabelle 4.11: Intrinsik und Extrinsik - Orthogonal ohne Initial Guess.

Intrinsik				Extrinsik					
f_x	f_y	c_x	c_y	r_x	r_y	r_z	t_x	t_y	t_z
338.87	-33.59	1,352.26	2,800.99	-67.45°	-11.7°	-25.01°	0.94	-91.61	-105.64

Equisolidangle mit Initial Guess



(a) Position A



(b) Position B



(c) Position C

Abbildung 4.12: Färbung mit dem equisolidangle Kameramodell mit Initial Guess (Reprojection Error: 6.1).

Bei der Färbung mit dem equisolidangle Modell sind deutliche Fehler beim Stitching und Verschiebungen vorhanden (s. Abbildung 4.12). Bei der Betrachtung der Parameter (s. Tabelle 4.12) fällt wiederum auf, dass die Rotation recht nahe an denen der perspektivischen wie beispielsweise in Tabelle 4.4 liegt. Die Translation hingegen weicht stärker ab und führt dazu, dass die Verschiebung insbesondere nach unten recht groß ausfällt.

Tabelle 4.12: Intrinsik und Extrinsik - Equisolidangle mit Initial Guess.

Intrinsik				Extrinsik					
f_x	f_y	c_x	c_y	r_x	r_y	r_z	t_x	t_y	t_z
3,091.29	3,227.05	1,288.99	1,982.54	77.68°	-56.10°	55.90°	1.90	19.96	-3.50

Equisolidangle ohne Initial Guess



(a) Position A



(b) Position B



(c) Position C

Abbildung 4.13: Färbung mit dem equisolidangle Kameramodell ohne Initial Guess (Reprojection Error: 623.3).

Bei der Färbung ohne Initial Guess zeigt sich bereits bei den Werten für die Brennweiten (s. Tabelle 4.13) und dem Reprojection Error von 623.3, dass dies zu keinem brauchbaren Ergebnis führen kann. Hier sind alle Werte weit von den erwarteten entfernt und somit sind die Ergebnisse der Kalibrierung und der damit erstellten Färbung unbrauchbar (s. 4.13).

Tabelle 4.13: Intrinsik und Extrinsik - Equisolidangle ohne Initial Guess.

Intrinsik				Extrinsik					
fx	fy	cx	cy	rx	ry	rz	tx	ty	tz
876.51	-32.15	1,159.59	2,876.28	-35.59°	3.74°	12.82°	0.16	33.82	80.37

4.2.3 Evaluierung mit dem Horizon Datensatz

Die Kalibrierung und Färbung der 3D-Punktwolke mit dem Horizon Datensatz hat gezeigt, dass Punktkorrespondenzen nur auf der Horizontlinie nicht ausreichen, um brauchbare Ergebnisse zu erzielen. So liegt der Reprojection Error (s. Abb. 4.3) im besten Fall bei 64.5 und im Schnitt bei 210. Bei zwei Modellen wird sogar überhaupt kein Ergebnis erreicht. Der große Fehler der Kalibrierung führt somit in allen Fällen dazu, dass die Färbung misslingt. Dies lässt sich durch mehrere Punkte begründen:

Anzahl der Korrespondenzen

Im Gegensatz zur Kalibrierung mit dem Schachbrettmuster oder auch mit dem Ground Datensatz ist die Anzahl der korrespondierenden Punkte hier sehr gering. Auch wenn insgesamt 87 Punkte auf der Horizontlinie gefunden wurden, so ist die Anzahl für jedes der 10 Bilder viel geringer, da jedes Foto lediglich einen kleinen Ausschnitt des Scans zeigt. Wie in Tabelle 4.14 zu

Tabelle 4.14: Anzahl der Korrespondenzen.

Foto	Anzahl der Korrespondenzen	
	Ground	Horizon
1	37	9
2	27	1
3	17	4
4	32	10
5	23	14
6	25	16
7	14	7
8	11	7
9	32	9
10	37	10
Gesamanzahl	255	87
Schnitt	25.5	8.7
Median	26	9

sehen ist, unterscheidet sich die Punktzahl für jedes Bild deutlich und liegt bei dem Horizon Datensatz zwischen nur einem und 16 Punkten. Da mindestens vier Punkte pro Bild vorhanden sein müssen, kann mit Bild 2 tatsächlich auch gar nicht kalibriert werden. Beim Ground Datensatz gibt es im Schnitt pro Bild 25.5 Korrespondenzen, beim Horizon Datensatz lediglich 8.7. Somit ist beim Horizon Datensatz auch nicht mit einem gleich guten, oder besseren Ergebnis zu rechnen. Die Anzahl dieser Korrespondenzen ist natürlich stark vom Ort abhängig, an dem der Scan und die Fotos aufgenommen werden. Somit lässt sich das Problem mit den wenigen Korrespondenzen auf der Horizontlinie womöglich umgehen oder zumindest verringern.

Genauigkeit der Korrespondenzwahl

Ein weiteres Problem liegt in der Genauigkeit der Korrespondenzen. Bei Objekten, die sehr nahe am Scanner und der Kamera liegen, gibt es eine sehr viel höhere Pixeldichte und auch viel mehr Messpunkte des Laserscanners. Somit lassen sich die Koordinaten im Scan und auch im Bild besser auswählen. Beispielsweise liegen die Messpunkte des 3D-Laserscanners aufgrund der großen Entfernung auf den Schornsteinen ganz links im Bild 4.15 bereits über 10 Zentimeter auseinander. In der Nähe des Scanners liegt die Distanz dagegen im Millimeterbereich.

Position der Korrespondenzen

Der wohl entscheidendste Faktor ist jedoch die Position der Punkte. Um eine Kamera zu kalibrieren, sollten immer Korrespondenzen gewählt werden, die möglichst alle Bereiche des Sensors abdecken. Nur so können die Parameter des Kameramodells gut bestimmt werden. Es ist offensichtlich, dass das Kameramodell nicht passend berechnet werden kann, wenn beispielsweise nur Korrespondenzen im mittleren Bereich des Kamerasensors genutzt werden. Denn dort sind auch bei einer Fischaugenkamera kaum bis gar keine radialen Verzerrungen vorhanden und somit geben diese Punkte keinerlei Informationen über die Stärke der Verzerrung am Bildrand. Daher lassen sich natürlich auch keine Distortionparameter passend berechnen.

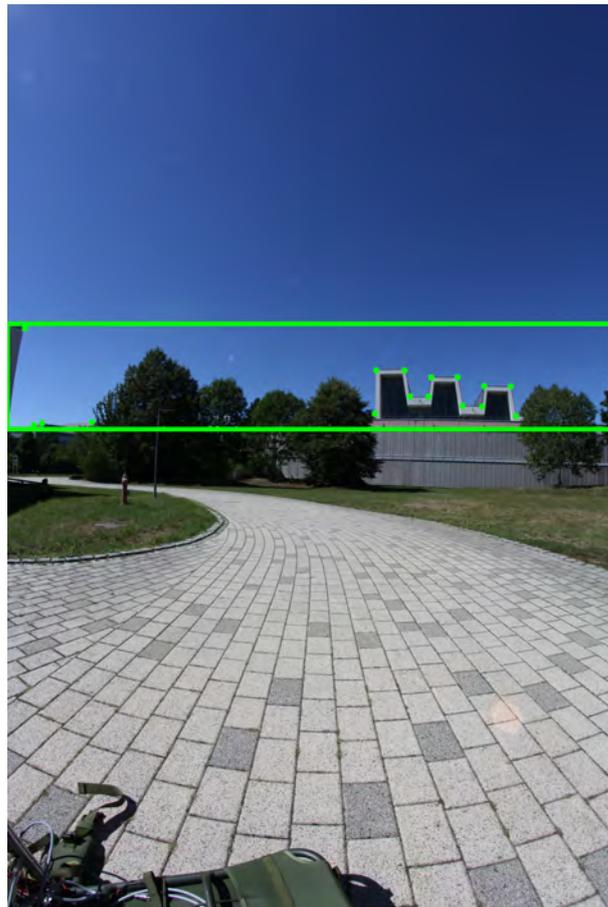


Abbildung 4.14: Punktkorrespondenzen in Photo Nummer 5.

Im fünften der für die Färbungen verwendeten Fotos 4.14 wird deutlich, dass selbst bei dem Bild mit den meisten Korrespondenzen (16) nur ein sehr geringer Teil des Sensors abgedeckt wird. Dadurch fehlen insbesondere Informationen für die Abbildungsfunktion im oberen und unteren Bereich des Bildes. Bei der Schachbrettkalibrierung kann dies durch andere Bilder ausgeglichen werden. In diesem Fall sind diese Bereiche aber in keinem Bild gut abgedeckt. Der obere Bereich beinhaltet immer den Himmel und bietet somit keinerlei verwertbare Punkte, der untere enthält

dagegen nie die Horizontlinie. Hier wird klar, dass der Ground Datensatz besser performt. Allein durch den Boden wird von ganz unten bis ungefähr zur Bildmitte der Sensor gut abgedeckt. Da die Verzeichnungen und die Abbildung annähernd radialsymmetrisch sind, lässt sich somit auch auf die nicht erreichten Bereiche schließen.

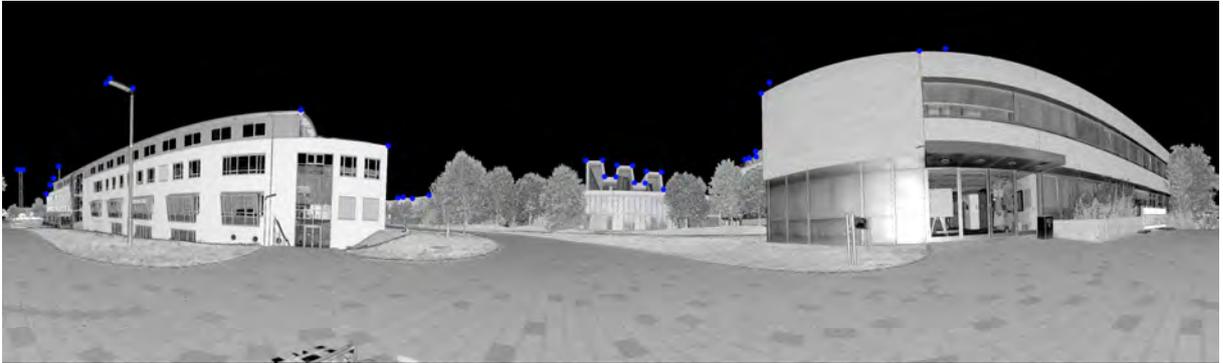


Abbildung 4.15: Beispiele für Punktkorrespondenzen des Horizon Datensatzes.

Kapitel 5

Automatische Korrespondenzenerkennung

Wie sich in Kapitel 4.2.3 gezeigt hat, sind die Ergebnisse einer Kalibrierung nur mit manuell gesuchten Punkten auf der Horizontlinie nicht zufriedenstellend. Daher wird im Folgenden eine Methode vorgestellt diese Punkte automatisch zu erfassen. Hierdurch können mehr Punkte gefunden werden, da sich diese nicht ausschließlich auf Kanten und Ecken befinden. Um dies zu erreichen, wird sowohl in den Fotos als auch in einem aus dem Laserscan generierten Panoramabild die Horizontlinie extrahiert. Der Bereich unter der Horizontlinie wird jeweils schwarz gefärbt, wodurch ein Binärbild entsteht. Anschließend werden darauf mit einem Featureextraktor interessante Punkte gesucht und zwischen Scan und Fotos gematched.

5.1 Automatische Horizontextraktion

Nachfolgend wird beschrieben wie die Horizontlinie sowohl in den Kamerabildern, als auch in der 3D-Punktwolke gefunden werden kann. Bei der 3D-Punktwolke ist dies eine triviale Aufgabe. Der Scan wird hierfür als Panoramabild gespeichert. Dies kann sich so vorgestellt werden, als würde im Zentrum des Koordinatensystems der Punktwolke ein Panoramafoto mit einer Kamera geschossen werden, welche die selben horizontalen und vertikalen Öffnungswinkel besitzt wie der Laserscanner. Somit ist gewährleistet, dass alle Punkte der 3D-Punktwolke in diesem Panoramafoto enthalten sind. Anders als bei einem normalen Foto werden hier nun aber nicht die RGB-Werte für jeden Pixel, welche die 3D-Punktwolke auch nicht beinhaltet, sondern die Koordinaten in XYZ-Richtung gespeichert. Pixel, für die kein Messpunkt vorliegt, werden mit $(0,0,0)$ gefüllt. Dies ist zum Beispiel für alle Pixel im Bereich des Himmels der Fall. Solch ein Panoramafoto ist in Grafik 5.1 zu sehen. Anschließend kann, um die Horizontlinie zu erhalten, jede Spalte des Bildes von oben nach unten durchgegangen und jeweils der erste Pixel, dessen Wert ungleich $(0,0,0)$ ist, gespeichert werden. Somit wird anschließend der Horizont erhalten. Da zum Teil beispielsweise Regentropfen oder sonstige Objekte oberhalb des Horizonts zu Messpunkten im Himmel führen, wurde ein Parameter *Nextpixels* eingeführt, welcher bestimmt, wie

viele Pixel nach dem ersten gefundenen Horizontwert ebenfalls ungleich $(0,0,0)$ sein müssen. Das Programm und das Ergebnis einer Horizontextraktion sind in Abb. 5.1 zu sehen. Das gezeigte Programm liest eine 3D-Punktwolke ein, wandelt diese in ein Panoramabild um und speichert anschließend die gezeigten Bilder und eine 3D-Punktwolke, welche lediglich die Messpunkte der Horizontlinie enthält. Den Horizont in einem Foto automatisch zu extrahieren erfordert hingegen einen größeren Aufwand.

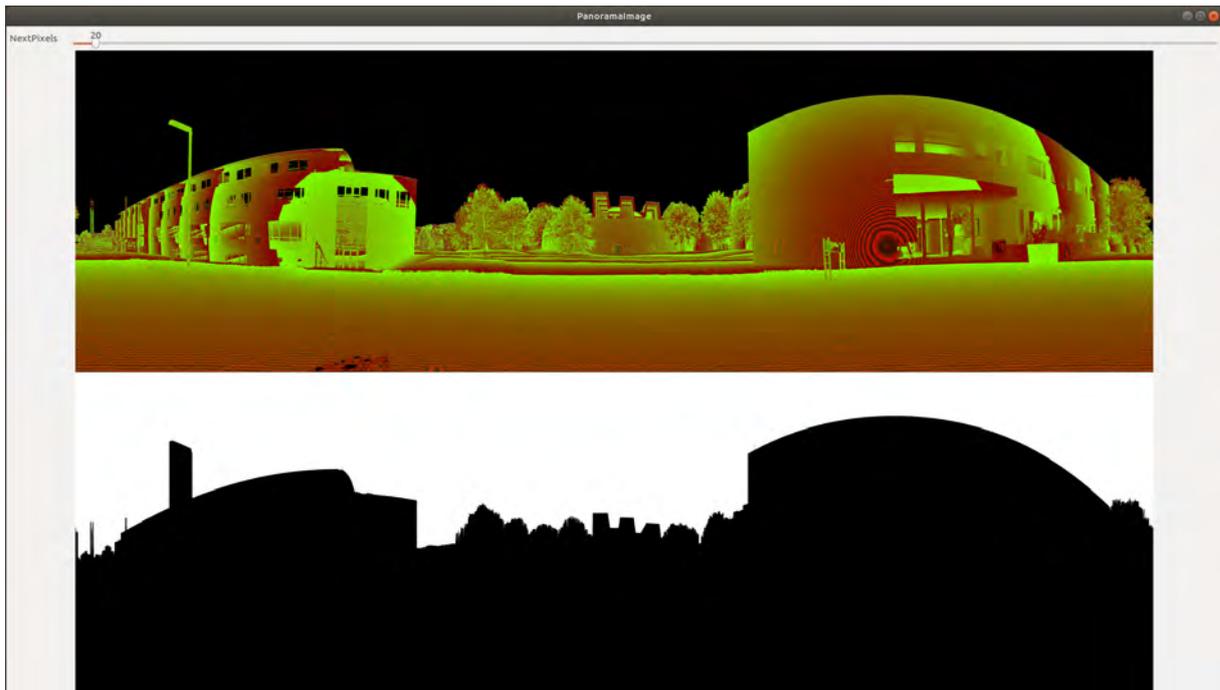


Abbildung 5.1: Programm zu Horizontextraktion aus einer 3D-Punktwolke.

Das hierfür erstellte Programm (s. Abb. 5.2) liest ein Foto ein und ermöglicht es, über eine Ähnlichkeitsfunktion, den Himmel zu entfernen. Hierbei wird mittels des *Floodfill*-Algorithmus zuerst über die per GUI einstellbaren Grenzen der Himmel schwarz gefärbt. Anschließend wird analog wie bei der 3D-Punktwolke der Horizont gefunden. Der Algorithmus funktioniert vom Prinzip her so, dass für einen initialen Pixel, hier standardmäßig der ganz oben in der Mitte des Bildes, für jeden benachbarten Pixel die Grenzen überprüft werden. Liegt der Pixelwert des benachbarten Pixels für alle Farbkanäle innerhalb der Grenzen, wird der Pixel schwarz gefärbt und für dessen benachbarten Pixel wird genauso vorgegangen. Somit werden Flächen, die sich ähnlich sind, schwarz markiert und erst große Farbänderungen sorgen für einen Abbruch. Im Bild ist ersichtlich, dass die Sonne und auch die Wolken im Hintergrund so noch entfernt werden, aber Gebäude und Pflanzen eine ausreichende Hürde darstellen. Auch hierbei gibt es wieder den Parameter *NextPixels*, der bestimmt, wie viele Pixel unter dem ersten gefundenen schwarzen Pixel ebenfalls schwarz sein müssen, um eventuell auftretendes Rauschen entfernen zu können. In Grafik 5.2 ist zu sehen, dass auch Bereiche unterhalb eines Objekts, wie beispielsweise der Straßenlaterne, schwarz gefärbt werden. Somit wird in diesem Verfahren die Horizontlinie und

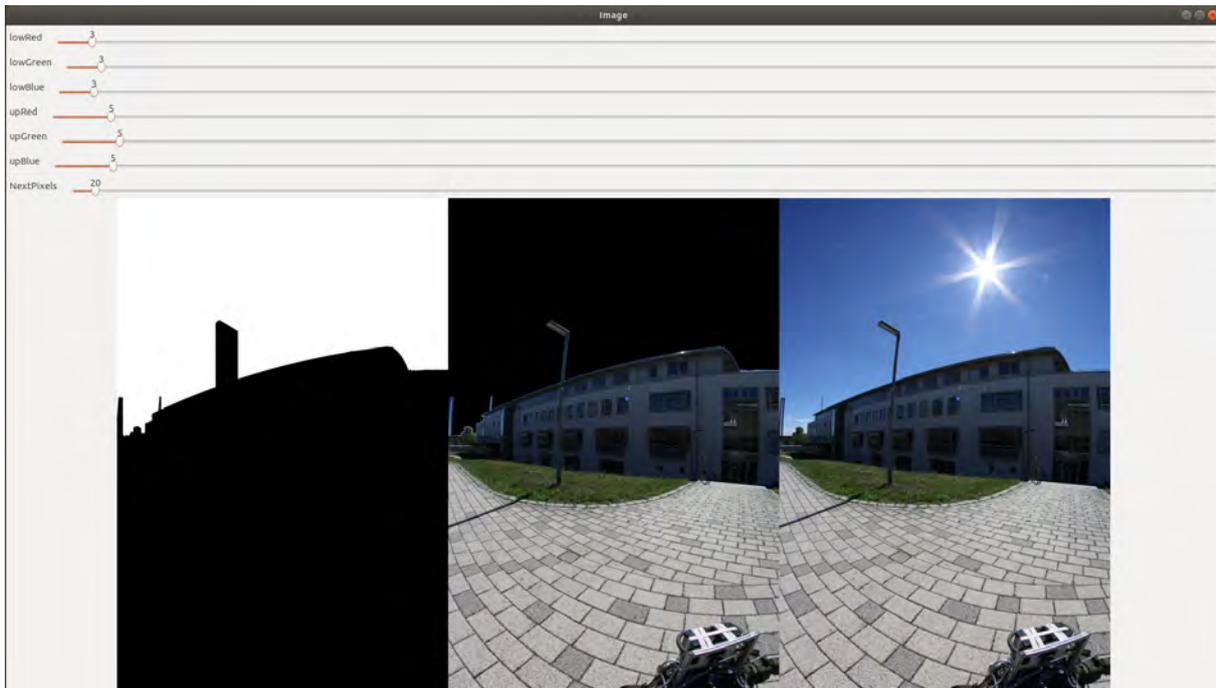


Abbildung 5.2: Horizontextraktion aus einem Foto.

nicht die Kontur der abgebildeten Objekte verwendet. Dies stellt kein Problem dar, da für den Scan genauso vorgegangen wird.

5.2 Automatische Extraktion der Keypoints

Hiermit sind nun sowohl für den Scan als auch für das Bild ein Binärbild vorhanden, welche in einem bestimmten Bereich übereinstimmen. Mithilfe von Featureextraktoren, in diesem Fall Sift (Scale-invariant feature Transform) und Surf (Speeded Up Robust Features) können Punkte mit einer auffälligen Umgebung gefunden und beschrieben werden. Die genaue Funktionsweise von Sift kann bei [13] und die von Surf bei [7] nachgelesen werden. Diese Punkte werden in den Binärbildern des Scans und des Fotos gesucht und anschließend miteinander verglichen. Sind sich die Beschreibungen dieser Punkte sehr ähnlich wird von einem Match gesprochen und die Punkte sollten das selbe Objekt oder die selbe Stelle zeigen.

5.3 Anwendung des Feature Mappers

Für das oben beschriebene Vorgehen wurde ein Programm geschrieben, welches den Nutzer durch diese Schritte führt. Es wurde innerhalb des *3DTK - The 3D Toolkit* [14] implementiert. Innerhalb des *3DTK* Stammverzeichnisses kann es folgendermaßen aufgerufen werden:

Listing 5.1: Programmaufruf

```

1 //Zeigt die Hilfe, in der alle Optionen aufgelistet werden an
2 ./bin/show scan_to_horizon -h
3
4 //Ein typischer Aufruf
5 ./bin/show scan_to_horizon -f rxp -t RIEGL -a SIFT ScanAndImageDir
   -O OutDir

```

Wurden Bilder und Scan gefunden, wird der Scan automatisch in ein Panoramabild konvertiert und der User sieht das Ergebnis der Horizontextraktion.

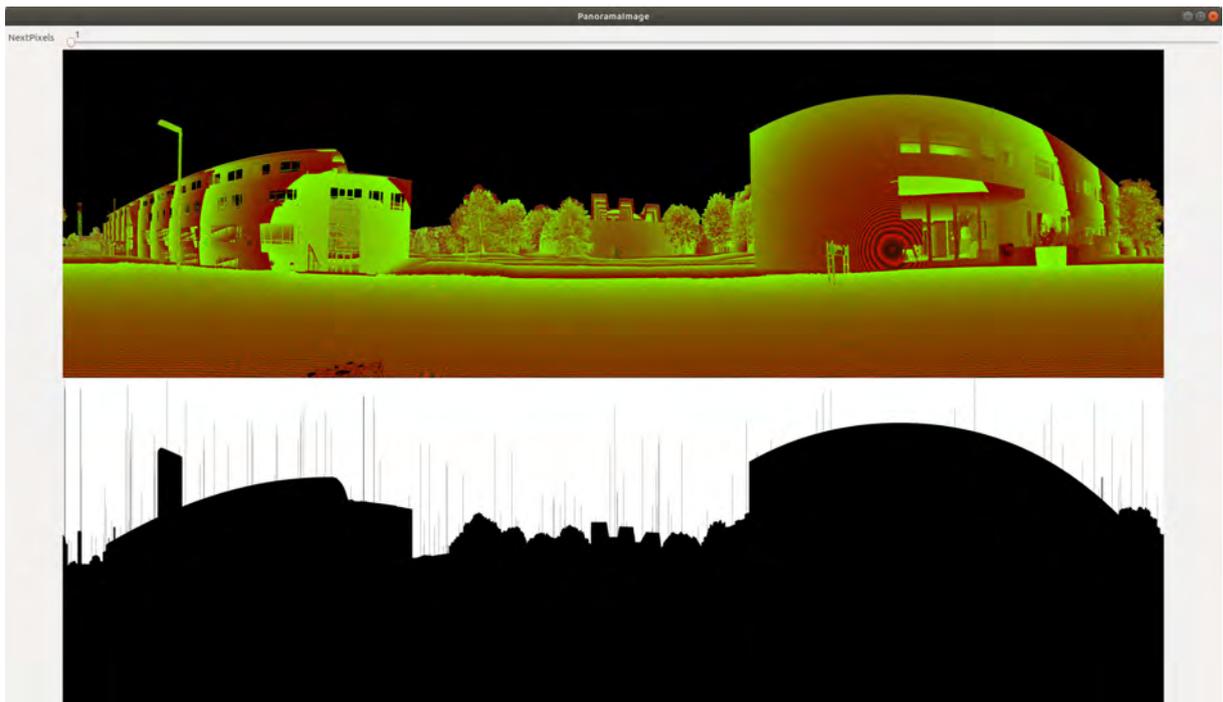


Abbildung 5.3: Binärbild des Laserscans.

Ist, wie in diesem Fall (s. Abb. 5.3), ein Rauschen vorhanden, so dass durch Messpunkte im Bereich des Himmels unerwünschte Linien entstehen, kann dies durch den Schieberegler und den Parameter *NextPixels* behoben werden. Solches Rauschen kann durch Regentropfen, Staubpartikel oder auch Vögel und Insekten hervorgerufen werden. Wird der Regler verschoben, kann durch einen beliebigen Tastendruck das Bild neu berechnet werden. In der Regel genügt hier ein Wert zwischen 8 und 15, um eine glatte Horizontlinie wie in Abb. 5.4 zu erhalten.

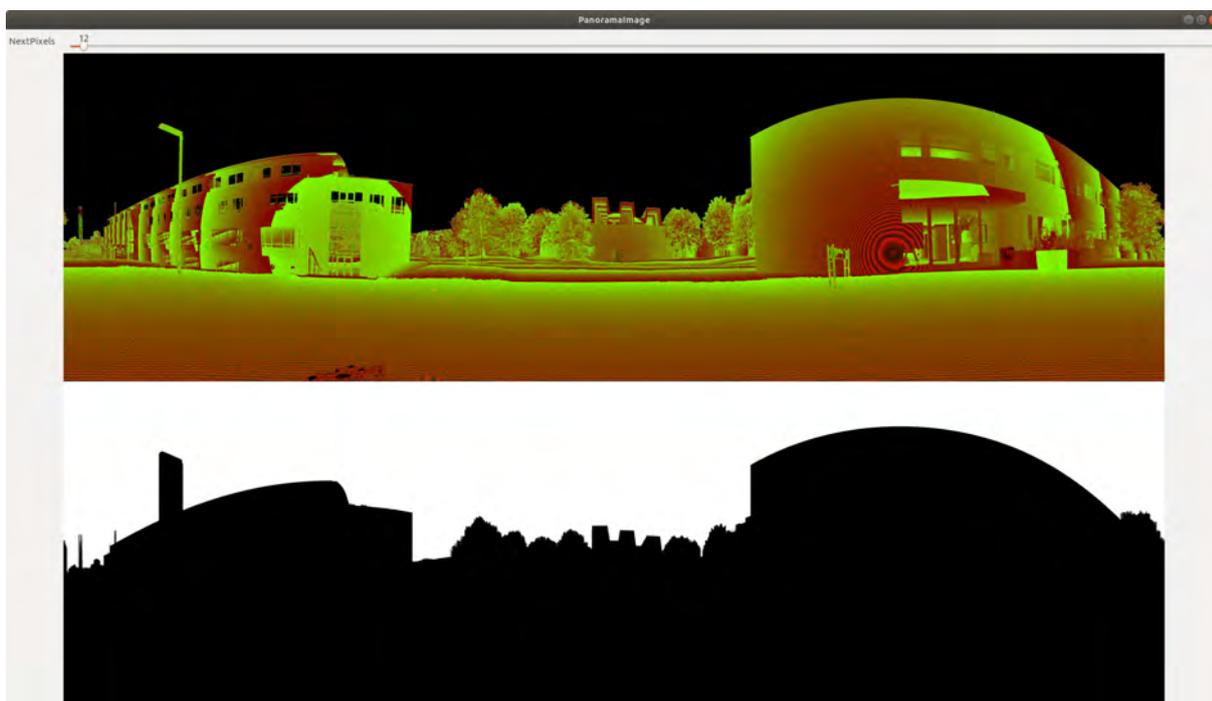


Abbildung 5.4: Binärbild des Laserscans mit Filterung.

Sobald das gewünschte Ergebnis erreicht ist, kann dieses mit der *s*-Taste bestätigt werden und das Binärbild des Fotos wird angezeigt.

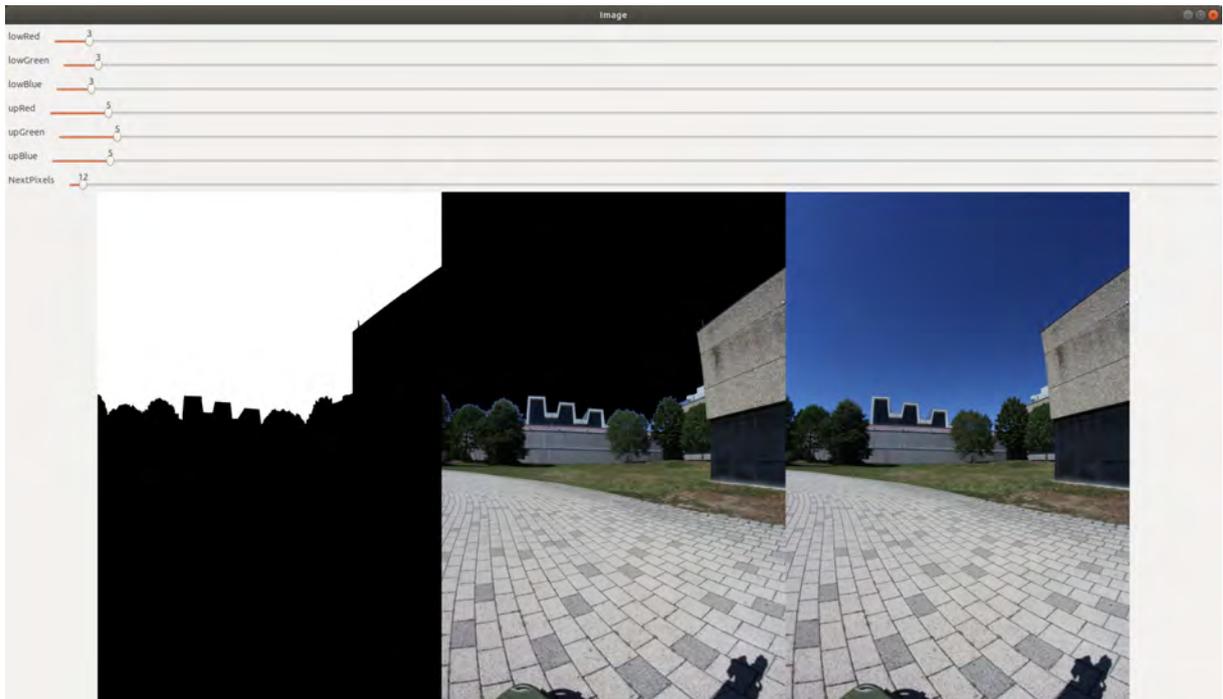


Abbildung 5.5: Binärbild des Fotos.

Anders als beim Scan wird hier ein drittes Bild eingeblendet (s. Abb. 5.5), welches den Zwischenschritt zeigt, bei dem der Himmel schwarz gefärbt ist. Dies erleichtert es, mithilfe der verschiedenen Regler (s. Abb. 5.5), die Farbgrößen so festzulegen, dass nicht mehr als der Himmel entfernt wird oder am Rand der Gebäude ein bläulicher Rand bleibt. Auch dieser Schritt lässt sich mit einem Druck auf die *s*-Taste bestätigen. Daraufhin werden die Features auf beiden Bildern berechnet und verglichen.

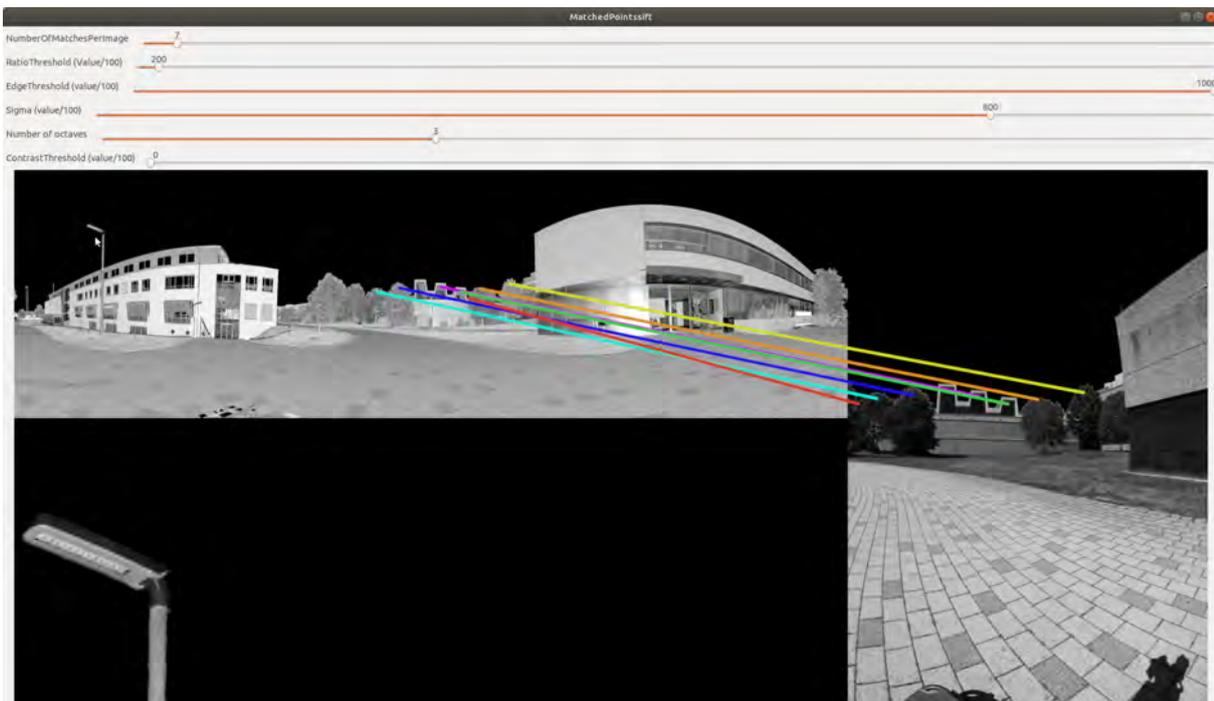


Abbildung 5.6: Matches mit Sift.

Im nächsten Schritt wird, falls vorhanden, das Reflectance-Image des Scans angezeigt. Ansonsten wird wie im ersten Schritt das Range-Image verwendet. Zusätzlich wird das Foto als Schwarz-Weiß-Bild eingeblendet und die besten Matches werden farbig verbunden. Über die Regler lassen sich die Anzahl der angezeigten Matches und die Parameter für den jeweiligen genutzten Featureextraktor anpassen und Matches werden nach einer Änderung automatisch neu berechnet.

Im Beispielbild 5.6 wurde Sift verwendet. Um zu überprüfen, ob die gefundenen Matches richtig sind, wurde eine Lupe implementiert. Diese zeigt unten links den Bereich um den Cursor stark vergrößert an. Die Features liegen, wie in Screenshot 5.6 zu sehen ist, in der Regel nicht auf der Horizontlinie sondern darüber oder darunter. Liegen sie darunter, wie in diesem Beispiel, stellt das kein Problem dar. Liegen sie darüber, also im Himmel, werden die Keypoints gerade nach unten gezogen, bis sie auf der Horizontlinie liegen. Dies ist nötig, da im Himmel kein 3D-Wert des Laserscanners vorhanden ist und diese Punkte somit keinen Messwert und dadurch keine Informationen besitzen, die für eine Kalibrierung nötig sind. Sind die gefundenen Matches zufriedenstellend, kann wiederum mit der *s*-Taste bestätigt werden. Daraufhin werden die Binärbilder, das Bild mit den eingezeichneten Matches und die gefundenen Punktpaare inklusive der 3D-Koordinaten im angegebenen Ausgabeverzeichnis gespeichert und das Programm ist beendet. Anstatt, wie oben beschrieben, mit den Binärbildern zu arbeiten, wurde zusätzlich eine Funktion implementiert, um auf einem Kantenbild die Features zu berechnen. Dies hat den Vorteil, dass das gesamte Bild unterhalb des Horizonts verwendet werden kann. Es erfordert aber einen weiteren Schritt und es erweist sich als schwierig die Kantenbilder so zu generieren,

dass sie im Scan und im Foto möglichst ähnlich sind. Wie in Screenshot 5.7 erkennbar ist, lassen sich somit die Kanten hervorheben, welche grundsätzlich im Bild und im Reflectanceimage trotz der verschiedenen Farbräume gleich sind. Mithilfe der Regler lassen sich hier die Parameter des Canny-Edge Algorithmus (mehr dazu siehe [9]) einstellen.



Abbildung 5.7: Aus Scan generiertes Kantenbild.

Anschließend werden die Features auf diesen Kantenbildern gesucht und gematched. In Screenshot 5.8 ist zu sehen, dass die Kanten sich zwischen Scan und Foto stark unterscheiden. Das liegt daran, dass bei dem Scanner die Reflektionsunterschiede zwischen den Objekten zu Kanten führen und im Foto die Farbunterschiede. So können beispielsweise im Foto auch Licht und Schatten zu Kanten führen, welche im Scan nicht vorhanden sind. Solche Differenzen ergeben schließlich auch die falschen Matches.

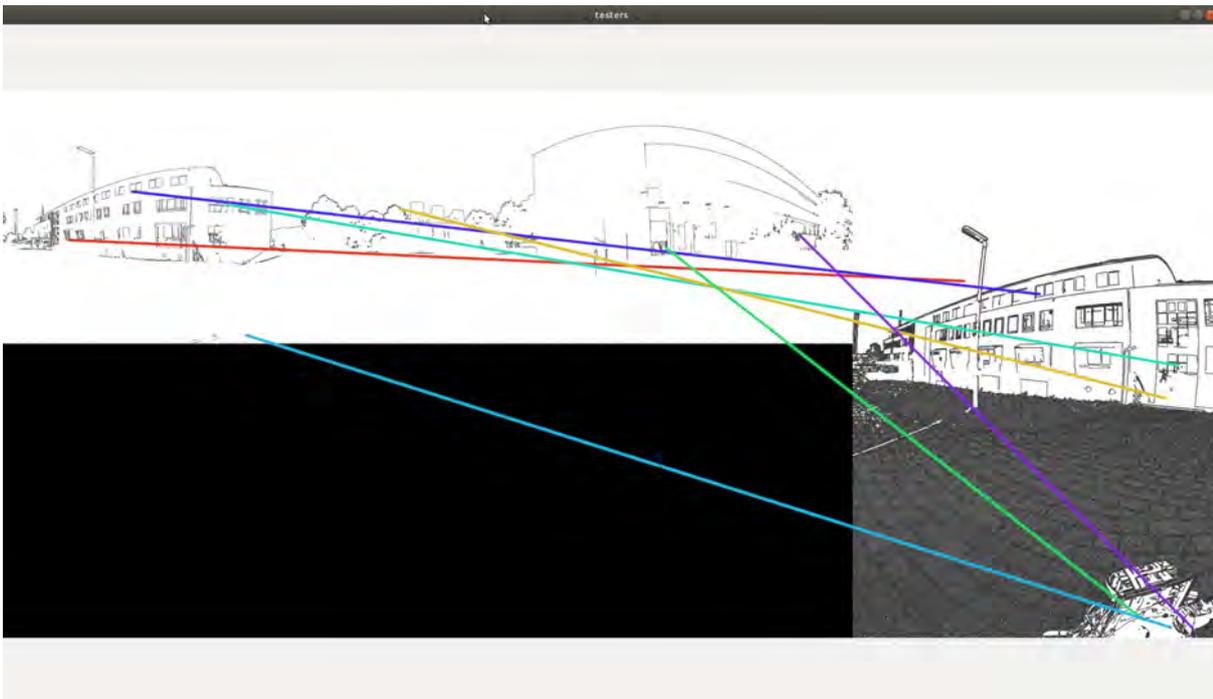


Abbildung 5.8: Mit Sift gefundene Matches auf den Kantenbildern.

5.4 Evaluation der automatisch extrahierten Korrespondenzen

Die Auswertung hat gezeigt, dass es durchaus möglich ist automatisch Korrespondenzen zu erkennen. Die Güte und Anzahl hängt aber sehr stark von der abgebildeten Umgebung und den gewählten Parametern ab. So ist es trotz der recht eindeutigen Horizontlinie in Abbildung 5.6 auch hier nicht möglich, sieben fehlerfreie Matches zu erhalten. Sobald ein fehlerhafter Match, wie hier der rote, enthalten ist, schlägt die Kalibrierung bereits fehl, beziehungsweise führt zu einem unbrauchbaren Ergebnis. Daher ist es auf jeden Fall nötig, zuvor eine Outlier-Detection manuell oder besser noch automatisch durchzuführen. Ein großes Problem ergibt sich auch dadurch, dass die Feature-Detektoren aufgrund der vielen Kanten und Ecken oft Features in Bäumen und anderer Vegetation finden. Diese sind in der Regel auf Grund dessen, dass die Fotos und der Scan nicht gleichzeitig entstehen und sich somit durch Wind die Blätter bewegen können, nicht ideal. Außerdem führen sie oft zu falschen Matches, da die Blätter zu sehr ähnlichen Umrissen führen. Bei Verwendung der Kantenbilder konnte leider kein zufriedenstellendes Ergebnis erreicht werden. Die Kanten unterscheiden sich einfach zu stark zwischen Scan und Bild, so dass es nicht möglich war Kantenbilder zu erzeugen, die sich ähnlich genug sind, um gute Matches zu erreichen. Es wäre aber auch möglich, dass dies bei einer anderen Szene wie beispielsweise im Gebäudeinneren besser funktionieren könnte. Es gibt auch weitere Ursachen, welche das Matching deutlich erschweren. So wurde auch hier mit dem Walimex Fisheye Objektiv gearbeitet. Wie in Abbildung 5.6 ersichtlich ist, führt dies durch die radiale Verzeichnung dazu, dass im Foto

die Hauswand rechts im Bild deutlich gekrümmt ist. Im Panoramabild des Scans ist dies nicht der Fall. Zusätzlich ist im Panoramabild ebenso eine starke Verzeichnung vorhanden, welche keine geraden Linien erhält. So sind dort Linien wie die oberen Kanten der Gebäude stark gekrümmt. Diese Faktoren führen dazu, dass das Matching deutlich erschwert wird.

Kapitel 6

Fazit und Ausblick

Die verschiedenen Kameramodelle wurden sowohl mit gewöhnlichen als auch mit Fischaugenobjektiven getestet. Es hat sich gezeigt, dass der Reprojection Error bei der Verwendung des Distortionmodells für alle getesteten Systeme mit dem perspektivischen Modell, welches beispielsweise auch in *OpenCV* als Standard hinterlegt ist, am kleinsten ist.

Desweiteren wurde festgestellt, dass das perspektivische Modell wohl am robustesten ist, da die Varianz und Standardabweichung für die verschiedenen Splits der Datensätze mit diesem am geringsten sind. Dennoch haben die anderen Modelle aufgrund ihrer speziellen Eigenschaften ihre Daseinsberechtigung. So wurde beispielsweise ermittelt, dass das orthogonale und das stereographische Modell die Randbereiche deutlich weniger in die Länge ziehen als das perspektivische und das equidistante Modell.

Die Versuche haben auch belegt, dass die verschiedenen Modelle für diverse Kameras zu unterschiedlich guten Ergebnissen führen. So liefert zum Beispiel das perspektivische Modell mit Distortionparametern (s. Tabellen 3.1 und 3.3) den besten Reprojection Error für das Walimex Fischaugenobjektiv, das equidistante hingegen liefert den besten Wert für die Kamera des Xiaomi Mi 9 SE Smartphones.

Es hat sich auch herausgestellt, dass das Distortionmodell eine große Rolle spielt. Ohne Distortionparameter liefert das perspektivische Modell nur noch für die Standardobjektive die besten Ergebnisse. Für die Fischaugenoptiken sind jeweils andere Projektionsmodelle vorzuziehen. Daraus lässt sich schließen, dass das Distortionmodell wohl für diese Modelle entsprechend angepasst werden muss. Dadurch ließe sich mit großer Wahrscheinlichkeit auch ein besseres Ergebnis erzielen als mit dem perspektivischen Modell.

Bei der Kalibrierung mithilfe eines Laserscans hat sich gezeigt, dass das perspektivische Modell den geringsten Reprojection Error aufweist. Doch auch hier konnte aufgezeigt werden, dass das equidistante Modell trotz des größeren Fehlers bei Verwendung eines Initial Guess eine etwas geringere Verschiebung bei der Färbung aufweist und eine genauere Extrinsik erzeugt. Kein anderes Modell als das perspektivische lieferte brauchbare Ergebnisse, wenn kein Initial Guess verwendet wurde. Dies spricht wiederum für die Robustheit dieses Modells und bezeugt,

dass dieses offenbar mit weniger Korrespondenzen und größeren Fehlern umgehen kann. Es wurde auch festgestellt, dass es sehr schwierig ist, eine gute Kalibrierung in Verbindung mit einem Laserscanner zu erhalten, wenn dafür keine speziellen Objekte mit eindeutigen Punkten innerhalb der aufgenommenen Szene platziert werden.

Bei der Kalibrierung mit Punkten auf dem Horizont stellte sich heraus, dass keine befriedigenden Ergebnisse erreicht werden konnten. Dies liegt voraussichtlich an der niedrigen Anzahl, der geringeren Genauigkeit der weit entfernten Korrespondenzen und an der mangelhaften Sensorabdeckung, da sich diese alle auf einer ähnlichen Höhe befinden und somit nur einen schmalen Streifen (s. Abb. 4.14) des Sensors abdecken. Insbesondere der obere und untere Bereich werden durch die fehlenden Punkte im Himmel und am Boden nie erfasst.

Es konnte auch festgestellt werden, dass die automatische Extraktion des Himmels sowohl aus dem 3D-Scan als auch aus den Fotos sehr gut funktioniert. Auch die automatische Extraktion von Korrespondenzen mittels Featureextraktor wie beispielsweise Sift ist grundsätzlich möglich. Hier ist es aber nötig, eine gute Outlier Detection zu implementieren oder zumindest das Programm so zu erweitern, dass manuell falsche Matches entfernt werden können. So lassen sich damit schnell Korrespondenzen auf der Horizontlinie automatisch bestimmen. Wie in dieser Arbeit herausgefunden wurde, eignen sich diese zwar nicht für eine vollständige Kalibrierung, könnten aber ausreichen, um die Position und Ausrichtung, also die Pose der Kamera, mit ausreichender Genauigkeit festzustellen. Hierfür gibt es beispielsweise in *OpenCV* bereits die Funktion *solvePnP*, welche mithilfe einer vorherigen Kalibrierung und mindestens vier Punktkorrespondenzen die Pose bestimmt. Damit könnte anschließend ohne erneute Kalibrierung eine Punktwolke eingefärbt werden.

Literaturverzeichnis

- [1] *Chromatische Aberration*. https://www.univie.ac.at/mikroskopie/1_grundlagen/optik/opt_linsen/5c_chromatisch.htm, . – [Online; accessed 21. Apr. 2021]
- [2] *Camera Calibration — OpenCV-Python Tutorials 1 documentation*. https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_calib3d/py_calibration/py_calibration.html. Version: Oct 2019. – [Online; accessed 21. Apr. 2021]
- [3] *Definition von Fisheyes*. <http://www.fisheyelens.de/html/def.html>. Version: Aug 2019. – [Online; accessed 21. Apr. 2021]
- [4] *GoPro Hero4 Silver Datenblatt*. https://www.digitaleyes.de/Produkte/ActionCams/GoPro_Hero4_Silver/8796781488002. Version: Feb 2021. – [Online; accessed 21. Apr. 2021]
- [5] *Mi Deutschland*. <https://www.mi.com/de/mi-9-se>. Version: Feb 2021. – [Online; accessed 21. Apr. 2021]
- [6] ABRAHAM, Steffen; FOERSTNER, Wolfgang: Fish-eye-stereo calibration and epipolar rectification. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 59 (2005), Nr. 5, 278-288. <http://dx.doi.org/https://doi.org/10.1016/j.isprsjprs.2005.03.001>. – DOI <https://doi.org/10.1016/j.isprsjprs.2005.03.001>. – ISSN 0924-2716
- [7] BAY, Herbert; TUYTELAARS, Tinne; VAN GOOL, Luc: SURF: Speeded Up Robust Features. In: LEONARDIS, Aleš (Hrsg.); BISCHOF, Horst (Hrsg.); PINZ, Axel (Hrsg.): *Computer Vision – ECCV 2006*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2006. – ISBN 978-3-540-33833-8
- [8] BORRMANN, Dorit: *Multi-modal 3D mapping - Combining 3D point clouds with thermal and color information*, Universität Würzburg, doctoralthesis, 2018. <http://dx.doi.org/10.25972/OPUS-15708>. – DOI 10.25972/OPUS-15708

- [9] CANNY, J.: A Computational Approach to Edge Detection. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8 (1986), Nr. 6, S. 679–698. <http://dx.doi.org/10.1109/TPAMI.1986.4767851>. – DOI 10.1109/TPAMI.1986.4767851
- [10] FOERSTNER, Wolfgang; WROBEL, Bernhard P.: *Photogrammetric Computer Vision: Statistics, Geometry, Orientation and Reconstruction*. 1st. Springer Publishing Company, Incorporated, 2016. – ISBN 3319115499
- [11] GAO, Wenliang; WANG, Kaixuan; DING, Wenchao; GAO, Fei; QIN, Tong; SHEN, Shaojie: Autonomous aerial robot using dual-fisheye cameras. In: *Journal of Field Robotics* 37 (2020), 02. <http://dx.doi.org/10.1002/rob.21946>. – DOI 10.1002/rob.21946
- [12] KRAUSE, Michael: *Walimex pro 8/3,5 Fisheye II APS-C Canon EF-S*. <https://www.walimex-online.de/objektive/foto-objektiv/foto-objektiv-8mm/walimex-pro-8/3-5-fisheye-ii-aps-c-canon-ef-s>. Version: Feb 2021. – [Online; accessed 21. Apr. 2021]
- [13] LOWE, D. G.: Object recognition from local scale-invariant features. In: *Proceedings of the Seventh IEEE International Conference on Computer Vision* Bd. 2, 1999
- [14] NUECHTER, Andreas; BORRMANN, Dorit; SCHAUER, Johannes: *3DTK - The 3D Toolkit*. <http://threedtk.de>. Version: Jan 2018. – [Online; accessed 21. Apr. 2021]
- [15] ZHANG, Z.: A flexible new technique for camera calibration. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (2000), Nr. 11, S. 1330–1334. <http://dx.doi.org/10.1109/34.888718>. – DOI 10.1109/34.888718

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 2021