Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

DIPLOMARBEIT

Vervollständigung von Stadtmodellen mittels Distanzkameras

Ausgeführt am
Institut für Photogrammetrie und Fernerkundung
der Technischen Universität Wien

unter der Anleitung von

Univ.Prof. Dipl.-Ing. Dr.techn. Norbert Pfeifer

und

Dipl.-Ing. Wilfried Karel

als verantwortlich mitwirkendem Assistenten

durch

Stefan Niedermayr
Matr.Nr.: 0325207
Adresse: Schumanngasse 70/8, 1170 Wien
E-mail: s.niedermayr@atn.nu

Wien, Februar 2010 .....

Danksagung

Im Folgenden möchte ich mich bei all jenen bedanken, die mich im Laufe des Studiums sowie beim Verfassen der Diplomarbeit begleitet und unterstützt haben.

Besonders bedanke ich mich bei Herrn Dipl.-Ing. Wilfried Karel für die hervorragende Betreuung und fachlich kompetente Unterstützung sowie für die zahlreichen Kommentare, die besonders in der Endphase, einen wesentlichen Beitrag zum Gelingen der Diplomarbeit geleistet haben.
Ich bedanke mich auch bei Herrn Univ.-Prof. Dr. Norbert Pfeifer für die Möglichkeit, ein sehr interessantes und spannendes Thema in meiner Diplomarbeit bearbeiten zu können sowie für die kompetente Begleitung vom Beginn bis zum Abschluss der Arbeit.
Außerdem danke ich der Firma Amisola Immobilien AG, Wien (DI (FH) Kössler) für die Erlaubnis, in der Fußgängerpassage im Palais Ferstel Aufnahmen für meine Diplomarbeit durchführen zu können.

Dir Johanna, danke ich für die Motivation und die außerfachliche Unterstützung während der letzten Studienjahre.
Meinen Eltern danke ich dafür, dass sie mir das Studium ermöglicht haben und mich in allen Belangen unterstützt und gefördert haben.
Weiters bedanke ich mich bei Stefan und Johanna für die Korrekturarbeiten in der intensiven Abschlussphase.
Nicht zuletzt bedanke ich mich besonders bei Nicole, Irene, Christine, Ulli, Stefan, Martin und Wolfgang, die mir unter anderem durch unsere „Nelson´s-Vorlesungen“ den Studienalltag erheitert haben.

Zusammenfassung

In letzter Zeit steigt der Bedarf an immer detaillierteren 3D-Stadtmodellen. Typischerweise werden zur Erfassung der Daten, aus denen in weiterer Folge ein Modell abgeleitet wird, flugzeuggetragenes Laserscanning und Luftbildphotogrammetrie eingesetzt. Zur Erstellung von hoch detaillierten Stadtmodellen sind zusätzlich auch terrestrische Methoden notwendig.
Beim sogenannten 'Range Imaging' (RIM) kommen Kameras zum Einsatz, deren Sensoren für alle Pixel gleichzeitig die Laufzeit eines emittierten Signals bestimmen, woraus die Distanzen zum Objekt unmittelbar bestimmbar sind. Gleichzeitig zur Distanz wird die Amplitude des reflektierten Signals aufgezeichnet, sodass neben dem Distanzbild ein Amplitudenbild erzeugt werden kann, wobei mehrere Bilder pro Sekunde aufgenommen werden können. Spezielle Kameras, welche auf Photomischdetektoren (PMD) basieren, werden mittlerweile kostengünstig in Serie produziert. PMD-Kameras weisen maximale Aufnahmedistanzen von etwa zehn Metern auf, weshalb sich diese potentiell für terrestrische Detailaufnahmen zur Erstellung von Stadtmodellen eignen.
In der vorliegenden Arbeit wird ein Verfahren vorgestellt, mithilfe dessen es möglich ist[1], aus den mit einer Distanzkamera gewonnenen Daten ein Objekt in der Größenordnung von mehreren Metern aufzuzeichnen, und daraus ein Modell abzuleiten. Dabei treten für Distanzkameras extrem große Distanzen auf, wodurch insbesondere die Distanzbilder ein starkes Rauschen aufweisen. Aufgrund dessen wird die Kamera während der Aufnahmen auf ein Stativ montiert, und eine Sequenz von Bildern mit gleicher Orientierung gemittelt (statische Bilder). Die Orientierung der Aufnahmen erfolgt in zwei Stufen. Zur näherungsweisen Orientierung werden die Amplitudenbilder und die Distanzbilder, die während der Bewegung der Kamera zwischen den statischen Bildern aufgezeichnet wurden (bewegte Bilder), verwendet. In den Amplitudenbildern, die eine höhere Präzision als die Distanzbilder aufweisen, werden mittels des Speeded-Up Robust Features(SURF)-Algorithmus Verknüpfungspunkte im Bildraum ermittelt. Die Distanzbilder werden verwendet, um diese Verknüpfungspunkte in den Objektraum zu übertragen und die Orientierung eines Bildes mittels eines räumlichen Rückwärtsschnittes aus jener des vorangehenden Bildes zu berechnen. Die Fein-Orientierung erfolgt über 3D-Punktwolken, die aus den statischen Distanzbildern durch eine Variante des Iterative Closest Point (ICP) - Algorithmus berechnet werden. Als Modell und Ergebnis dient eine Triangulation der fein-orientierten, vereinigten Punktwolken.

[1] Nach den vorliegenden Informationen wurde zum ersten Mal ein Objekt mit Ausmaßen von mehreren Metern mit einer Distanzkamera aufgenommen und daraus ein Modell abgeleitet.

Zur beispielhaften Anwendung dieses Verfahrens wird ein Teil einer Einkaufspassage aufgezeichnet und ausgewertet. Das Ergebnis nach der Modellierung ist optisch ansprechend, weist jedoch einen erheblichen Unterschied in der Skalierung zwischen Objekt und Modell auf. Zudem sind lokale Deformationen im Modell vorhanden.

Abstract

Nowadays, the need for detailed 3D city-models is ever-increasing. Usually, for the acquisition of data that these models are based on, airborne laser scanning and aerial photogrammetry are used. For the creation of highly detailed city-models, terrestrial methods are needed to complement these aerial techniques.
So-called „range imaging“ (RIM), uses special cameras, whose sensors simultaneously determine the time-of-flight of an emitted signal for all pixels. From that time-of-flight, the distance to the object can be derived directly. At the same time, the amplitude of the reflected signal is recorded, both with a rate of several frames per second. Especially cameras which are based on photonic mixer devices (PMD) are nowadays mass-produced in a cost effective way. PMD cameras can measure distances up to about ten meters. For that reason, these cameras are potentially suitable for the terrestrial capture of data in order to complement city-models.
This thesis presents, a method to survey and model an object with extents of some meters using range cameras[2]. For current range cameras, the distances to be observed in this measurement setup are very large. For this reason, especially the pictures containing the distance information show a significant amount of noise. Thus, the range camera is mounted on a tripod during the survey, and a sequence of pictures with the same orientation (static pictures) is averaged. The orientation of the pictures is carried out in two steps. For the approximate orientation the pictures containing the amplitude information and the distance information, which were collected between static pictures (moved pictures), are used. In the pictures containing the amplitude information, which show a higher precision than the pictures containing the distance information, tie-points are determined in image space using the Speeded-Up Robust Features(SURF)-algorithm. The images containing the distance information are used to transform these tie-points to object space. Using spatial resections, image orientations are then computed one by one, based on the point correspondences with the previously oriented image. The fine-orientation is

[2]According to the author´s best knowledge, this is the first time, that an object with a dimension of several meters recorded and modelled using a range camera.

carried out using 3D-pointclouds, which are derived from static images using a variant of the Iterative Closest Point (ICP)-algorithm. The triangulation of the fine-oriented unified point-clouds serves as model and result. To test the presented approach, a part of a shopping arcade is surveyed, and the resulting model is analysed. The model is optically appealing, but features significant absolute differences from the object. Furthermore, local deformations are present.

Inhaltsverzeichnis

1 Einleitung **8**

1.1 Stadtmodelle 8

1.2 Zielsetzung 10

1.3 Aufbau der Arbeit 10

2 Distanzkameras **12**

2.1 Messverfahren 13

2.1.1 PMD - Photonic Mixer Device
Photomischdetektor 14

2.1.2 Heterodyne Mischung 18

2.1.3 SPAD - single-photon avalanche diodes
Einzelphotonendetektion mittels Lawinenfotodiode 18

2.1.4 MDSI - multiple double short time integration
Mehrfache Kurzzeit-Doppelbelichtung 19

2.2 Vergleich der Distanzkamera zur klassischen Photogrammetrie und zum Laserscanning 20

2.3 PMD-Daten 22

2.4 Kalibrierung 25

2.5 Anwendungen 26

2.5.1 Anwendung im Vermessungswesen 27

3 Methodik **28**

3.1 Vorgeschlagener Arbeitsablauf 28

3.2 Datenaufbereitung 30

3.2.1 Segmentierung in statische und bewegte Sequenzen 31

3.2.2 Amplitudennormalisierung 32

3.3 Orientierung 34
3.3.1 Definition eines globalen Koordinatensystems 34
3.3.2 Näherungsweise Orientierung 35
3.3.2.1 Verknüpfungspunkte 35
3.3.2.2 Ermittlung der Objektkoordinaten 41
3.3.2.3 Berechnung der äußeren Orientierung mittels Rückwärtsschnitt 43
3.3.3 Fein-Orientierung 46
3.3.3.1 Normalvektorenberechnung und Filterung 46
3.3.3.2 ICP-Algorithmus 48
3.4 Modellierung 54
3.4.1 Geomagic Studio 54
3.4.2 Bearbeitungsschritte vor der Triangulation 55
3.4.3 Triangulation 56

4 Aufnahme und Auswertung der Passage im Palais Ferstel 57
4.1 Palais Ferstel 57
4.2 Aufnahme der Passage 59
4.2.1 Verwendete Geräte 60
4.2.2 Planung der Aufnahme 62
4.2.3 Durchführung der Aufnahme 63
4.3 Auswertung 65
4.3.1 Datenvorprozessierung 65
4.3.2 Bestimmung der näherungsweisen Orientierung 67
4.3.3 Fein-Orientierung 76
4.3.4 Erstellung des Modells 86
4.4 Analyse der Modellqualität 90
4.4.1 Abstand Punktwolke zur Modelloberfläche 90
4.4.2 Äußere Genauigkeit 90
4.5 Diskussion und Ausblick 92

A Anhang 95
A.1 Differentiale der Kollinearitätsgleichung 95
A.2 Differentiale der Ähnlichkeitstransformation 96

Abbildungsverzeichnis 98

Literaturverzeichnis **99**

Lebenslauf **106**

Kapitel 1

Einleitung

1.1 Stadtmodelle

Unter einem 3D-Stadtmodell wird im Allgemeinen das digitale 3D-Abbild von Geländeoberflächen und den Gebäuden sowie auch Verkehrswege, Wasserkörper, Vegetation und Stadtmöblierung eines urbanen Gebietes verstanden [Kada, 2007].
Typische Anwendungsfelder von 3D-Stadtmodellen sind im Bereich der Raum- und Stadtplanung zu finden. Dort werden Stadtmodelle beispielsweise zur Visualisierung von Bauvorhaben und zur Erhaltung von als wichtig empfundenen Sichtachsen (z.B. Wien: Schloss Schönbrunn - Leopoldsberg, Karlsplatz-Stephansdom) verwendet [Stadtentwicklung Wien]. Eine wesentliche Rolle spielen dreidimensionale Modelle bei der Berechnung zur Ausleuchtung von Funkräumen, für die Planung von Antennenstandorten für Mobilfunknetze und zur Berechnung der Lärmausbreitung. Weitere Aufgabenstellungen, bei denen der Einsatz von 3D-Stadtmodellen sinnvoll ist, wären die Optimierung von Rettungseinsätzen und Berechnungen von Abschattungen zur optimalen Positionierung von Solaranlagen[1]. Stadtmodelle werden zunehmend im touristischen Bereich oder für realitätsnahe Computerspiele eingesetzt. [Koppers, 2002; Kada, 2007; Jansa u. Stanek, 2003].
Eine der Massenanwendungen von 3D-Modellen wird bei künftigen Navigationsgeräten zu finden sein. Zum Beispiel sind in der aktuellen Version der Fußgängernavigationssoftware *City Explorer 3* von Nokia 3D-Ansichten wichtiger Gebäude integriert. Gerade für die Fußgänger-

[1] Laut einer EU-Richtlinie müssen Gebäude ab 2019 ihre Energie selbst produzieren. Dazu werden unter anderem auch Solaranlagen im städtischen Bereich zum Einsatz kommen.

navigation sind detaillierte Stadtmodelle notwendig, in welchen zum Beispiel auch Passagen als wichtige Fußwege modelliert sind.
Für die unterschiedlichen Anwendungen werden Stadtmodelle mit entsprechendem Detaillierungsgrad benötigt. In weiterer Folge wird nun in erster Linie auf die Modellierung von Gebäuden eingegangen. Gebäudemodelle können in fünf *Level of Detail* (LoD) eingeteilt werden [Gröger et al., 2005]:

- LoD 0: Regionalmodell
- LoD 1: Klötzchenmodell ohne Dachstrukturen
- LoD 2: Strukturmodell mit differenzierten Dachstrukturen
- LoD 3: Architekturmodell mit fein differenzierten Dachstrukturen und Fassaden
- LoD 4: Innenraummodell

Für LoD 0, LoD 1 und großteils auch für LoD 2 sind luftgestützte Aufnahmeverfahren ausreichend, um daraus die entsprechenden Modelle ableiten zu können. Für ein Architekturmodell (LoD 3) werden in jedem Fall auch terrestrische Aufnahmen benötigt.
Bei der Wahl der Erfassungsmethode ist auch auf eine wirtschaftliche Auswertung zu achten. Um einen hohen Automatisierungsgrad bei LoD 1 und LoD 2 zu erreichen, kommen auch Kombinationen von photogrammetrischen Methoden mit Laserscannerdaten oder eine Kombination von Grundriss (GIS-Daten) und Laserscannerdaten zum Einsatz [Jansa u. Stanek, 2003]. Einige Methoden zur Gebäuderekonstruktion sind in Koppers [2002] zusammengefasst. In Dorninger u. Pfeifer [2008] werden LoD 2 Modelle aus Airborne-Laser-Scanner(ALS)-Daten abgeleitet.
Für hoch-detaillierte Stadtmodelle (LoD 3) werden, wie bereits erwähnt, auch terrestrische Aufnahmen benötigt. Dabei werden in erster Linie terrestrische Kameras und terrestrische Laserscanner eingesetzt. In Forkert et al. [2005] wird ein System vorgestellt, welches diese terrestrischen Aufnahmen auf eine preiswerte Art ermöglicht. Der sogenannte *CityGRID Scanner* ist ein fahrzeuggestütztes System bei dem ein 3D-Laserscanner, zwei hochauflösende Digitalkameras und GPS zum Einsatz kommen. Es können jedoch nur Bereiche aufgenommen werden, welche mit dem Fahrzeug erreichbar sind. Für ein detailliertes Stadtmodell ist es aber wünschenswert, dass auch Bereiche aufgenommen werden, die nicht direkt von der Straße aus einsehbar sind. Dazu können terrestrische Laserscanner verwendet werden. Für die Aufnahme, zum Beispiel bei der Aufzeichnung eines Arkadenhofs, sind jedoch viele Standpunkte notwendig. Zudem sind die relativ hohen Anschaffungskosten und das hohe Gewicht von terrestrischen Laser-

scannern zu beachten. Bei photogrammetrischen Aufnahmen ist die Auswertung von LoD 3 Modellen nicht ausreichend automatisierbar.

1.2 Zielsetzung

In den letzten Jahren wurde eine neue Technologie entwickelt, die potentiell für terrestrische Aufnahmen eingesetzt werden kann. Dazu werden sogenannte Distanzkameras verwendet. Dies wird als 'Range Imaging' (RIM) bezeichnet und ist eine relativ junge Technologie, mit der es möglich ist, für alle Pixel eines Bildes die Distanz zum Objekt simultan zu messen. Somit kann aus jedem Distanzbild unmittelbar eine 3D-Punktwolke abgeleitet werden.
Ziel dieser Arbeit ist es, die Verwendungsmöglichkeit von RIM für die Vervollständigung von LoD 3–Gebäudemodellen als Bestandteil von Stadtmodellen anhand einer beispielhaften Einkaufspassage zu prüfen. Das entwickelte Verfahren sollte einen möglichst hohen Automatisierungsgrad aufweisen. Die Orientierung der Bilder sollte ohne zusätzliche Information (z.B. Grundrissplan, GPS, IMU, ...) erfolgen. Deswegen sind vor der abschließenden Fein-Orientierung Näherungswerte für die Orientierung aus den Bildern selbst zu bestimmen.

1.3 Aufbau der Arbeit

Die Arbeit ist in zwei Teile gegliedert. Im ersten Teil werden die Methoden vorgestellt, im zweiten Teil wird die Anwendung des entwickelten Verfahrens auf die Passage im Palais Ferstel samt Analyse der Ergebnisse beschrieben.

1. Teil:
In dieser Arbeit wird RIM verwendet. Diese junge Technik ist noch relativ unbekannt und wird daher in Kapitel 2 vorgestellt. RIM erzeugt Bilder, die im Zuge der Auswertung orientiert werden müssen. Zur Fein-Orientierung (Abschnitt 4.3.3) wird ein Algorithmus verwendet, der näherungsweise orientierte Daten benötigt. Da das entwickelte Verfahren ohne zusätzliche Sensoren auskommen soll, muss diese genäherte Orientierung aus den RIM-Daten selbst gewonnen werden, was in Abschnitt 3.3.2 beschrieben wird. Dazu sind für die aufgenommenen RIM-Daten einige Bearbeitungsschritte im Vorfeld notwendig, auf diese wird in Abschnitt 3.2

eingegangen. Ergebnis der Fein-Orientierung sind orientierte Punktwolken, welche in Abschnitt 3.4 zu einem triangulierten Oberflächenmodell vereinigt werden.

2. Teil:

Der zweite Teil beinhaltet eine beispielhafte Anwendung des entwickelten Verfahrens (Abschnitt 4). Dazu wird die Passage im Palais Ferstel (Abschnitt 4.1) mit einer Distanzkamera aufgenommen (Abschnitt 4.2). Die Auswertung dieser Aufnahme sowie eine Analyse der Zwischenergebnisse wird in Kapitel 4.3 beschrieben. Zum Schluss erfolgt eine Analyse der Modellqualität (Abschnitt 4.4) und eine Diskussion über das entwickelte Verfahren sowie ein Ausblick.

Kapitel 2

Distanzkameras

In diesem Kapitel werden unterschiedliche Kameratypen und deren Funktionsweisen erklärt, ein Vergleich zur klassischen Photogrammetrie und zum Laserscanning vorgenommen, sowie mögliche Anwendungen von Distanzkameras erläutert.
Wie eingangs schon erwähnt, wird bei den sogenannten RIM Kameras die Distanz zum Objekt für alle Pixel eines Bildes simultan gemessen. Abhängig vom verwendeten Messverfahren ist dies sogar mit Videobildraten und höher möglich. Zusätzlich zur Distanz wird meist die Amplitude des reflektierten Signals aufgezeichnet, sodass neben dem Distanzbild ein Amplitudenbild erzeugt werden kann (siehe Abbildung 2.1).

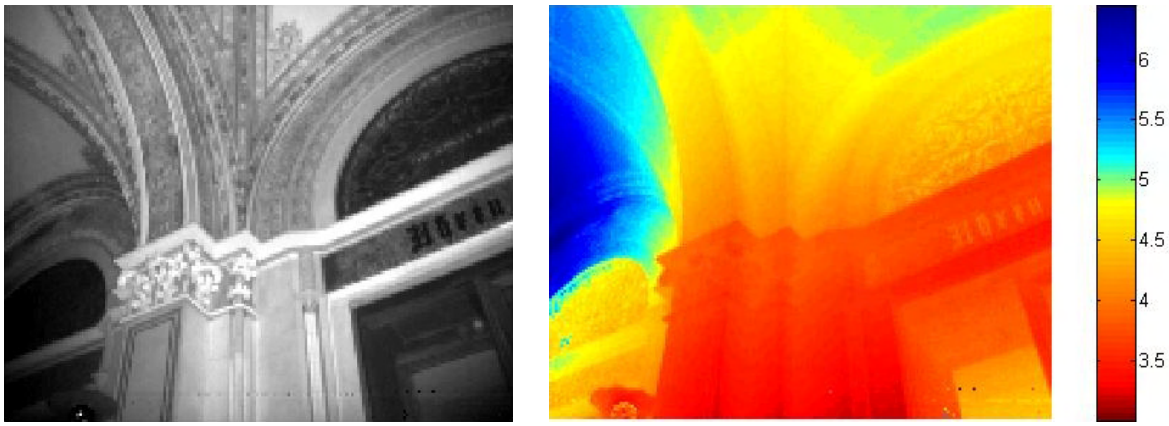


Abbildung 2.1: Amplitudenbild (links), Distanzbild (rechts): Einheit der Skala [m]

2.1 Messverfahren

Distanzkameras zählen zu den aktiven Sensoren. Zwei wesentliche Komponenten sind bei allen RIM-Typen vorhanden (siehe Abbildung 2.2). Dies ist zum einen die Beleuchtungseinheit, die ein typischerweise im nahen Infrarotbereich liegendes Signal aussendet. Dieses wird am Objekt reflektiert und vom Objektiv auf die zweite Komponente, den Bildsensor, projiziert. Dieser demoduliert für alle Pixel das Signal und bestimmt aus diesem die Laufzeit.

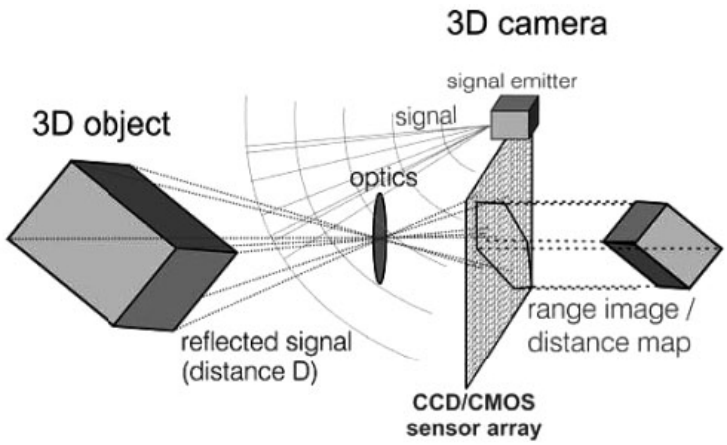


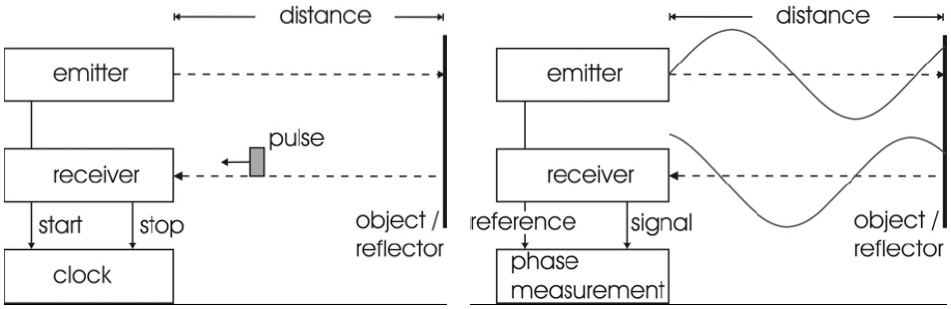
Abbildung 2.2: Grundprinzip von allen RIM-Typen: die Beleuchtungseinheit sendet ein Signal aus, dieses wird am Objekt reflektiert und vom Objektiv auf den Bildsensor projiziert. Dieser demoduliert für alle Pixel das Signal und bestimmt aus diesem die Laufzeit. Quelle: [Kahlmann u. Ingensand, 2007b]

Der Signalemitter besteht entweder aus einem Leuchtdioden(LED)-Feld oder einem Laser. Die ausgesandte Trägerwelle besitzt eine Wellenlänge im Bereich des langwelligen sichtbaren Lichtes bis nahes Infrarot. Die Distanzmessverfahren können nach Art des Messsignals eingeteilt werden in:

- *Impulsverfahren:* Bei diesem Verfahren wird das Signal in Form eines Pulses ausgesandt. Eine Messung beginnt mit dem Aussenden des Signals und endet mit der Detektion des Echos (siehe Abbildung 2.3(a)). Dieser Vorgang wird in der Regel mehrfach wiederholt und die Messungen gemittelt [Karel, 2008].

- *Phasenvergleichsverfahren:* Hier wird das amplitudenmodulierte Messsignal kontinuierlich ausgesandt (z.B.: sinusoidale Modulation). Die Messung erfolgt über mehrere Peri-

oden. Die Laufzeit wird indirekt über die Phasenlage des empfangenen Signals bestimmt (siehe Abbildung 2.3(b)).



(a) Prinzip Impulsverfahren: das Signal wird in Form eines Pulses ausgesandt. Eine Messung beginnt mit dem Aussenden des Signals und endet mit der Detektion des Echos

(b) Prinzip Phasenvergleichsverfahren: ein amplitudenmoduliertes Messsignal wird kontinuierlich ausgesandt. Die Messung erfolgt über mehrere Perioden. Die Laufzeit wird indirekt über die Phasenlage des empfangenen Signals bestimmt.

Abbildung 2.3: Prinzipien der Distanzmessung. Quelle: [Kahlmann et al., 2006]

Es werden in Folge einige Kameratypen vorgestellt. Die weitere Arbeit bezieht sich auf die PMD Kamera (Abschnitt 2.1.1).

2.1.1 PMD - Photonic Mixer Device

Photomischdetektor

PMD Kameras arbeiten nach dem Phasenvergleichsverfahren. Bei diesem wird die Distanz indirekt über die Verschiebung der Phase zwischen dem emittierten und dem am Sensor detektierten Signal bestimmt. Als Signalemitter wird meist ein LED-Array verwendet, welches ein amplitudenmoduliertes nahes Infrarot (NIR) Signal aussendet. Beispielsweise wird dazu eine sinusoidale Modulation mit der Wellenlänge λ_{mod} verwendet. Die Bestimmung der Phasenlage am Sensor erfolgt durch vier sequentielle Intensitätsmessungen (siehe Abbildung 2.4). Dabei wird der Intensitätswert $c(\tau_i)$ des empfangenen Lichts an den vier um $T_{mod}/4$ verschobenen Abtastpunkten $\tau_{0..3}$ ermittelt. Der Intensitätswert $c(\tau_i)$ ist nicht direkt für einen Abtastpunkt bestimmbar, sondern durch Integration über einen endlich kurzen Zeitraum.

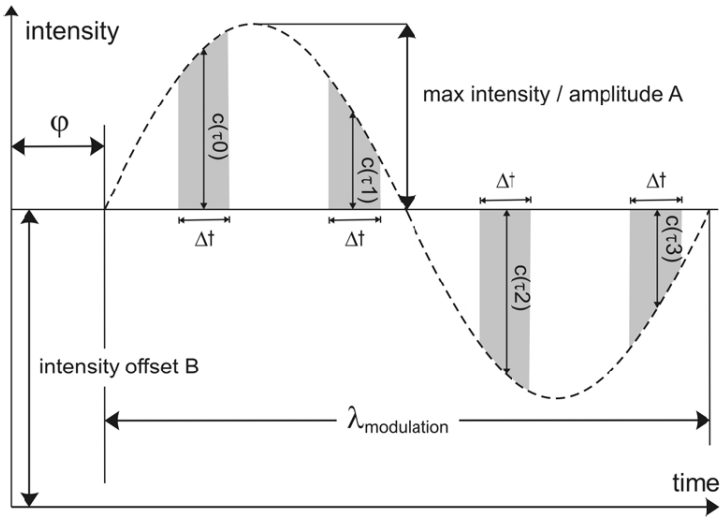


Abbildung 2.4: Prinzip der Bestimmung der Phasenlage ϕ bei PMD-Kameras: die Intensitätswerte $c(\tau_i)$ des empfangenen Lichts werden an den vier um $\frac{\pi}{2}$ verschobenen Abtastpunkten $\tau_{0..3}$ ermittelt. Der Intensitätswert $c(\tau_i)$ wird durch Integration über den Zeitraum Δt bestimmt. Quelle: [Kahlmann et al., 2006]

Abbildung 2.5 zeigt eine vereinfachte schematische Darstellung eines einzelnen Sensorelements (Pixel) von PMD[vision]®. In der Mitte befinden sich die lichtdurchlässigen Photogates (gelb). Rechts und links davon sind die Auslesedioden, welche mit der Ausleseelektronik verbunden sind, angeordnet. Darunter befindet sich ein Halbleiter, welcher wiederum auf ein Substrat aufgetragen ist. Eintreffende Photonen reißen aus dem Halbleiter Elektronen-Loch-Paare. Durch Anlegen einer Spannung an den Photogates kann die Bewegungsrichtung der Elektronen im Substrat beeinflusst werden. So werden je nach angelegter Spannung die photogenerierten Ladungsträger zum linken oder zum rechten Ausleseregister geleitet. Die angelegte Spannung schwingt mit der gleichen Frequenz (homodyne Mischung) wie die Amplitude des ausgesandten Lichts [Ringbeck u. Hagebeuker, 2007a]. Im Sensor wird somit der Wert $c(\tau_0)$ durch Integration über die erste Hälfte der Periode der Schwingung und der Wert $c(\tau_2)$ über die zweite Hälfte ermittelt. Die beiden anderen Werte $c(\tau_1)$ und $c(\tau_3)$ werden bestimmt, indem die angelegte Spannung um $T_{mod}/4$ versetzt wird und die Messungen wiederholt werden.

Unter der idealisierten Annahme, dass das optische Echo und die Referenzspannung sinusförmig sind und das Abtastungsintervall unendlich kurz ist, kann die Phasenlage ϕ mittels Gleichung (2.1) und daraus die Distanz d (2.2), sowie die Amplitude A mit der Gleichung (2.3) berechnet werden.

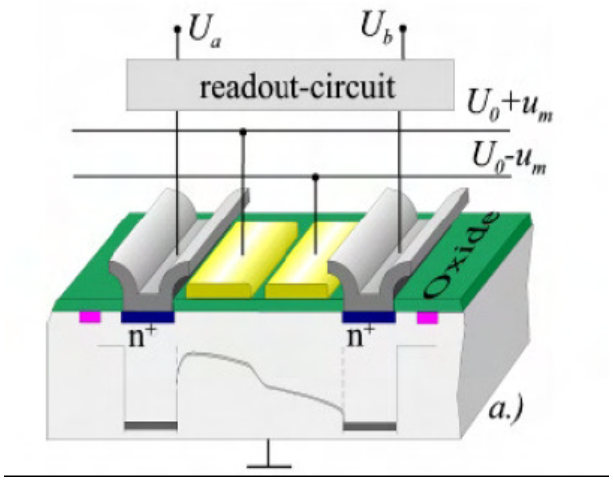


Abbildung 2.5: Schematische Darstellung eines PMD-Pixelelements: in der Mitte befinden sich die lichtdurchlässigen Photogates (gelb), rechts und links davon sind die Auslesedioden angeordnet und darunter befinden sich die Halbleiterschicht und das Substrat. Quelle: [Ringbeck u. Hagebeuker, 2007a]

$$\phi = \arctan\left(\frac{c(\tau_0) - c(\tau_2)}{c(\tau_1) - c(\tau_3)}\right) \tag{2.1}$$

$$d = \frac{1}{2} * \lambda_{mod} * \frac{\phi}{2\pi} \tag{2.2}$$

$$A = \frac{\sqrt{(c(\tau_0) - c(\tau_2))^2 + (c(\tau_1) - c(\tau_3))^2}}{2} \tag{2.3}$$

Ein wesentlicher Qualitätsparameter der Distanzmessung ist die Amplitude A.

Die Abtastung des Signals wird für mehrere Perioden wiederholt. Dadurch verbessert sich das Signal-Rausch-Verhältnis. Die gesamte Dauer, in der die Aufzeichnung des Signals erfolgt, wird als Integrationszeit bezeichnet. Diese ist vergleichbar mit der Belichtungszeit von herkömmlichen Kameras. Die Integrationszeit sollte an die Aufnahmesituation angepasst sein, da es sonst bei zu langer Integrationszeit zu Sättigungserscheinungen bei nahen, stark reflektierenden Objekten kommt (Überbelichtung). Ist die Integrationszeit zu kurz, tritt bei Objektpunkten mit einer größeren Distanz bzw. geringem Reflexionsvermögen ein geringes Signal-Rausch-Verhältnis auf, wodurch die Präzision in der Rekonstruktion des Phasenwinkels bzw. der Objektdistanz ebenfalls geringer wird (siehe Abschnitt 2.3).

Da die Intensität der Sonnenstrahlung im nahen Infrarot jene der Beleuchtungseinheit bei weitem übertrifft, muss zur Verwendung von PMD Kameras bei Außenaufnahmen dieses Fremd-

licht bereits im Sensor unterdrückt werden, da sich ansonsten das Signal-Rausch-Verhältnis als verschwindend klein ergibt bzw. sofortige Sättigung des Sensors eintritt. Eine Maßnahme zur Verbesserung des Signal-Rausch-Verhältnisses ist die Verwendung eines optischen Bandpassfilters, der nur Wellen im Bereich der Wellenlänge der Trägerfrequenz passieren lässt (NIR). Da zur Bestimmung der Phasenlage nur Differenzen der Abtastwerte $c(\tau_0)$ und $c(\tau_2)$ bzw. $c(\tau_1)$ und $c(\tau_3)$ benötigt werden, können diese bereits im Sensor berechnet werden. Weil nun das Hintergrundlicht in $c(\tau_{0..3})$ nicht moduliert ist wird dies bei der Differenzbildung eliminiert, und die Sättigung des Sensors verhindert. Dadurch sind größere Belichtungszeiten, die zu einer Qualitätssteigerung führen, möglich.
Bei Sensoren, die nach dem Phasenvergleichsverfahren arbeiten ist die Distanzbestimmung nur innerhalb der halben Modulationswellenlänge eindeutig $d_{max} = \lambda_{mod}/2$. Abhilfe bietet hier die Verwendung verschiedener Modulationsfrequenzen. Dabei werden größere Wellenlängen zur Bestimmung der Mehrdeutigkeiten verwendet, die abschließende Feinmessung erfolgt über die kleinste Wellenlänge.

Anbieter von PMD-Kameras

In Tabelle 2.1 werden drei hochauflösende Kameras der Hersteller PMDTechnologies (CamCube 2.0), MESA Imaging (SR4000) und Canesta™ (DP200) miteinander verglichen. Der PMD[vision]® CamCube 2.0 besitzt die derzeit größte Auflösung von 204 x 204 Pixel. Von dem Hersteller Canesta™ sind für das aktuelle Modell keine Daten verfügbar, deshalb wird hier die Kamera DP200 vorgestellt.

(a) CamCube [PMDTechnologies GmbH]

(b) SR4000 [MESA Imaging AG]

(c) DP200 [Canesta, Inc.]

Abbildung 2.6: PMD-Kameras

Am Institut für Photogrammetrie und Fernerkundung ist ein Swissranger™ SR-3000 verfügbar. Es ist das Vorgängermodell des Swissranger™ SR-4000.

Parameter		CamCube 2.0	SR4000	DP200
Anzahl der Pixel		204x204	176x144	64x64
Blickfeld	[°]	40x40	43,6x34,6	21,5x21,5 / 40,5x40,5 / 70x70
max. Bildrate	[fps]	25	54	30
Trägerwellenlänge	[nm]	870	850	785
Modulationsfrequenz	[MHz]	20	30 (29 / 31)	13 / 26 / 52 / 104
Distanzbereich	[m]	0,3-7	0,3-5	0,5-20 / 0,3-10 / 0,3-6
Abmessung	[mm]	modular	65x65x68	125x60x63
Gewicht	[g]	-	470	-

Tabelle 2.1: Vergleich von drei PMD Kameras

2.1.2 Heterodyne Mischung

Bei dieser Technik wird ein amplitudenmoduliertes Signal ausgesandt und vom Objekt reflektiert. Bevor das Signal zum Bildsensor gelangt, durchläuft es einen sehr schnellen elektronischen Verschluss, der mit einer geringfügig anderen Frequenz als die Beleuchtungseinheit betrieben wird (heterodyne Mischung). Dadurch entsteht hinter dem Verschluss eine Schwebung. Wegen der geringen Frequenz (im Bereich weniger Hertz) der Schwebung kann ein herkömmlicher CCD Chip verwendet werden, um die Phasenlage ϕ zu rekonstruieren. Daraus lässt sich die Distanz d zum Objekt für jedes einzelne Pixel bestimmen (Gleichung 2.2). Bei einer Entfernung von 5 m und einer Auflösung von 512 x 512 Pixel kann eine Präzision kleiner 1 mm erreicht werden [Dorrington et al., 2007]. Jedoch liegt die zeitliche Auflösung bei dieser Präzision nur im Bereich einiger Sekunden.

2.1.3 SPAD - single-photon avalanche diodes

Einzelphotonendetektion mittels Lawinenfotodiode

Dieser Kameratyp arbeitet nach dem Impulsverfahren, wobei die Distanz direkt über Laufzeitmessungen bestimmt wird.

$$D = \frac{1}{2} c \cdot T \tag{2.4}$$

$c \ldots$ Lichtgeschwindigkeit, $D \ldots$ Distanz, $T \ldots$ Laufzeit

Jedes Pixel-Element der Kamera besteht im Wesentlichen aus einer Lawinenfotodiode (avalanche photodiode), die ein präzises Stoppsignal an eine hochgenaue Uhr liefert. Diese wird beim Aussenden eines kurzen Laserpulses initialisiert. Die Lawinenfotodioden arbeiten im sogenann-

ten Geiger-Modus, das heißt sie werden mit einer Spannung, die geringfügig über der Durchbruchsspannung liegt, betrieben. Dadurch wird die Fotodiode extrem empfindlich und ist in der Lage, einzelne Photonen des ausgesandten Laserpulses zu detektieren und dabei ein sehr präzises Stoppsignal zu liefern. Die Messung wird mit einer Frequenz von einigen MHz wiederholt, und in einem Histogramm ausgewertet. Daraus lässt sich die Laufzeit sehr genau bestimmen und es wird eine Präzision (bei einer Distanz von 3 m) von 1,8 mm erreicht. Distanzkameras, die nach diesem Prinzip arbeiten, liegen bisher nur als Prototypen vor [Karel, 2008].

2.1.4 MDSI - multiple double short time integration

Mehrfache Kurzzeit-Doppelbelichtung

Bei diesem Verfahren wird die Laufzeit nicht direkt gemessen, sondern über die Belichtung, welche im Bildsensor bei unterschiedlichen Verschlusszeiten ermittelt wird.
Der von einem Laser ausgesandte und am Objekt reflektierte Lichtpuls wird beim ersten Grauwertbild nicht über die gesamte Pulsbreite aufgezeichnet, sondern während er den Verschluss passiert, abgeschnitten. Somit sind die Grauwerte von der Objektdistanz abhängig, d. h. nähere Abtastpunkte ergeben potentiell hellere Pixel (siehe Abbildung 2.7).

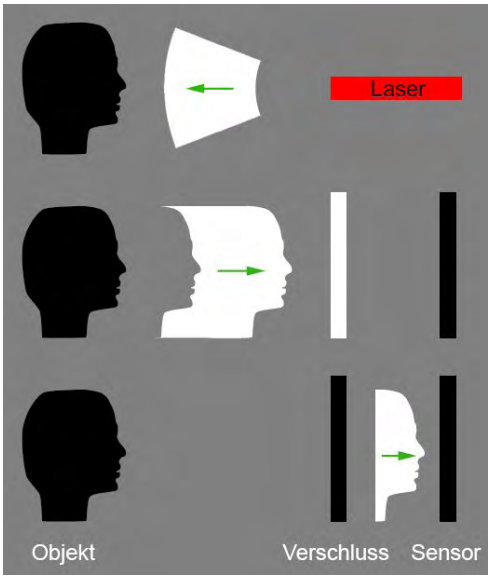


Abbildung 2.7: Prinzip MDSI: vom Laser wird ein divergierender Lichtpuls ausgesandt und am Objekt reflektiert. Das Echo wird bei Passieren des Verschlusses von diesem abgeschnitten, dadurch ist die Dicke des Echos von der Objektdistanz abhängig, was zu entsprechenden Grauwerten am Sensor führt. Quelle: [Karel, 2008]

Nun hängen die Grauwerte noch vom Hintergrundlicht und der unterschiedlichen Reflektivität am Objekt ab. Zur Elimination des Hintergrundlichtes wird ein Bild mit gleicher Belichtungszeit, aber ohne Lichtpuls aufgezeichnet und vom Bild mit Lichtpuls subtrahiert.

$$G_{1,reduziert} = G_{1,ges} - G_{1,hg} \tag{2.5}$$

$G_{1,hg}$ ... erstes Bild ohne Lichtpuls, $G_{1,ges}$ ... erstes Bild mit Lichtpuls

Zur Beseitigung des störenden Einflusses der objektabhängigen Reflektivitäten wird ein weiteres Bild mit einer Belichtungszeit, bei welcher der vollständige Puls abgebildet wird, aufgezeichnet. Es ist wieder die erste Reduktion anzubringen.

$$G_{2,reduziert} = G_{2,ges} - G_{2,hg} \tag{2.6}$$

$G_{1,reduziert}$ ist von der Laufzeit und von der Reflektivität abhängig. $G_{2,reduziert}$ ist nur von der Reflektivität abhängig, deshalb ist es möglich durch Quotientenbildung die Laufzeit zu bestimmen.

$$\frac{G_{1,reduziert}}{G_{2,reduziert}} = \frac{T_{Puls} - T_{Laufzeit}}{T_{Puls}} \tag{2.7}$$

2.2 Vergleich der Distanzkamera zur klassischen Photogrammetrie und zum Laserscanning

Die dreidimensionale Aufnahme des Objektraums ist eine der Hauptaufgaben der Photogrammetrie. Insbesondere in der terrestrischen Photogrammetrie kommen digitale Kameras zum Einsatz. Eine Aufnahme definiert ein Bündel von Richtungen, dieses wird als Strahlenbündel bezeichnet. Die Rekonstruktion eines 3D Objektpunktes erfolgt mittels eines räumlichen Vorwärtsschnittes. Dazu sind mindestens zwei Aufnahmen notwendig, die eine ausreichende Schnittqualität der Strahlenbündel gewährleisten müssen. Außerdem ist die Photogrammetrie auf eine kontrastreiche Oberfläche angewiesen, um Korrespondenzen bestimmen zu können. Dies kann bei manueller Auswertung sehr aufwendig sein. Es gibt zwar in letzter Zeit Ansätze die Auswertung zu automatisieren, diese sind jedoch rechenintensiv und die Zuverlässigkeit ist bei komplexen Aufnahmen nicht immer gewährleistet [Kraus, 2004].
Mit dem Laserscanning hat sich in den letzten Jahren eine neue Technik entwickelt, mit der es möglich ist, eine große Anzahl von Punkten automatisch zu messen. Das Laserscanning zählt

zu den aktiven Sensoren, wobei die Distanz zu einem Objektpunkt direkt mit Hilfe eines Laserstrahls gemessen wird. Der Strahl wird in unterschiedliche Richtungen abgelenkt (z.B. durch Rotationsprismen). Dadurch können auch bereits mit einem Standpunkt ohne homologe Punkte Objektpunkte im dreidimensionalen Raum geometrisch rekonstruiert werden. Die Messung der Punkte erfolgt aber sequentiell, das heißt jeder Punkt besitzt einen anderen Aufnahmezeitpunkt. Bewegt sich der Scanner während der Aufnahmen, so ist es notwendig die Orientierung des Scanners extern zu beobachten. Dies wird meist mittels Satellitenpositionierungssystem (z.B. GPS) und einer inertialen Messeinrichtung (IMU) bewerkstelligt [Kraus, 2004]. Auch bei bewegten Objekten bringt die sequentielle Aufnahme große Nachteile.

Distanzkameras verbinden nun einige Vorteile dieser beiden Techniken. Einerseits erfolgt die Erfassung der Punkte simultan, wodurch auch bewegte Szenen aufgenommen werden können. Wird die Kamera bewegt, so kann ihre Orientierung aus den Bildern bestimmt werden. Andererseits wird die Distanz zum Objekt direkt gemessen, weshalb keine Punktkorrespondenzen nötig sind und somit keine Oberflächentextur notwendig ist. Außerdem ist es bereits mit einer Aufstellung möglich, 3D-Koordinaten zu erhalten. Dies reduziert natürlich auch die Anzahl der erforderlichen Aufstellungen, um komplexe Objekte zu erfassen. Im Vergleich zum Laserscanning haben Distanzkameras auch den Vorteil, dass sie keine bewegten Bauteile besitzen und dadurch eine sehr kompakte Bauweise (siehe Abbildung 2.8) bzw. niedrige Anschaffungskosten aufweisen. Es sei jedoch erwähnt, dass bei den jetzigen Modellen der Distanzkameras ein Notebook zur Aufzeichnung der Daten und eine externe Stromversorgung notwendig sind. Außerdem ist die Präzision der Distanzbestimmung bei aktuellen Modellen noch um Größenordnungen niedriger als im Falle des Laserscannings. Dies schmälert aus momentaner Sicht den Vorteil, der sich aus der kompakten Bauweise ergibt.

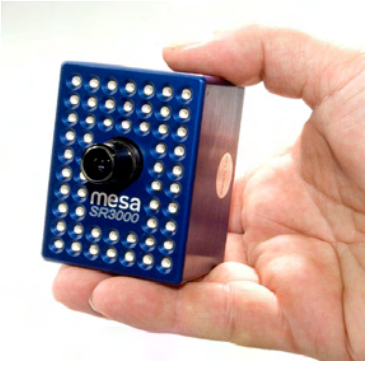


Abbildung 2.8: SR-3000: aus der Abbildung ist die Größe der Kamera ersichtlich. Quelle: [Mechat u. Büttgen, 2007]

Die Nachteile der Distanzkameras gegenüber der klassischen Photogrammetrie und des Laserscannings sind hauptsächlich die eingeschränkte Reichweite aufgrund der einzuhaltenden Augensicherheit, die geringe räumliche Auflösung, die schlechtere Präzision (im Zentimeterbereich), sowie große systematische Fehler. Es ist zu erwarten, dass sich die Technologie in den nächsten Jahren weiterentwickelt und sich dadurch die Bedeutung dieser Nachteile verringern wird.

2.3 PMD-Daten

Dieser Abschnitt beschäftigt sich mit den Charakteristiken der Messdaten von PMD-Kameras (siehe Abschnitt 2.1.1). Dabei wird insbesondere auf die Fehler in den Distanzbeobachtungen eingegangen.
Im Sensor der Distanzkamera werden Intensitätswerte $c(\tau)$ gemessen. Daraus können mittels der Gleichungen 2.1 und 2.2 die Distanz- bzw. Amplitudenbilder bestimmt werden (siehe Abschnitt 2.1.1). Diese Umrechnung geschieht beim SR-3000 in Echtzeit, sodass direkt Distanz- und Intensitätsbilder aufgezeichnet werden. Die Bilder haben eine Auflösung von 144 x 176 Pixel. Insbesondere die Distanzbilder weisen erhebliche systematische und stochastische Fehler auf. Auf diese wird im Folgenden kurz eingegangen.

Systematische Fehler

Die Bestimmung der Phasenlage (Gleichung 2.1) entspricht einer Korrelation des empfangenen Signals mit einem Referenzsignal (siehe Abschnitt 2.1.1). Bei Gleichung 2.1 wird vorausgesetzt, dass beide Signale sinusförmig sind und die gleiche Frequenz haben [Rapp, 2007]. Diese Voraussetzung erfüllt der SR-3000 nicht, weshalb es, aufgrund von ungeraden harmonischen Oberschwingungen im modulierten Signal, zu zyklischen systematischen Fehlern bei der Distanzmessung kommt (siehe Abbildung 2.9).

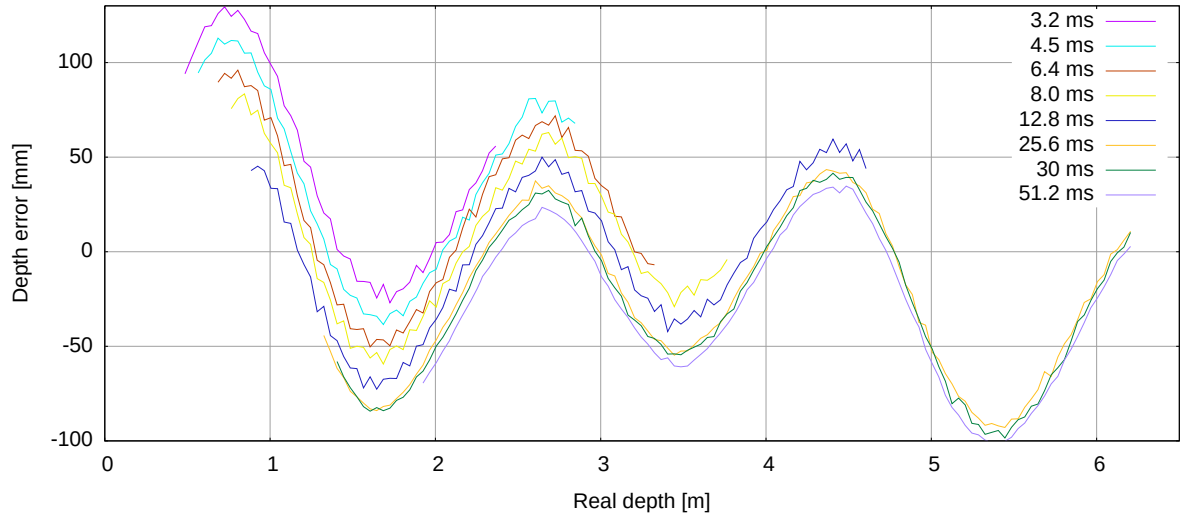


Abbildung 2.9: Linearitätsfehler des SR-3000 für unterschiedliche Integrationszeiten im zentralen Pixel. Quelle: [Rapp, 2007]

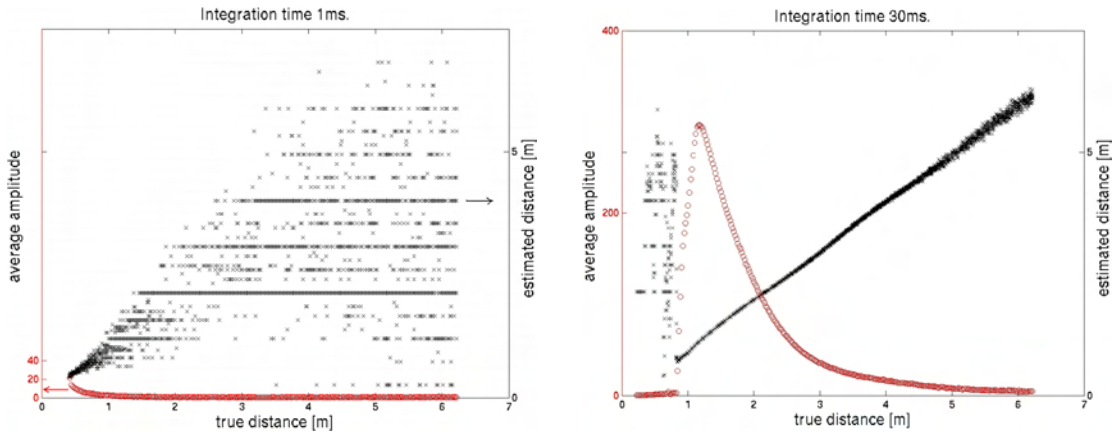


Abbildung 2.10: Quantisierungsfehler: dargestellt ist die gemessene Distanz (schwarz) und mittlere Amplitude (rot) zur nominalen Distanz bei Integrationszeiten von 1ms (links) und 30ms (rechts). Je niedriger die Amplitude ist, desto mehr horizontale Linien treten hervor. Die horizontalen Linien sind auf Quantisierungsfehler zurückzuführen. In der rechten Abbildung sind bei geringer nominaler Distanz große Fehler bei der gemessenen Distanz zu beobachten. Diese sind auf Sättigung (Überbelichtung) des Sensors zurückzuführen. Quelle: [Frank et al., 2009]

Wie aus Abbildung 2.9 ersichtlich, ergeben sich für verschiedene Integrationszeiten unterschiedliche Distanzoffsets.

Die gemessene Distanz ist auch von der Temperatur abhängig. Zum einen hat das Aufwärmen des Sensors beim Einschalten einen Einfluss, nach ein paar Minuten stabilisiert sich die Distanzmessung. Zum anderen hat die Umgebungstemperatur einen Einfluss auf die Distanzbeobachtungen [Kahlmann u. Ingensand, 2007b].
Ein weiterer Einflussfaktor, der Distanzbeobachtungen systematisch beeinflusst, ist die Amplitude des empfangenen Signals. Die Amplitude ist neben der Reflektivität des Objekts auch von der Entfernung selbst abhängig. Der theoretische Zusammenhang zwischen der Distanz d und der Amplitude A wird in Rapp [2007] beschrieben. Näherungsweise gilt folgender Zusammenhang:

$$A \sim \frac{1}{d^2} \tag{2.8}$$

Das empfangene Signal $c(\tau)$ wird digitalisiert, dabei treten Rundungsfehler (Quantisierungsfehler) auf. Dies führt insbesondere bei geringen Amplituden zu Fehlern bei der Berechnung der Phase ϕ und somit auch zu Fehlern in der berechneten Distanz d (siehe linke Abbildung 2.10) [Frank et al., 2009].

Wie aus Abbildung 2.10 ersichtlich sind für größere Distanzen längere Integrationszeiten not-

wendig, da es sonst zu Quantisierungsfehlern kommt. Da bei der Aufnahme, welche im Rahmen dieser Arbeit durchgeführt wurde, die Distanzen für Distanzkameras groß sind, wurde die maximale Integrationszeit verwendet. Im Nahbereich kann es, bei längeren Integrationszeiten zu Sättigung (Überbelichtung) kommen (siehe rechte Abbildung 2.10). Dies tritt bei der Aufnahme in dieser Arbeit kaum auf.
Ein Fehler, bei dem die gesamte Nachbarschaft von Pixeln beeinflusst wird und der insbesondere bei großen Distanzunterschieden im Bild auftritt, ist der sogenannte Streulichteinfluß (Scattering). Im Idealfall wird in jedem Pixel ausschließlich Strahlung vom entsprechend korrespondierenden Abtastfleck am Objekt empfangen. Jedoch kommt es durch kamerainterne Streuung und Reflexionen dazu, dass die vom Objekt reflektierte Strahlung nicht ausschließlich vom korrespondierenden Pixel empfangen wird, sondern auch die benachbarten Pixel beeinflusst werden. Dadurch wird die Distanzinformation der Pixel verfälscht. Das Scattering wirkt sich besonders stark auf den Phasenwinkel und damit auf die Distanzmessung aus [Mure-Dubois u. Hügli, 2007].
Durch mehrfache Reflexionen am Objekt kommt es ebenfalls zu Fehlern bei der Distanzmessung. Dies wird als Multipath bezeichnet [Mure-Dubois u. Hügli, 2007].

Zufällige Fehler
Wie bereits erwähnt ist die Amplitude ein wesentlicher Qualitätsparameter der Distanzmessung. Mittels des Fehlerfortpflanzungsgesetzes lässt sich der Zusammenhang zwischen der Varianz der Distanzmessung d und der Amplitude A ableiten. Dies wird in Rapp [2007] durchgeführt:

$$var(d) \sim \frac{1}{A^2} \tag{2.9}$$

Dieser Zusammenhang ist für diese Arbeit von besonderer Bedeutung, da hier hauptsächlich große Distanzen und dementsprechend geringe Amplituden auftreten. Dies ist auch der Grund dafür, dass in den Distanzbildern ein erhebliches Rauschen auftritt und es daher notwendig ist, diese zu mitteln.

Die Präzision der Distanzmessung hängt damit auch von der Pixelposition ab. Sie nimmt von der Bildmitte zum Bildrand hin ab, da die Amplitude aufgrund des Lichtabfalls durch das Objektiv und die schwächere Beleuchtung zum Rand hin abnimmt.

Die Präzision der Amplitudenbeobachtungen ist wesentlich besser als jene der Distanzbeobachtungen. Dies ist aus dem Mittelwert der Quotienten aus der Standardabweichung der Amplitudenbeobachtungen σ_{amp} durch die Standardabweichung der Distanzbeobachtungen σ_{dist}

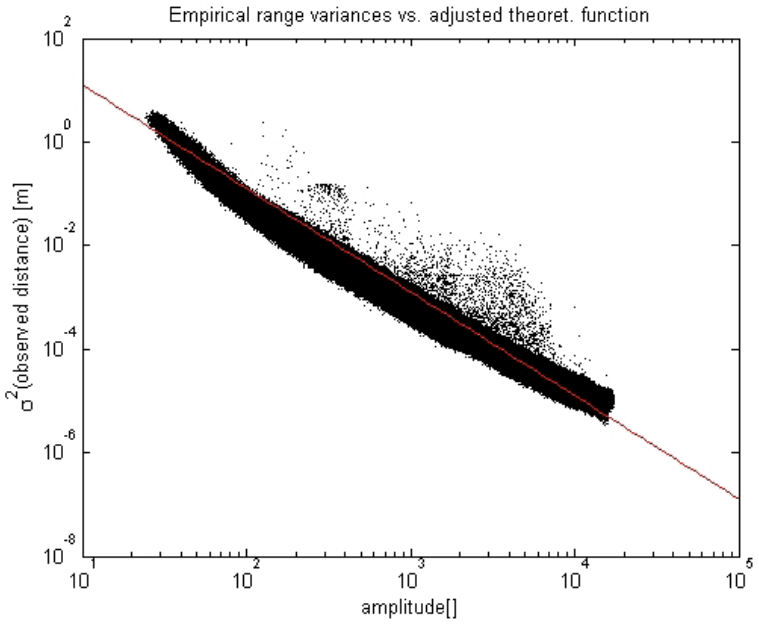


Abbildung 2.11: Empirische Varianz der Distanzbeobachtungen bei unterschiedlichen Amplituden: bei niedrigerer Amplitude, steigt die empirische Varianz der Distanzbeobachtungen. Einheit der Amplitude: der Messbereich geht von 0 bis $2^{16} - 1$. Quelle: [Karel, 2009]

ersichtlich. Dazu wird für alle Pixel einer statischen Sequenz die empirische Standardabweichung σ_{amp} und σ_{dist} bestimmt. Anschließend wird der Quotient $q_{präz} = \sigma_{amp}/\sigma_{dist}$ für jedes Pixel berechnet. Beispielhalft beträgt der Mittelwert der Quotienten $\bar{q}_{präz}$ bei einer mittleren Objektdistanz von 5,2 Meter und einer mittleren Reflektivität: $\bar{q}_{präz} = 0,11$.

2.4 Kalibrierung

Wie aus dem vorherigen Abschnitt zu entnehmen ist, weisen die Distanzbeobachtungen von PMD- Kameras zum Teil erhebliche systematische Fehler auf, weshalb eine Kalibrierung zum Erreichen der entsprechenden absoluten Genauigkeiten zwingend notwendig ist. Die Qualität der verwendeten Kalibrierung ist entscheidend für die Ergebnisse der für die Auswertung verwendeten Methoden. So konvergiert zum Beispiel die Fein-Orientierung mittels Iterative Closest Point (ICP) Algorithmus (Abschnitt 3.3.3.2) nur mit entsprechend kalibrierten Daten. Die Ergebnisse der Orientierung und der Modellierung hängen wesentlich von der Effizienz der Kalibrierung ab.

In dieser Arbeit wird die Kalibrierung von Karel u. Pfeifer [2009] verwendet. Diese basiert

auf einem Distanzkalibrierungsmodell, das sich aus folgenden Einflussfaktoren zusammensetzt: Beobachtete Distanz, die Amplitude, die Integrationszeit sowie die Position im Sensor.
Nicht modelliert werden Scattering- und Multipath- Einflüsse (siehe Abschnitt 2.3). Insbesonders Multipath ist bei Aufnahmen, die zur Vervollständigung von Stadtmodellen dienen, zu erwarten. Dies führt im Ergebnis zu systematischen Fehlern (siehe Abschnitt 4.3.3). Der Temperatureinfluss wird bei der verwendeten Kalibrierung ebenfalls nicht berücksichtigt. Zudem ist die Kalibrierung bei extrem niedrigen Amplituden nur bedingt einsetzbar.

2.5 Anwendungen

Die Anwendungen der Distanzkameras werden, dem aktuellen Entwicklungsstand entsprechend, vor allem durch die oben genannten Schwächen (geringe Reichweite und der geringen geometrischen Auflösung) eingeschränkt.
Die Kooperationen der Kamerahersteller mit der Automobilbranche deuten darauf hin, dass sich in diesem Bereich breite Anwendungsmöglichkeiten, hauptsächlich im Gebiet der aktiven und passiven Sicherheit ergeben werden. Die möglichen Anwendungen beziehen sich zum einen auf den Außenraum des Fahrzeugs, wie zum Beispiel automatisches Erkennen von Objekten und Warnung vor Hindernissen bzw. automatische Notbremsung, sowie Einparkassistent [Acharya et al., 2008]. Zum anderen gibt es auch Anwendungsmöglichkeiten für den Innenraum wie zum Beispiel: Erkennung der Belegung des Sitzplatzes zur Anpassung bzw. Deaktivierung des Airbags, Müdigkeits-Erkennung bzw. Aufmerksamkeits-Erkennung usw. [PMDTechnologies GmbH, 2009].
Ein weiteres Anwendungsfeld der Distanzkameras liegt in der autonomen Navigation von Robotern, außerdem ist ein Einsatz im Bereich der Industrieroboter denkbar. Hier können Distanzkameras zum Beispiel Roboter beim Segmentieren und Erkennen von Gegenständen unterstützen [Kristensen et al., 2001]. In der Logistik ergeben sich Anwendungsmöglichkeiten wie schnelle Volumendetektion in Paketverteilzentren, Erkennung von freien Plätzen in Lagern oder Schüttgutvermessung [Ringbeck u. Hagebeuker, 2007b]. Die bisher erwähnten Anwendungen befinden sich noch in der Versuchsphase oder sind denkbare Anwendungen. In der Unterhaltungsindustrie hingegen stehen Distanzkameras kurz vor der Markteinführung. Dabei ergeben sich neue Möglichkeiten, zum Beispiel in der Bedienung von Fernsehgeräten mittels Gesten. Auch die Steuerung von Computerspielen kann durch Körperbewegungen erfolgen.

2.5.1 Anwendung im Vermessungswesen

Aufgrund der in Abschnitt 2.2 genannten Vorzüge der Distanzkameras ergeben sich potentielle Anwendungsgebiete im Vermessungswesen. Durch die simultane, fortlaufende Messung von Punktwolken erscheinen Anwendungen, ähnlich denen von Laserscannern, als prädestiniert für Distanzkameras. Auch in dieser Arbeit wird die Einsatzfähigkeit eines Swissrangers ™ SR-3000 in einer für Laserscanner typischen Anwendung gezeigt (siehe Kapitel 4). Wegen der Möglichkeit der Aufnahme mit bewegter Kamera, der Erfassung in sich bewegter Szenen und der niedrigeren Anschaffungskosten scheinen Distanzkameras eine anzustrebende Alternative zu Laserscannern zu sein. Jedoch wird im Vermessungswesen eine äußere Genauigkeit zumindestens im cm - Bereich gefordert, was mit den gegenwärtigen Distanzkameras nur eingeschränkt erreicht wird. Diesbezüglich sollte die Weiterentwicklung der Technologie und der Kalibrierungsansätze eine Verbesserung bringen. Anwendungen wie die Erfassung von Felsstürzen, Muren- und Lawinenabgängen, bei denen die simultane Distanzmessung der Kameras ein großer Vorteil wäre, liegen derzeit aufgrund der eingeschränkten Reichweite nicht im Bereich des Möglichen. Hier wäre aber der Einsatz in Modellsimulationen zu diskutieren. Potentielle Einsatzmöglichkeiten als Alternative zum Laserscanner liegen im Nahbereich, wie zum Beispiel bei der Aufnahme zur Modellierung kleinerer Objekte und der Erfassung von Innenräumen, bei der städtischen Detailaufnahme usw. [Karel, 2008].

Kapitel 3

Methodik

In diesem Kapitel werden die Methoden, die bei dem in dieser Arbeit entwickelten Verfahren verwendet werden, vorgestellt. Abschnitt 3.1 gibt einen Überblick über den Ablauf des Verfahrens. Anschließend werden in Abschnitt 3.2 die zur Auswertung notwendigen Datenvorprozessierungsschritte beschrieben. Die zur Orientierung der Aufnahmen verwendeten Methoden sind Inhalt des Abschnittes 3.3. Am Ende des Kapitels wird in Abschnitt 3.4 auf die Triangulation und die dazu verwendete Software eingegangen.

3.1 Vorgeschlagener Arbeitsablauf

Potentiell gibt es viele Möglichkeiten mittels einer PMD-Kamera Daten aufzunehmen und auszuwerten, um in weiterer Folge ein Modell ableiten zu können. In dieser Arbeit wird ein Ablauf vorgestellt, der ohne zusätzliche Information (z.B. Grundrissplan, GPS, IMU, ...) weitgehend automatisch die Orientierung der Aufnahmen bestimmt. Für die einzelnen Schritte des vorgestellten Verfahrens wurden jene Methoden ausgewählt, die angesichts der geringen Präzision der Distanzbeobachtungen aufgrund der großen Objektdistanzen (PMD–Datencharakteristika siehe Abschnitt 2.3) als aussichtsreich erschienen, und deren Umsetzbarkeit anhand der Aufnahmen des Palais Ferstel überprüft.

Die PMD-Daten weisen insbesondere bei zunehmender Distanz bzw. geringer Amplitude und im Randbereich des Distanzbildes ein erhöhtes Rauschen auf. Es ist daher notwendig, mehrere Bilder mit gleicher Orientierung zu mitteln, um Bilder mit einer höherer Präzision zu erhalten.

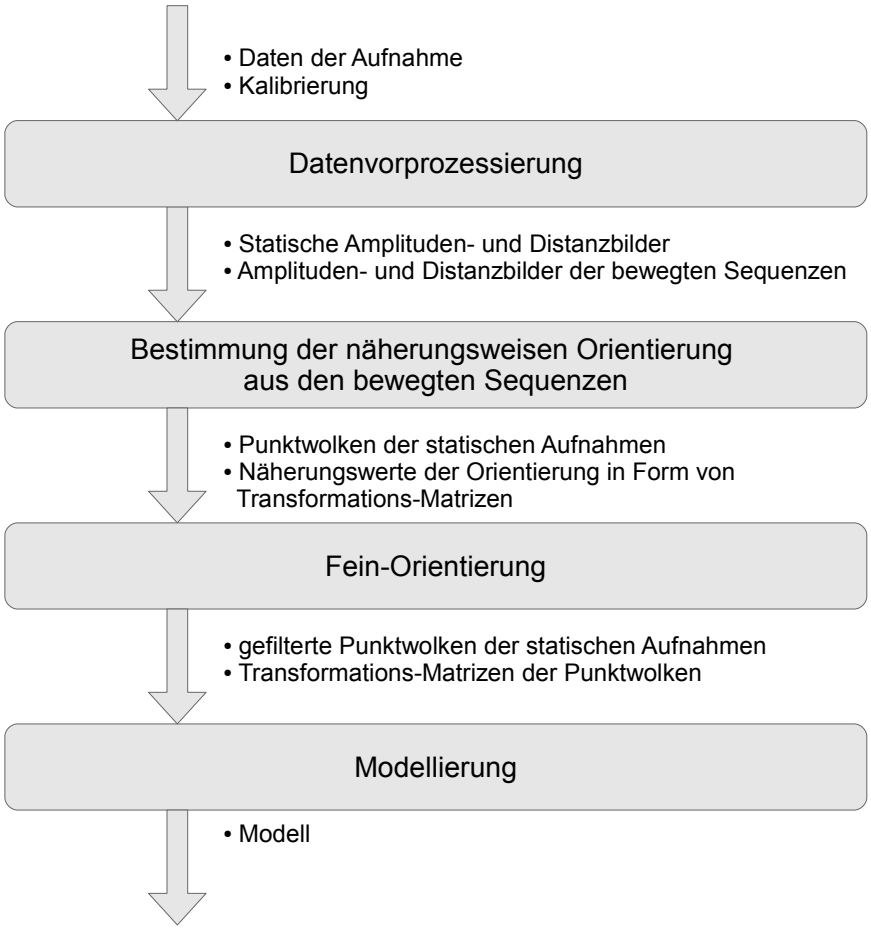


Abbildung 3.1: Überblick des gesamten Ablaufs der Auswertung vom Einlesen der aufgenommenen Daten bis zum fertigen Modell

Diese gemittelten Bilder werden in weiterer Folge als *statische Bilder* bezeichnet. Zur Aufzeichnung der statischen Sequenz ist ein Stativ notwendig und die Orientierung der Kamera wird während der Aufnahme eines *statischen Bildes* nicht verändert. Zwischen den *statischen Bildern* werden ebenfalls Bilder aufgenommen, da diese zur näherungsweisen Orientierung benötigt werden. Das heißt die Aufnahme erfolgt kontinuierlich, wobei sich statische und bewegte Sequenzen abwechseln. Bei der Auswertung muss das aufgenommene Entfernungsvideo in statische und bewegte Sequenzen segmentiert werden. Die näherungsweise Orientierung erfolgt über Punktkorrespondenzen, die in den Amplitudenbildern der Distanzkamera bestimmt werden. Grundsätzlich könnten zur Bestimmung der relativen Orientierung zweier Aufnahmen geometrische Primitive (Ebenen bzw. andere Flächen) in der Punktwolke verwendet werden. Dies kommt bei Methoden zur automatischen Registrierung von Laserscannerdaten zum Ein-

satz. Für die bewegten Sequenzen der Distanzkameraaufnahme sind diese Methoden weniger geeignet, da die Distanzbilder von Einzelaufnahmen und die daraus berechneten Punktwolken eine geringere Präzision aufweisen. Dies gilt im besonderen für die Aufnahmen der Passage, da hier für die PMD-Kamera extreme Distanzen auftreten. Zur Fein-Orientierung werden die aus den statischen Distanzbildern berechneten Punktwolken verwendet. Zum Schluss erfolgt die Triangulation der orientierten Punktwolken. In Abbildung 3.1 ist der Ablauf der Auswertung zusammengefasst.

3.2 Datenaufbereitung

Wie im vorigen Abschnitt beschrieben, muss das aufgenommene Entfernungsvideo in statische und bewegte Sequenzen segmentiert werden, worauf in Abschnitt 3.2.1 eingegangen wird. Bei der näherungsweisen Orientierung werden Amplitudenbilder verwendet. Die Amplitudennormalisierung dieser Bilder wird in Abschnitt 3.2.2 beschrieben. Die Distanzbilder werden bereits im Vorfeld der Datenaufbereitung kalibriert. Dies ist aufgrund großer systematischer Fehler in den Distanzbildern notwendig (siehe Abschnitt 2.4). Abbildung 3.2 zeigt den Ablauf der Datenvorprozessierung.

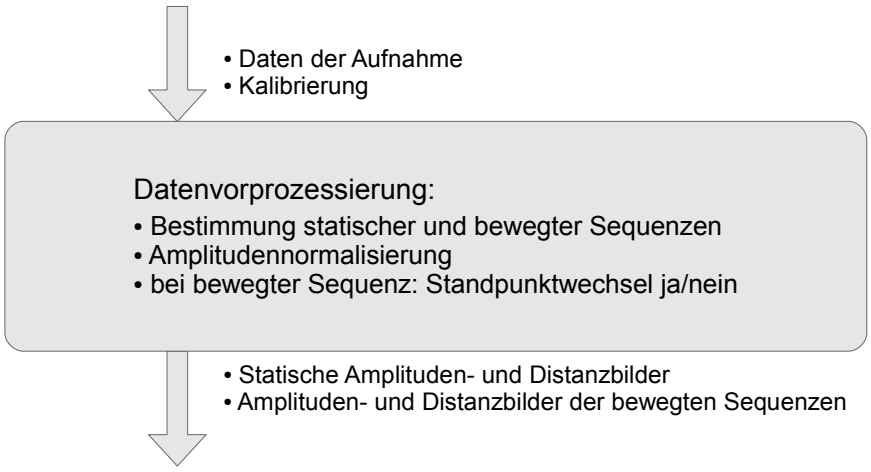


Abbildung 3.2: Ablauf der Datenvorprozessierung

3.2.1 Segmentierung in statische und bewegte Sequenzen

Ein erster einfacher Ansatz zur Bestimmung der Grenzen zwischen statischen und bewegten Sequenzen ist die Berechnung des Mittelwertes aller Grauwerte pro Amplitudenbild. Eine statische Sequenz liegt nun dort vor, wo für zwei aufeinanderfolgende Bilder die Differenz der mittleren Grauwerte der beiden Bilder für eine gewisse Anzahl von Bildern unter einem Schwellwert liegt. Wenn sich der Mittelwert von aufeinander folgenden Bildern stark ändert, das heißt die Ableitung über dem Schwellwert liegt, handelt es sich um eine Sequenz von bewegten Bildern. Der Nachteil dieses Ansatzes ist, dass zum Beispiel bei einer Änderung der Beleuchtung der Integrationszeit während einer statischen Sequenz fälschlich eine Bewegung detektiert wird. Dieses Problem kann verhindert werden, indem anstatt der Ableitung des Mittelwerts, der Korrelationskoeffizient von zwei aufeinander folgenden Amplitudenbildern verwendet wird. Der 2D - Korrelationskoeffizient r berechnet sich aus den Grauwerten G zweier benachbarter Amplitudenbilder:

$$r = \frac{\sum_m \sum_n \left(G1_{mn} - \overline{G1}\right)\left(G2_{mn} - \overline{G2}\right)}{\sqrt{\sum_m \sum_n \left(G1_{mn} - \overline{G1}\right)^2 \sum_m \sum_n \left(G2_{mn} - \overline{G2}\right)^2}} \tag{3.1}$$

Dabei ist $\overline{G}$ der Mittelwert aller Grauwerte eines Bildes.
Sind die Bilder annähernd gleich, so wird der Korrelationskoeffizient r annähernd $+1$. Ist r für eine bestimmte Anzahl von Bildern annähernd $+1$, so liegt eine statische Sequenz vor. Der Vorteil ist nun, dass der Korrelationskoeffizient radiometrisch automatisch skaliert wird und dadurch unempfindlich gegenüber Schwankungen in den mittleren Grauwerten ist. Noch besser ist die Verwendung der Differenz der Korrelationskoeffizienten, da in diesem Fall der Wertebereich, in dem eine statische Sequenz vorliegt, enger gefasst werden kann. Somit ist die Abgrenzung zwischen statischer und bewegter Sequenz noch schärfer möglich.

Bei den Distanzbildern ist mit Ausreißern in den Distanzwerten zu rechnen. Deswegen wird bei der Mittelung einer statischen Sequenz von Distanzbildern ein gestutztes Mittel verwendet. Das heißt, für die Distanzwerte eines Pixels der aufeinanderfolgenden Bilder einer statischen Sequenz und einem vorgegebenen Anteil an Ausreißern α werden jene Beobachtungen, die kleiner als das $(\alpha/2)$-Quantil oder größer als das $(1 - \alpha/2)$-Quantil sind, verworfen.

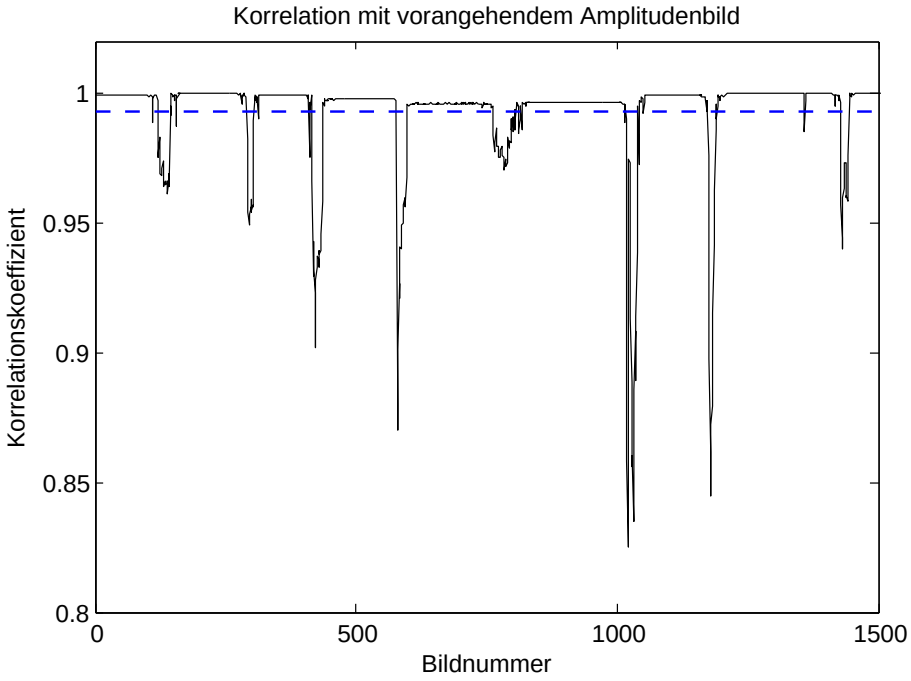


Abbildung 3.3: Korrelation mit vorangehendem Amplitudenbild: bei konstanten Korrelationskoeffizienten (für eine bestimmte Anzahl von Bildern) über dem Schwellwert (blau strichlierte Linie) liegen statische Sequenzen vor.

3.2.2 Amplitudennormalisierung

Bei den Distanzkameras handelt es sich, wie bereits erwähnt, um aktive Sensoren. Das heißt es wird eine Strahlung ausgesandt, am Objekt reflektiert, und am Sensor der Kamera wieder empfangen. Objekte, die eine größere Distanz zur Kamera aufweisen bzw. näher am Bildrand liegen werden im Amplitudenbild bei gleichem Reflexionsvermögen des Objektes dunkler abgebildet (siehe Säulen in Abbildung 3.4 links). Der in weiterer Folge bei der Berechnung der Verknüfungspunkte verwendete Scale-Invariant Feature Transform (SIFT) bzw. Speeded Up Robust Features (SURF) Algorithmus (Kapitel 3.3.2.1) ist zwar weitgehend invariant gegenüber lokal affinen Änderungen der Grauwerte, jedoch sind die eben beschriebenen Effekte nicht affin. Deswegen sollen die Amplitudenbilder vor der Anwendung des SIFT/SURF-Algorithmus normalisiert werden. Die normalisierten Amplitudenbilder ergeben sich als genähert proportional zur richtungsabhängigen, lokalen Reflektivität am Objekt.

Korrektur des distanzabhängigen Amplitudenabfalls

Laut Gleichung 2.8 ist die Amplitude indirekt proportional zum Quadrat der Objektdistanz. Nun ist aber eben diese Distanz beobachtet und es kann das Amplitudenbild $A(m,n)$ mit Hilfe des Distanzbildes $D(m,n)$ vom Einfluss der Objektdistanz befreit werden [Oprisescu et al., 2007].

$$A_{norm,dist}(m,n) = A(m,n)D(m,n)^2 \tag{3.2}$$

Dabei ergibt sich jedoch das Problem, dass insbesonders Distanzbilder aus den bewegten Sequenzen verrauscht sind, und dieses Rauschen in der anzubringenden Korrektur quadriert eingeht. Zusätzlich nimmt das Rauschen mit der Distanz noch zu (siehe 2.3). Deshalb ist es notwendig, das Rauschen der Distanzbilder vor der Amplitudenkorrektur zu reduzieren. Dies kann zum Beispiel mit einem Wiener-Filter (WF) erfolgen, welcher als adaptiver Filter verwendet wird. Die Filterparameter „Mittelwert“ und „Standardabweichung“ werden aus den benachbarten Pixel bestimmt und somit für jeden Bereich eines Bildes angepasst. Stellen mit größerer Varianz werden weniger stark geglättet, als jene mit geringerer Varianz. Dadurch bleiben Kanten gut erhalten und das Rauschen wird unterdrückt [The MathWorks Inc., 2007].
Mit dem gefilterten Distanzbild D_{WF} kann nun das korrigierte Amplitudenbild berechnet werden:

$$A_{norm,dist}(m,n) = A(m,n)D_{WF}(m,n)^2 \tag{3.3}$$

Abbildung 3.4: Korrektur des distanzabhängigen Amplitudenabfalls: links Amplitudenbild vor der Korrektur, rechts nach der Korrektur.

Radiale Korrektur

Zusätzlich zum distanzabhängigen Amplitudenabfall kann auch ein Abfall der Amplitude von der Bildmitte zum Bildrand beobachtet werden. Dies liegt zum einen am allgemeinen Lichtabfall im Objektiv einer Kamera. Dieser folgt dem so genannten $cos^n\tau$-Gesetz, wobei τ der

Bildwinkel ist. Für eine Kamera mit einer einzigen dünnen Linse gilt theoretisch n=4 [Kraus, 2004]. Zum anderen kommt bei einer Distanzkamera noch hinzu, dass die Ausleuchtung des Objektes nicht homogen ist, was im Amplitudenbild zu einem zusätzlichen Lichtabfall zum Rand hin führt.

3.3 Orientierung

Dieses Kapitel soll den prinzipiellen Ablauf von kalibrierten statischen Bildern und bewegten Bildsequenzen zu gefilterten orientierten Punktwolken darstellen. Der erste Schritt ist die Datumsfestlegung (Abschnitt 3.3.1) In Abschnitt 3.3.2 wird der Ablauf der näherungsweisen Orientierung und die dazu verwendeten Methoden beschrieben. Die Fein-Orientierung erfolgt mittels der Punktwolken, die aus den statischen Distanzbildern berechnet werden. Auf den verwendeten ICP-Algorithmus und die dazu notwendige Normalvektorenberechnung und Filterung wird in Abschnitt 3.3.3 eingegangen.

3.3.1 Definition eines globalen Koordinatensystems

Um von den einzelnen Aufnahmen zu einem gesamten Modell zu gelangen ist es notwendig, diese in ein übergeordnetes Koordinatensystem (Objektkoordinatensystem) zu bringen. Beispielsweise kann das Landeskoordinatensystem, auf das sich das Stadtmodell bezieht, ein solches übergeordnetes Koordinatensystem vorgeben. Die Überführung vom lokalen Koordinatensystem in das übergeordnete Koordinatensystem kann über Passpunkte erfolgen.
Es ist aber nicht immer der Fall, dass das übergeordnetes Koordinatensystem ein bereits bekanntes Objektkoordinatensystem ist. In diesem Fall muss das Bezugskoordinatensystem erst definiert werden, was als Datumsfestlegung bezeichnet wird. Bei einem räumlichen, kartesischen Koordinatensystem müssen dazu drei Translationen und drei Drehwinkel festgelegt werden. Bei Distanzkameras muss der Maßstab bei der Datumsfestlegung nicht einbezogen werden, da sich dieser bereits durch die Distanzmessung ergibt.
Die Festlegung des Datums kann zum Beispiel erfolgen, indem ein beliebiger Neupunkt fingierte Objektkoordinaten (X_1, Y_1, Z_1) erhält. Dadurch werden die drei Translationen festgelegt. Durch die weiteren fingierten globalen Koordinaten in zwei Koordinatenrichtungen (z.B.: Y_2, Z_2) eines weiteren Neupunktes, wird die Richtung von einer der drei Koordinatenachsen festgelegt. Dies entspricht der Festlegung zweier Drehwinkel. Der verbleibende Drehwinkel wird

durch die Festlegung einer Objektkoordinate in einer Koordinatenrichtung (z.B.: Z_3) eines weiteren Punktes, der nicht auf der Geraden der ersten beiden Punkte liegt, definiert. Damit sind alle Freiheitsgrade des Koordinatensystems fixiert.
Eine andere Möglichkeit der Datumsfestlegung ist die Vorgabe der äußeren Orientierung (Position des Projektionszentrums und Stellung des Bildes im Objektkoordinatensystem), beispielsweise für die erste Aufnahme.

3.3.2 Näherungsweise Orientierung

Für den nachfolgenden ICP Algorithmus (siehe Abschnitt 3.3.3.2) ist es notwendig, die näherungsweise Orientierung der *statischen Bilder* zu kennen. Dies erfolgt nach der Datumsfestlegung (siehe Abschnitt 3.3.1) mit Hilfe der bewegten Sequenz, welche zwischen den jeweiligen statischen Aufnahmen aufgezeichnet wird. Bei dem in dieser Arbeit vorgestellten Ansatz wird die Orientierung ausgehend vom ersten Bild der bewegten Sequenz für alle Bilder der Sequenz bestimmt. Dazu werden zuerst Punktkorrespondenzen in den Amplitudenbildern gesucht, und anschließend für das erste Bild der Sequenz mit Hilfe der Distanzinformation die Objektkoordinaten der Verknüpfungspunkte bestimmt. Nun kann für das nächste Bild mittels eines räumlichen Rückwärtsschnittes, ausgehend von den bekannten Objektkoordinaten der Verknüpfungspunkte, die Orientierung ermittelt werden. Dabei werden als Näherungswerte für die Orientierung die Werte des vorherigen Bildes verwendet. Bei vom gleichen Standpunkt aufgenommenen Sequenzen wird der gemeinsame Drehpunkt als Exzenter beim Rückwärtsschnitt einbezogen. Anschließend werden die Objektkoordinaten von den in diesem Bild neuen Verknüpfungspunkten berechnet. Dies wird nun sukzessive für alle weiteren Bilder durchgeführt bis die Orientierung des letzten Bildes der bewegten Sequenz bestimmt ist. Diese Orientierung entspricht nun auch jener des *statischen Bildes*, welches nach der bewegten Sequenz aufgezeichnet wurde. Der prinzipielle Ablauf der näherungsweisen Orientierung wird in Abbildung 3.5 zusammengefasst.

3.3.2.1 Verknüpfungspunkte

Zur Bestimmung von potenziell geeigneten Verknüpfungspunkten kann der *Scale-Invariant Feature Transform* (SIFT) bzw. der *Speeded Up Robust Features* (SURF) Algorithmus verwendet werden. Die Algorithmen bestimmen für markante Punkte (Feature-Punkte) in den Grau-

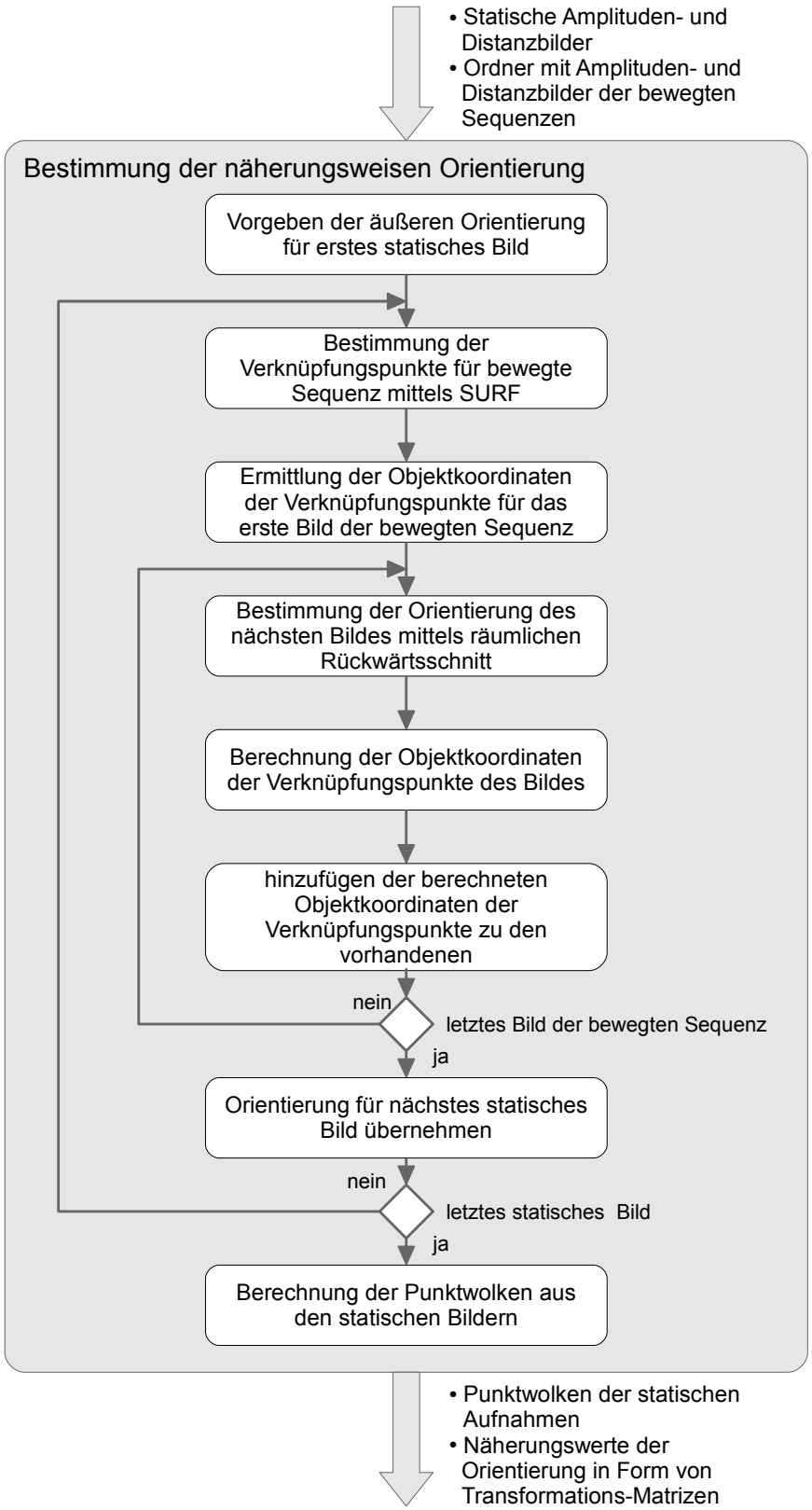


Abbildung 3.5: Ablauf der näherungsweisen Orientierung

wertbildern[1] Merkmalsvektoren (Feature-Deskriptoren) mit denen es möglich ist, Punktkorrespondenzen zu finden. Die beiden Algorithmen werden in Folge in Abschnitt 3.3.2.1.1 bzw. in Abschnitt 3.3.2.1.2 vorgestellt.

3.3.2.1.1 Scale-Invariant Feature Transform (SIFT)

Der Ansatz kommt aus dem Bereich der Computer Vision. Er wurde 1999 von Lowe [1999] veröffentlicht. Die bei der *Scale-Invariant Feature Transform* verwendeten Feature-Deskriptoren sind invariant gegenüber Skalierung und Rotation. Weiters ermöglichen sie ein robustes Matchen bei einer Veränderung der Beleuchtung und des Aufnahmeortes. Die Berechnung der *Features* erfolgt in folgenden Schritten [Lowe, 2004]:

1. Suche von Extremwerten im *scale space*
2. Subpixel-genaue Lokalisierung der Feature-Punkte
3. Bestimmung der Orientierung der Feature-Punkte
4. Berechnung der Feature-Deskriptoren

Suche von Extremwerten im *scale space*

Der erste Schritt ist die Suche nach Feature-Punkten, welche auch bei unterschiedlichen Ansichten des gleichen Objektes mit hoher Wahrscheinlichkeit wieder gefunden werden. Um eine Skalierungsinvarianz zu erreichen, werden die Bilder in den *scale space* transformiert [Lowe, 1999]. Der *scale space* eines Bildes ist definiert als Funktion $L(x, y, \sigma)$, welche sich mittels einer Faltung einer zweidimensionalen Gaußfunktion (Gauß-Filter) $G(x, y, \sigma)$ mit unterschiedlichen Skalierungsparametern σ und dem Bild $I(x, y)$ ergibt [Lowe, 2004]:

$$L(x, y, \sigma) = G(x, y, \sigma) ** I(x, y) \tag{3.4}$$

Dabei bedeutet $**$ eine Faltung in beide Koordinatenrichtungen (x und y). Die zweidimensionale Gaußfunktion $G(x, y, \sigma)$ berechnet sich für unterschiedliche Skalierungsparameter σ aus:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3.5}$$

[1] in dieser Arbeit werden Amplitudenbilder als Grauwertbilder für den SIFT bzw. SURF Algorithmus verwendet.

Um nun stabile Feature-Punkte zu erhalten führt Lowe [1999] eine Extremwertsuche im *scale space* von *difference-of-Gaussians* (DoG) $D(x, y, \sigma)$ durch, wobei sich $D(x, y, \sigma)$ als Faltung der Differenz zweier Gauß-Filter mit dem Bild $I(x, y)$ ergibt:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * *I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \tag{3.6}$$

D kann dabei direkt aus den Differenzen von zwei geglätteten Bildern L, die einen um den Faktor k verschiedenen Skalierungsparameter σ haben, berechnet werden.

Die Suche nach Extremwerten erfolgt pixelweise in einer 26er-Nachbarschaft. Dabei befinden sich die 8 benachbarten Pixel im gleichen DoG-Bild und jeweils 9 Pixel im *scale space* im darüber bzw. darunter liegenden DoG-Bild. Ist das aktuelle Pixel nun größer bzw. kleiner als alle umgebenden Pixel, so liegt ein potentieller Kandidat für einen Feature-Punkt vor.

Der Vorteil bei der Suche im *scale space* liegt nun darin, dass der Extremwert bei einer Skalierung, welche der Struktur des Objekts angepasst ist, gefunden wird. Das heißt der Feature-Punkt ist vom Objekt abhängig und weniger vom Bildmaßstab.

Subpixel-genaue Lokalisierung der Feature-Punkte

Für die oben gefundenen potentiellen Feature-Punkte wird nun in weiterer Folge deren Position im Bild und die Skalierung bestimmt. Anschließend werden die Punkte mit einem schlechten Kontrast bzw. jene, die auf einer Kante liegen, eliminiert.

Zur Bestimmung der Subpixelposition und Skalierung des Extremwertes kann eine Taylor-Entwicklung zweiten Grades der DoG Funktion $D(x, y, \sigma)$ verwendet werden, deren Extremum durch Ableitung leicht gefunden werden kann. Der Funktionswert an der Position des Feature-Punktes $D(\hat{x})$ kann verwendet werden, um Punkte mit schwachem Kontrast zu finden. Liegt $|D(\hat{x})|$ unter einem Schwellwert (Lowe [2004] schlägt 0,03 vor), so wird dieser Punkt verworfen.

In einem weiteren Schritt sollen Punkte, die auf einer Kante liegen, detektiert werden. Auch diese Punkte sind keine geeigneten Feature-Punkte, da sie in Richtung der Kante schlecht bestimmt sind. Zum Auffinden solcher Punkte kann die Hesse-Matrix H verwendet werden [Lowe, 2004]. Die Eigenwerte von H sind proportional zu den Hauptkrümmungen der Funktion. Die Hauptkrümmungen sind in Richtung einer Kante kleiner, normal darauf aber größer. Daraus lässt sich

ein Kriterium ableiten, ab wann ein Punkt zu einer Kante gezählt wird:

$$\frac{trace(H)^2}{det(H)} < \frac{(r+1)^2}{r} \tag{3.7}$$

Lowe [2004] verwendet als Grenzwert $r = 10$. Das bedeutet, dass jene Feature-Punkte, bei denen das Verhältnis zwischen den beiden Hauptkrümmungen größer 10 ist, eliminiert werden.

Bestimmung der Orientierung der Feature-Punkte

Durch die Bestimmung einer Orientierung für die zuvor erhaltenen Feature-Punkte, kann anschließend der Feature-Deskriptor relativ zu dieser bestimmt werden, wodurch die Invarianz gegenüber einer Bildrotation erreicht wird. Zur Bestimmung der Orientierung eines Feature-Punktes verwendet Lowe [2004] den nun folgenden Ansatz.
Aus dem geglätteten Bild $L(x, y, \sigma)$ wird eine Gradientenrichtung $\theta(x, y)$ und ein Gradientenbetrag (Magnitude) $m(x, y)$ bestimmt.

$$m(x,y) = \sqrt{(L(x+1,y,\sigma) - L(x-1,y,\sigma))^2 + (L(x,y+1,\sigma) - L(x,y-1,\sigma))^2} \tag{3.8}$$

$$\theta(x,y) = arctan\frac{L(x,y+1,\sigma) - L(x,y-1,\sigma)}{L(x+1,y,\sigma) - L(x-1,y,\sigma)} \tag{3.9}$$

Für $L(x, y, \sigma)$ wird der Skalierungsparameter σ, den der Feature-Punkt besitzt, verwendet. Mittels $m(x, y)$ und $\theta(x, y)$ wird für einen Bereich um den Feature-Punkt ein Orientierungshistogramm erstellt. Lowe [2004] verwendet eine Klassenbreite von 10°, dadurch entstehen für den Vollkreis 36 Orientierungsklassen. Die Zuteilung zu einer Klasse erfolgt über die Gradientenrichtung $\theta(x, y)$. Der Gradientenbetrag $m(x, y)$ wird mittels einer Gaußfunktion gewichtet und zur entsprechenden Klasse addiert. Dadurch werden Gradienten in der Nähe des Feature-Punktes stärker berücksichtigt. Der Feature-Punkt erhält nun die Orientierung des Maximalwerts des Histogramms.

Berechnung der Feature-Deskriptoren

Bisher wurden für die Feature-Punkte die Position, eine Skalierung und eine Orientierung ermittelt. Nun soll ein Deskriptor für jeden Feature-Punkt bestimmt werden, der es ermöglicht diesen eindeutig zu beschreiben, sodass in weiterer Folge Punktkorrespondenzen in den Bildern bei unterschiedlichen Ansichten und Beleuchtungen gefunden werden können. Dazu werden wiederum die Gradientenrichtungen und die Gradientenbeträge für die Pixel in der Umgebung des

Feature-Punktes bestimmt. Um nun Rotationsinvarianz zu erreichen erfolgt die Berechung der Gradienten und in weiterer Folge die Bestimmung des Deskriptors nicht im Bildkoordinatensystem, sondern in einem Koordinatensystem, welches in Richtung der Orientierung gedreht ist. Die Gradientenbeträge werden wiederum mittels einer Gaußfunktion gewichtet, sodass Punkte in der Nähe des Feature-Punktes stärker berücksichtigt werden. Der Deskriptor wird aus Orientierungshistogrammen aufgebaut. Abbildung 3.6 zeigt einen 2x2 Deskriptor, wobei ein Orientierungshistogramm von einer 4x4 px^2 Region bestimmt wird. Ein Histogramm besteht aus 8 Orientierungsklassen. Es sei erwähnt, dass die Zuordnung zu einer Klasse nicht streng über die Klassengrenzen erfolgt, sondern ein Gradient wird gewichtet zu mehreren Klassen hinzugefügt. Dies dient der Vermeidung von Randeffekten. Versuche zeigten, dass ein 4x4 Deskriptor optimal zur Beschreibung eines Feature-Punktes ist [Lowe, 2004]. Damit besitzt ein Feature-Deskriptor 4x4x8 = 128 Elemente.
Der Deskriptor wird schließlich normiert, wodurch dieser invariant gegenüber konstanten Helligkeitsunterschieden wird.

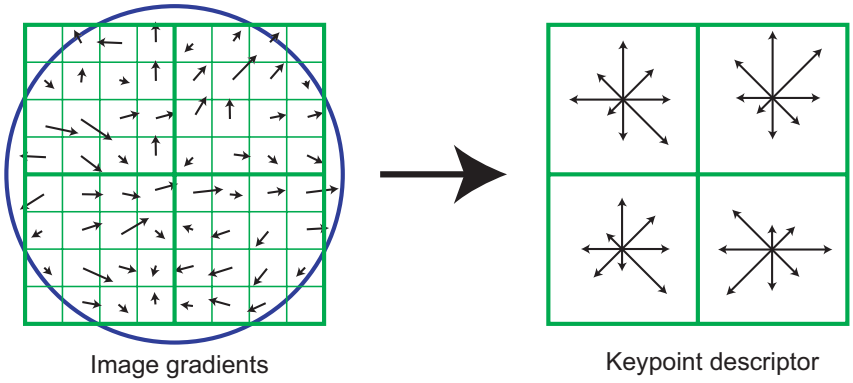


Abbildung 3.6: Feature-Deskriptor: die rechte Abbildung zeigt einen 2x2 Deskriptor, ein Orientierungshistogramm wird von einer 4x4 Region bestimmt und besteht aus 8 Klassen. Quelle: [Lowe, 2004]

Matching von Feature-Punkte in zwei Bildern

Als Ähnlichkeitsmaß zweier Feature-Punkte dient die euklidische Distanz ihrer Deskriptoren. Dabei werden Paare mit der geringsten euklidischen Distanz gesucht, diese sind potentielle Punktkorrenspondenzen. Damit Fehlzuordnungen verringert werden, schlägt Lowe [2004] die Verwendung des Quotienten aus nächster Nachbarschaft und übernächster Nachbarschaft

(kleinste und zweit kleinste euklidischen Distanz) vor. Ist dieser größer als 0,8, so wird die Punktkorrespondenz verworfen. Es können auch noch weitere Kriterien verwendet werden. Ist zum Beispiel im Vorfeld bekannt, dass sich der Aufnahmeort und die Aufnahmerichtung zwischen den beiden Bildern nicht sehr stark verändert, können auch die Position, Skalierung und/oder Orientierung der Feature-Punkte als Kriterium verwendet werden (siehe Abschnitt 4.3.2).

3.3.2.1.2 Speeded Up Robust Features (SURF)

Der SURF Algorithmus ist dem SIFT Algorithmus sehr ähnlich. Er besteht im wesentlichen aus drei Schritten [Bay et al., 2008]:

1. *Interest Point Detection*: Dies erfolgt bei SURF mittels der Hesse-Matrix $H(x, y, \sigma)$, wobei σ wiederum der Skalierungsparameter ist. Beim SIFT Algorithmus entspricht dies der Suche von Extremwerten im *scale space*.

2. *Interest Point Description*: Die Beschreibung der *Interest Points* (Feature-Punkte) erfolgt wiederum über einen Feature-Vektor (Feature-Deskriptor), welcher sich aus den benachbarten Pixeln berechnet. Die Bestimmung des Deskriptors erfolgt mit Hilfe von Haar-Wavelets, wozu der SIFT-Algorithmus die Gradienten der umgebenden Pixel verwendet.

3. *Matching*: Der letzte Schritt ist die Suche nach homologen Bildpunkten. Dies erfolgt über die zuvor bestimmten Deskriptoren. Laut Bay et al. [2008] ist bei SURF eine Deskriptorlänge von 64 ausreichend. Im Gegensatz dazu ist bei SIFT die optimale Länge des Deskriptorvektors 128, weshalb auch bei SURF die Suche nach Punktkorrespondenzen wesentlich schneller verläuft als bei SIFT. Bay et al. [2006] schlagen vor, dass der Quotient aus nächster und übernächster Nachbarschaft kleiner als 0.7 sein muss, damit die Punktkorrespondenz akzeptiert wird.

Der Vorteil von SURF gegenüber SIFT liegt in der wesentlich kürzeren Rechenzeit. Zudem ist laut Bay et al. [2008] diese Methode auch robuster.

3.3.2.2 Ermittlung der Objektkoordinaten

Bei den Distanzkameras ist es bei bekannter Orientierung des Bildes möglich mittels der für jedes Pixel gemessenen Distanzinformation direkt von den Bildkoordinaten auf die Objektkoordinaten zu schließen. Dieser Zusammenhang ist in Abbildung 3.7 dargestellt und wird in

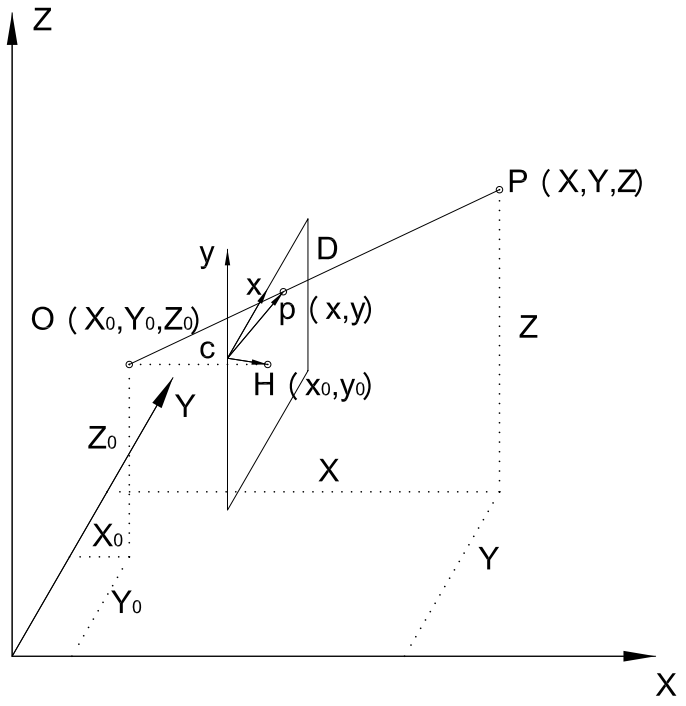


Abbildung 3.7: Zusammenhang Bildkoordinaten und Objektkoordinaten: über die Distanz D kann bei bekannter innerer Orientierung (Kamerakonstante c und Lage des Hauptpunktes $H(x_0, y_0)$) und äußerer Orientierung (Objektkoordinaten des Projektionszentrums $O(X_0, Y_0, Z_0)$ und Drehmatrix R) von den Bildkoordinaten $p(x, y)$ auf die Objektkoordinaten $P(X, Y, Z)$ zurückgeschlossen werden.

Gleichung 3.10 mathematisch beschrieben. Die gemessene Distanz bezieht sich auf das Projektionszentrum.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + \boldsymbol{R} \begin{pmatrix} x - x_0 \\ y - y_0 \\ -c \end{pmatrix} \frac{D}{\sqrt{(x - x_0)^2 + (y - y_0)^2 + c^2}} \tag{3.10}$$

$\boldsymbol{X}$	Objektkoordinaten des Objektpunktes
$\boldsymbol{X}_0$	Objektkoordinaten des Projektionszentrums
$\boldsymbol{R}$	Drehmatrix
$\boldsymbol{x}$	Bildkoordinaten des Objektpunktes
$\boldsymbol{x}_0$	Bildkoordinaten des Bildhauptpunktes
c	Kamerakonstante
D	gemessene Distanz

3.3.2.3 Berechnung der äußeren Orientierung mittels Rückwärtsschnitt

Eine einfache Möglichkeit, die genäherte Orientierung zu berechnen wäre eine überbestimmte räumliche Ähnlichkeitstransformation (mit Maßstab = 1) der Verknüpfungspunkte. Jedoch würde dies auch die Distanzbeobachtungen des zu orientierenden Bildes berücksichtigen. Da aber bei der näherungsweisen Orientierung die Bilder der bewegten Sequenz verwendet werden und diese ein hohes Distanzrauschen aufweisen, wird in dieser Arbeit der räumliche Rückwärtsschnitt verwendet da dieser die Distanzbeobachtungen des zu orientierenden Bildes nicht benötigt (die Distanzbeobachtungen werden nur zur Berechnung der Objektkoordinaten der Verknüpfungspunkte von bereits orientierten Bildern verwendet). Die Orientierung mittels des räumlichen Rückwärtsschnittes wird nun in weiterer Folge beschrieben.
Mittels eines räumlichen Rückwärtsschnittes kann die äußere Orientierung der Kamera, sprich deren Position und Stellung im Raum, bestimmt werden. Dazu muss, neben mindestens drei Objektpunkten mit den zugehörigen Bildkoordinaten, auch die innere Orientierung bekannt sein. Die Elemente der inneren Orientierung bestehen aus der Kamerakonstante c und den Bildkoordinaten des Bildhauptpunktes x_0 und y_0.
Der Zusammenhang zwischen den Objektpunkten und den Bildpunkten wird über die Kollinearitätsgleichungen (Gleichung 3.11) mathematisch beschrieben [Kraus, 1996].

$$\begin{aligned} x &= \bar{x} + v_x = x_0 - c\frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \\ y &= \bar{x} + v_x = y_0 - c\frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \end{aligned} \tag{3.11}$$

Dabei sind X, Y und Z die Koordinaten der Objektpunkte. Bei $\bar{x}$ und $\bar{y}$ handelt es sich um die beobachteten Bildkoordinaten mit ihren Verbesserungen v_x und v_y. r_{11}, r_{12}, usw. sind die Elemente der Drehmatrix **R**. Diese legt die Stellung des Bildes im Objektkoordinatensystem fest und enthält die Unbekannten α, ν und κ [2]. X_0, Y_0 und Z_0 sind die Koordinaten des Aufnahmestandpunktes und somit ebenfalls Unbekannte.
Die Lösung eines überbestimmten, räumlichen Rückwärtsschnittes kann mittels einer Ausgleichung nach vermittelnden Beobachtungen erfolgen. Dazu müssen die Gleichungen linearisiert

[2] Bei terrestrischen Bildern kommen, um numerische Probleme zu vermeiden, in erster Linie die Drehwinkel α, ν und κ zum Einsatz (siehe [Kraus, 1996]). Bei der in dieser Arbeit verwendeten Implementierung wird auch die Drehmatrix mit den Winkeln ω, ϕ und κ verwendet, da diese zum Beispiel bei der Aufnahme der Decke numerische Vorteile hat. Der Einfachheit halber beziehen sich die Formeln in der Arbeit auf die Drehwinkeln α, ν und κ.

werden, weshalb zur Bestimmung der sechs Unbekannten Näherungswerte benötigt werden. Bei den in dieser Arbeit verwendeten Entfernungsvideos ist die Änderung der Position und der Stellung der Kamera zwischen zwei Aufnahmen gering, deshalb können als Näherungswerte die Werte des vorherigen Bildes verwendet werden. Die linearisierten Verbesserungsgleichungen erhält man durch Ableitung der Gleichung 3.11 nach den Unbekannten X_0, Y_0, Z_0, α, ν und κ.

$$\begin{aligned} v_x &= a_{11}dX_0 + a_{12}dY_0 + a_{13}dZ_0 + a_{14}d\alpha + a_{15}d\nu + a_{16}d\kappa - (\bar{x} - x^0) \\ v_y &= a_{21}dX_0 + a_{22}dY_0 + a_{23}dZ_0 + a_{24}d\alpha + a_{25}d\nu + a_{26}d\kappa - (\bar{y} - y^0) \end{aligned} \tag{3.12}$$

Die Koeffizienten a sind die Differentiale der Kollinearitätsgleichung. Siehe Gleichung A.1 im Anhang A. x^0 und y^0 sind die gerechneten Bildkoordinaten, die sich aus der Gleichung 3.11 mit Hilfe der Näherungswerte ergeben. Die Differentiale lassen sich zur Designmatrix $\boldsymbol{A}$ zusammenfassen. Für die beobachteten Bildkoordinaten kann eine Gewichtung eingeführt werden, die zur Gewichtsmatrix $\boldsymbol{P}$ zusammengefasst werden. In dieser Arbeit werden jene Bildpunkte, die bereits mehrmals gefunden wurden, höher gewichtet. Dies trägt zur Stabilität der näherungsweisen Orientierung bei (siehe Abschnitt 4.3.2).
Die Bestimmung der Unbekannten kann, wie bereits erwähnt, mittels Ausgleichung nach vermittelnden Beobachtungen erfolgen.

$$\hat{\boldsymbol{x}} = \left(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{l} \tag{3.13}$$

Dabei ist hier $\boldsymbol{l}$ der reduzierte Beobachtungsvektor, das heißt $\boldsymbol{l}$ ist die Differenz aus den beobachteten, und den mit 3.11 berechneten Bildkoordinaten. Weiters sind $\hat{\boldsymbol{x}}$ die Parameterzuschläge zu den Näherungswerten [Kraus, 2004].
Die verwendeten Bildpunkte sind die in Abschnitt 3.3.2.1 ermittelten Verknüpfungspunkte. Der Berechnungsablauf der zugehörigen Objektpunkte ist in Abschnitt 3.3.2.2 beschrieben. Bei der Bestimmung der Verknüpfungspunkte werden einige Punkte in den Bildern falsch zugeordnet. Dadurch stimmen für diese Punkte auch die berechneten Objektkoordinaten nicht zu den Bildkoordinaten. Diese grob falschen Beobachtungen verfälschen das Ergebnis der Ausgleichung. Nun sind aber wesentlich mehr Punkte vorhanden, als zur Bestimmung des räumlichen Rückwärtsschnittes notwendig sind. Deswegen ist es möglich mittels einer robusten Ausgleichung diese grob falschen Beobachtungen zu eliminieren. Die in dieser Arbeit verwendete ro-

buste Ausgleichung funktioniert derart, dass die Kleinste-Quadrate-Ausgleichung in mehreren Iterationsschritten angewendet wird und von Iterationsschritt zu Iterationsschritt eine Gewichtsmodifikation durchgeführt wird. Dabei werden jene Beobachtungen mit größeren Verbesserungen, also solche die potentiell grob falsch sind, geringer gewichtet. Somit wird der Einfluss der grob falschen Beobachtungen mit jedem Iterationsschritt geringer. Dieses Verfahren kann eingesetzt werden, wenn die grob falschen Beobachtungen nicht überwiegen und ist umso erfolgreicher, je größer die Redundanz ist [Kraus, 1996].
Wie eingangs bereits erwähnt bzw. in Kapitel 4 noch ausführlicher beschrieben, erfolgt die Aufnahme mittels eines Stativs. Dabei werden von einem Standpunkt aus Aufnahmen in unterschiedliche Richtungen durchgeführt. Die Kamera wird dazu um einen Punkt abseits des Projektionszentrums gedreht, der für alle Aufnahmen eines Standpunktes derselbe ist. Diese Information kann nun in der vermittelnden Ausgleichung berücksichtigt werden, indem die zusätzlichen Beobachtungen mit einer sehr großen Gewichtung hinzugefügt werden. Die Koordinaten $\boldsymbol{x}_{MP}$ des Drehpunktes sind im Kamerakoordinatensystem für alle Aufnahmen immer die gleichen. Die Objektkoordinaten des Drehpunktes $\boldsymbol{X}_{MP}$ ändern sich, wie bereits erwähnt, ebenfalls nicht. Die beiden Koordinatensysteme können über eine Ähnlichkeitstransformation miteinander verknüpft werden.

$$\boldsymbol{x}_{MP} = \boldsymbol{R}^T \left(\boldsymbol{X}_{MP} - \boldsymbol{X}_0 \right) \tag{3.14}$$

Die linearisierten Verbesserungsgleichungen ergeben sich wiederum durch Ableitung der Gleichung 3.14 nach den Unbekannten X_0, Y_0, Z_0, α, ν und κ.

$$\begin{aligned} v_x &= b_{11}dX_0 + b_{12}dY_0 + b_{13}dZ_0 + b_{14}d\alpha + b_{15}d\nu + b_{16}d\kappa - (\bar{x} - x^0) \\ v_y &= b_{21}dX_0 + b_{22}dY_0 + b_{23}dZ_0 + b_{24}d\alpha + b_{25}d\nu + b_{26}d\kappa - (\bar{y} - y^0) \\ v_z &= b_{31}dX_0 + b_{32}dY_0 + b_{33}dZ_0 + b_{34}d\alpha + b_{35}d\nu + b_{36}d\kappa - (\bar{z} - z^0) \end{aligned} \tag{3.15}$$

Die Koeffizienten b können aus dem Anhang A Gleichung A.2 entnommen werden. Die Einführung dieser zusätzlichen Bedingung trägt wesentlich zur Stabilität bei der Berechnung der näherungsweisen Orientierung bei.

3.3.3 Fein-Orientierung

Zur Fein-Orientierung wird der ICP (*Iterative Closest Point*) Algorithmus verwendet. Dieser ermöglicht die Orientierung der aus den statischen Distanzbildern berechneten Punktwolken ohne signalisierte Verknüpfungspunkte. Der ICP wird in Abschnitt 3.3.3.2 beschrieben. Für den ICP werden Normalvektoren der Punkte der Punktwolken benötigt. Zudem ist es notwendig, die Punktwolken vor der Fein-Orientierung zu filtern, d.h. einen Teil der Punkte zu verwerfen. Die Normalvektorenberechnung und Filterung wird in Abschnitt 3.3.3.1 behandelt.

3.3.3.1 Normalvektorenberechnung und Filterung

Die Normalvektoren werden zum einen für den ICP-Algorithmus benötigt, zum anderen sind sie Grundlage für einige Filterungsarten (Korona-Filter, Rauheits-Filter).

Normalvektorenberechnung
Die Bestimmung der Flächennormale n kann durch eine ausgleichende Ebene, welche die umgebenden Punkte eines gegebenen Datenpunktes hinreichend annähert, erfolgen. Diese Ebene wird durch den Ausgangspunkt und einige Datenpunkte d in dessen Umgebung bestimmt [Haring, 2007]. Als Datenpunkte können die in einer 3D-Punktwolke k nächstgelegenen Punkte verwendet werden, oder auch die Punkte innerhalb einer gewissen Distanz. Bei Distanzbildern kann auch die Ordnung der Punkte in der Bildmatrix berücksichtigt werden.
Die Orientierung des Normalvektors ist auch bei fester Auswahl der Datenpunkte und Normierung des Vektors im Allgemeinen zweideutig: der Gegenvektor eines gültigen Normalvektors ist ebenfalls ein gültiger Normalvektor. Diese Zweideutigkeit kann behoben werden, indem eine Orientierung in Richtung des Aufnahmezentrums vorgegeben wird. Dazu ist gegebenenfalls ein Vorzeichenwechsel notwendig.

$$n = \begin{cases} -n & \text{falls } n^T d_0 > 0 \\ n & \text{sonst} \end{cases} \tag{3.16}$$

d_0 ist der Vektor vom Aufnahmezentrum zum Datenpunkt d.

Filterung
Die in diesem Abschnitt behandelte Filterung wird auf die 3D-Punktwolke im Zuge der Berechnung der Normalvektoren angewendet. Diese wird mittels Gleichung 3.10 aus den Distanzbil-

dern bestimmt.
Durch die Filterung sollen einerseits Punkte, die außerhalb des Eindeutigkeitsbereiches der Kamera liegen, verworfen werden. Andererseits sollen auch Punkte, die in einer nicht glatten Oberfläche liegen entfernt werden, da in diesem Bereich der Normalvektor nicht zuverlässig abgeschätzt werden kann und es dadurch beim 3D-Matching zu falschen Ergebnissen kommen kann [Haring, 2007].
Folgende Filter kommen im Rahmen dieser Arbeit zum Einsatz:

- *Rauheits-Filter*: Im Zuge der Abschätzung des Normalvektors werden drei Eigenvektoren berechnet, aus diesen lässt sich ein Kriterium für die „Rauheit“ der Oberfläche im Bereich des Punktes für den der Normalvektor bestimmt wurde, aufstellen [Haring, 2007]:

$$r = \sqrt{\frac{I3}{I1}} \tag{3.17}$$

  r ist ein Parameter für die Rauheit, $I3$ der kleinste und $I1$ der größte Eigenwert. Punkte die außerhalb eines vorgegebenen Wertebereichs für r liegen, werden verworfen.

- *Distanz-Filter*: Es kann eine minimale und eine maximale Distanz zwischen Kamera und Objektpunkt angegeben werden. Punkte mit Distanzen außerhalb dieser Werte werden gefiltert. Punkte mit einer geringen Distanz sind, bei den in dieser Arbeit verwendeten Daten, häufig Punkte deren Abstand zur Kamera gerade nicht mehr im Eindeutigkeitsbereich der Kamera liegen und deshalb nicht die Objektoberfläche wiedergeben.

- *Amplituden-Filter*: Durch die Vorgabe eines minimalen Intensitätswertes werden die Punkte, welche diesen Wert unterschreiten, gefiltert. Punkte mit einem geringen Amplitudenwert weisen ein deutlich höheres Rauschen auf (siehe Abschnitt 2.3). Zudem haben Punkte, die sich nicht mehr im Eindeutigkeitsbereich der Distanzkamera befinden eine geringe Amplitude und können somit auch teilweise mittels des Amplituden-Filters eliminiert werden.

- *Korona-Filter*: Bei diesem Filter werden jene Punkte entfernt, deren Normalvektor n mit dem Vektor vom Aufnahmezentrum zum Datenpunkt d_0 einen Winkel einschließt, der einen vorgegebenen Schwellwert überschreitet. Das heißt es werden jene Punkte gefiltert, die bei einem schleifenden Winkel aufgezeichnet wurden, da bei diesen Punkten der Fehler bei der Distanzmessung größer ist. Dieser Filter eliminiert auch Punkte auf bzw. in der Umgebung von Kanten. Bei diesen Punkten tritt häufig ein Fehler auf, der als *mixed pixels*

bezeichnet wird. Dabei stammt das von einem Pixel empfangene Signal von zwei unterschiedlichen Flächen, wodurch es zu grob falschen Distanzmessungen kommen kann.

Die Anwendung dieser Filter und deren Auswirkungen auf die Aufnahmen werden in Abschnitt 4.3.3 beschrieben.

3.3.3.2 ICP-Algorithmus

Der ICP (*Iterative Closest Point*) Algorithmus ermöglicht eine Orientierung von Punktwolken ohne signalisierte Verknüpfungspunkte. Wie der Name schon sagt, handelt es sich dabei um ein iteratives Verfahren. In dieser Arbeit werden dazu als Startwerte die in Abschnitt 3.3.2 bestimmten Näherungswerte verwendet.
Es sollen nun zunächst die Voraussetzungen und das Grundprinzip für die Orientierung von zwei Punktwolken vorgestellt werden. Anschließend werden verschiedene Varianten des ICP Algorithmus präsentiert. Schließlich erfolgt der Übergang auf die Orientierung von mehr als zwei Punktwolken.

Grundprinzip des ICP

Eine ursprüngliche Aufgabe des ICP Algorithmus ist die Orientierung von gescannten Daten eines Objekts in Bezug auf ein Modell, welches bereits im übergeordneten Koordinatensystem vorliegt [Besl u. McKay, 1992]. Diese Aufgabenstellung tritt zum Beispiel bei der Qualitätskontrolle von Bauteilen, welche nach einem CAD-Modell gefertigt wurden, auf. Diese Bauteile werden nach der Fertigung gescannt und sollen mit dem CAD-Modell verglichen werden. Dazu müssen die gescannten Daten ins Modellkoordinatensystem transformiert werden. Aufgrund dieser ursprünglichen Anwendung ergeben sich die Begriffe Modell-Punktwolke und Daten-Punktwolke, welche in Folge verwendet werden.
Im ersten Schritt wird zu jedem Datenpunkt jener Modellpunkt gesucht, der den kürzesten euklidischen Abstand hat. Dadurch entsteht eine große Menge von Punktepaaren, mit denen die Transformation der Datenpunkte in Richtung des Modellkoordinatensystems, zum Beispiel mit einer stark überbestimmten räumlichen Ähnlichkeitstransformation, erfolgen kann [Kraus, 2004]. Dadurch nähern sich die beiden Punktwolken einander an. Mit den transformierten Datenpunkten und den ursprünglichen Modellpunkten erfolgt in gleicher Weise der nächste Iterationsschritt. Dies wird solange durchgeführt, bis ein Abbruchkriterium erfüllt ist.
Damit der ICP Algorithmus zum gewünschten Ergebnis konvergiert, müssen folgende Voraus-

setzungen erfüllt sein. Zum einen müssen hinreichend gute Näherungswerte vorhanden sein, damit der Algorithmus zum globalen Minimum konvergiert und nicht bei einem lokalen Minimum „steckenbleibt“. Zum anderen müssen die Objekte ausreichend unregelmäßige Oberflächenstrukturen aufweisen. Würden beispielsweise die beiden Punktwolken nur in einer Ebene liegen, die angenommen parallel zur xy-Ebene ist, so bleiben bei der Orientierung drei Freiheitsgrade übrig. Das heißt, die Datenpunktwolke kann beliebig in x- und y-Richtung verschoben und um die z-Achse gedreht werden, ohne dass sich das Ergebnis der Fehlerfunktion wesentlich ändert[3]. Weitere geometrische Grundformen und deren Freiheitsgrade sind in Abbildung 3.8 dargestellt.

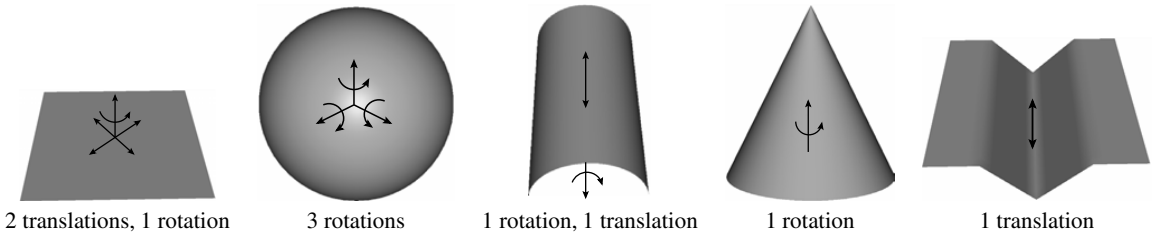


Abbildung 3.8: geometrische Grundformen und deren Freiheitsgrade. Quelle: [Gelfand et al., 2003]

Varianten des ICP Algorithmus

Seit der Veröffentlichung des ursprünglichen ICP Algorithmus ([Besl u. McKay, 1992; Chen u. Medioni, 1991]) wurde eine Vielzahl von Varianten publiziert. Rusinkiewicz u. Levoy [2001] klassifizieren die Varianten nach folgenden 6 Gesichtspunkten:

1. Selektion der Punkte
2. Bestimmung der Korrespondenzen
3. Gewichtung der korrespondierenden Punktepaare
4. Kriterien zur Verwerfung von Punktepaaren
5. Verwendete Fehlermetrik
6. Minimierung der Fehlermetrik

[3] vorausgesetzt die Punkte sind zufällig angeordnet bzw. es wird die Punkt-zu-Ebene Methode verwendet

Selektion der Punkte

Bei der ursprünglichen Version des ICP Algorithmus werden alle Datenpunkte verwendet. Dies hat den Nachteil, dass insbesondere bei großen Datenmengen die Rechenzeit sehr groß werden kann. Deswegen ist es in diesen Fällen notwendig, die Punktwolken auszudünnen. Eine sehr einfache Möglichkeit die Datenmenge zu reduzieren besteht darin, dass nur jeder k-te Punkt weiter verwendet wird. Es handelt sich hierbei um eine regelmäßige Ausdünnung. Diese hat den Nachteil, dass kleine Strukturen verschwinden können. Eine weitere einfache Möglichkeit ist die zufällige Ausdünnung. Dabei wird eine bestimmte Anzahl von zufällig gewählten Punkten beibehalten. Dies ermöglicht eine Ausdünnung auf eine vorgegebene Punktanzahl. Jedoch werden auch bei dieser Methode kleinräumige Strukturen nur mangelhaft wiedergegeben. Zudem ist die zufällige Ausdünnung nicht reproduzierbar [Haring, 2007].

Ein weiteres Verfahren, welches sich besonders bei Ebenen mit kleinräumigen Strukturen bewährt hat, wurde von Rusinkiewicz u. Levoy [2001] vorgestellt. Bei diesem Ansatz, der als *normal-space sampling* bezeichnet wird, werden für alle Punkte deren Normalvektoren auf der Einheitskugel betrachtet. Die Einheitskugel wird in gleich große Bereiche unterteilt und für jeden Bereich wird eine maximale Anzahl von Punkten festgelegt. Übersteigt in einem Bereich die Punkteanzahl diesen Wert, so werden die Punkte in diesem Bereich reduziert. Dadurch erfolgt eine Reduktion jener Punkte, deren Normalvektoren eine gängige Richtung besitzen während jene Punkte, deren Normalvektoren eine selten vorkommende Richtung haben, erhalten bleiben.

Bestimmung der Korrespondenzen

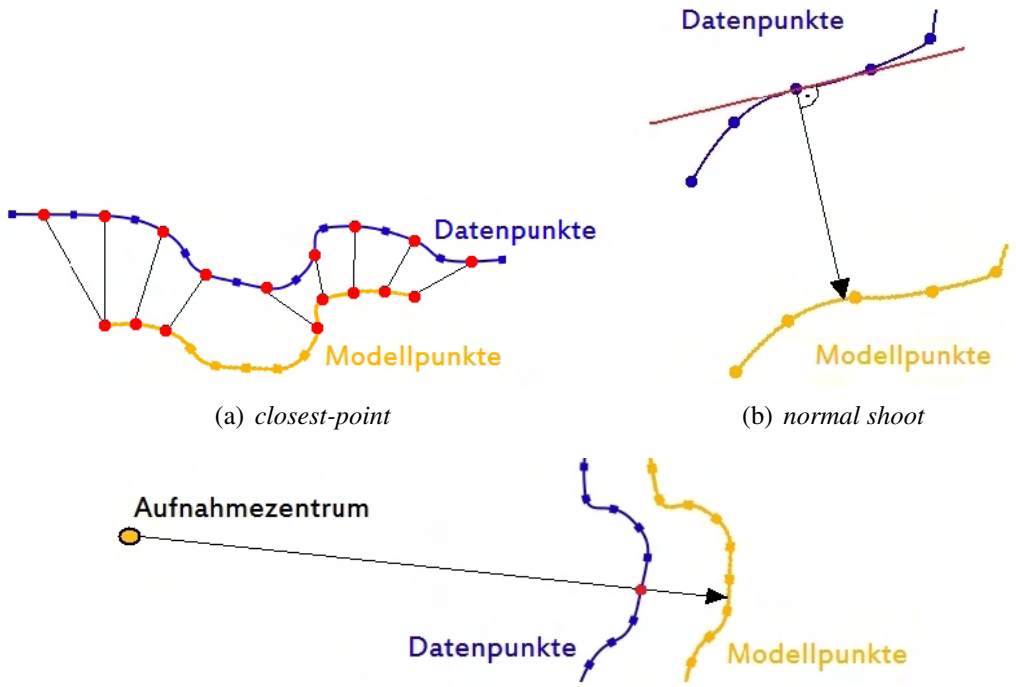
Bei der klassischen Methode, dem *closest-point* Algorithmus, wird zu einem Datenpunkt der nächstgelegene Modellpunkt ausgewählt [Besl u. McKay, 1992].

Bei der *normal shoot* Methode wird für einen Datenpunkt die Flächennormale (siehe Abschnitt 3.3.3.1) abgeschätzt, und deren Schnittpunkt mit der Modellfläche bestimmt. Der Schnittpunkt ist der zum Datenpunkt korrespondierende Punkt [Chen u. Medioni, 1991].

Eine weitere Methode, die als *reverse calibration* bezeichnet wird, projiziert einen Datenpunkt ausgehend vom Aufnahmezentrum des Sensors auf die Modellfläche.

Gewichtung der korrespondierenden Punktepaare

Die einfachste Methode besteht darin, dass alle Punktepaare das gleiche Gewicht erhalten. Die Gewichtung kann aber auch in Abhängigkeit von der Distanz zwischen den Punkten m_i und d_i



(a) *closest-point* (b) *normal shoot*

(c) *reverse calibration*

Abbildung 3.9: Bestimmung der Korrespondenzen: (a) *closest-point* Methode: es wird zu einem Datenpunkt der nächstgelegene Modellpunkt ausgewählt. (b) *normal shoot* Methode: der zum Datenpunkt korrespondierende Punkt ist der Schnittpunkt der Flächennormale des Datenpunktes mit der Modellfläche. (c) *reverse calibration* Methode: der korrespondierende Punkt ist der vom Aufnahmezentrum ausgehend auf die Modellfläche projizierte Datenpunkt. Quelle: [Klauser, 2009]

erfolgen.

$$Gewicht = 1 - \frac{Dist(m_i, d_i)}{Dist_{max}} \tag{3.18}$$

Eine weitere Möglichkeit einer unterschiedlichen Gewichtung ist die Verwendung des Skalarprodukts der an den Punkten abgeschätzten Normalvektoren. Dies führt bei Ebenen mit kleinräumigen Strukturen zu einer schnelleren Konvergenz [Rusinkiewicz u. Levoy, 2001].

$$Gewicht = n_{m_i} \cdot n_{d_i} \tag{3.19}$$

Kriterien zur Verwerfung von Punktepaaren

Bei der ursprünglichen Version des ICP Alogrithmus werden alle Punktepaare verwendet. Häufig überdecken sich Datenpunkte und Modellpunkte nicht vollständig, weshalb die Korrespondenzen von außerhalb des Überlappungsbereiches liegenden Punkte verworfen werden müssen.

Zudem beeinflussen Ausreißer eine Kleinste-Quadrate-Ausgleichung erheblich und sollen im Vorfeld eliminiert werden. Dazu können verschiedene Kriterien verwendet werden. Ein häufig verwendetes Kriterium ist die Distanz. Dabei wird entweder ein gewisser Prozentsatz der Paare, zum Beispiel jene mit der größten Distanz, oder diejenigen Punktepaare, deren Distanz einen bestimmten Schwellwert übersteigt, verworfen. Der Schwellwert kann auch ein von der Standardabweichung der Distanz abhängiger Grenzwert sein. Weiters können auch Kriterien, die von den benachbarten Punkten abhängig sind, verwendet werden. Beispielsweise kann dies ein Grenzwert für den Absolutbetrag der Differenz der Distanz zweier benachbarter Punktepaare, oder ein maximal zulässiger Winkel zwischen benachbarten Normalvektoren sein [Haring, 2007].

Fehlermetrik und deren Minimierung

Bei den meisten ICP-Varianten werden folgende Fehlermetriken verwendet:

- Punkt-zu-Punkt Methode
- Punkt-zu-Ebene Methode

Bei beiden Methoden werden die Abstandsmaße nach der Methode der Kleinsten-Quadrate minimiert [Haring, 2007].

Bei der Punkt-zu-Punkt Methode wird als Abstandsmaß der euklidische Abstand zwischen zwei korrespondierenden Punkten d_i und m_i verwendet. Die zu minimierende Fehlerfunktion $E_{PP}(\boldsymbol{R}, \boldsymbol{t})$ in Abhängigkeit von der Rotationsmatrix $\boldsymbol{R}$ und der Verschiebung $\boldsymbol{t}$ lautet [Haring, 2008]:

$$E_{PP}(\boldsymbol{R}, \boldsymbol{t}) = \sum_i \|\boldsymbol{R}d_i + \boldsymbol{t} - m_i\|^2 \rightarrow min \tag{3.20}$$

Bei der Punkt-zu-Ebene Methode wird als Abstandsmaß der Abstand des Datenpunktes d_i zur Tangentialebene vom korrespondierenden Modellpunkt m_i verwendet. Die Fehlerfunktion $E_{PE}(\boldsymbol{R}, \boldsymbol{t})$ ist wieder bezüglich $\boldsymbol{R}$ und $\boldsymbol{t}$ zu minimieren.

$$E_{PE}(\boldsymbol{R}, \boldsymbol{t}) = \sum_i \left[(\boldsymbol{R}d_i + \boldsymbol{t} - m_i)^T n_{m_i} \right]^2 \rightarrow min \tag{3.21}$$

Die Normalvektoren n_{m_i} an den Modellpunkten m_i müssen vorab abgeschätzt werden (siehe Bestimmung der Korrespondenzen).

Die Minimierung der Fehlerfunktion erfolgt in beiden Methoden iterativ. In jedem Iterations-

schritt k werden die Punktkorrespondenzen ermittelt, die Fehlerfunktion minimiert und anschließend die Transformation (mit $\boldsymbol{R}^{(k)}$ und $\boldsymbol{t}^{(k)}$) auf die Datenpunkte $d_i^{(k-1)}$ angewendet.
Die Punkt-zu-Ebene Methode weist gegenüber der Punkt-zu-Punkt Methode folgende Vorteile auf [Haring, 2008]: deutlich raschere Konvergenz, weniger anfällig bei inhomogenen Punktdichten, liefert genauere Ergebnisse[4] und Ebenen können leichter aneinander „gleiten“. Der Hauptnachteil ist, dass die Punkt-zu-Ebene Methode anfälliger gegenüber schlechter Näherungswerte ist.

ICP bei mehr als zwei Punktwolken

Der bisher beschriebene ICP Algorithmus bezieht sich auf die Orientierung von zwei Punktwolken, in der Praxis wird jedoch meist die Orientierung von mehr als zwei Punktwolken gefordert. Eine einfache Möglichkeit wäre die sequentielle Bestimmung der Orientierung der Punktwolken. Das heißt, es würde ausgehend von einer Punktwolke die Orientierung der anderen Punktwolken durch sukzessives Anwenden des ICP Algorithmus auf zwei benachbarte Punktwolken bestimmt. Dies hat den Nachteil, dass sich die Fehler bei der paarweisen Orientierung von Punktwolke zu Punktwolke aufsummieren.

Zudem hängt das Ergebnis von der Orientierungsreihenfolge und der Ausgangspunktwolke ab. Deswegen wurden Methoden entwickelt, die eine Verallgemeinerung des ICP Algorithmus ermöglichen, um die Orientierung von mehr als zwei Punktwolken simultan bestimmen zu können. Ein Überblick der Methoden wird in der Dissertation von Haring [2007] gegeben.
Der Ansatz von Neugebauer [1997] beruht auf der Punkt-zu-Ebene Methode. Die Punktkorrespondenzen werden für alle überlappenden Punktwolkenpaare (J,K) bestimmt. Zur simultanen Berechnung der Transformationsparameter wird die Gesamtsumme aller paarweisen Fehlerquadratsummen minimiert.

$$E_{Gesamt} = \sum E_{J,K} \tag{3.22}$$

Auf diesem Ansatz beruht auch die Implementierung von Haring [2007], welche bei dieser Arbeit verwendet wird.

[4] Punkt-zu-Punkt Korrespondenzen existieren de facto nicht, da die Punkte mehr oder weniger willkürlich auf der Oberfläche angeordnet sind.

3.4 Modellierung

Ziel der Modellierung ist es, aus den orientierten Punktwolken ein geometrisches 3D-Oberflächenmodell zu erzeugen. In diesem Kapitel werden das Werkzeug und die notwendigen Schritte, um zu einem Modell zu gelangen, vorgestellt. Die Durchführung der Modellierung der Aufnahme der Passage im Palais Ferstel und die dabei gewonnenen Erkenntnisse werden im Detail in Abschnitt 4.3.4 beschrieben.
Zur Modellierung wird die Software *Geomagic Studio* 10 verwendet. Alternativ wäre dies beispielsweise auch mit der Open Source Software MeshLab [MeshLab, 2009] möglich, diese erwies sich jedoch in der Version 1.2.0 als zu instabil. Am Beginn dieses Kapitels wird *Geomagic Studio* vorgestellt. Anschließend werden die Bearbeitungsschritte vor der Triangulation und das Prinzip einer Triangulation beschrieben.

3.4.1 Geomagic Studio

Die Software *Geomagic Studio* wurde in erster Linie für Aufgaben im Bereich des Reverse Engineerings entwickelt. Beim Reverse Engineering werden aus einem bestehenden Objekt die Konstruktionselemente rekonstruiert, das heißt aus einem realen Objekt wird ein CAD-Modell erstellt. Typische Anwendungsgebiete dieser Software liegen im Bereich des Produktdesigns, wie zum Beispiel Erstellung einer Konstruktionsgrundlage auf Basis eines Tonmodelles. *Geomagic Studio* eignet sich aber auch für geodätische Aufgaben. In erster Linie wird hier die Software zur Erstellung von Modellen aus Punktwolken von terrestrischen Laserscanneraufnahmen verwendet. Da aus den Daten einer Distanzkamera ebenfalls Punktwolken berechnet werden können, kann die Modellierung ebenfalls, ähnlich den Laserscanneraufnahmen, mit *Geomagic Studio* durchgeführt werden.
In *Geomagic* durchlaufen die Daten bis zum fertigen Modell mehrere Phasen:

- *Punktphase*: In dieser Phase liegen die Daten als Punktwolken vor. Die Punktwolken werden orientiert, gefiltert, geglättet, reduziert und zusammengeführt.
- *Polygonphase*: Als Übergang zur Polygonphase wird eine Dreiecksvermaschung durchgeführt. Das Polygonmodell wird in dieser Phase editiert, damit eine geschlossene saubere Oberfläche entsteht.

- *Oberflächenphase*: Aus dem Polygonmodell wird in dieser Phase ein CAD taugliches Oberflächenmodell erstellt. Dies kann entweder ein aus NURBS(Non-Uniform Rational B-Spline)-Flächen bestehendes Modell sein (Shape-Phase), oder ein Modell aus zusammengefügten geometrischen Grundformen (Fashion-Phase).

Im Rahmen dieser Arbeit wurde auf die Erzeugung eines CAD taugliches Oberflächenmodells verzichtet, da zur reinen Visualisierung eine Triangulation ausreichend ist. Die einzelnen Werkzeuge in den Bearbeitungsschritten können meist über Parameter gesteuert werden. Leider ist häufig die Funktionsweise der Werkzeuge und die Bedeutung der Parameter nicht dokumentiert. Deswegen ist das Ergebnis bezüglich der Parameter schwerer interpretierbar und Vergleiche mit den Ergebnissen anderer Implementierungen lassen sich schwer ziehen. Außerdem ist es bei einem solchen proprietären Programmpaket nicht möglich, die einzelnen Funktionen zu verändern oder zu erweitern, wie es zum Beispiel für die Einbeziehung des exzentrischen Drehpunktes in die Orientierung notwendig wäre.

3.4.2 Bearbeitungsschritte vor der Triangulation

Die Orientierung der *statischen Bilder* wurde bereits im Kapitel 3.3 behandelt. An dieser Stelle sei nochmals erwähnt, dass es aufgrund der großen Anzahl von Aufnahmen unbedingt notwendig ist, diese zumindest weitgehend vollautomatisch zu orientieren. Dies ist für die Distanzkameraaufnahmen mit *Geomagic Studio* nicht möglich. Für die Verbesserung der Orientierung kann Geomagic eingesetzt werden. Dazu stehen die Werkzeuge *Manuelle Registrierung* und *Globale Registrierung* zur Verfügung (siehe Abschnitt 4.3.3).
Ein weiterer vorbereitender Schritt für die Modellierung ist die Filterung der Punkte. Diese wurde ebenfalls bereits im Rahmen der Datenaufbereitung (siehe Abschnitt 3.3.3.1) behandelt. *Geomagic* bietet hier zum einen eine Möglichkeit zur Filterung von Ausreißern, zum anderen ist es mit der Software möglich, komfortabel Punkte manuell zu löschen (siehe Abschnitt 4.3.4).
Weiters erfolgt vor der Berechnung der Triangulation noch ein Glättung der Daten und eine Ausdünnung der Punktwolken. Schließlich werden die einzelnen Punktwolken vereinigt und eventuell nochmals ausgedünnt.

3.4.3 Triangulation

Bei der Dreiecksvermaschung (Triangulation) werden jeweils drei benachbarte Punkte der Punktwolke zu einer Dreiecksfläche verbunden, sodass die ursprüngliche 3D-Oberfläche möglichst gut approximiert wird. Für die Bestimmung der Dreiecksvermaschung gibt es eine Vielzahl von Triangulationsverfahren. In Gassmann u. Franke [2005] wird ein Überblick über diese Verfahren gegeben.
Eine häufig verwendete Methode (Inkrementelle Methoden) zur 3D-Dreiecksvermaschung läuft folgendermaßen ab: Es wird mit einem beliebigen Dreieck begonnen und der Dreiecksverband anhand lokaler Kriterien sukzessive Punkt für Punkt erweitert. Dabei ist wichtig, dass durch das Hinzufügen eines neuen Punktes das bereits vorhandene Dreiecksnetz nur lokal umgestaltet wird [Kraus, 2000].
Zur Triangulation der Punktwolke aus der Aufnahme der Passage wird in dieser Arbeit *Geomagic Studio* verwendet. Leider ist nicht bekannt welches Triangulationsverfahren dabei zur Anwendung kommt.
In *Geomagic Studio* 10 kann die Triangulation, eine Glättung der Daten und eine Ausdünnung der Punktwolken in einem Schritt mit dem Werkzeug *Vereinigen* erfolgen.
Nach der Durchführung der Triangulation ist es noch notwendig diese zu bearbeiten (siehe Abschnitt 4.3.4).

Kapitel 4

Aufnahme und Auswertung der Passage im Palais Ferstel

Dieses Kapitel beinhaltet die beispielhafte Anwendung des entwickelten Verfahrens. Es wird die im Zuge dieser Arbeit durchgeführte Aufnahme und die anschließende Auswertung der Passage im Palais Ferstel beschrieben. Zuerst soll in Abschnitt 4.1 der Aufnahmeort vorgestellt werden. In Abschnitt 4.2 wird die Planung und Durchführung der Aufnahme beschrieben. Dabei wird insbesonders auf die verwendeten Geräte, die herrschenden Bedingungen und den Ablauf bei der Aufnahme eingegangen. Anschließend werden in Abschnitt 4.3 die einzelnen Schritte der Auswertung und die dabei gemachten Erfahrungen beschrieben und die Zwischenergebnisse präsentiert. Dem folgt eine Analyse der Modellqualität (Abschnitt 4.4). Am Ende des Kapitels steht in Abschnitt 4.5 eine Diskussion des vorgestellten Verfahrens sowie ein Ausblick.

4.1 Palais Ferstel

Das Palais Ferstel, genauer gesagt ein Teil der prunkvollen Einkaufspassage im Palais, wurde zur Erprobung des vorgestellten Verfahrens zur Vervollständigung von Stadtmodellen ausgewählt, da die Passage ein Beispiel für einen öffentlich zugänglichen Bereich ist, der nicht von einem Fahrzeug aus aufgenommen werden kann (siehe Abschnitt 1.1) und somit eine Aufgabenstellung darstellt, die potentiell wirtschaftlich mittels einer Distanzkamera gelöst werden könnte.

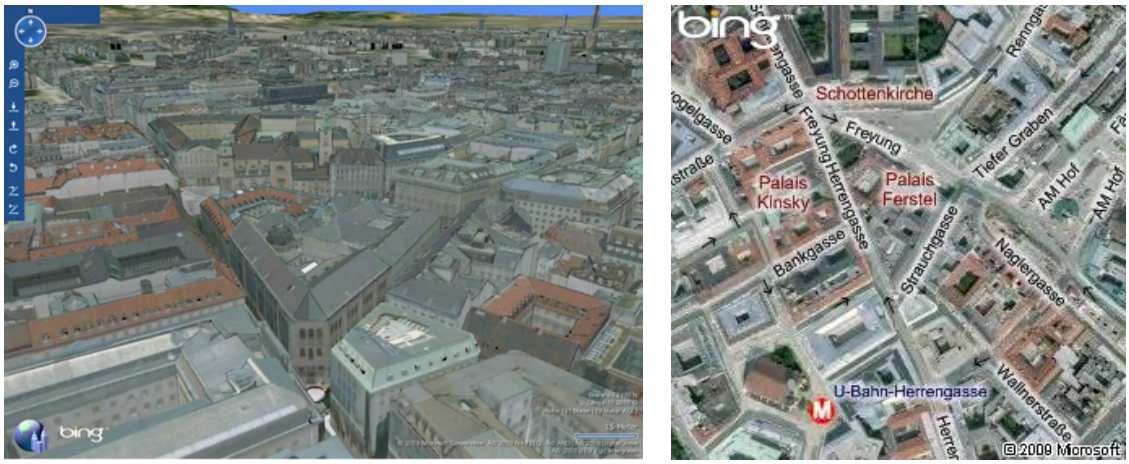


Abbildung 4.1: Lage des Palais Ferstel: das Palais Ferstel befindet sich zwischen Freyung, Strauchgasse und Herrengasse im ersten Wiener Gemeindebezirk. Quelle: [Microsoft Corporation, 2009]

Das Palais Ferstel befindet sich zwischen Freyung, Strauchgasse und Herrengasse im 1. Wiener Gemeindebezirk (Abbildung 4.1).

Die Geschichte des Palais Ferstel (Abbildung 4.2) in seiner heutigen Gestalt geht zurück ins 19. Jahrhundert. Das Gebäude wurde zwischen 1856 und 1860 vom Architekten Heinrich von Ferstel erbaut. Durch die Kombination der venezianisch-florentinischen Trecento Arichitektur mit dem Werksteinbau gilt das Palais als eines der prachtvollsten Bauwerke der Gründerzeit und erwirbt bei der Eröffnung 1860 den Ruf des modernsten Hauses Wiens.

Von 1860 bis zur Eröffnung des Börsengebäudes am Ring im Jahre 1877 waren die offiziellen Wiener Börsensäle im Erdgeschoss des Palais Ferstel untergebracht. Auch die Österreichisch-Ungarische Nationalbank hatte in den ersten Jahren des Bestehens des Gebäudes ihren Sitz im Palais Ferstel. Nach Auszug der Börse eröffneten die Gebrüder Pach im Erdgeschoss das Café Central. Dieses wurde zum Treffpunkt bedeutender Künstler und Literaten. Im zweiten Weltkrieg erleidet das Palais in Folge von Bombentreffern schwere Beschädigungen. Nach dem Krieg wird das Gebäude vorerst nur notdürftig renoviert. Erst 1978 beginnen die eigentlichen Renovierungsarbeiten. 1982 eröffnet auch wieder das Café Central vorerst im Arkadenhof und seit 1986 befindet sich das Kaffee wieder im Säulensaal.
Heute können die Festsäle des Palais Ferstel für Veranstaltungen gebucht werden. Zudem befindet sich im Gebäude eine der elegantesten Einkaufspassagen der Stadt, die den Platz der Freyung mit der Herrengasse verbindet. Teile dieser Passage wurden im Rahmen der Diplomarbeit aufgezeichnet [Palais Events, 2009; Burgenkunde, 2009].

Abbildung 4.2: Historisches Bild des Palais Ferstel. Quelle: [Historisches Bild]

4.2 Aufnahme der Passage

Wie zuvor bereits erwähnt, wurde ein Teil der Passage im Palais Ferstel mittels der Distanzkamera SR-3000 dreidimensional erfasst. Die Aufnahme erfolgte mit Hilfe eines Stativs, sodass es möglich war, mehrere Bilder mit derselben äußeren Orientierung aufzunehmen und diese im Anschluss zu mitteln. Zwischen diesen statischen Aufnahmen zeichnet die Kamera die Schwenkbewegung auf. Das Stativ ist während der ganzen Aufnahme auf einen fahrbaren Untersatz gestellt, der für die Aufnahmen von einem Standpunkt aus fixiert werden kann. Die Bewegung zwischen zwei Standpunkten wird ebenfalls aufgezeichnet. Mittels dieser Aufnahmen in der Bewegung werden in späterer Folgen Näherungswerte für die äußere Orientierung der *statischen Bilder* automatisiert bestimmt (siehe Abschnitt 4.3). Die genaue Vorgehensweise bei der Aufnahme wird in Abschnitt 4.2.3 beschrieben. In Abschnitt 4.2.1 werden zunächst kurz die verwendeten Geräte vorgestellt. Anschließend wird in Abschnitt 4.2.2 auf die Planung der Aufnahme eingegangen.

4.2.1 Verwendete Geräte

Bei der Aufnahme der Passage kommen im wesentlichem folgende Komponenten zum Einsatz:

- Distanzkamera Swissranger™ SR-3000
- Stativ mit einem fahrbaren Untersatz sowie je eine Halterung für die Distanzkamera und für das Notebook
- Bleiakkumulator 12V
- Standard Notebook mit 4 Reserveakkus
- Software für die Aufnahme (Eigenentwicklung am Institut für Photogrammetrie und Fernerkundung (IPF) Wien)

In Tabelle 4.1 werden die Daten der verwendete Distanzkamera Swissranger™ SR-3000 zusammengefasst.

Parameter		SR-3000
Anzahl der Pixel		176x144
Blickfeld	[°]	47,5x39,6
max. Bildrate	[fps]	25
Trägerwellenlänge	[nm]	850
Modulationsfrequenz	[MHz]	20
Eindeutigkeitsbereich	[m]	7,5
Abmessung	[mm]	50x67x43,3

Tabelle 4.1: Daten Swissranger™ SR-3000

Das Stativ ermöglicht eine Drehung um die vertikale Achse sowie ein Kippen der Kamera (verändern der Zenitdistanz). Dabei wird die Kamera um einen gemeinsamen Drehpunkt bewegt, welcher bei der Auswertung eine wesentliche Rolle spielt. Der fahrbare Untersatz wurde eigens für das Stativ konstruiert und ermöglicht eine gleichmäßige Bewegung des Statives zwischen den Standpunkten. Zudem sind die Rollen fixierbar, damit sich das Stativ am Standpunkt nicht unbeabsichtigt bewegen kann. Die Halterung des Notebooks besteht aus zwei gebogenen Gewindestangen, die ins Stativ geschraubt werden. Die Stromversorgung der Kamera erfolgt mittels eines Bleiakkus aus einem Motorrad. Mit diesem kann die Distanzkamera einige Stunden betrieben werden. Damit die Mobilität des gesamten Aufbaus gewährleistet ist, wird der

Abbildung 4.3: Photo mit den Komponenten: zu sehen ist das Stativ auf fahr- und fixierbarem Untersatz, darauf montiert das Notebook samt Kamera, daran hängend der Akkumulator.

Akku mittels einer Schnur am Stativ befestigt (siehe Abbildung 4.3).
Zum Aufzeichnen der Daten kommt ein Notebook zum Einsatz. Die Distanzkamera wird über die USB-Schnittstelle mit dem Notebook verbunden.
Die Aufzeichnung der Daten erfolgt über eine am IPF entwickelte Aufnahmesoftware für den Swissranger™ SR-3000. Die Software ermöglicht es, Parameter wie zum Beispiel die Integrationszeit und die Modulationsfrequenz einzustellen. Während der Aufzeichnung können die Aufnahmen bereits als 3D-Punktwolke betrachtet werden. Die Anzahl der aufgenommenen Bilder ist im Prinzip nur durch die Größe der Festplatte des Notebooks beschränkt.

4.2.2 Planung der Aufnahme

Ein einzelnes Bild des Swissrangers™ SR-3000 weist ein erhebliches Distanzrauschen auf. Um grob falsche Distanzbeobachtungen zu elimininieren, werden vom gleichen Objektausschnitt mehrere Bilder aufgezeichnet und anschließend wird ein robuster Mittelwert für die Distanzmessung der Pixel bestimmt (siehe Abschnitt 3.2.1). Außerdem zeigten Aufnahmen, die im Vorfeld der Aufzeichnung der Passage im Palais Ferstel durchgeführt wurden, dass die Distanzkamera nicht schnell bewegt werden darf, da sonst die Bewegungsunschärfe in den Bildern zu groß ist.

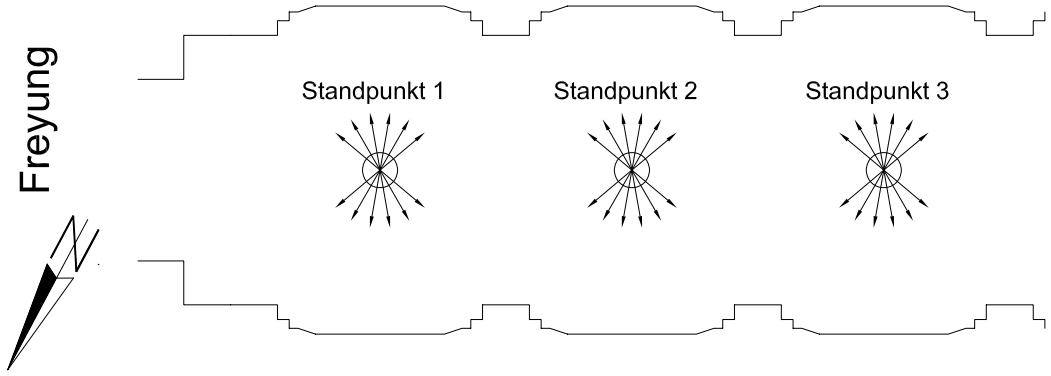


Abbildung 4.4: Skizze des Grundrisses der ersten drei Bögen vom östlichen Teil der Passage mit den 3 Standpunkten. Zu sehen sind auch die Blickrichtungen bei der Aufnahme.

Abbildung 4.4 zeigt eine Skizze des Grundrisses der ersten drei Bögen des östlichen Teils der Passage. Auf Basis dieser Skizze erfolgte die Planung der Aufnahme. Einschränkend bei der Aufnahmeplanung wirkt, dass die Distanzkamera nur einen eingeschränkten Öffnungswinkel besitzt und die Aufnahme von der Mitte eines Bogens aus erfolgen musste, damit Passanten die Möglichkeit haben, ungehindert vorbeizugehen und somit nicht im Bildmaterial aufscheinen. Zudem ist die Breite der Passage vorgegeben, sodass die Punktwolke einer Aufnahme nur einen kleinen Ausschnitt des Objektes wiedergibt. Dies ist, insbesondere bei Bildern die horizontal und im rechten Winkel zur Achse der Passage aufgenommen worden sind heikel, da hier die Punktwolke nur wenige geometrische Merkmale abbildet und somit die Fein-Orientierung der Punktwolke mittels ICP schlecht bestimmt ist. Deswegen ist bei der Planung bereits darauf zu achten, dass diese Aufnahmekonstellation möglichst vermieden wird bzw. dass ausreichend geometrische Features abgebildet werden (siehe Abbildung 3.8) und zusätzlich eine minimale

Überdeckung von mindestens 50 Prozent gewählt wird. Diese hohe Überdeckung ist auch notwendig, da eine gewisse Anzahl an Bild-/Abtastpunkten erforderlich ist, um eine geometrische Struktur abzubilden. Diese notwendige Anzahl an Pixel macht bei der geringen Auflösung des SR-3000 einen beträchtlichen Teil des Sensors aus. Um die Aufnahmen mittels ICP über geometrische Strukturen verknüpfen zu können, müssen diese in beiden Bildern abgebildet sein.
Insgesamt wurden zu zwei unterschiedlichen Zeitpunkten Aufnahmen in der Passage durchgeführt. Bei der zweiten Aufnahme wurden die oben genannten Erfahrungen aus der ersten Aufnahme berücksichtigt. Abbildung 4.5 zeigt die Planung der Blickrichtungen der Kamera für einen Standpunkt.
Von einem Standpunkt aus werden auch der nächste und der vorhergehende Bogen (Vor- und Rückblick) aufgenommen. Dadurch können die Standpunkte besser zueinander orientiert werden. Die Aufnahme startet mit Blickrichtung zur Decke, in weiterer Folge wird die Kamera in einer Vertikalebene in 20 Grad Schritten bis zu einem Zenitwinkel von 120 Grad gekippt, anschließend um die vertikale Achse des Stativs um 20 Grad gedreht und wiederum nach oben gekippt. Pro Seite werden so für jeden Standpunkt 6 vertikale Streifen aufgenommen. Im Bereich des Deckengewölbes ist die Überdeckung durch den eben beschriebenen Ablauf natürlich wesentlich größer als 50 Prozent. Es wäre daher nicht notwendig bei allen Streifen den oberen Bereich der Passage aufzunehmen. Die vollständige Aufnahme aller Streifen hat aber zum einen den Vorteil, dass bei der Aufnahme für einen Streifen immer die gleichen Vertikalwinkel verwendet werden können und so Fehler (Lücken) vermieden werden. Zum anderen ist bei der Fein-Orientierung, bei nicht ausreichender näherungsweiser Orientierung, ein manueller Eingriff leichter möglich, da zwei benachbarte, vollständige Streifen besser zueinander orientiert werden können (siehe Abschnitt 4.3.3). Insgesamt sind für jeden Standpunkt 84 (!) Aufnahmen notwendig. Dies entspricht bei 3 Bildern pro Sekunde sowie 50 Bildern pro statischem Bild und durchschnittlich ebensovielen pro Schwenk (bewegte Sequenz), einer theoretischen Aufnahmezeit von 45 Minuten pro Standpunkt. Während des Standpunktwechsels werden in der Bewegung ebenfalls Bilder aufgezeichnet. Diese dienen genauso wie jene Bilder, die während der Schwenkbewegung aufgezeichnet werden, zur automatisierten Bestimmung der Näherungswerte der Stellung der Distanzkamera während der statischen Aufnahmen.

4.2.3 Durchführung der Aufnahme

Wie bereits erwähnt, erfolgten zwei Aufnahmen eines Teils der Passage im Palais Ferstel, zu unterschiedlichen Zeitpunkten.

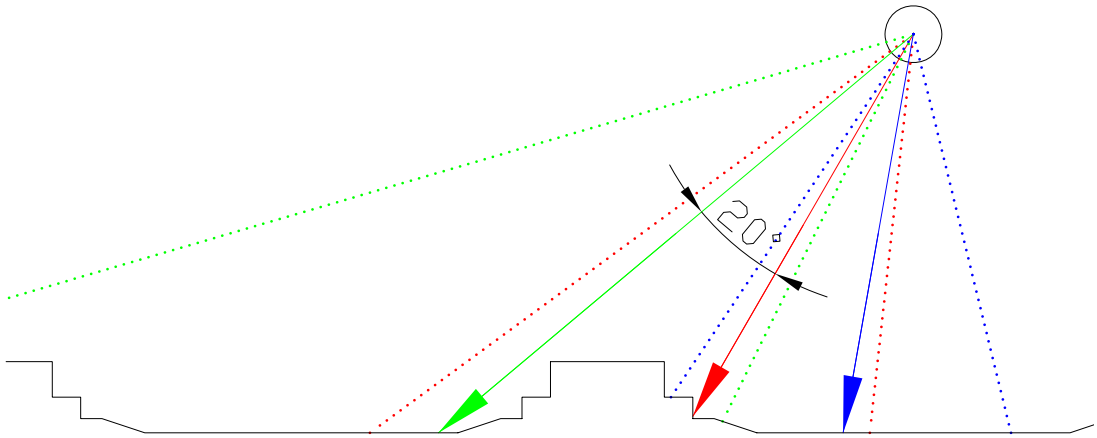


Abbildung 4.5: Überdeckung bei den Aufnahmen: zu sehen ist ein Teil des Grundrisses, ein Standpunkt, drei Blickrichtungen der Kamera (Pfeile) sowie die zugehörigen Blickfelder der Kamera (punktierte Linien gleicher Farbe). Die Überdeckung der Bilder soll mindestens 50% betragen, dadurch ergibt sich bei der in der Skizze dargestellten Aufnahmegeometrie zwischen zwei Aufnahmen ein Winkel von ca. 20 Grad.

Die erste Aufnahme wurde am 9.4.2009 durchgeführt. Erfasst wurde bei dieser Aufnahme der erste Bogen und Teile des zweiten Bogens am östlichen Ende der Passage. Insgesamt wurden 90 Standbilder aufgenommen. Bei der Aufnahme herrschte künstliche Beleuchtung und eine Temperatur von ca. 15°C. Die erste Aufnahme erfolgte nicht auf Basis der in 4.2.2 beschriebenen Planung, der Ablauf ist jedoch ähnlich wie oben beschrieben. Probleme gab es bei dieser Aufnahme mit der Akkulaufzeit des Notebooks, weshalb der zweite Bogen nicht vollständig aufgezeichnet werden konnte.
Am 14.10.2009 wurde die zweite Aufnahme durchgeführt, bei der neben den ersten beiden auch der dritte Bogen aufgenommen wurde. Die Temperatur bei der Aufnahme lag bei ca. 5°C und es war wiederum künstliches Licht vorhanden. Im Zuge dieser Aufnahme wurden 270 statische Bilder aufgenommen. Insgesamt wurden 46400 Einzelbilder (Distanz- und Amplitudenbilder) aufgezeichnet. Die Durchführung der Aufnahme entspricht in etwa dem in der Planung festgelegten Ablauf, jedoch wurden teilweise mehr Bilder pro statischem Bild aufgezeichnet und beim dritten Standpunkt wurde ein vertikaler Streifen nicht aufgenommen. Es sei ebenfalls erwähnt, dass teilweise Passanten im Blickfeld der Kamera befanden.

4.3 Auswertung

Der theoretische Ablauf der Auswertung ist bereits in Kapitel 3 beschrieben worden. In diesem Abschnitt wird gezeigt, wie die Auswertung der Aufnahme der Passage im Palais Ferstel erfolgt. Dabei soll im Detail auf die Durchführung der Auswertung sowie auf die verwendeten Funktionen und die dazu benötigten Parameter eingegangen werden. Die Auswertung erfolgt in erster Linie mit MATLAB® Version 7.4 und *Geomagic*® STUDIO® Version 10.

4.3.1 Datenvorprozessierung

Bevor mit der Orientierung begonnen werden kann, müssen die aufgenommenen Daten aufbereitet und in statische und bewegte Sequenzen segmentiert werden (siehe Abschnitt 3.2). Der Ablauf der Datenvorprozessierung ist in Abbildung 3.2 zusammengefasst.

Im ersten Schritt werden die Daten in MATLAB eingelesen und die Grenzen der statischen und bewegten Sequenzen bestimmt. Dazu wird der 2D - Korrelationskoeffizient r (siehe Gleichung 3.1) von zwei benachbarten Amplitudenbildern ermittelt. Ist nun der Betrag der Differenz von benachbarten Korrelationskoeffizienten für mindestens 30 aufeinander folgende Bilder kleiner als *maxkorraend*, so liegt eine statische Sequenz vor. Für *maxkorraend* wird der empirische Schwellwert von 0.0005 verwendet. Überschreitet die Differenz den Grenzwert, so ist ab diesem Bild eine bewegte Sequenz vorhanden bis wieder ein Block von mindestens 30 Bildern den Schwellwert unterschreitet. Ist eine bewegte Sequenz kleiner als 5 Bilder so wird diese zur statischen Sequenz gezählt, da eine bewegte Sequenz von dieser Länge nicht zu erwarten ist. Dies könnte zum Beispiel auftreten, wenn ein Passant kurz im Blickfeld der Kamera ist oder bei einer sonstigen kurzen Störung. Damit in der statischen Sequenz keine Bilder der Bewegung vorhanden sind, werden die ersten und die letzten 3 Bilder der statischen Sequenz der bewegten Sequenz zugeordnet.
Um ein statisches Amplitudenbild zu erhalten, wird von den Amplitudenbildern, welche innerhalb einer statischen Sequenz liegen, das arithmetische Mittel gebildet. Bei den Distanzbildern wird zuerst die am IPF entwickelte Kalibrierung angebracht (siehe Abschnitt 2.4). Wie bereits in Abschnitt 3.2.1 erwähnt, wird zur Mittelung der Distanzbilder von einer statischen Sequenz ein gestutztes Mittel verwendet. Es werden für jedes Pixel jeweils zehn Prozent der kürzesten und längsten Distanzen verworfen, aus den restlichen 80 % wird der arithmetische Mittelwert gebildet. So werden grob falsche Beobachtungen verworfen.

Der letzte Schritt bei der Datenvorprozessierung ist die Amplitudennormalisierung der Amplitudenbilder. Dazu wird zunächst mittels der MATLAB-Funktion *imadjust* allgemein der Kontrast der Bilder verbessert. Anschließend erfolgt die in Abschnitt 3.2.2 vorgestellte, distanzabhängige Korrektur der Amplitudenbilder. Zunächst werden die Distanzbilder mittels eines Wiener-Filters geglättet. Dabei wird in einer 3 x 3 Umgebung der lokale Mittelwert und die Standardabweichung bestimmt (siehe Abschnitt 3.2.2). Die Verbesserung der Amplitudenbilder kann in weiterer Folge nach Gleichung 3.3 erfolgen. Bei den Bildern zeigt sich jedoch, dass folgende Korrekturfunktion ein besseres Resultat erzielt:

$$A_{norm,dist}(m,n) = A(m,n)D_{WF}(m,n)^{1.5} \tag{4.1}$$

Die Korrektur des radialen Lichtabfalls erfolgt mit Hilfe der Matrix K_{radial}, wobei der Parameter $\tau = 8$ empirisch ermittelt wurde.

$$A_{norm,amp}(m,n) = A(m,n)K_{radial}(m,n) \tag{4.2}$$

$$K_{radial}(m,n) = \left[\frac{\sqrt{(m-x_0)^2 + (n-y_0)^2 + c^2}}{c}\right]^{\tau} \tag{4.3}$$

Die Kamerakonstante c und Lage des Hauptpunktes x_0, y_0 sind die Elemente der inneren Orientierung. Mit der Matrix K_{radial} konnte der radiale Lichtabfall verringert werden (siehe Abbildung 4.6).

Abbildung 4.6: Korrektur des radialen Lichtabfalls: links Amplitudenbild vor der Korrektur, rechts nach der Korrektur. In der rechten Abbildung ist der radiale Lichtabfall geringer.

Erkennung der Standpunktwechsel

Für den in weiterer Folge verwendeten räumlichen Rückwärtsschnitt ist es notwendig bei den bewegten Sequenzen zu unterscheiden, ob es sich um eine bewegte Sequenz handelt die während

eines Standpunktwechsels aufgezeichnet wurde, oder ob diese bei einem Schwenk der Kamera am gleichen Standpunkt entstanden ist. Diese Unterscheidung erfolgt über die Länge der bewegten Sequenz. Ist diese länger als 300 Bilder handelt es sich potentiell um einen Standpunktwechsel. Diese Angabe sollte jedoch kontrolliert werden, da es zum Beispiel durch einen ungeplanten Schwenk, der sich vom Boden der einen Seite zur Decke der anderen Wand erstreckt, auch bei einer reinen Schwenkbewegung der Kamera, zu einer längeren bewegten Bildfolge kommen kann.

4.3.2 Bestimmung der näherungsweisen Orientierung

Nach der Datenvorprozessierung und der Erkennung der Standpunktwechsel müssen im nächsten Schritt die Näherungswerte der Orientierung für die statischen Bilder bestimmt werden. Wie schon in Abschnitt 3.2.1 beschrieben erfolgt dies mittels der bewegten Sequenzen. In Abbildung 3.5 ist der Ablauf für die Bestimmung der näherungsweisen Orientierung grafisch dargestellt.

Vorgabe der Orientierung

Zuerst wird für das erste statische Bild die äußere Orientierung vorgegeben. In dieser Arbeit wurde für die Rotationsmatrix R die Einheitsmatrix und für die Translation t der Nullvektor verwendet.

Verknüpfungspunkte

Für die näherungsweise Orientierung müssen anschließend Verknüpfungspunkte in den Amplitudenbildern der bewegten Sequenz bestimmt werden. Dies erfolgt mittels des SURF Algorithmus (siehe Abschnitt 3.3.2.1.2). Genauer gesagt wird das SURFmax MATLAB Interface von Petter Strandmark verwendet [Strandmark, 2008].

Im ersten Schritt werden mittels der Funktion *surfpoints* die Bildpositionen mit den dazugehörigen Deskriptoren von potentiellen Verknüpfungspunkten bestimmt. Die Funktion gibt auch für jeden Punkt die Skalierung der Keypoints zurück. Diese werden in weiterer Folge noch verwendet, um falsche Punktkorrespondenzen zu reduzieren. Die wichtigsten Parameter die bei *surfpoints* verwendet wurden sind:

- *upright*=1: Mit diesem Parameter wird die Rotationsinvarianz von SURF deaktiviert, da diese für die Aufnahmen nicht notwendig ist.

- *doubleImageSize*=1: Die Auflösung der Bilder wird verdoppelt, um durch die Glättung der Bilder im Algorithmus nicht zu viele Keypoints zu verlieren.

Anschließend erfolgt über die Funktion *surfmatch* die Suche nach korrespondierenden Punkten (siehe Abbildung 4.7) im aktuellen (Bild A) und den 10 vorhergehenden Bildern (Bild B). Dies erfolgt mittels den zu den Punkten zugehörigen Deskriptoren, genauer gesagt über die euklidische Distanz der Deskriptoren (siehe Abschnitt 3.3.2.1.2). Der Parameter *thresh* gibt den Grenzwert für den Quotienten aus geringster und zweitgeringster euklidischer Distanz an. Ist dieser Quotient größer als 0.7 wird die Punktkorrespondenz verworfen. Da diese Bilder unmittelbar aufeinander folgend aufgenommen wurden, können zusätzlich noch weitere Kriterien definiert werden, um falsche Zuordnungen zu erkennen. Zum einen dürfen sich die Skalierungen eines Punktes nur maximal um den Faktor 2 unterscheiden, zum anderen darf sich die Position des Verknüpfungspunktes in den beiden Bildern nicht um mehr als 15 Pixel ändern (*Positionskriterium*). Außerdem werden Punktkorrespondenzen deren Distanz (aus dem Distanzbild entnommen) sich um mehr als 3 cm unterscheidet ebenfalls verworfen (*Distanzkriterium*). Der Schwellwert des *Distanzkriteriums* erhöht sich mit der Anzahl der Bilder die zwischen Bild A und Bild B liegen ($Schwellwert(i) = 3cm * \sqrt{i}$ wobei $i - 1$ die Anzahl der Bilder zwischen Bild A und Bild B ist). Wird ein Punkt aus Bild B zu mehr als einem Punkt im aktuellen Bild zugeordnet, werden diese Punktkorrespondenzen ebenfalls verworfen. Die Werte für die eben genannten Kriterien wurden empirisch bestimmt. Dazu wurden die Schwellwerte der Kriterien variiert und sowohl das Endergebnis der näherungsweisen Orientierung betrachtet, als auch die verworfenen und die beibehaltenen Punktkorrespondenzen dargestellt und die Auswirkungen der Kriterien überprüft.

Nach dieser Filterung erfolgt eine Überprüfung ob der korrespondierende Punkt bereits für Bild B als Verknüpfungspunkt gespeichert ist. Das heißt es wird überprüft, ob bereits zu einem früheren Zeitpunkt eine Punktkorrespondenz zwischen Bild B und einem vorhergehenden Bild existiert hat und ob diesem Punkt schon eine Nummer zugeordnet wurde. Diese Nummer wird für den Punkt in Bild A übernommen. Dies trägt zur Stabilität bei, da dadurch mehrere Bilder über Punkte mit denselben Objektkoordinaten verknüpft werden. Nachdem die korrespondierenden Punkte die für Bild B bereits als Verknüpfungspunkte (Punkt hatte im Bild B bereits eine Nummer) gespeichert sind und in Bild A hinzugefügt wurden, erfolgt die Suche nach Verknüpfungspunkten, die auch in den Bildern B neu sind. Diese erhalten eine neue Nummer und werden in beiden Bildern hinzugefügt. Die Suche nach bereits vorhandenen und neuen Verknüpfungspunkten wird nacheinander ausgeführt, da zuerst in allen 10 vorhergehenden Bildern

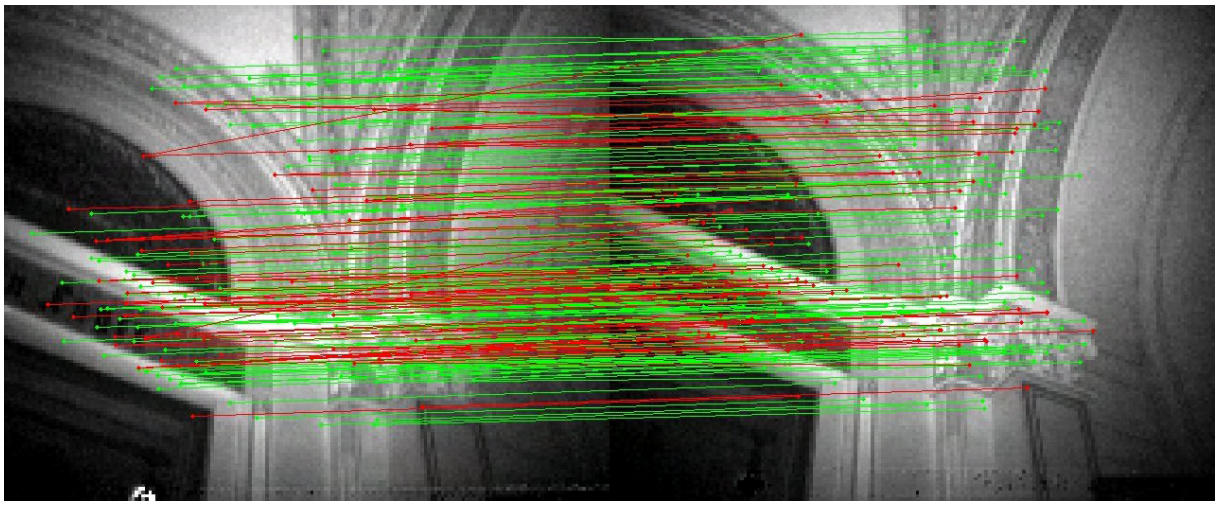


Abbildung 4.7: Punktkorrespondenzen: zu sehen sind 2 Amplitudenbilder einer bewegten Sequenz. Die grünen Linien zeigen die Punktkorrespondenzen, die mittels *surfmatch* gefunden wurden, die rote Linien zeigen Korrespondenzen die ebenfalls mittels *surfmatch* gefunden, aber durch Kriterien (*Positionskriterium*, *Distanzkriterium*,...) verworfen wurden.

überprüft wird ob der korrespondierende Punkt bereits in einen der Bilder als Verknüpfungspunkt gespeichert ist. Nur wenn der Punkt in keinem der 10 vorhergehenden Bilder als Verknüpfungspunkt vorhanden ist wird dieser in Bild A und in Bild B als neuer Verknüpfungspunkt hinzugefügt.

Die Abbildung 4.8 zeigt die mittlere Anzahl der gefundenen Punktkorrespondenzen in Abhängigkeit der Anzahl der Bilder die zwischen den Bildern A und B, in welchen die Korrespondenzen gesucht werden, liegen. Die Anzahl der Punktkorrespondenzen fällt leicht mit dem zeitlichen Abstand der Bilder. Im Mittel werden 120,8 Punktkorrespondenzen zwischen zwei Bildern gefunden. Die roten Balken zeigen die Anzahl der Punktkorrespondenzen nach dem Anwenden der oben genannten Kriterien. Dabei fällt auf, dass die Anzahl der Punkte die nach der Filterung übrig bleiben, bei geringerem zeitlichen Abstand der Bilder zu Beginn niedriger ist. Der Grund dafür liegt in der Anpassung des *Distanzkriteriums* an den zeitlichen Abstand der Bilder. Durchschnittlich werden ca. die Hälfte der Punktkorrespondenzen gefiltert (von den 120 Punktkorrespondenzen bleiben 64,4 nach der Filterung). Da aber nicht nur Punktkorrespondenzen zwischen zwei Bildern, sondern zwischen mehreren Bildern gesucht werden, existieren pro Bild im Mittel 172,4 Verknüpfungspunkte.

Das Kriterium der maximalen Positionsunterschiede der Verknüpfungspunkte in den beiden Bildern wird nicht an die Anzahl der Bilder zwischen Bild A und Bild B angepasst. Dadurch kann

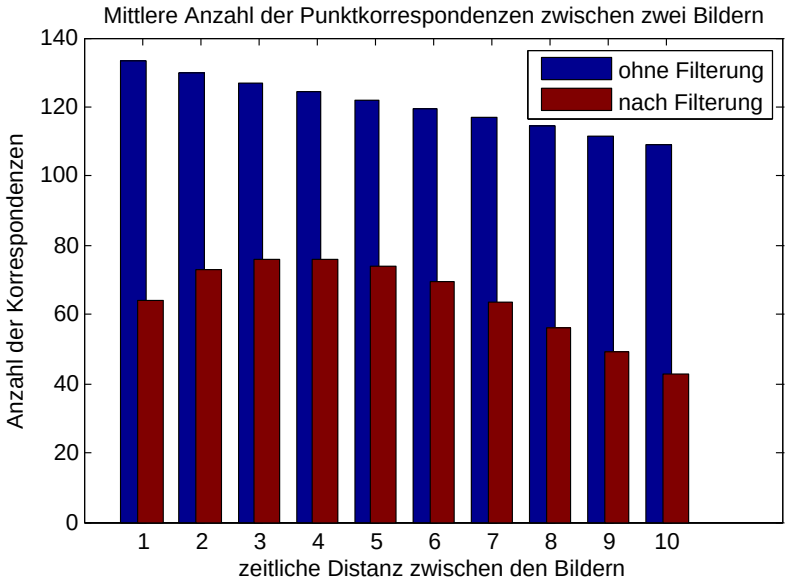


Abbildung 4.8: Mittlere Anzahl der Punktkorrespondenzen in Abhängigkeit von der Anzahl der Bilder die zwischen den Bildern A und B liegen: die blauen Balken zeigen die mittlere Anzahl aller gefundenen Korrespondenzen zwischen zwei Bildern, die Anzahl der Punktkorrespondenzen fällt leicht mit dem Abstand der Bilder. Die roten Balken zeigen die Korrespondenzen nach der Filterung, die Anzahl der Punktkorrespondenzen steigt mit dem zeitlichen Abstand der Bilder. Der Grund dafür liegt in der Anpassung des *Distanzkriteriums* an diesen Abstand.

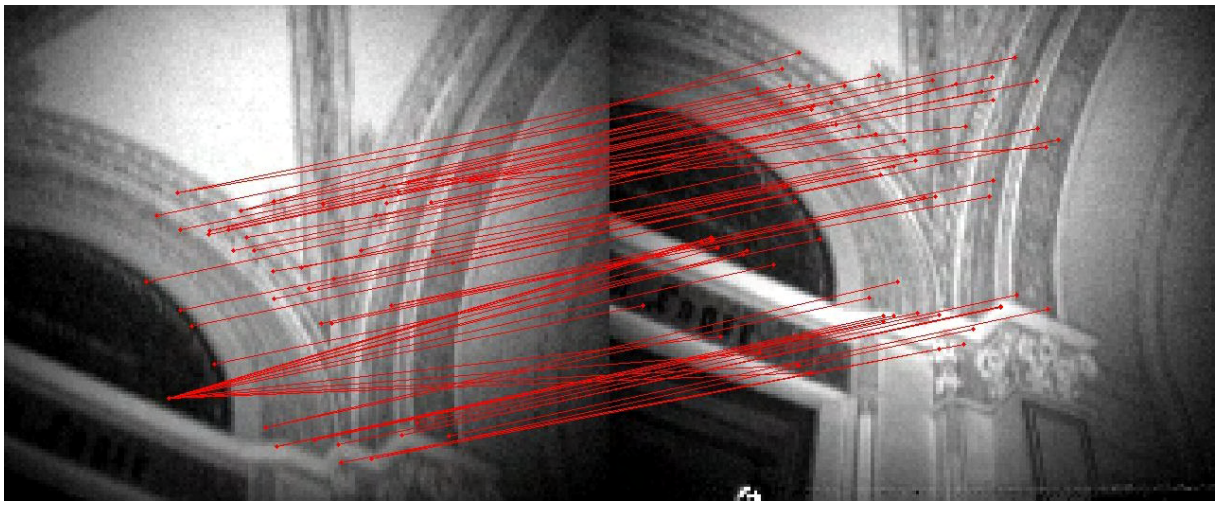


Abbildung 4.9: Punktkorrespondenzen: zu sehen sind 2 Amplitudenbilder einer bewegten Sequenz. Die Linien zeigen die Punktkorrespondenzen die mittels *surfmatch* gefunden wurden. Diese sind alle rot, da durch das *Positionskriterium* alle Korrespondenzen gefiltert werden.

es, bei einer schnelleren Bewegung bei der Aufnahme und bei größer werdender Anzahl von Bildern zwischen Bild A und Bild B dazu kommen, dass alle Punktkorrespondenzen gefiltert werden (siehe Abbildung 4.9). Eine Vergrößerung der Schranke so wie beim Distanzkriterium würde beim Positionskriterium dazu führen, dass die Anzahl der nicht gefilterten falschen Punktkorrespondenzen zu groß wird. Eine mögliche Lösung wäre ein adaptiver Schwellwert, mit einer an die Bewegung angepassten oberen und unteren Schranke. Beispielsweise könnten alle Positionsunterschiede der Verknüpfungspunkte in ein Histogramm eingetragen und ein oberer und unterer Schwellwert, der sich nach dem Maximum im Histogramm richtet, berechnet werden. Die Verbesserung des Positionskriteriums würde die Stabilität der näherungsweisen Orientierung erhöhen, da insbesondere bei Bildern mit schnelleren Bewegungen während der Aufnahme ohnedies aufgrund der Bewegungsunschärfe die Anzahl der Punktkorrespondenzen gering ist. Durch eine starre Schranke werden zusätzlich noch richtige Punktkorrespondenzen gefiltert. Dies kann beim anschließenden Rückwärtsschnitt zu Singularitäten führen.

Ermittlung der Objektkoordinaten der Verknüpfungspunkte

Nachdem die Verknüpfungspunkte in den Bildern einer bewegten Sequenz bestimmt sind, werden für das erste Bild der Sequenz die Objektkoordinaten für die Verknüpfungspunkte berechnet. Dies erfolgt gemäß Gleichung 3.10. Die dazu benötigte Orientierung stammt aus dem vorherigen statischen Bild.

Räumlicher Rückwärtsschnitt

Die Bestimmung der Orientierung des nächsten Bildes basiert auf den eben bestimmten Objektkoordinaten der Verknüpfungspunkte und wird mit einem räumlichen Rückwärtsschnitt durchgeführt. Dazu wird eine am IPF implementierte MATLAB-Funktion (*adjPhoRob.m*) verwendet. Die Funktion bestimmt einen räumlichen Rückwärtsschnitt durch eine robuste Ausgleichung nach vermittelten Beobachtungen (siehe Abschnitt 3.3.2.3). Die Funktion benötigt eine Liste mit Bildkoordinaten x, y mit den zugehörigen Genauigkeiten (Gewichtung) σ_x, σ_y, sowie die Objektkoordinaten der Verknüpfungspunkte. Die Gewichtung der Punkte erfolgt in Abhängigkeit von der Anzahl der bisherigen Verwendungen eines Punktes. Ist ein Punkt bereits in einem oder mehreren vorherigen Bildpaaren verwendet worden, so erhält dieser ein größeres Gewicht. Die Wahl der Gewichtung hat einen wesentlichen Einfluss auf das Ergebnis. Abbildung 4.10 zeigt das Ergebnis der näherungsweisen Orientierung bei unterschiedlichen Gewichtungsfunktionen für den Standpunkt 2. Folgende Gewichtungsfunktion liefert bei den verwendeten Daten

das beste Ergebnis (Abbildung 4.10(d)):

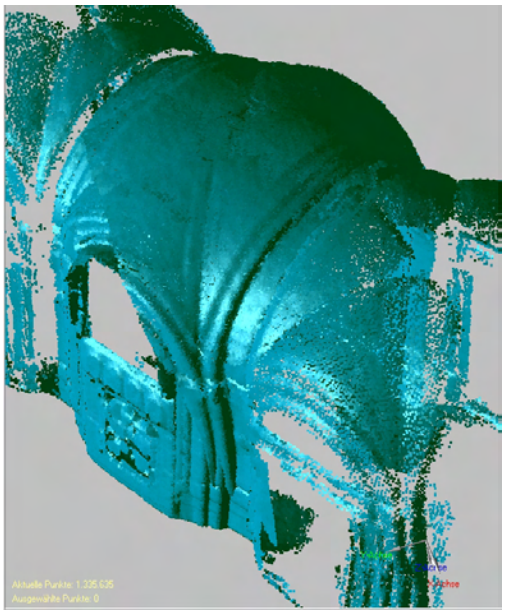
$$\frac{1}{\hat{\sigma_{x,y}}} = Gewicht = \frac{ln(anz)}{\sigma_{x,y}} \tag{4.4}$$

anz...Anzahl der bisherigen Verwendung des Punktes

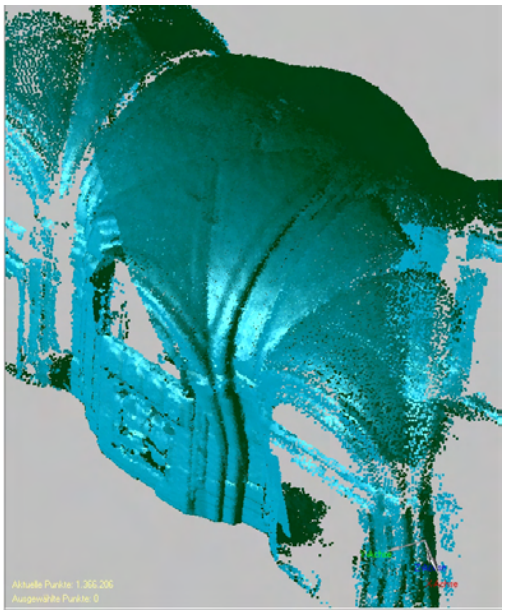
Im Mittel wird ein Verknüpfungspunkt 3,4 mal verwendet. Abbildung 4.11 zeigt die Häufigkeit der Wiederverwendung von Verknüpfungspunkten für die gesamte Aufnahme. Der maximal auftretende Wert ist 98.

Die Funktion *adjPhoRob.m* benötigt weiters Näherungswerte für die äußere Orientierung. Dazu werden die Werte des vorherigen Bildes übernommen.
Ein weiterer Schritt, der zur Stabilität bei der Berechnung der näherungsweisen Orientierung beiträgt, ist die Einführung der zusätzlichen Beobachtung des gemeinsamen Drehpunktes für einen Standpunkt. Dazu ist es notwendig zu unterscheiden, ob es sich um eine bewegte Sequenz handelt, die während eines Standpunktwechsels aufgezeichnet wurde, oder ob diese bei einem Schwenk der Kamera am gleichen Standpunkt registriert wurde (siehe Abschnitt 4.3.1). Für die letztgenannte bewegte Sequenz wird die Funktion *adjPhoRob.m* adaptiert und als zusätzliche Beobachtung die Position des Drehpunktes im Kamerakoordinatensystem mit einer großen Gewichtung hinzugefügt (siehe Abschnitt 3.3.2.3 sowie Anhang A).
Ist die Bestimmung der Orientierung für ein Bild mittels des räumlichen Rückwärtsschnittes aufgrund von Singularitäten nicht möglich, wird die Orientierung des vorherigen Bildes übernommen. Dies tritt bei der Auswertung der zweiten Aufnahme der Passage 15 mal auf. Insgesamt wurden 19192 Orientierungen bestimmt. Bei der robusten Ausgleichung werden durchschnittlich 122,8 Verknüpfungspunkte verwendet. Davon werden im Mittel 8,2 Punkte durch den Algorithmus als Ausreißer detektiert und verworfen.
Nachdem die Orientierung für das Bild ermittelt ist, werden wiederum die Objektkoordinaten der in diesem Bild neuen Verknüpfungspunkte bestimmt. Von diesen ausgehend wird die Orientierung des nächsten Bildes berechnet. Dies wird für die gesamte bewegte Sequenz sukzessive durchgeführt. Am Ende einer bewegten Sequenz ergibt sich die Orientierung des nächsten statischen Bildes.
Die in diesem Abschnitt genannten Schritte werden für jede bewegte Sequenz durchgeführt, sodass am Ende für alle 270 statischen Bilder eine näherungsweise äußere Orientierung vorliegt.

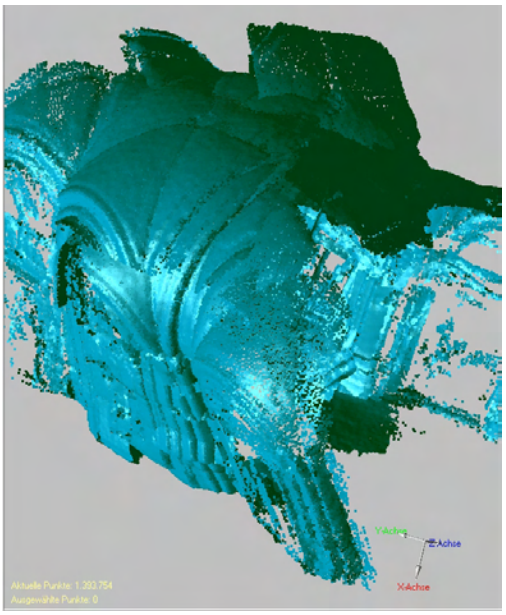
Das Ergebnis der näherungsweisen Orientierung ist insbesonders für Standpunkt 2 sehr zufriedenstellend (siehe Abbildung 4.12(b)). Bei Standpunkt 1 ist die Vororientierung im Bereich des



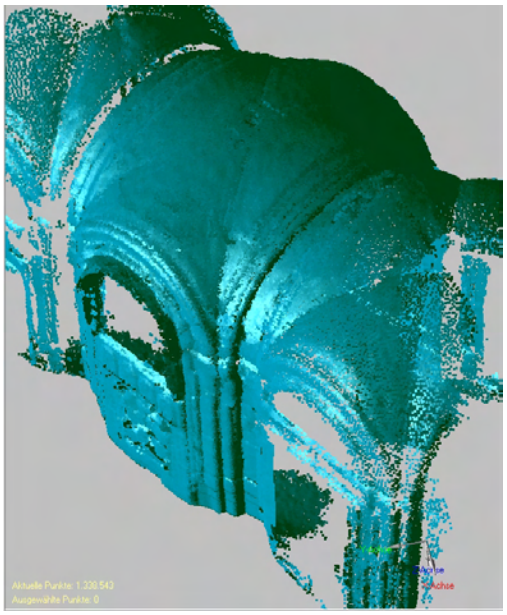
(a) $Gewicht = 1/\sigma_{x,y}$



(b) $Gewicht = \sqrt{anz}/\sigma_{x,y}$



(c) $Gewicht = anz^2/\sigma_{x,y}$



(d) $Gewicht = ln(anz)/\sigma_{x,y}$

Abbildung 4.10: Gewichtung der Verknüpfungspunkte: Ergebnis der näherungsweisen Orientierung bei unterschiedlichen Gewichtungsfunktionen für den Standpunkt 2. Die Gewichtung der Punkte erfolgt in Abhängigkeit von der Anzahl (anz) wie oft der Punkt bereits verwendet wurde. Das beste Ergebnis liefert: $Gewicht = ln(anz)/\sigma_{x,y}$.

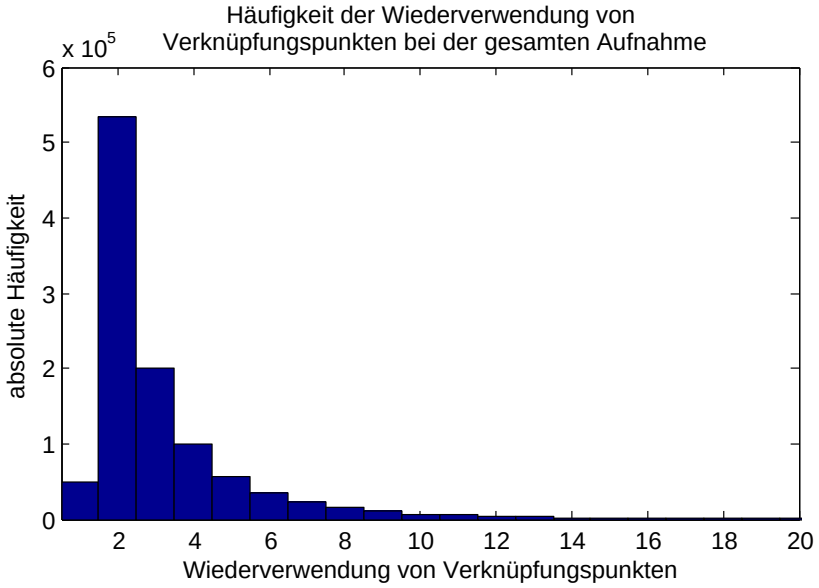
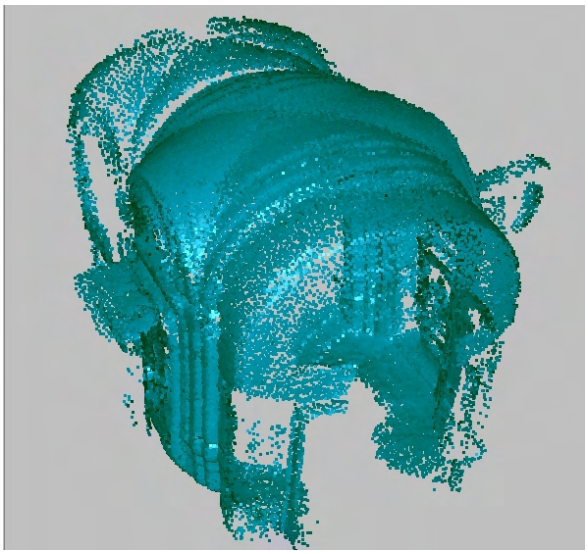
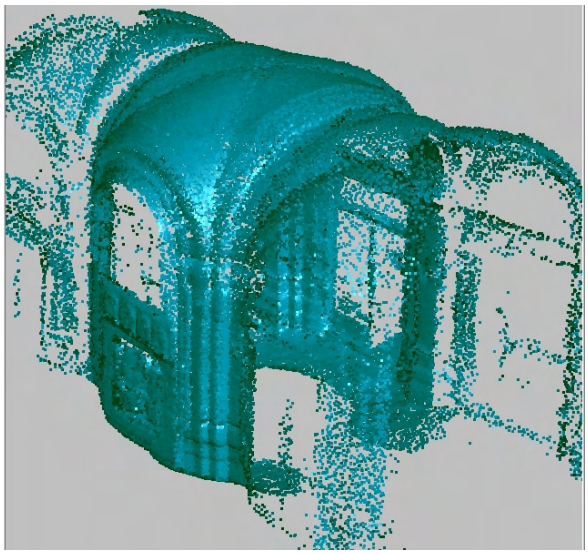


Abbildung 4.11: Häufigkeit der Wiederverwendung von Verknüpfungspunkten für die gesamte Aufnahme. Zur besseren Lesbarkeit wurde das Histogramm bei 20 Wiederverwendungen abgeschnitten. Der maximal auftretende Wert ist 98, im Mittel wird ein Verknüpfungspunkt 3,4 mal verwendet.

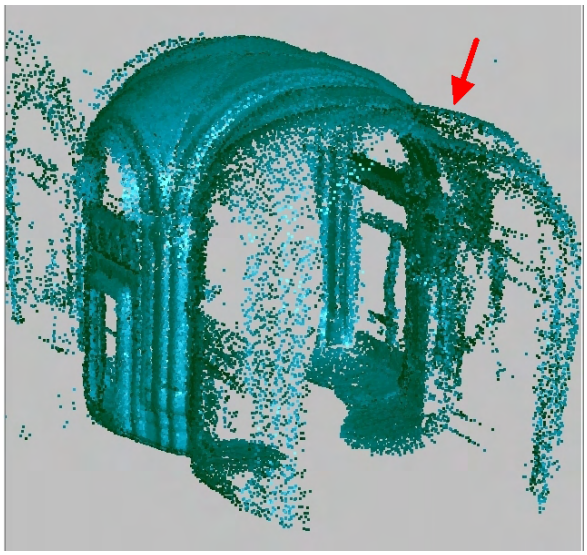
Deckengewölbes nicht optimal (siehe Abbildung 4.12(a)). Hier traten bei der Berechnung des räumlichen Rückwärtsschnittes auch Singularitäten auf. Diese sind teilweise auf zu ruckartige Bewegungen beim Schwenken der Kamera zurückzuführen. Die näherungsweise Orientierung für den Standpunkt 3 ist bei den Aufnahmen der nördlichen Wand, hier insbesonders wieder beim Deckengewölbe, ebenfalls nicht optimal. Beim Ablauf der Aufnahme wurde in diesem Bereich ein Fehler gemacht. Es wurden Teile dieser Seite irrtümlich doppelt aufgenommen, dadurch war es notwendig eine Schwenkbewegung vom Boden dieser Seite zur Decke der anderen Seite zu machen. Solche langen Schwenkbewegungen sind für die Stabilität der näherungsweisen Orientierung nicht optimal und möglicherweise eine Ursache für die etwas schlechtere Vororientierung von Standpunkt 3 gegenüber Standpunkt 2. Abbildung 4.12(d) zeigt die Orientierung der Standpunkte zueinander. Es ist ein deutlicher Knick zwischen Standpunkt 1 und Standpunkt 2 zu erkennen. Die Näherungswerte der Orientierung der Standpunkte zueinander sind jedoch für die anschließende Fein-Orientierung mittels *Geomagic* Studio ausreichend (siehe Abschnitt 4.3.3).



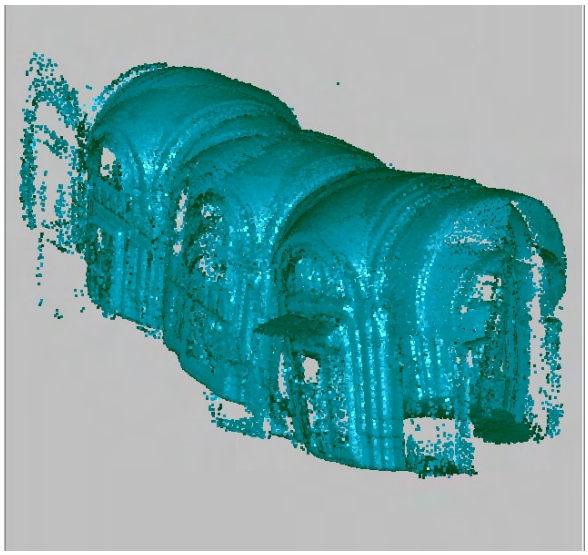
(a) Standpunkt 1



(b) Standpunkt 2



(c) Standpunkt 3



(d) Gesamt

Abbildung 4.12: Ergebnis der näherungsweisen Orientierung: (b) für Standpunkt 2 ist das Ergebnis der näherungsweisen Orientierung sehr zufriedenstellend. (a) Bei Standpunkt 1 sind im Bereich des Deckengewölbes Punktwolken zu erkennen, die quer zur Achsrichtung der Passage auseinander geschoben sind. Bei der Abbildung des Standpunktes 3 (c) ist zu erkennen, dass beim Deckengewölbe insbesondere beim Rückblick die Punktwolken verschoben sind. Aus Abbildung (d) ist eine Verdrehung der einzelnen Standpunkte zueinander erkennbar.

4.3.3 Fein-Orientierung

Nach der Bestimmung der näherungsweisen Orientierung erfolgt die Fein-Orientierung. In diesem Abschnitt wird die Durchführung der Fein-Orientierung beschrieben. Dabei wird auf die aufgetretenen Probleme im Zusammenhang mit den Aufnahmen der Distanzkamera eingegangen. Die Fein-Orientierung wird zum einen mittels eines ICP-Algorithmus, der am IPF implementiert wurde und in weiterer Folge als *TLS-ICP* bezeichnet wird, durchgeführt. Zum anderen erfolgt die Fein-Orientierung mittels *Geomagic* Studio. *TLS-ICP* hat den Vorteil, dass dieser die Berücksichtigung des exzentrischen Drehpunktes in der Ausgleichung ermöglicht. Die Fein-Orientierung der Punktwolken erfolgt zunächst für jeden Standpunkt separat. Anschließend werden die Standpunkte zueinander orientiert und abschließend erfolgt die Fein-Orientierung für alle Punktwolken gemeinsam. Vor der Fein-Orientierung müssen die einzelnen Punktwolken gefiltert, und die zu den Punkten zugehörigen Normalvektoren abgeschätzt werden.

Filterung der Punktwolken und Abschätzung der Normalvektoren

Der theoretische Hintergrund dazu wurde bereits in Abschnitt 3.3.3.1 erläutert. Die Filterung der Punktwolken und die Abschätzung der Normalvektoren erfolgt mit der am IPF entwickelten Software-Bibliothek *TLSptcl* [Nothegger u. Dorninger, 2007]. Der Nachteil von *TLSptcl* in Bezug auf die Verwendung bei Aufnahmen von Distanzkameras ist, dass dieser im unstrukturierten Objektraum (Punktwolke) arbeitet und nicht die vorhandene Struktur der Bildmatrix ausnützt. Die Funktion benötigt die Punktwolken der Aufnahmen mitsamt den zugehörigen Amplitudenwerten.

Für die Abschätzung eines Normalvektors werden die 64 (doppelter Default-Wert des *TLSptcl*) nächstgelegenen Punkte in der Punktwolke verwendet.

TLSptcl ermöglicht es nach der Normalvektorenberechnung die Punktwolke zu filtern (Punkte verwerfen). Dazu können Merkmale die bereits vor der Normalvektorenberechnung vorhanden sind (Amplitude, Position und abgeleitete Größen) und solche Größen, die bei der Berechnung der Normalvektoren anfallen, verwendet werden. Dies sind zum einen die Normalvektoren selbst, zum anderen aber auch die Eigenvektoren, die im Zuge der Abschätzung des Normalvektors berechnet werden. In dieser Arbeit werden folgende Filter verwendet: Distanz-, Amplituden-, Korona- und Rauheits-Filter.

Beim Distanz-Filter wird eine minimale Distanz von einem Meter vorgegeben. Objektpunkte weisen bei der verwendeten Aufnahmekonfiguration durchwegs wahre Distanzen von über 2

Metern auf. Distanzbeobachtungen von unter einem Meter müssen daher entweder grob falsch sein (z.B. gesättigte bzw. überbelichtete Pixel), oder die entsprechenden wahren Objektdistanzen überschreiten den Eindeutigkeitsbereich. Diese Punkte werden gefiltert (siehe Abbildung 4.13(b)).

Als besonders wichtig erweist sich der Amplituden-Filter, da die Qualität der Distanzmessung sehr stark von der Amplitude abhängt (siehe Abschnitt 2.3) und somit jene Punkte gefiltert werden, die eine unpräzise Distanzinformation besitzen. Diese Punkte sollen im Vorfeld des ICP gefiltert werden, da der Algorithmus selbst die punktweise unterschiedlichen Präzisionen nicht berücksichtigt, sondern alle Punkte gleich gewichtet. Zudem ist die verwendete Kalibrierung für extrem niedrige Amplituden nicht geeignet, sodass die entsprechenden Distanzbeobachtungen auch größere systematische Restfehler beinhalten würden. Der Schwellwert von 11000 wurde empirisch ermittelt, und ist von der Amplitudennormalisierung[1] und dem Wertebereich (von 0 bis $2^{16} - 1$) der Amplituden abhängig. In Abbildung 4.13(c) ist die Wirkung des Filters gut erkennbar: gefiltert werden vor allem Punkte, die sich im Inneren der Passage befinden. Bei diesen Punkten liegen die Distanzen außerhalb des eindeutigen Bereichs, daher werden diese nahe beim Projektionszentrum rekonstruiert (im 'Inneren' der Passage). Die Amplitude dieser Punkte ist aufgrund der großen wahren Distanz gering. Weiters werden Punkte im Bereich der Schaufenster und in Bodennähe beim Vor- und Rückblick gefiltert.

Mittels des Korona-Filters werden vor allem Punkte, die sich außerhalb des Eindeutigkeitsbereiches der Kamera befinden, gefiltert. In Abbildung 4.13(d) scheint es, als würden die Punkte zum Aufnahmezentrum hingezogen. Dadurch stehen die Normalvektoren mehr oder weniger im rechten Winkel zum Aufnahmevektor[2] und werden mittels des Korona-Filters detektiert. Der Schwellwert des Filters beträgt 80 Grad.

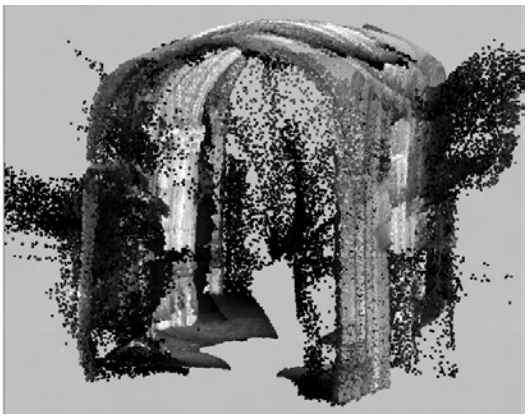
Der Rauheits-Filter entfernt jene Punkte, in deren Umgebung keine glatte Oberfläche vorhanden ist. So werden vor allem Punkte im Bereich der Fenster (siehe Abbildung 4.13(e)) und des eisernen Tores (Standpunkt 1) gefiltert. Der Wertebereich für den Parameter r (siehe Abschnitt 3.3.3.1) ist: $0 \leq r \leq 0,15$. Punkte, die außerhalb des Wertebereichs für r liegen werden verworfen.

Fein-Orientierung mit TLS-ICP

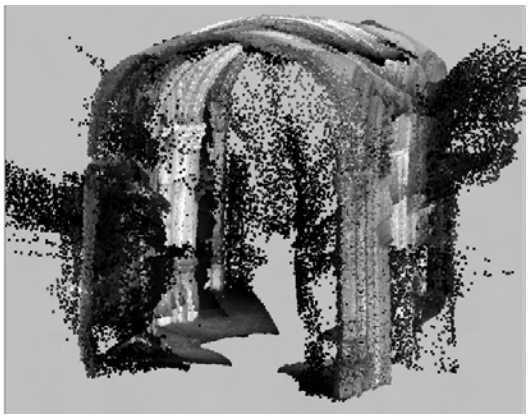
TLS-ICP arbeitet nach der Punkt-zu-Ebene-Methode, weshalb zur Berechnung Normalvekto-

[1] In dieser Arbeit wurde die Amplitudennormalisierung der Amplitudenbilder bereits am Beginn der Auswertung durchgeführt. Besser wäre es für die Amplitudenfilterung die Original-Amplituden zu verwenden.

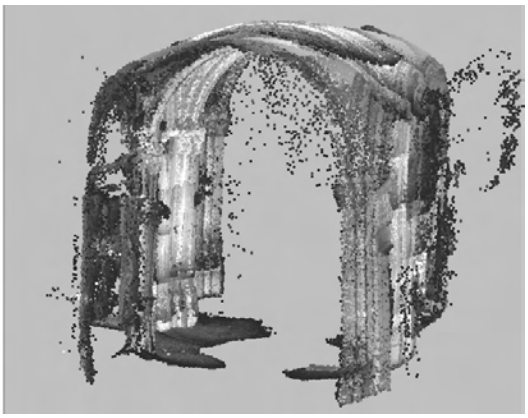
[2] Vektor vom Aufnahmezentrum zum Punkt



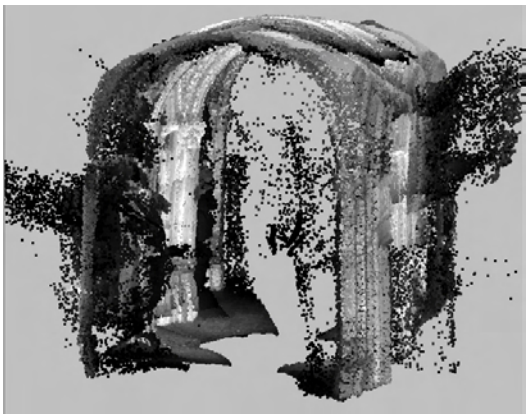
(a) Ohne-Filter



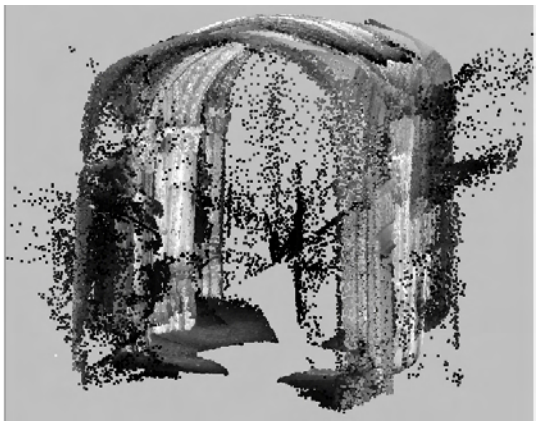
(b) Distanz-Filter



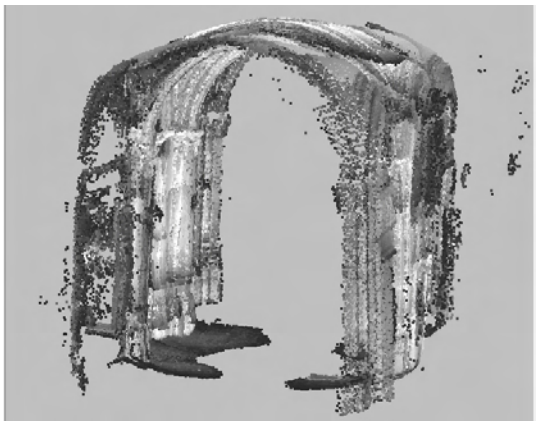
(c) Amplituden-Filter



(d) Korona-Filter



(e) Rauheits-Filter



(f) Alle-Filter

Abbildung 4.13: Filterung der Punktwolken: erkennbar sind die Auswirkungen der einzelnen verwendeten Filter auf die Punktwolken des Standpunktes 2. (b) Distanz-Filter: Punkte in der Umgebung des Standpunktes der Kamera werden gefiltert. (c) Amplituden-Filter: hat die größte Wirkung. (d) Korona-Filter: detektiert unter anderem Punkte im Mehrdeutigkeitsbereich der Kamera. (e) Rauheits-Filter: in erster Linie werden Punkte im Bereich der Schaufenster verworfen. (f) Alle-Filter: zeigt das Ergebnis nach Anwendung aller genannten Filter.

ren benötigt werden. Wie eingangs bereits erwähnt, können mit *TLS-ICP* Exzenterkoordinaten ausgeglichen werden, das heißt es kann, wie schon bei der näherungsweisen Orientierung, bei Aufnahmen mit dem Stativ von einem Standpunkt aus der gemeinsame Drehpunkt berücksichtigt werden. Dies erfolgt durch Transformation der Koordinaten des gemeinsamen Drehpunktes vom Objektkoordinatensystem ins Kamerakoordinatensystem. Anschließend wird eine Distanz der Länge Null (Exzenterbeobachtung) zwischen den transformierten Koordinaten und den Exzenterkoordinaten im Kamerakoordinatensystem beobachtet. Beim *TLS-ICP* ohne zusätzliche Beobachtungen werden Null-Distanzen zwischen korrespondierenden Objektpunkten und lokal geschätzten Ebenen beobachtet. Dabei gehen sämtliche Korrespondenzen gleich gewichtet in die Ausgleichung ein, weshalb sich das stochastische Modell a priori nicht auf die Lage der resultierenden Orientierungsparameter auswirkt. Exzenterbeobachtungen führen jedoch eine weitere Beobachtungsklasse ein, deren Gewichtung relativ zu jener der Objektkorrespondenzen entscheidend für das Ergebnis ist. Wird die Exzenterbeobachtung zu stark gewichtet, konvergiert *TLS-ICP* nicht, es verschieben sich einzelne Punktwolken-Gruppen gegenseitig und die Gesamtpunktwolke wird stark deformiert (siehe Abbildung 4.14 links). Ist die Gewichtung zu gering bzw. werden die Exzenterkoordinaten nicht berücksichtigt, verschieben sich die einzelnen Punktwolken in größerem Ausmaß, sodass der ICP zum falschen lokalen Minimum der Fehlerfunktion konvergiert. In Abbildung 4.14 rechts sind die Punktwolken der Wände ineinander geschoben. Für die Aufnahmen erwies sich eine 10 mal stärkere Gewichtung der Exzenterbeobachtung als optimal.

Die Filterung der Punktwolken im Vorfeld spielt ebenfalls eine wesentliche Rolle. Bei zu starker Filterung gehen kleine, aber für die Konvergenz des ICP wichtige Strukturen, verloren. (siehe Absatz: Filterung der Punktwolken).
Weitere Parameter des *TLS-ICP*, deren Werte empirisch angepasst wurden sind *pointPairDistanceThreshold* und *normalVecPairScalarProductThreshold*. Der erste Parameter gibt die maximale Distanz zwischen zwei Punkten an, damit diese als Punktkorrespondenz verwendet werden. Der zweite Wert ist der minimal zulässige Wert des Skalarproduktes der Normalvektoren einer Punktkorrespondenz. Verwendet wird $pointPairDistanceThreshold = 0,05$ und $normalVecPairScalarProductThreshold = 0,80$ (dies entspricht einem maximal zulässigen Winkel zwischen den Normalvektoren von $36,9$ Grad).
Probleme ergeben sich insbesonders im Bereich der Wände, da aufgrund des eingeschränkten Blickwinkels der Kamera nur kleinere Bereiche aufgenommen werden und diese eine zu geringe waagrechte Oberflächenstruktur (siehe Abbildung 3.8) aufweisen, um die Verschiebung in senkrechter Richtung (besser) bestimmen zu können. Aufgrund dessen liefert der ICP in die-

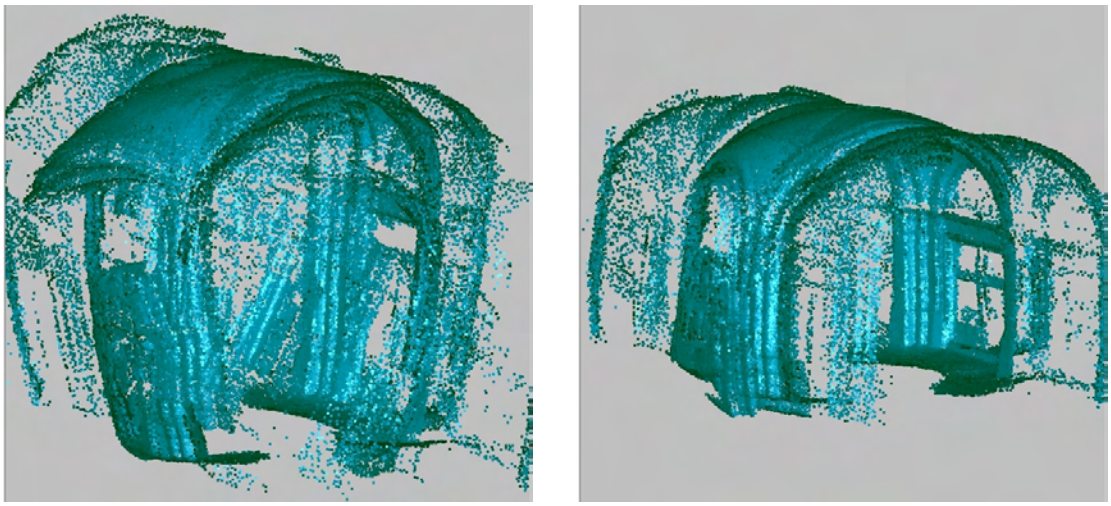
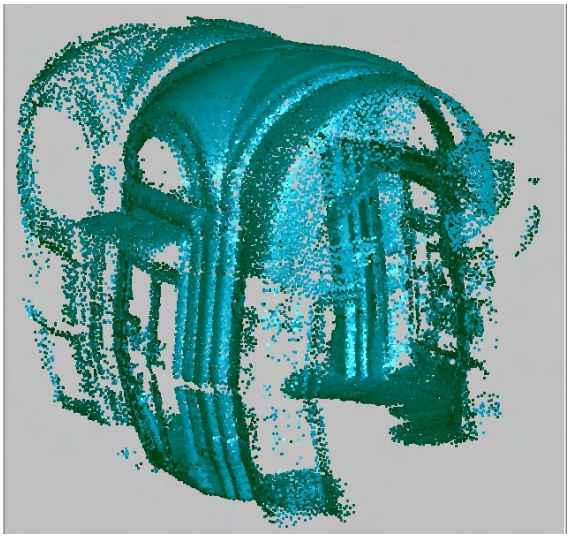


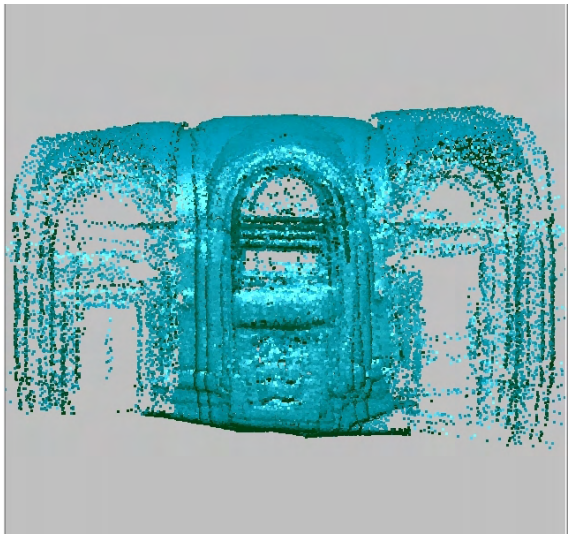
Abbildung 4.14: Orientierung mit *TLS-ICP* bei unterschiedlicher Gewichtung der Exzenterkoordinaten: links ist das Ergebnis bei einer hohen Gewichtung der Exzenterkoordinaten (Verhältnis der Gewichte von Exzenter- und Objektkorrespondenzen: 100:1) zu sehen. Dabei sind einzelne Punktwolken-Gruppen gegeneinander verschoben und die Gesamtpunktwolke ist stark deformiert. Die rechte Abbildung illustriert das Ergebnis der Berechnung ohne Berücksichtigung der Exzenterkoordinaten. Die Punktwolken der Wände sind ineinander geschoben.

sem Bereich, vor allem ohne Berücksichtigung des gemeinsamen Drehpunktes, kein stabiles Ergebnis (siehe Abbildung 4.14 rechts bzw. Abbildung 4.15). Eine Verbesserung würde hier vermutlich die zusätzliche Berücksichtigung der vorhandenen Textur bringen. Schwierigkeiten bereiten auch der Vor- und Rückblick zum benachbarten Gewölbe. Die Distanzen in diesen Aufnahmen sind für die verwendete Kamera sehr groß, und der Schnittwinkel zwischen Objektoberfläche und Kameraachse ist in diesem Bereich schleifend. Dadurch weisen die Aufnahmen eine geringere Signalamplitude und somit ein erhöhtes Rauschen sowie größere systematische Restfehler auf. Zudem ist die Auflösung am Objekt durch die größere Distanz geringer. Dadurch ist das Ergebnis des ICP in diesem Bereich nicht zufriedenstellend (siehe Abbildung 4.15(c)).

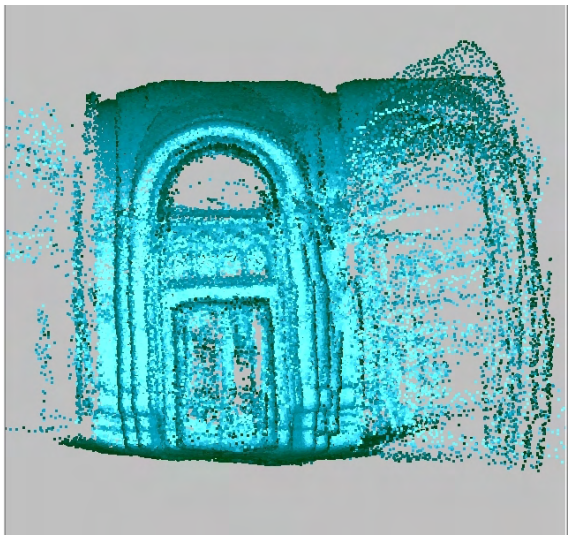
Teilweise ist das Ergebnis des *TLS-ICP* bei der Fein-Orientierung der Distanzkameraaufnahmen als noch nicht zufriedenstellend einzustufen. Beim Deckengewölbe über dem ersten Standpunkt erzielt die Fein-Orientierung eine deutliche Verbesserung der ursprünglichen Orientierung, welche bei Verwendung von *Geomagic* nicht ohne manuellen Eingriff erreicht werden konnte (siehe unten). Im Bereich der Wände kommt es aber zu einem unerwünschten auseinanderdriften der Aufnahmen (siehe Abbildung 4.15(a)). Zu Deformationen kommt es bei den Standpunkten 2 und 3 beim Vorausblick bzw. Rückblick und zusätzlich kommt es zu Verschiebungen in den Wänden (siehe Abbildung 4.15(b) und Abbildung 4.15(c)).



(a) Standpunkt 1



(b) Standpunkt 2



(c) Standpunkt 3

Abbildung 4.15: Ergebnis der Orientierung mit *TLS-ICP* für die 3 Standpunkte (Verhältnis der Gewichte von Exzenter- und Objektkorrespondenzen: 10:1): Bei Standpunkt 1 sind die Punktwolken des unteren Teils der in der Abbildung zum Betrachter näheren (südlichen) Wand nach unten verschoben. Die Ursache dafür ist die zu geringe Oberflächenstruktur. Bei den Standpunkten 2 und 3 sind zum einen wieder unerwünschte Verschiebungen der Punktwolken in den Wänden zu sehen, zum anderen treten insbesondere beim Vor- und Rückblick Deformationen auf. Eine Ursache ist die geringere Qualität der Distanzbeobachtungen in den Aufnahmen des Vor- und Rückblickes und die geringe Auflösung am Objekt in diesen Bereichen.

Es wurde auch versucht, statt der näherungsweisen Orientierung das Ergebnis der unten beschriebenen Fein-Orientierung mittels *Geomagic* als Startwert für den *TLC-ICP* zu verwenden. Auch diese Vorgehensweise führte zu keinem besseren Ergebnis. Daraus folgend liegt die Ursache des teils nicht zufriedenstellenden Ergebnisses des *TLS-ICP* nicht bei den Näherungswerten für die Orientierung. Eine Verbesserung des *TLS-ICP* könnte erreicht werden, indem wie oben bereits vorgeschlagen, die Amplitudenwerte beim ICP mit berücksichtigt werden. Weiters ließe sich der *TLS-ICP* erweitern indem die Tatsache berücksichtigt wird, dass die Kamera nicht um deren optische Achse gedreht wird. Eine Verbesserung der Kalibrierung würde das Konvergenzverhalten des ICP-Algorithmus verbessern. Auch eine höhere Auflösung (kleinere Oberflächenstrukturen werden besser abgebildet), sowie ein größerer Blickwinkel (mehr Oberflächenstrukturen werden in einer Aufnahme abgebildet) würde zu besseren Ergebnissen führen. Zusätzlich wäre der Einfluss von Multipath und Scattering auf das Ergebnis zu untersuchen. Der Einfluss des Multipath ist besonders stark im Bereich zwischen Boden und Wand zu erkennen.

Fein-Orientierung mit Geomagic Studio

Das Programmpaket *Geomagic* Studio 10 wurde bereits im Kapitel Modellierung in Abschnitt 3.4.1 vorgestellt. Mit Hilfe des Werkzeugs *Globale Registrierung* kann mit der Software ebenfalls die Fein-Orientierung der Punktwolken durchgeführt werden. Ein großer Vorteil von *Geomagic* ist die grafische Oberfläche und die Darstellung der Punktwolken. Dadurch ist ein einfacher manueller Eingriff, zum Beispiel durch Gruppenbildung bei den Punktwolken, möglich. Jedoch wird der gemeinsame Drehpunkt der Aufnahmen von einem Standpunkt nicht berücksichtigt, außerdem sind Details der zur Verfügung gestellten Funktionen in *Geomagic* nicht dokumentiert.

In einem ersten Schritt werden die gefilterten Punktwolken mit den zugehörigen Transformationsmatrizen der näherungsweisen Orientierung in *Geomagic* eingelesen. Danach werden die zu einem Standpunkt gehörenden Aufnahmen zu je einer Gruppe zusammengefasst. Zunächst wird die Fein-Orientierung für jeden Standpunkt separat bestimmt. Für Standpunkt 2 ist es möglich, direkt die Fein-Orientierung durchzuführen. Dazu wird eine Aufnahme vom Bereich des Deckengewölbes fixiert und mit dem Werkzeug *Globale Registrierung* die Orientierung durchgeführt.

Beim Standpunkt 1 ist das Ergebnis der *globalen Registrierung* im Bereich des Deckengewölbes weniger gut. Hier konnte mit *TLS-ICP* ein besseres Resultat erzielt werden. In *Geomagic* ist dazu ein manueller Eingriff notwendig. Wie in Abschnitt 4.2.2 beschrieben, erfolgt die Aufnahme in vertikalen Streifen. Da die näherungsweise Orientierung sukzessive vom ersten Bild ausge-

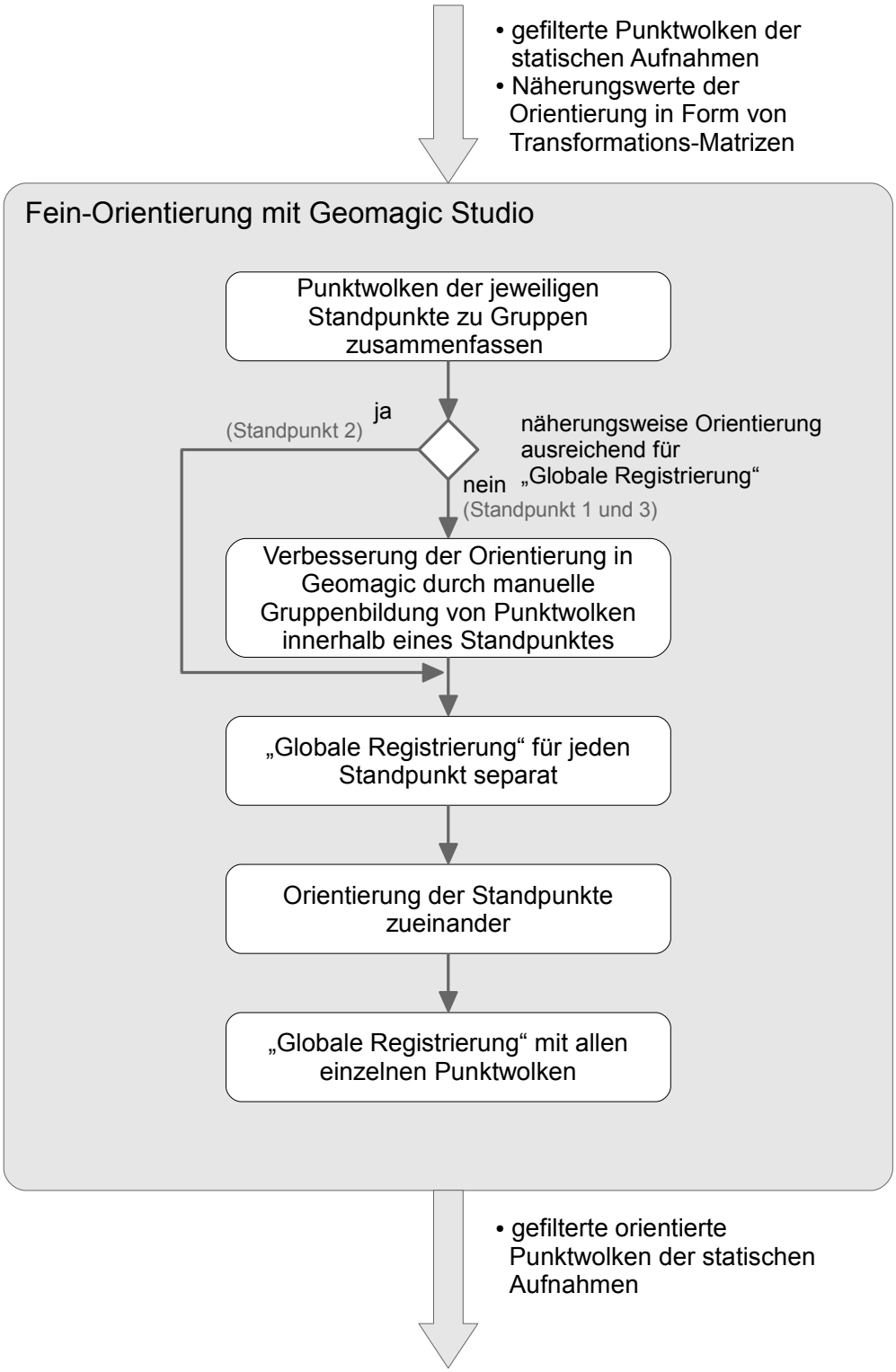
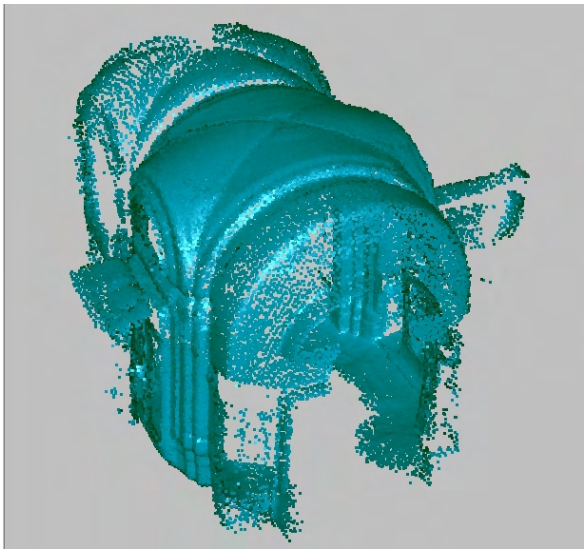


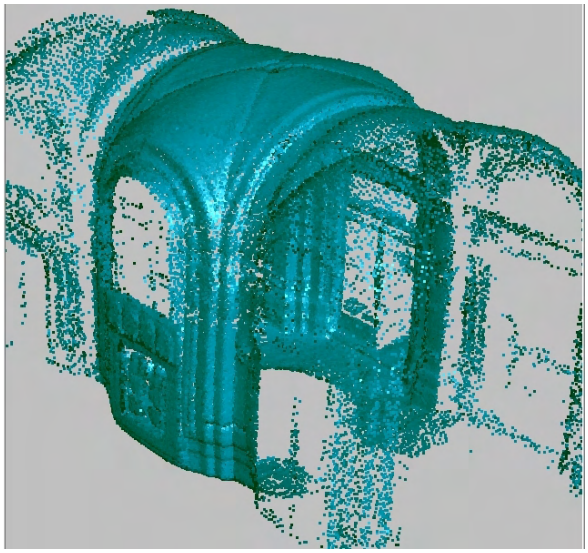
Abbildung 4.16: Ablauf der Fein-Orientierung in *Geomagic*

hend erfolgt, summieren sich auch die Fehler der Orientierung auf. Dem kann nun teilweise entgegengewirkt werden, indem benachbarte Streifen zueinander orientiert werden. Dazu werden die Aufnahmen von einem Streifen zu einer Gruppe zusammengefasst. Ist innerhalb eines Streifens zwischen zwei benachbarten Punktwolken ein größerer Orientierungsfehler zu erkennen, wird der Streifen wiederum in Gruppen aufgespaltet. Eine Gruppe beinhaltet die Punktwolken „oberhalb“ des Orientierungsfehlers, in der anderen Gruppe werden die Aufnahmen „unterhalb“ zusammengefasst. In weiterer Folge wird jene Gruppe, die am besten zum nächsten Streifen passt fixiert, und die andere Gruppe mittels des Werkzeugs *Globale Registrierung* zur fixierten Gruppe orientiert. Eine andere Möglichkeit die „obere“ und die „untere“ Punktwolkengruppe zueinander und zu den benachbarten Streifen anzupassen besteht darin, den benachbarten Streifen zu fixieren und die beiden Gruppen an diesen zu orientieren. Bei Standpunkt 1 war eine Korrektur des ersten Streifens notwendig. Im Anschluss kann die Orientierung der Streifen zueinander verbessert werden. Dazu wird ausgehend vom ersten Steifen immer der nächste fixiert und mittels der *Globalen Registrierung* die Orientierung zu diesem verbessert. In weiterer Folge wird der erste und zweite Streifen zu einer neuen Gruppe zusammengefasst und die Orientierung dieser wird neuerlich zum nächsten Streifen, der wiederum fixiert wird, verbessert. Dies wird für die erste Wand des Standpunktes durchgeführt. Für die gegenüberliegende Wand wird derselbe Algorithmus in umgekehrter Reihenfolge, d.h. beginnend mit dem letzten Streifen, angewandt. Abschließend wird die Orientierung der Punktwolken der beiden Hälften zueinander verbessert, wobei die zweite Hälfte fixiert wird. Nach der Verbesserung der näherungsweisen Orientierung erfolgt die simultane Fein-Orientierung aller Punktwolken des Standpunktes, wobei wiederum eine Aufnahme vom Bereich des Deckengewölbes fixiert wird.
Die Fein-Orientierung des Standpunktes 3 wird auf die gleiche Weise durchgeführt wie jene von Standpunkt 1, da hier ebenfalls die näherungsweise Orientierung nicht ausreichend genau ist, um direkt die Fein-Orientierung durchführen zu können.
Bei dem Werkzeug *Globale Registrierung* wurde immer die sogenannte *Rutschkontrolle* verwendet, um größere unerwünschte Verschiebungen im Bereich der Wände zu vermeiden.

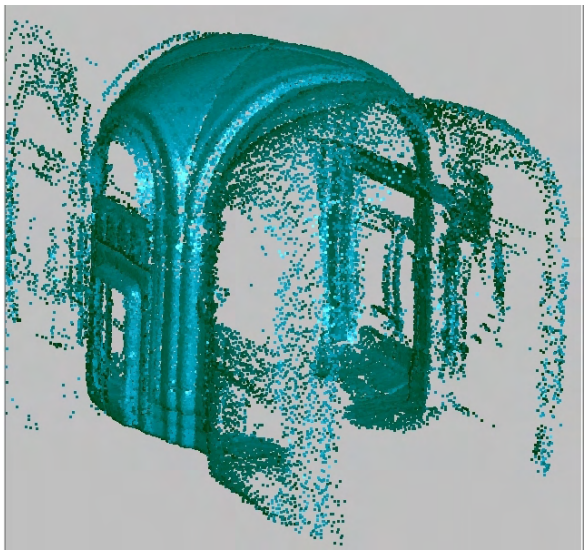
Nach der Fein-Orientierung der einzelnen Standpunkte müssen diese noch zueinander orientiert werden. Dazu werden zuvor die Punkte, die sich in der vom Standpunkt weiter entfernten Hälfte des Vor- und Rückblickes befinden gelöscht, da diese wegen größerer systematischer Restfehler zu einer Verschlechterung der Gesamt-Orientierung beitragen würden. Anschließend wird der Standpunkt 2 fixiert und die beiden anderen, wieder unter Verwendung der *Globalen Registrierung*, relativ zu diesem orientiert. In einem letzten Schritt werden alle Punktwolken der einzelnen Aufnahmen nochmals mittels der *Globalen Registrierung* zueinander orientiert



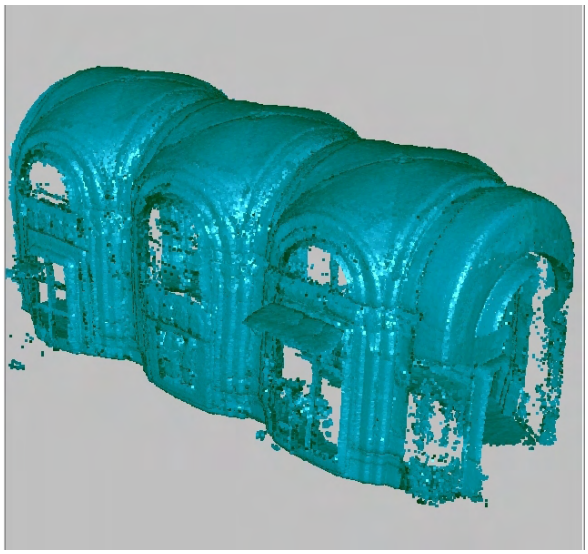
(a) Standpunkt 1



(b) Standpunkt 2



(c) Standpunkt 3



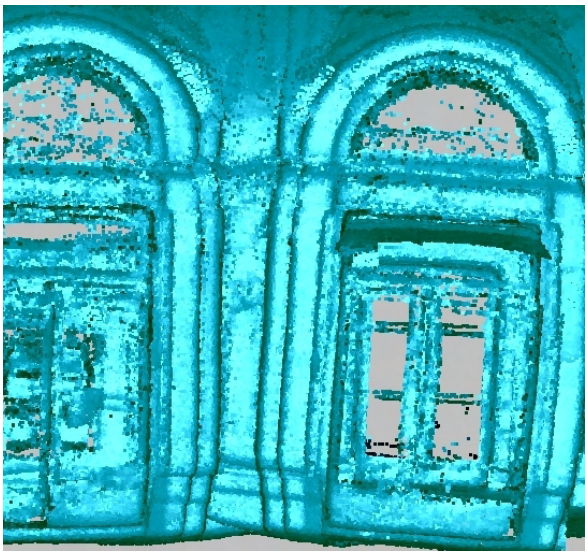
(d) Gesamt

Abbildung 4.17: Ergebnis der Fein-Orientierung mit *Geomagic*: zu sehen sind die separat orientierten Punktwolken der drei Standpunkte (vor der Gesamt-Orientierung), sowie das Endergebnis der Fein-Orientierung (d). Bei den Ergebnissen der Orientierung der einzelnen Standpunkte konnten in den Abbildungen (a)-(c) keine größeren Deformationen beobachtet werden. Beim Gesamtergebnis ist zu sehen, dass der dritte (in der Abbildung der linke) Standpunkt gegenüber den anderen leicht verdreht ist.

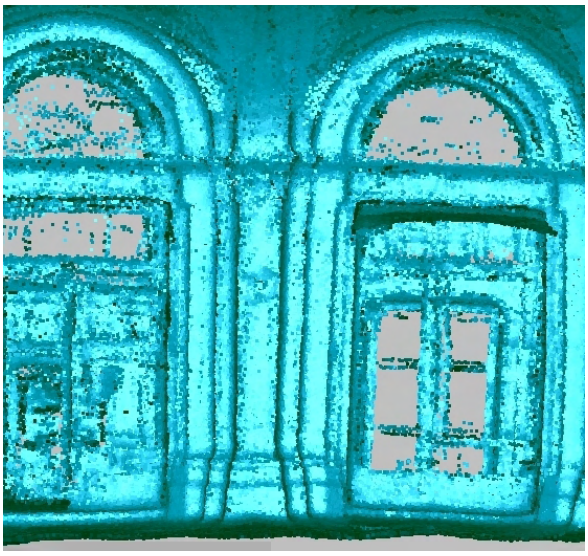
(Gesamt-Orientierung).
Das Ergebnis der Fein-Orientierung durch *Geomagic* Studio ist in Abbildung 4.17 zu sehen. Bei den einzelnen Standpunkten können stellenweise kleinere Sprünge am Übergang der Punktwolken beobachtet werden (siehe Abbildung 4.18(c)). Dies könnte auf systematischer Restfehler die nach der Kalibrierung in den einzelnen Aufnahmen noch vorhanden sind, zurückzuführen sein. Weiters kommt es zu unerwünschten Verschiebungen von einzelnen Punktwolken (Abbildung 4.3.3 zeigt Verschiebungen im Standpunkt 1) oder eines ganzen Bereiches (in Abbildung 4.18(c) von Standpunkt 2 ist die linke Säule der Wand länger als die rechte). Dies ist auf die mangelnde Bestimmtheit aller Freiheitsgrade im Bereich der Wände (siehe oben) zurückzuführen. Zusätzlich können allgemein beim Übergang der Wände zum Boden größere Deformationen in den einzelnen Punktwolken beobachtet werden. Die Ursache dafür sind in erster Linie Multipatheinflüsse bei den Distanzmessungen. Bei Standpunkt 3 ist die nördliche Wand etwas verdreht. Dies wird insbesondere nach dem gegenseitigen Orientieren der Standpunkte sichtbar (siehe Abbildung 4.18(a)). Generell treten an den Grenzen zwischen den Punktwolken unterschiedlicher Standpunkte nach der gegenseitigen Orientierung Sprünge auf. Diese werden durch die *Globale Registrierung* aller Punktwolken verringert (siehe Abbildung 4.18(b)). Bei den beiden Bögen im Bereich der Decke können auch nach der *Globalen Registrierung* noch größere Unterschiede bei den Punktwolken von den verschiedenen Standpunkten festgestellt werden (Abbildung 4.18(d)). Eine Ursache dafür liegt darin, dass bei der Aufnahme von der dem Standpunkt gegenüberliegenden Seite des Bogens von diesem nur die Kontur sichtbar ist, und es dadurch bei den Distanzmessungen zu *mixed pixels* (siehe Abschnitt 3.3.3.1) kommt. Insgesamt ist das Ergebnis der Fein-Orientierung mittels *Geomagic* besser als jenes bei der Orientierung mit *TLS-ICP*. Der Grund dafür dürfte die in *Geomagic* verwendete *Rutschkontrolle* sein.

4.3.4 Erstellung des Modells

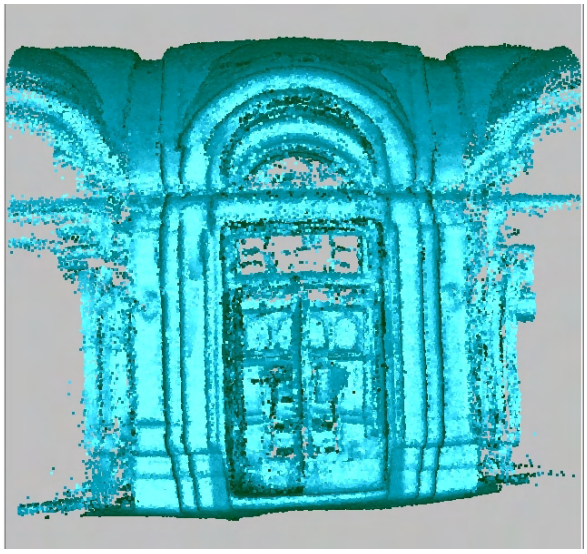
Die Erstellung des Modells erfolgt, wie schon die Fein-Orientierung, mit dem in Abschnitt 3.4.1 vorgestellten Programmpaket *Geomagic* Studio. Das aus dieser Arbeit resultierende Modell wird vor allem zu Visualisierungszwecken abgeleitet. Dazu wird eine Vereinigung und Dreiecksvermaschung der Punktwolken durchgeführt.
Vor der Triangulation sind einige Bearbeitungsschritte notwendig, die mit *Geomagic* durchgeführt werden. Wie in Abschnitt 4.3.3 beschrieben, kommt es im Bereich der Wände zu unerwünschten Verschiebungen von einzelnen Punktwolken (siehe Abbildung 4.3.3). Aufgrund



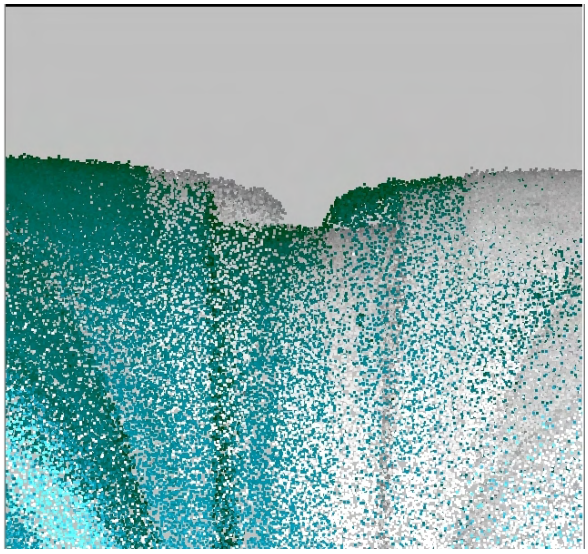
(a) Standpunkt 2 und 3 vor Gesamt-Orientierung



(b) Standpunkt 2 und 3 nach Gesamt-Orientierung



(c) Standpunkt 2 nach Gesamt-Orientierung



(d) Detail, Bogen im Bereich der Decke

Abbildung 4.18: Details Orientierung mit *Geomagic*: die Abbildung (a) zeigt den Übergang der Standpunkte 1 und 2 nach der gegenseitigen Orientierung der Standpunkte und vor der Gesamt-Orientierung. Insbesondere im Bereich des Bodens sind größere Sprünge zwischen den beiden Standpunkten zu sehen. Diese Deformationen sind nach der Gesamt-Orientierung geringer (b). In (c) ist der Standpunkt 2 nach der Gesamt-Orientierung zu sehen. Es können stellenweise kleinere Sprünge am Übergang der einzelnen Punktwolken beobachtet werden. Zudem ist in dieser Abbildung zu sehen, dass die linke Säule der Wand länger als die rechte ist. Die Ursache dafür ist die zu geringe Oberflächenstruktur in waagrechter Richtung. Bei dem Bogen im Bereich der Decke (d) sind die Punktwolken von den verschiedenen Standpunkten unterschiedlich eingefärbt. Dabei sind größere Unterschiede zwischen den Punktwolken zu sehen. Eine Ursache dafür dürften *mixed pixels* in den Distanzbeobachtungen sein.

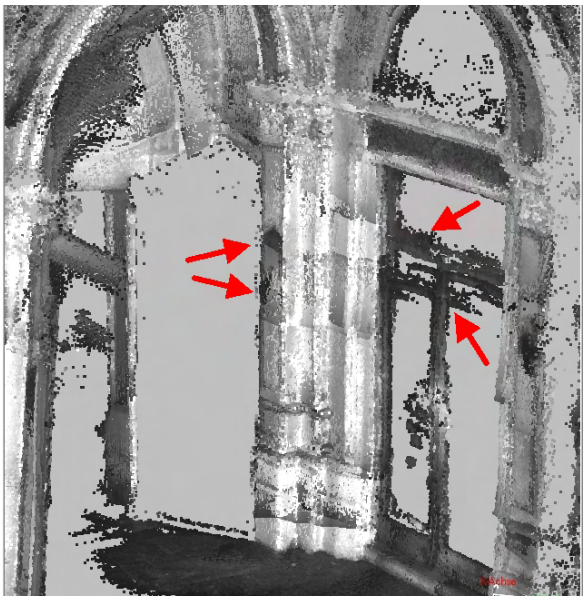


Abbildung 4.19: Detail Orientierung mit *Geomagic*: zu sehen ist eine unerwünschte Verschiebung einer Punktwolke in der ersten Säule des Standpunktes 1. Diese tritt trotz in *Geomagic* aktivierter *Rutschkontrolle* auf. Die Ursache ist die zu geringe Oberflächenstruktur in vertikaler Richtung.

der großen Überdeckung (mindestens 50%) bei den Aufnahmen ist es möglich, für die Modellierung einzelne verschobene Punktwolken zu entfernen. Weiters müssen offensichtliche Ausreißer, die nicht im Vorfeld gefiltert wurden, gelöscht werden. Dazu stellt *Geomagic* das Werkzeug *Getrennte Komponenten Auswählen* zur Verfügung. Dieses liefert jedoch kein zufriedenstellendes Ergebnis. Deswegen erfolgt das Entfernen dieser Punkte durch manuelles Selektieren und Löschen. Zusätzlich werden die vollständigen Vor- und Rückblicke entfernt (teilweise wurde dies bereits bei der Orientierung der Standpunkte zueinander durchgeführt), da diese eine geringere Qualität aufweisen und so zu einer Verschlechterung der Triangulation führen würden. Die Punkte im Bereich des Bogens zwischen den Standpunkten, die wie oben beschrieben (siehe Abbildung 4.18(d)) von der einen Seite nur mangelhaft aufgenommen worden sind, wurden ebenfalls mit den Vor- und Rückblicken gelöscht. An manchen Stellen sind systematische Abweichungen zwischen den Punktwolken vorhanden. Dadurch sind beim triangulierten Modell raue Oberflächen, wie zum Beispiel im Bereich des Deckengewölbes in Abbildung 4.20 vorhanden. Eine Ursache dafür könnte darin liegen, dass systematische Restfehler in den Distanzbeobachtungen nach deren Kalibrierung übrig bleiben. Im Zuge der Vorbereitung zur Triangulation wurden daher diese Bereiche teilweise manuell in den Punktwolken gelöscht. Es wäre aufgrund der großen Überdeckung auch möglich, generell die Ränder der Punktwolken zu entfernen. Eine

andere Ursache für die Fehler dürften Unzulänglichkeiten beim Ergebnis der Fein-Orientierung sein (siehe Abschnitt 4.3.3).
Nach diesen vorbereitenden Arbeitsschritten werden in *Geomagic* mit dem Werkzeug *Vereinigen* in einem Schritt die einzelnen Punktwolken zusammengeführt, reduziert, geglättet und die Dreiecksvermaschung berechnet. Als Parameter können eine lokale und globale Rauschverringerung in den Stufen *Min*, *Mittel*, *Max* und *Aus* angegeben werden. Aufgrund der Daten der Distanzkamera bzw. der vorliegenden Orientierung der Punktwolken zueinander werden die Parameter auf *Max* gesetzt. Weiters kann die Anzahl der Zieldreiecke bzw. der Punktabstand angegeben werden. Da auch die Textur von der Anzahl der Dreiecke abhängig ist, wurde hier keine Reduzierung vorgenommen. Abbildung 4.21(b) zeigt das Ergebnis der Triangulation nach der Anwendung von *Vereinigen*.

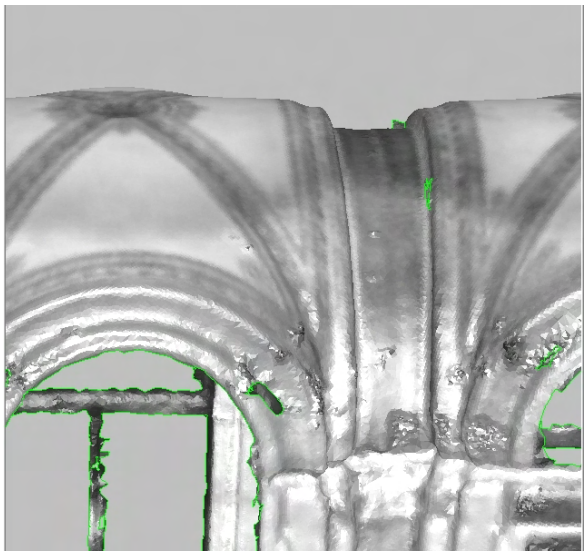


Abbildung 4.20: Detail Triangulation: zeigt ein Detail der triangulierten Oberfläche, die direkt aus den orientierten Punktwolken mit *Geomagic* berechnet wurde. Im Bereich der Bögen sind raue Oberflächen zu sehen. Eine Ursache dafür könnten systematische Restfehler im Randbereich der kalibrierten Distanzbilder sein. Diese Bereiche werden teilweise manuell gelöscht.

Abbildung 4.21 zeigt das Ergebnis der Triangulation. Wie aus Abbildung 4.21(a) und Abbildung 4.20 ersichtlich, sind im Übergangsbereich der Standpunkte lokale Deformationen im Modell vorhanden. Diese treten insbesondere zwischen Standpunkt 1 und 2 bei der südlichen Wand und weiters im Bereich des Bogens zwischen den Standpunkten auf. Dort sind die beiden Seiten des Bogens nicht parallel. Die Ursachen dafür liegen in Orientierungsfehlern und möglicherweise auch in systematischen Restfehlern, die nach der Kalibrierung noch vorhanden sind.

Die in Abschnitt 4.3.3 beschriebenen kleineren Sprünge am Übergang der Punktwolken konnten innerhalb eines Standpunktes bei der Bearbeitung der Punktwolken vor der Triangulation (manuelles entfernen von Teilen von einigen Punktwolken im Bereich des Deckengewölbes) weitgehend beseitigt werden. Wie bereits in Abschnitt 4.3.3 ausgeführt, wurden auch nach der Triangulation beim Übergang der Wände zum Boden größere Deformationen festgestellt. Aus der Abbildung 4.21(a) ist ersichtlich, dass die Höhe von Standpunkt 1 zu Standpunkt 3 hin etwas abnimmt und der Standpunkt 3 gegenüber den anderen Standpunkten leicht verdreht ist.

4.4 Analyse der Modellqualität

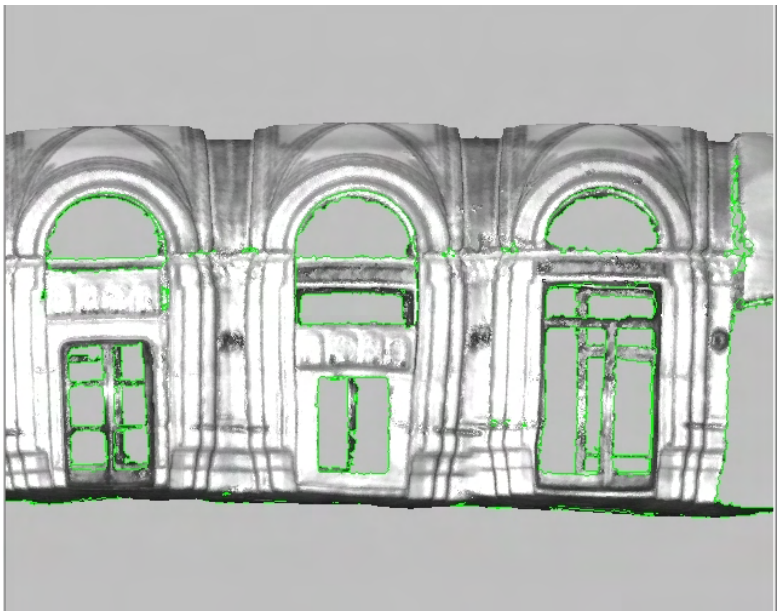
In diesem Abschnitt wird eine Analyse der Qualität des Modells durchgeführt. Zunächst werden die Abstände der Punktwolken zur Modelloberfläche bestimmt. Die Beurteilung der äußeren Genauigkeit erfolgt über Sperrmaße, die mit dem Modell verglichen werden.

4.4.1 Abstand Punktwolke zur Modelloberfläche

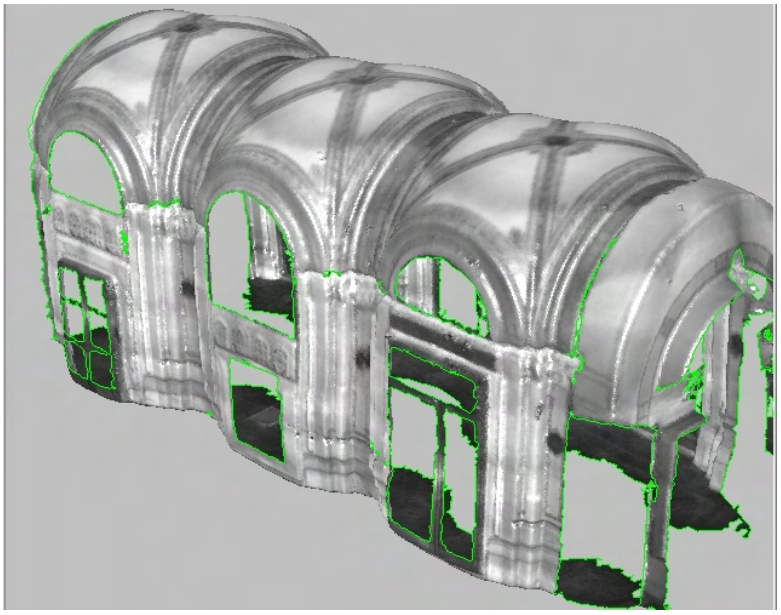
Die Ermittlung des Abstandes der Punkte zur Modelloberfläche erfolgt mit dem *Geomagic* Werkzeug *3D-Vergleich*. Im Vorfeld des *3D-Vergleiches* werden die einzelnen Punktwolken vereinigt. Abbildung 4.22 zeigt das Ergebnis des *3D-Vergleiches* für die gesamte Aufnahme. Dabei wird die Modelloberfläche in Abhängigkeit der Punktabstände von der Modelloberfläche eingefärbt. Der durchschnittliche Abstand der Punkte zur Modelloberfläche beträgt $\pm 1cm$, der maximale Abstand ist $+12,5cm$ und $-21cm$ und die Standardabweichung beträgt $1,3cm$. Größere Abweichungen sind im Bereich von Kanten zu sehen. Eine Ursache dafür liegt in der Glättung, die bei der Anwendung des Werkzeugs *Vereinigen* durchgeführt wird. Weiters ist aus Abbildung 4.22 links ersichtlich, dass beim Deckengewölbe kreisförmige Bereiche mit größeren Abweichungen vorhanden sind. Das Modell könnte in diesem Bereichen noch verbessert werden, indem die Randbereiche der Punktwolken stärker gefiltert werden.

4.4.2 Äußere Genauigkeit

Zur Beurteilung der äußeren Genauigkeit wurden Sperrmaße am Objekt und im Modell gemessen und miteinander verglichen, dadurch können Maßstabsfehler detektiert werden. Dabei kann



(a) Modell Seitenansicht



(b) Modell Schrägansicht

Abbildung 4.21: Trianguliertes Modell: zu sehen ist das Ergebnis der Triangulation, die mit *Geomagic* durchgeführt wurde. Im Übergangsbereich der Standpunkte sind lokale Deformationen im Modell vorhanden. Diese sind insbesondere zwischen Standpunkt 1 und 2 in (a) zu sehen. Weiters ist aus (a) ersichtlich, dass die Höhe von Standpunkt 1 zu Standpunkt 3 hin etwas abnimmt und Standpunkt 3 gegenüber den anderen Standpunkten leicht verdreht ist.

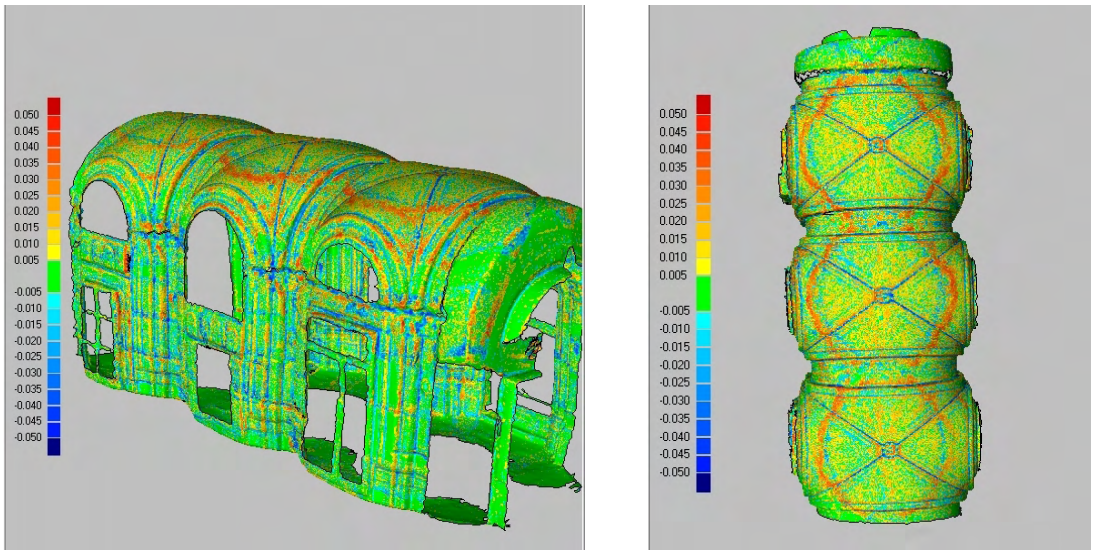


Abbildung 4.22: Abstand Punktwolke zur Modelloberfläche: zu sehen ist das Ergebnis eines 3D-Vergleiches, der mit *Geomagic* durchgeführt wurde. Die Modelloberfläche ist in Abhängigkeit des Abstandes der Punkte von der Modelloberfläche eingefärbt. Punkte, die in das Innere des Modells abweichen, besitzen negative Residuen und Punkte die nach außen abweichen ein positives. Der Wertebereich der Skala reicht von $-0,05m$ bis $+0,05m$. Der durchschnittliche Abstand der Punkte zur Modelloberfläche beträgt $\pm 1cm$, der maximale Abstand $+12,5cm$ und $-21cm$, die Standardabweichung ist $1,3cm$. Größere Abweichungen sind im Bereich von Kanten zu sehen. Eine Ursache dafür liegt in der Glättung, die im Zuge der Triangulation durchgeführt wird. Weiters sind im Deckengewölbe kreisförmige Bereiche mit größeren Abweichungen ersichtlich.

beobachtet werden, dass das Modell erheblich größer ist. In Längsrichtung der Passage ist das Modell um ca. 5% länger und quer zur Passage beträgt die Abweichung ca. 8% (ca. $30cm$). Eine mögliche Ursache dafür ist, dass die verwendete Kalibrierung die Außentemperatur nicht brücksichtigt. Die Daten für die Bestimmung der Kalibrierung wurden bei Zimmertemperatur aufgenommen, während bei der Aufnahme der Passage die Temperatur ca. 5°C betrug. In Abbildung 4.23 sind die Sperrmaße eingetragen. Die Messung erfolgte in einer Höhe von ca. $1,8m$.

4.5 Diskussion und Ausblick

Allgemein gesprochen ist die Aufnahme von Objekten zur Vervollständigung von Stadtmodellen mittels einer Distnazkamera zum gegenwärtigen Zeitpunkt nur bedingt zu empfehlen. Die

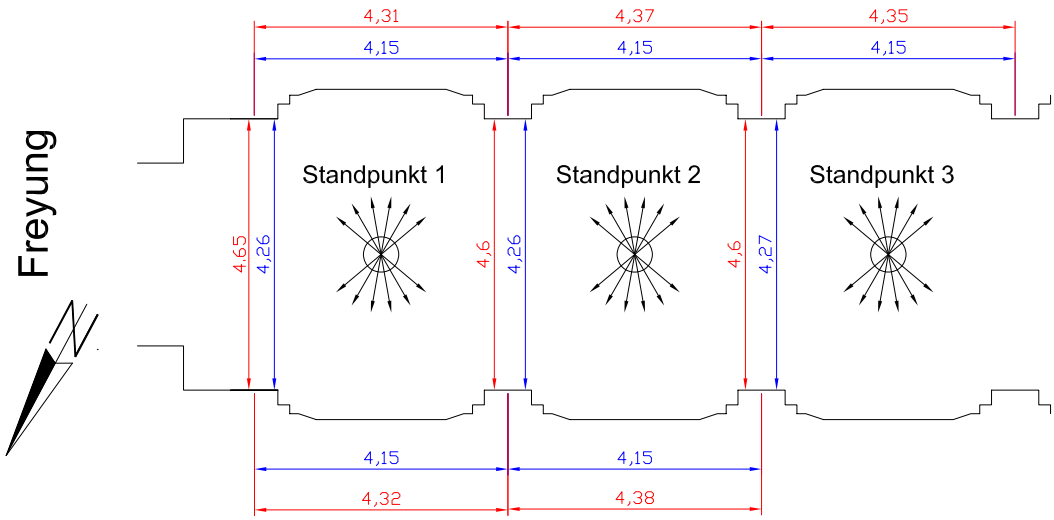


Abbildung 4.23: Äußere Genauigkeit: in der Skizze sind die gemessenen Sperrmaße am Objekt (blau) und im Modell (rot) eingezeichnet. Es ist ersichtlich, dass das Modell erheblich größer ist.

Reichweite der Distanzkameras ist relativ gering. Bei der Aufnahme der Passage treten für die RIM-Kameras sehr große Distanzen auf, wodurch die Präzision der Aufnahmen stark abnimmt (siehe Abschnitt 2.3). Es ist daher notwendig, die Aufnahmen unter Verwendung eines Stativs durchzuführen und mehrere Bilder mit gleicher äußerer Orientierung zu mitteln. Die Aufnahme der Passage erfolgt ausschließlich aus der Mitte, damit Passanten die Möglichkeit haben, ungehindert vorbei zu gehen und nicht im Bildmaterial aufscheinen. Zudem ist das Blickfeld der Kamera eingeschränkt. Deswegen werden nur relativ kleine Objektbereiche pro Bild aufgenommen. Dies und die hohe Überdeckung der Aufnahmen (siehe Abschnitt 4.2.2) führten dazu, dass zur Erstellung des Modells über 250 (!) statische Bilder notwendig waren. Daher ist zum gegenwärtigen Zeitpunkt die Aufnahme mittels einer Distanzkamera von einem wirtschaftlichen Standpunkt aus noch nicht sinnvoll möglich. Die Distanzkameras befinden sich noch in der Entwicklung. Es ist daher zu hoffen, dass bei künftigen Kameramodellen die eben genannten Probleme verringert werden.

Der in dieser Arbeit vorgestellte Ansatz zur näherungsweisen Orientierung der Aufnahmen ergibt insbesondere für den Standpunkt 2 (siehe Abbildung 4.12(b)) ein sehr zufriedenstellendes Ergebnis. Dies wurde unter der Voraussetzung erreicht, dass bereits während der Aufnahme schnelle Bewegungen der Kamera möglichst vermieden wurden. Im Abschnitt 4.3.2 wurde

bereits erwähnt, dass bei einer schnelleren Bewegung eine Verbesserung des Filters der Verknüpfungspunkte notwendig wäre. Da die Aufnahmen mit Hilfe eines Stativs durchgeführt wurden, konnte bei der näherungsweisen Orientierung der gemeinsame Drehpunkt für die Aufnahmen eines Standpunktes berücksichtigt werden. Dies trägt wesentlich zur Stabilität des Ansatzes bei.

Die Fein-Orientierung wurde in dieser Arbeit einerseits mit *TLS-ICP*, andererseits mit *Geomagic* durchgeführt, wobei *TLS-ICP* kein zufriedenstellendes Ergebnis lieferte. Eine Verbesserung des Ergebnisses könnte durch die zusätzliche Verwendung der Amplitudenbilder erreicht werden. Es könnten zum Beispiel in den statischen Bildern, so wie bei der näherungsweisen Orientierung, ebenfalls Verknüpfungspunkte gesucht werden, und als zusätzliche Beobachtungen im *TLS-ICP* berücksichtigt werden. Weiters könnten auch die Amplitudenwerte aller Punkte, zum Beispiel zur unterschiedlichen Gewichtung der Punkte im *TLS-ICP* verwendet werden. Die Verwendung von *Geomagic* führte bei der Fein-Orientierung zu einem akzeptablen Ergebnis.

In Abbildung 4.21(b) ist die Triangulation der orientierten Aufnahmen der ersten drei Bögen am östlichen Ende der Passage zu sehen. Dies ist das Ergebnis der ersten größeren Aufnahme, die mit einer Distanzkamera durchgeführt wurde. Dieses Resultat ist besonders aus dem Blickwinkel der eingehend diskutierten Eigenschaften der Distanzkamera zu betrachten. Zusammenfassend kann das Ergebnis als zufriedenstellend bewertet werden. Die Ursachen des erheblichen Unterschiedes in der Skalierung zwischen Objekt und dem Modell sind in weiterer Folge noch genauer zu untersuchen, da es dadurch beim Eingliedern der Daten in ein größeres Modell zu Spannungen kommen kann. Dies ließe sich aber auch teilweise durch Einpassen des Modells (z.B. Einführen von Skalierungsparametern) in einen Datenbestand lösen.

Generell sind noch weitere Untersuchungen bezüglich der Genauigkeit, des aus Distanzkameraaufnahmen erstellten Modells notwendig, um eine Aussage über dessen Verwendbarkeit in einem übergeordneten Stadtmodell tätigen zu können. Mit dieser Arbeit konnte jedoch gezeigt werden, dass es möglich ist, mit einer Distanzkamera ein Objekt in der Größenordnung der Passage aufzunehmen und daraus ein Modell abzuleiten. Dabei wurden für die näherungsweise Orientierung sowohl die Amplitudenbilder (zur Bestimmung der Verknüpfungspunkte) als auch die Distanzbilder (zur Berechnung der Objektkoordinaten der Verknüpfungspunkte) verwendet. Für die Fein-Orientierung wurden die aus den Distanzbildern abgeleiteten Punktwolken herangezogen. Weiterführend wäre es sinnvoll, bei der Fein-Orientierung die Amplitudenbeobachtungen in Kombination mit den Distanzbeobachtungen zu verwenden.

Anhang A

Anhang

A.1 Differentiale der Kollinearitätsgleichung

$$
\begin{aligned}
a_{11} &= \frac{\partial x}{\partial X_0} = -\frac{c}{N^2}\left(r_{13}Z_x - r_{11}N\right) \\
a_{12} &= \frac{\partial x}{\partial Y_0} = -\frac{c}{N^2}\left(r_{23}Z_x - r_{21}N\right) \\
a_{13} &= \frac{\partial x}{\partial Z_0} = -\frac{c}{N^2}\left(r_{33}Z_x - r_{31}N\right)
\end{aligned}
$$

$$
\begin{aligned}
a_{14} = \frac{\partial x}{\partial \alpha} &= \frac{c}{N}\left\{\left(X - X_0\right) r_{21} - \left(Y - Y_0\right) r_{11} + \frac{Z_x}{N}\Big[-\left(X - X_0\right) r_{23} + \left(Y - Y_0\right) r_{13}\Big]\right\} \\
a_{15} = \frac{\partial x}{\partial \nu} &= -\frac{c}{N}\Big\{\sin\kappa\Big[\left(X - X_0\right) r_{13} + \left(Y - Y_0\right) r_{23} + \left(Z - Z_0\right) r_{33}\Big] - \\
&\quad \frac{Z_x}{N}\Big[\left(X - X_0\right)\sin\alpha\ \cos\nu - \left(Y - Y_0\right)\cos\alpha\ \cos\nu - \left(Z - Z_0\right)\sin\nu\Big]\Big\} \\
a_{16} &= \frac{\partial x}{\partial \kappa} = -\frac{c}{N} Z_y
\end{aligned}
\tag{A.1}
$$

$$a_{21} = \frac{\partial y}{\partial X_0} = -\frac{c}{N^2}\left(r_{13}Z_y - r_{12}N\right)$$
$$a_{22} = \frac{\partial y}{\partial Y_0} = -\frac{c}{N^2}\left(r_{23}Z_y - r_{22}N\right)$$
$$a_{23} = \frac{\partial y}{\partial Z_0} = -\frac{c}{N^2}\left(r_{33}Z_y - r_{32}N\right)$$

$$a_{24} = \frac{\partial y}{\partial \alpha} = \frac{c}{N}\left\{\left(X - X_0\right)r_{22} - \left(Y - Y_0\right)r_{12} + \frac{Z_y}{N}\Big[-\left(X - X_0\right)r_{23} + \left(Y - Y_0\right)r_{13}\Big]\right\}$$
$$a_{25} = \frac{\partial y}{\partial \nu} = -\frac{c}{N}\Big\{\cos\kappa\Big[\left(X - X_0\right)r_{13} + \left(Y - Y_0\right)r_{23} + \left(Z - Z_0\right)r_{33}\Big] -$$
$$\frac{Z_y}{N}\Big[\left(X - X_0\right)\sin\alpha\ \cos\nu - \left(Y - Y_0\right)\cos\alpha\ \cos\nu - \left(Z - Z_0\right)\sin\nu\Big]\Big\}$$
$$a_{26} = \frac{\partial y}{\partial \kappa} = \frac{c}{N}Z_x$$

Dabei wurden für die Gleichungen 3.11 für die Zähler die Abkürzungen Z_x und Z_y sowie für den Nenner die Abkürzung N eingeführt.

A.2 Differentiale der Ähnlichkeitstransformation

$$b_{11} = -r_{11} \qquad b_{21} = -r_{12} \qquad b_{31} = -r_{13}$$
$$b_{12} = -r_{21} \qquad b_{22} = -r_{22} \qquad b_{32} = -r_{23}$$
$$b_{13} = -r_{31} \qquad b_{23} = -r_{32} \qquad b_{33} = -r_{33}$$

$$b_{14} = -r_{21}\left(X_{MP} - X_0\right) + r_{11}\left(Y_{MP} - Y_0\right)$$
$$b_{15} = \sin\alpha\sin\nu\sin\kappa\left(X_{MP} - X_0\right) - \cos\alpha\sin\nu\sin\kappa\left(Y_{MP} - Y_0\right) + \cos\nu\sin\kappa\left(Z_{MP} - Z_0\right)$$
$$b_{16} = r_{12}\left(X_{MP} - X_0\right) + r_{22}\left(Y_{MP} - Y_0\right) + r_{32}\left(Z_{MP} - Z_0\right) \tag{A.2}$$

$$b_{24} = -r_{22}\left(X_{MP} - X_0\right) + r_{12}\left(Y_{MP} - Y_0\right)$$
$$b_{25} = \sin\alpha\sin\nu\cos\kappa\left(X_{MP} - X_0\right) - \cos\alpha\sin\nu\cos\kappa\left(Y_{MP} - Y_0\right) + \cos\nu\cos\kappa\left(Z_{MP} - Z_0\right)$$
$$b_{26} = -r_{11}\left(X_{MP} - X_0\right) - r_{21}\left(Y_{MP} - Y_0\right) - r_{31}\left(Z_{MP} - Z_0\right)$$

$$b_{34} = -r_{23}\left(X_{MP} - X_0\right) + r_{13}\left(Y_{MP} - Y_0\right)$$
$$b_{35} = \sin\alpha\cos\nu\left(X_{MP} - X_0\right) - \cos\alpha\cos\nu\left(Y_{MP} - Y_0\right) + \sin\nu\left(Z_{MP} - Z_0\right)$$
$$b_{36} = 0$$

Abbildungsverzeichnis

2.1 Amplitudenbild und Distanzbild 12
2.2 Grundprinzip von allen RIM-Typen [Kahlmann u. Ingensand, 2007b] 13
2.3 Prinzipien der Distanzmessung [Kahlmann et al., 2006] 14
2.4 Prinzip der Bestimmung der Phasenlage bei PMD-Kameras [Kahlmann et al., 2006] 15
2.5 Schematische Darstellung eines PMD-Pixelelements [Ringbeck u. Hagebeuker, 2007a] 16
2.6 PMD-Kameras. [PMDTechnologies GmbH; MESA Imaging AG; Canesta, Inc.] 17
2.7 Prinzip MDSI [Karel, 2008] 19
2.8 SR-3000 [Mechat u. Büttgen, 2007] 21
2.9 Linearitätsfehler des SR-3000 [Rapp, 2007] 22
2.10 Quantisierungsfehler [Frank et al., 2009] 23
2.11 Empirische Varianz der Distanzbeobachtungen [Karel, 2009] 25

3.1 Ablauf der Auswertung 29
3.2 Ablauf der Datenvorprozessierung 30
3.3 Korrelation mit vorangehendem Amplitudenbild 32
3.4 Korrektur des distanzabhängigen Amplitudenabfalls 33
3.5 Ablauf der näherungsweisen Orientierung 36
3.6 Feature-Deskriptor [Lowe, 2004] 40
3.7 Zusammenhang Bildkoordinaten und Objektkoordinaten 42
3.8 Geometrische Grundformen und deren Freiheitsgrade [Gelfand et al., 2003] . . 49
3.9 Bestimmung der Korrespondenzen [Klauser, 2009] 51

4.1 Lage des Palais Ferstel [Microsoft Corporation, 2009] 58
4.2 Historisches Bild des Palais Ferstel. Quelle: [Historisches Bild] 59

4.3 Photo mit den Komponenten 61
4.4 Grundriss Passage 62
4.5 Überdeckung bei den Aufnahmen 64
4.6 Korrektur des radialen Lichtabfalls 66
4.7 Punktkorrespondenzen zwischen zwei Amplitudenbildern 69
4.8 Mittlere Anzahl der Punktkorrespondenzen 70
4.9 Punktkorrespondenzen - Positionskriterium 70
4.10 Gewichtung der Verknüpfungspunkte 73
4.11 Häufigkeit der Wiederverwendung von Verknüpfungspunkten 74
4.12 Ergebnis der näherungsweisen Orientierung 75
4.13 Filterung der Punktwolken 78
4.14 *TLS-ICP* bei unterschiedlicher Gewichtung der Exzenterkoordinaten 80
4.15 Ergebnis *TLS-ICP* für die 3 Standpunkte 81
4.16 Ablauf der Fein-Orientierung in *Geomagic* 83
4.17 Ergebnis der Fein-Orientierung mit *Geomagic* 85
4.18 Details Orientierung mit *Geomagic* 87
4.19 Detail Orientierung mit *Geomagic* 88
4.20 Detail Triangulation 89
4.21 Trianguliertes Modell 91
4.22 Abstand Punktwolke zur Modelloberfläche 92
4.23 Äußere Genauigkeit 93

Literaturverzeichnis

[Acharya et al. 2008] ACHARYA, S. ; TRACEY, C. ; RAFII, A.: System design of time-of-flight range camera for car park assist and backup application. In: *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops CVPRW '08*, 2008, S. 1–6

[Bay et al. 2008] BAY, H. ; ESS, A. ; TUYTELAARS, T. ; VAN GOOL, L.: Speeded-Up Robust Features (SURF). In: *Computer Vision and Image Understanding* 110 (2008), S. 346–359

[Bay et al. 2006] BAY, H. ; TUYTELAARS, T. ; VAN GOOL, L.: SURF: Speeded Up Robust Features. In: *Computer Vision - ECCV 2006, Part I* 3951 (2006), S. 404–417

[Besl u. McKay 1992] BESL, P. ; MCKAY, N.: A method for Registration of 3D Shapes. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (1992), Nr. 2, S. 239–256

[Burgenkunde 2009] BURGENKUNDE: *Homepage Burgenkunde*. `http://burgenkunde.at/`. Version: 2009. – [Datum des Zugriffs: 01.12.2009]

[Canesta, Inc.] CANESTA, INC.: *DP200*. `http://www.roeder-johnson.com/RJDocs/CADP200Camera.jpg`. – [Datum des Zugriffs: 18.10.2009]

[Chen u. Medioni 1991] CHEN, Y. ; MEDIONI, G.: Object modeling by registration of mulitple range images. In: *Proc. IEEE International Conference on Robotics and Automation* Bd. 3, 1991, S. 2724–2729

[Dorninger u. Pfeifer 2008] DORNINGER, P. ; PFEIFER, N.: A comprehensive automated 3D approach for building extraction, reconstruction and regularization from airborne laser scanning point clouds. In: *Sensors* 8 (2008), Nr. 11, S. 7323–7343

[Dorrington et al. 2007] DORRINGTON, A. A. ; CREE, M. J. ; PAYNE, A. D. ; CONROY, R. M. ; CARNEGIE, D. A.: Achieving sub-millimetre precision with a solid-state full-field heterodyning range imaging camera. In: *Measurement Science and Technology* 18 (2007), S. 2809–2816

[Forkert et al. 2005] *Kapitel* Der Einsatz von Fahrzeug gestützten 3D Laserscanning für kommunale Anwendungen. In: FORKERT, G. ; HARING, A. ; SÜKAR, G.: *Angewandte Geoinformatik 2005 - Beiträge zum 17. AGIT-Symposium Salzburg*. Wichmann Verlag, 2005, S. 171–175

[Frank et al. 2009] FRANK, M. ; PLAUE, M. ; RAPP, H. ; KÖTHE, U. ; JÄHNE, B. ; HAMPRECHT, F.: Theoretical and Experimental Error Analysis of Continuous-Wave Time-Of-Flight Range Cameras. In: *Optical Engineering* 48 (2009), Nr. 1, S. 013602

[Gassmann u. Franke 2005] GASSMANN, F. ; FRANKE, K.-H.: *Verfahren zur Triangulation ungeordneter 3D-Punktmengen*. 3D-NordOst 2005, 8. Anwendungsbezogener Workshop zur Erfassung, Modellierung, Verarbeitung und Auswertung von 3D-Daten. `www.zbs-ilmenau.de/pdf/Ga_Fr_Triangulation_3dNO.pdf`. Version: 2005. – [Datum des Zugriffs: 20.04.2009]

[Gelfand et al. 2003] GELFAND, N. ; IKEMOTO, L. ; RUSINKIEWICZ, S. ; LEVOY, M.: Geometrically stable sampling for the ICP algorithm. In: *Proc. 4th International Conference on 3-D Digital Imaging and Modeling 3DIM 2003*, 2003, S. 260–267

[Gröger et al. 2005] GRÖGER, G. ; BENNER, J. ; DÖRSCHLAG, D. ; DREES, R. ; GRUBER, U. ; LEINEMANN, K. ; LÖWNER, M. O.: Das interoperable 3D-Stadtmodell der SIG 3D. In: *Zeitschrift für Geodäsie, Geoinformation und Landmanagement (zfv)* 6 (2005), S. 343–353

[Haring 2007] HARING, A.: *Die Orientierung von Laserscanner- und Bilddaten bei der fahrzeuggestützten Objekterfassung*, Technische Universität Wien, Diss., 2007

[Haring 2008] HARING, A.: *Orientierung von terrestrischen Laserscanner-Daten*. Vorlesung Photogrammetrie und Laserscanning, TU-Wien. `http://www.ipf.tuwien.ac.at/np/Lehre/download/VO_TLSori.pdf`. Version: 2008. – [Datum des Zugriffs: 03.10.2009]

[Historisches Bild] HISTORISCHES BILD: *Das Bank- und Börsengebäude in der Herrengasse (Palais Ferstel)*. `http://upload.wikimedia.org/wikipedia/commons/1/11/Rudolf_von_Alt_Palais_Ferstel.jpg`. – [Datum des Zugriffs: 01.12.2009]

[Jansa u. Stanek 2003] JANSA, J. ; STANEK, H.: Ableitung von Stadtmodellen aus Laser-Scanner- Daten, Grundrissplänen und photographischen Aufnahmen. In: *Österreichische Zeitschrift für Vermessung und Geoinformation (VGI)* 91 (2003), S. 262–270

[Kada 2007] KADA, M.: *Zur maßstabsabhängigen Erzeugung von 3D-Stadtmodellen*, Universität Stuttgart, Diss., 2007

[Kahlmann u. Ingensand 2007a] KAHLMANN, T. ; INGENSAND, H.: Increased accuracy of 3D range imaging camera by means of calibration. In: *Proc. 8th Conference on Optical 3D Measurements*, 2007, S. 101–108

[Kahlmann u. Ingensand 2007b] KAHLMANN, T. ; INGENSAND, H.: Range Imaging Metrologie: Einführung, Untersuchungen und Weiterentwicklungen. In: *Allgemeine Vermessungsnachrichten (AVN)* 11-12 (2007), S. 384–388

[Kahlmann et al. 2006] KAHLMANN, T. ; REMONDINO, F. ; INGENSAND, H.: Calibration for increased accuracy of the range imaging camera Swissranger. In: *Proc. ISPRS Commission V Symposium 'Image Engineering and Vision Metrology'* Bd. XXXVI(5). Dresden, Germany, 2006, S. 136–141

[Karel 2008] KAREL, W.: Range Imaging: 3D-Punktwolken in Echtzeit. In: *Österreichische Zeitschrift für Vermessung und Geoinformation* 96 (2008), S. 12–26

[Karel 2009] KAREL, W.: *Photogrammetrie + Laserscanning = Range Imaging*. Vortrag: 10. Österreichischer Geodätentag, Schladming; 06.10. - 08.10, 2009

[Karel u. Pfeifer 2009] KAREL, W. ; PFEIFER, N.: Range camera calibration based on image sequences and dense comprehensive error statistics. In: *Three Dimensional Imaging Metrology* Bd. 7239, SPIE, 2009 (proc. SPIE)

[Klauser 2009] KLAUSER, A.: *ICP Algorithmus*. Vortrag Seminar der Geowissenschaften, TU - Wien, 2009

[Koppers 2002] KOPPERS, L.: *Generierung von Objekten für 3D-Stadtmodelle*, Universität der Bundeswehr München, Diss., 2002

[Kraus 1996] KRAUS, K.: *Photogrammetrie, Band 2, Verfeinerte Methoden und Anwendungen*. 3. Dümmler, 1996

[Kraus 2000] KRAUS, K.: *Photogrammetrie, Band 3, Topographische Informationssysteme*. Dümmler, 2000

[Kraus 2004] KRAUS, K.: *Photogrammetrie, Band 1, Geometrische Informationen aus Potographien und Laserscanneraufnahmen*. 7. de Gruyter Lehrbuch, 2004

[Kristensen et al. 2001] KRISTENSEN, S. ; ESTABLE, S. ; KOSSOW, M. ; BRÖSEL, R.: Bin-picking with a solid state range camera. In: *Robotics and Autonomous Systems* 35 (2001), S. 143–151

[Lange 2000] LANGE, R.: *3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology*, University of Siegen, Diss., 2000

[Lindeberg 1994] LINDEBERG, T.: Scale-space theory: A basic tool for analysing structures at different scales. In: *Journal of Applied Statistics* 21 (1994), S. 225–270

[Lowe 1999] LOWE, D. G.: Object Recognition from Local Scale-Invariant Features, 1999, S. 1150–1157

[Lowe 2004] LOWE, D. G.: Distinctive Image Features from Scale-Invariant Keypoints. In: *International Journal of Computer Vision* 60 (2004), Nr. 2, S. 91–110

[Mechat u. Büttgen 2007] MECHAT, M.-A. E. ; BÜTTGEN, B.: Realization of multi-3D-TOF camera environments based on codedbinary sequences modulation. In: *Proc. Conference on Optical 3-D Measurement Techniques*. Zurich, Switzerland, 2007, S. 26–33

[MESA Imaging AG] MESA IMAGING AG: *SwissRanger SR4000*. `http://www.mesa-imaging.ch/`. – [Datum des Zugriffs: 18.01.2010]

[MeshLab 2009] MESHLAB: *Visual Compiting Lab - ISTI - CNR*. `http://meshlab.sourceforge.net/`. Version: 2009. – [Datum des Zugriffs: 04.05.2009]

[Microsoft Corporation 2009] MICROSOFT CORPORATION: *Bing - Maps*. `http://www.bing.com/maps/`. Version: 2009. – [Datum des Zugriffs: 14.12.2010]

[Mure-Dubois u. Hügli 2007] MURE-DUBOIS, J. ; HÜGLI, H.: Real-time scattering compensation for time-of-flight camera. In: *Proc. ICVS Workshop on Camera Calibration Methods for Computer Vision Systems - CCMVS2007*, 2007

[Neugebauer 1997] NEUGEBAUER, P.: Geometrical Cloning of 3D Objects via Simultaneous Registration of Multiple Range Images. In: *Proc.International Conference on Shape Modeling and Applications*, 1997, S. 130–139

[Nischwitz et al. 2007] *Kapitel* Scale Space Filtering. In: NISCHWITZ, A. ; FISCHER, M. ; HABERÄCKER, P.: *Computergrafik und Bildverarbeitung: Alles für Studium und Praxis*. Vieweg+Teubner Verlag, 2007, S. 663–681

[Nothegger u. Dorninger 2007] NOTHEGGER, C. ; DORNINGER, P.: Automated Modeling of Surface Detail from Point Clouds of Historical Objects. In: *21st CIPA Symposium, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXVI-5/C53*. Athens, Greece, 2007, S. 538–543

[Oprisescu et al. 2007] OPRISESCU, S. ; FALIE, D. ; CIUC, M. ; BUZULOIU, V.: Measurements with ToE Cameras and Their Necessary Corrections. In: *Proc. International Symposium on Signals, Circuits and Systems ISSCS 2007* Bd. 1, 2007, S. 1–4

[Palais Events 2009] PALAIS EVENTS: *Homepage Palais Events*. `http://www.palaisevents.at/`. Version: 2009. – [Datum des Zugriffs: 01.12.2009]

[PMDTechnologies GmbH] PMDTECHNOLOGIES GMBH: *PMD[vision] CamCube*. `http://upload.wikimedia.org/wikipedia/de/9/98/PMDvision_CamCube.jpg`. – [Datum des Zugriffs: 18.10.2009]

[PMDTechnologies GmbH 2009] PMDTECHNOLOGIES GMBH: *Homepage PMDTechnologies*. `http://www.pmdtec.com/`. Version: 2009. – [Datum des Zugriffs: 06.06.2009]

[Rapp 2007] RAPP, H.: *Experimental and Theoretical Investigation of Correlating TOF-Camera Systems*, Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Diplomarbeit, 2007

[Ringbeck u. Hagebeuker 2007a] RINGBECK, T. ; HAGEBEUKER, B.: A 3D time of flight camera for object detection. In: *Proc. 8th Conference on Optical 3-D Measurement Techniques*. Zürich, 2007

[Ringbeck u. Hagebeuker 2007b] RINGBECK, T. ; HAGEBEUKER, B.: Dreidimensionale Objekterfassung in Echtzeit - PMD Kameras erfassen pro Pixel Distanz und Helligkeit mit Videoframerate. In: *Allgemeine Vermessungsnachrichten* 7 (2007), S. 263–270

[Rusinkiewicz u. Levoy 2001] RUSINKIEWICZ, S. ; LEVOY, M.: Efficient variants of the ICP algorithm. In: *Proc. 3rd International Conference on 3-D Digital Imaging and Modeling*, 2001, S. 145–152

[Stadtentwicklung Wien] STADTENTWICKLUNG WIEN: *Sichtachsen*. `http://www.wien.gv.at/stadtentwicklung/grundlagen/hochhauskonzept/sichtachsen.html`. – [Datum des Zugriffs: 05.02.2010]

[Strandmark 2008] STRANDMARK, P.: *Homepage SURFmex*. `http://www.maths.lth.se/matematiklth/personal/petter/surfmex.php`. Version: 2008

[The MathWorks Inc. 2007] THE MATHWORKS INC.: *Programmhilfe MATLAB 7.4 - R2007a*. 2007

Lebenslauf

Persönliche Daten

Name:	Stefan Niedermayr
geboren am:	30.07.1982 in Linz

Schulbildung

1988-1989:	Vorschule in Ottensheim
1989-1993:	Volksschule in Ottensheim
1993-1997:	Hauptschule in Ottensheim
1997-2002:	Höhere Technische Bundeslehranstalt für Maschinenbau - Automatisierungstechnik in Neufelden
Juni 2002:	Reifeprüfung
2003:	Präsenzdienst
10/2003:	Inskription „Vermessung und Geoinformation“ an der TU Wien
01/2006:	1. Diplomprüfung