

Optische Methoden zur Form- und Positionserkennung von Körpern in Werkzeugmaschinen

Diplomarbeit

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA
PHYSIKALISCH-ASTRONOMISCHE FAKULTÄT



seit 1558

Jena, den 29. Oktober 2008

Eingereicht von: Mirko Schmidt
Geboren: am 26.02.1983 in Berlin

Der Autor steht Ihnen für Fragen und Anregungen gerne unter der email-Adresse SCHMIDT3D@GMX.DE zur Verfügung.

1. Gutachter:

Prof. Dr.-Ing habil. Peter Adam

2. Gutachter:

PD Dr. Wolfgang Fried

Tag der Verleihung des Diploms:

.....

Optische Methoden zur Form- und Positionserkennung von Körpern in Werkzeugmaschinen

Diese Arbeit untersucht optische Methoden zur Erzeugung von Tiefenbildern. Aufgrund des speziellen Einsatzgebietes im Arbeitsraum von Werkzeugmaschinen, wurde ausgehend vom Prinzip der stereoskopischen Tiefendatengewinnung ein Verfahren entwickelt, welches die schnelle Erzeugung von Tiefenbildern bei minimalem Platzbedarf des Messkopfes ermöglicht. Durch die Wahl einer sehr kleinen stereoskopischen Basis lässt sich das Auftreten des Korrespondenzproblems verhindern, und dadurch die Tiefeninformation in einer zu der Anzahl der Pixel linearen Zeit rekonstruieren. Das dem Verfahren zugrunde liegende mathematische Modell wird ausführlich diskutiert, mit dem Ergebnis, Anforderungen an die verwendete Kamera, die Lichtquelle und den Messaufbau formulieren zu können. In Experimenten wurden anhand der Vermessung von Testobjekten diese Vorhersagen bestätigt. Die Arbeit schließt mit einer ausführlichen Diskussion der Resultate.

Optical methods for acquiring the figure and position of solids in the working space of tooling machines

This thesis investigates optical methods for depth image acquisition. Due to the special application field in the working space of tooling machines, a method for acquiring depth images with very low spatial requirements was developed, based on the principle of stereoscopic depth image acquisition. By choosing a very small stereoscopic basis, the occurrence of the correspondence problem can be avoided, and therefore the depth information can be computed in to the number of pixels linear time. The underlying mathematical model is discussed in detail, enabling the formulation of requirements on the camera, the light source and the mechanical setup. Experimental measurements on test-objects were used to approve these predictions. The thesis closes with a detailed discussion of the results.

Inhaltsverzeichnis

1	Vorbereitung	3
1.1	Einleitung	3
1.2	Aufgabenstellung	3
1.3	Erzeugung der Raumdaten	4
1.4	Aufnahmemethoden & Tiefenbilder	4
1.5	Kamerabasierte Bilddatengewinnung	6
1.5.1	Die zentralprojektive Abbildung	6
1.5.2	Das Modell der dünnen Linse	6
1.5.3	Koordinatensysteme	7
1.5.4	Kamerakalibrierung	9
2	Stereoskopie bei minimaler Basis	11
2.1	Einleitung: Probleme Stereoskopischer Verfahren	11
2.2	Die Idee der Stereoskopie bei minimaler Basis	12
2.3	Abschätzung der Verschiebung s	13
2.4	Verwendung einer Marke	16
3	Möglichkeiten und Grenzen	17
3.1	Der Fehler der Intensitätsmessung	17
3.1.1	Sensorrauschen: ΔI	17
3.1.2	Das Signal-Rausch-Verhältnis	18
3.2	Kameraparameter bei idealer Optik und Umgebung	19
3.2.1	Der Fehler der sichtbaren Verschiebung Δs	19
3.2.2	Abhängigkeit des Fehlers der Tiefenbestimmung Δz von Δs	20
3.2.3	Wahl von x_a und s_M	20
3.2.4	$B_{d,min}$ und Tiefenschärfe	21
3.3	Umgebungseinflüsse	24
3.3.1	Schwankungen der Helligkeit der Lichtquelle	24
3.3.2	Mechanische Stabilität des Systems	25
3.4	Einflüsse einer nicht-idealen Optik: Abbildungsfehler des Objektivs	30
4	Messungen	35
4.1	Das Bildaufnahmesystem	35
4.2	Voruntersuchungen	36
4.2.1	Bestimmung des maximalen Signal-Rausch-Verhältnisses	36
4.2.2	Ermittlung der Parameter zur Verzeichnungskorrektur	37

4.3	Aufbau zur Aufnahme geeigneter Grauwertbilder	37
4.3.1	Karten	38
4.4	Datenauswertung	38
4.5	Analyse der Genauigkeit der ermittelten Tiefenwerte	41
4.5.1	Untersuchung der Unabhängigkeit der Tiefenwerte von der Position der abgebildeten Objekte im Messraum	41
4.5.2	Bestimmung der Tiefenauflösung Δz abhängig von z	46
4.5.3	Untersuchung der Abhängigkeit des Fehlers der Tiefenbestimmung Δz vom Signal-Rausch-Verhältnis der Kamera SNR_I	48
4.6	Geschwindigkeit	50
5	Zusammenfassung und Vergleich	53

1 Vorbereitung

1.1 Einleitung

Werkzeugmaschinen sind für moderne Produktionsanlagen besonders als numerisch gesteuerte CNC-Maschinen zu einem unentbehrlichen Hilfsmittel geworden. Die Möglichkeit, universelle Werkzeuge und Spannmittel für die Bearbeitung verschiedenster Werkstücke einsetzen zu können, bietet eine enorme Flexibilität, die besonders bei für die Fertigung kleiner Serien zur Notwendigkeit geworden ist.

Bei der Umrüstung der Maschinen auf neue Bearbeitungsfälle treten jedoch gelegentlich Fehler bei der Eingabe der Geometrie der im Arbeitsraum befindlichen Objekte - insbesondere Spannmitteln und Werkstücken - in die Maschinensteuerung auf. Diese Fehler können zu Kollisionen führen, die unter Umständen kostenintensive Maschinenschäden zur Folge haben, da z. B. Führungen, Kugelgewindetriebe, Spindellagerungen, Messsysteme und Antriebe Defekte erleiden können.

Daher ist es der Wunsch von Hersteller und Anwender der Werkzeugmaschine, ein System zur Prüfung der Konsistenz von dem in der Maschine hinterlegten Modell des Arbeitsraums mit der Realität zur Verfügung zu haben, welches rechtzeitig vor drohenden Kollisionen warnen und gegebenenfalls automatisch entsprechende Gegenmaßnahmen einleiten kann.

1.2 Aufgabenstellung

Ziel dieser Arbeit ist das Finden und Demonstrieren einer geeigneten Methode der optische Erfassung von Objekten im Arbeitsraum von Werkzeugmaschinen. Die Messungen sollen am Ende des Einrichtprozesses ausgeführt werden und zur Ermittlung von Raumdaten dienen, die zur Kollisionsvermeidung eingesetzt werden können.

Die zu vermessenden Objekte weisen aufgrund der speziellen Umgebung bestimmte Besonderheiten auf. So werden hauptsächlich metallische Objekte untersucht, deren Oberflächen stark reflektierend sein können. Starke Periodizitäten sind denkbar, die sich bei der Verwendung von stereoskopischen Verfahren negativ auswirken könnten. Rückstände von Kühl- und Schmiermitteln sowie Abrieb und Späne könnten die Erscheinung (Form und Farbe) der Oberflächen verändert haben.

Da die Messung innerhalb eines abgeschlossenen Raumes und außerhalb der Bearbeitungsphasen erfolgt, ist zu erwarten, dass die Objekte still stehen und relativ konstante Beleuchtungsverhältnisse herrschen. Die Dauer der Datenaufnahme und Umrechnung in 3D-Daten soll wenige Sekunden betragen. Die gewünschte räumliche Auflösung liegt bei

1mm in jeder Dimension. Die zu erwartenden Objekte innerhalb des Maschinenarbeitsraums sind definierte geometrische Strukturen. Es lässt sich jedoch nicht voraussetzen, dass jedes Objekt deutlich „sichtbare“ geometrische Kanten aufweist, da z. B. gekrümmte Oberflächen und Kavitäten denkbar sind.

1.3 Erzeugung der Raumdaten

Die Erzeugung von dreidimensionalen Raumdaten des zu vermessenden Arbeitsraums lässt sich in mehrere Phasen unterteilen. Zunächst erfasst ein Sensor auf optischem Wege einen Ausschnitt der Oberfläche des Arbeitsraums. Je nach Methode lässt sich in einem oder mehreren Schritten die Lage der sichtbaren Bildpunkte im Koordinatensystem des Sensors ermitteln. Diese Sensordaten werden mit Hilfe von Informationen über die Position und Lage des Sensors im Raum zu einer Punktwolke verrechnet, die die Gesamtheit aller (sichtbaren) Punkte im Raum darstellt. Die Beschreibung dieser Punktwolke kann nun durch geeignete Algorithmen komprimiert werden, indem bestimmte Regionen im Raum zum Beispiel durch geometrische Objekte beschrieben werden. Diese vereinfachten Raumdaten eignen sich zur weiteren Verarbeitung durch die Routinen der Kollisionsvermeidung.

Unabhängig von der verwendeten Methode wird der Sensor im Allgemeinen nicht die gesamte Oberfläche des Arbeitsraums auf einmal erfassen können, da es zum Beispiel möglich ist, dass einige Objekte andere verdecken. Es müssen also die Sensordaten von verschiedenen Punkten des Arbeitsraums aus aufgenommen werden. Ein Ansatz wäre, mehrere Sensoren fest in verschiedenen Positionen zu installieren. Dies würde aber eine Verringerung des zur Verfügung stehenden Arbeitsraumes bedeuten und es ließe sich nicht die vollständige Erfassung jeder möglichen Szene garantieren.

Eine andere Möglichkeit besteht darin, mit nur einem Sensor zu arbeiten, der innerhalb des Arbeitsraums bewegt wird und so von verschiedenen Punkten aus Messerte aufnehmen kann. Die Werkzeugmaschinen, auf denen das System eingesetzt werden soll, verfügen über eine Mechanik, welche in der Lage ist beliebige Werkzeuge mit hoher Geschwindigkeit und Präzision innerhalb des Arbeitsraums zu bewegen. Es bietet sich daher an, diese Mechanik für die Bewegung des Sensors zu nutzen.

Die so gewonnenen Punktwolken müssen anschließend zu einer einzigen registriert werden, die ein vollständiges Modell des vermessenen Arbeitsraums darstellt.

1.4 Aufnahmemethoden & Tiefenbilder

Es existiert eine Vielzahl optischer Sensoren zur Positionsbestimmung, die sich je nach angewandter Methode und geplantem Einsatzbereich unterscheiden. Die Messung kann entlang eines Strahls erfolgen (z. B. Laser-Messgeräte bei Geschwindigkeitskontrolle im Straßenverkehr) oder Daten zu Punkten liefern, die in einer bestimmten Ebene des Raums liegen. Da diese Sensoren nur ein- bzw. zweidimensionale Daten liefern, wäre ihre Verwendung für die vollständige Erfassung eines Raums nur durch die Kombination

sehr vieler Einzelmessungen möglich. Der Sensor müsste während der Messwertaufnahme durch den Raum bewegt oder geschwenkt werden („Scannen“).

Einen wesentlichen Geschwindigkeitsvorteil versprechen Verfahren, welche dreidimensionale Daten liefern. Hier ist sofort vollständig die Position aller in einem bestimmten Raumwinkel liegenden Punkte bekannt. Ein solches System würde also Daten ähnlich den Bildern einer Kamera liefern, jedoch neben der Lage in horizontaler und vertikaler Richtung auch Informationen über die Entfernung der Bildpunkte liefern. Diese Bilder werden als Tiefenbilder bezeichnet.

Drei Prinzipien kommen für die Gewinnung von Tiefenbildern hauptsächlich zum Einsatz: Interferometrische Methoden basieren auf der indirekten Messung der Differenz der Phase zwischen einem vom Objekt reflektierten Lichtstrahl und einem Referenzstrahl. Sie liefern schnell und sehr präzise Entfernungsdaten. Die hohe Empfindlichkeit ermöglicht eine Genauigkeit der Entfernungsmessung im Bereich der Wellenlänge λ des verwendeten Lichtes (einige Hundert nm). Das bedeutet aber auch, dass sich störende Einflüsse bereits in dieser Größenordnung stark negativ auf die Messung auswirken. Das ganze System muss also mechanisch so stabil sein, dass Verformungen in dieser Größenordnung weitgehend vermieden werden. Temperaturstabilität und Vibrationsfreiheit sind aber bei Werkzeugmaschinen im Allgemeinen nicht vorauszusetzen. Außerdem müssen die Oberflächen der interferometrisch zu vermessenden Objekte so glatt sein, dass ihre Rauigkeiten unterhalb der Wellenlänge liegen. Diese Bedingung wird nur in Ausnahmefällen erfüllt sein, daher lässt sich die Verwendung interferometrischer Methoden für die gegebene Problemstellung ausschließen.

Die Messung der Zeit, die das Licht benötigt, um von einer Quelle zum Objekt und von diesem auf einen Detektor zu gelangen, ermöglicht ebenfalls eine Entfernungsbestimmung. Solche Flugzeit-basierten Methoden (engl. *time-of-flight* = *TOF* depth sensor) werden seit Jahrzehnten erfolgreich bei Messungen in einer Dimension über große Distanzen ($> 10\text{m}$) eingesetzt (zum Beispiel Geschwindigkeitskontrolle im Straßenverkehr oder Messung des Abstandes zwischen Erde und Mond). Um eine möglichst präzise Messung zu ermöglichen, wird häufig nicht die Laufzeit eines einzelnen Pulses erfasst, sondern die Intensität der Lichtquelle moduliert und die Differenz der Phase der Intensität zwischen gesendetem und detektiertem Licht ausgewertet. Auch Systeme zur direkten Erfassung von Tiefenbildern basierend auf diesem Prinzip wurden vorgestellt (z. B. in [OLK⁺04] oder [GYB04]). Da hier aber die Detektion der Phase in jedem Pixel mit sehr hoher zeitlicher Auflösung notwendig ist, erreichen sie bisher geringe Genauigkeiten: Der absolute Fehler der Tiefenmessung liegt im Bereich von Millimetern bis Zentimetern bei Sensorgrößen von einigen Hundert bis Zehntausend Pixeln.

Die Abstandsbestimmung bei triangulationsbasierten Methoden erfolgt durch die Analyse des Dreiecks, welches durch die Sehstrahlen und die Verbindungslinie (Basis) zweier optischer Systeme erzeugt wird. Es werden aktive und passive Verfahren unterschieden. Aktive Verfahren verwenden eine Kamera und eine Beleuchtungseinheit, welche ein definiertes Muster auf das zu vermessende Objekt projiziert. Um eine eindeutige Beziehung zwischen den Sehstrahlen beider optischen Systeme herstellen zu können, ist für die meisten Anwendungen die Verwendung nur eines Musters nicht ausreichend.

Bei passiven Verfahren werden zwei Kamerabilder eingesetzt. Hier ist die Zuordnung

zusammengehöriger Sehstrahlen nur mit Hilfe rechenintensiver Algorithmen möglich, die die Bilder nach korrespondierenden Regionen absuchen. Das Finden der zusammengehörigen Regionen wird als Lösung des Korrespondenzproblems bezeichnet.

Zur Erhöhung der erreichbaren räumlichen Auflösung wurden die triangulationsbasierten Verfahren teilweise so erweitert, dass mit mehr als zwei optischen Systemen gearbeitet werden kann. Es existieren passive Verfahren mit mehreren Kameras, aktive Verfahren mit mehreren Projektoren und eine Vielzahl von Mischformen.

Es gibt eine Reihe weiterer Methoden zur optischen Gewinnung von Tiefenbildern, die bestimmte Kenntnisse über die zu vermessenden Objekte oder spezielle Optiken voraussetzen. Zu nennen wären hier zum Beispiel Verfahren, die im Englischen als *depth-from-motion*, *depth-from-shape* und *depth-from-focus* bezeichnet werden. Bei *depth-from-shape*-Verfahren wird mithilfe von Modellen der zu erkennenden Objekte deren Erscheinung in der Abbildung (abhängig von ihrer Position im Raum) vorhergesagt und mit den Bilddaten verglichen.

Bei dem Ansatz des *depth-from-motion*-Verfahrens werden Videosequenzen analysiert und aus der Bewegung korrespondierender Bildausschnitte Schlüsse auf deren Position gezogen. Eine Anwendung dieses Verfahrens sind etwa 3D-Scanner, bei welchen das zu vermessende Objekt auf einem Teller vor einer Kamera rotiert. Hierbei wird vorausgesetzt, dass das zu vermessende Objekt starr ist.

Depth-from-focus basiert auf entfernungsabhängigen Bildschärfe einer Abbildung und wertet mindestens zwei unterschiedliche fokussierte Aufnahmen aus. Hohe Genauigkeiten lassen sich mit diesem Verfahren besonders bei optischen Systemen erreichen, die über eine geringe Tiefenschärfe verfügen (z. B. Mikroskope).

1.5 Kamerabasierte Bilddatengewinnung

1.5.1 Die zentralprojektive Abbildung

Das mathematische Modell der durch Kameras realisierten physikalischen Bildgewinnung ist die *Zentralprojektion*. Sie wird folgendermaßen erhalten: Man fixiert einen Raumpunkt C als Projektionszentrum und eine Projektionsebene E . Als Bild P' eines Raumpunktes P (der ungleich C ist) wird der Schnittpunkt des Projektionsstrahls PC mit der Ebene E definiert. Die dem folgenden Unterabschnitt zugeordnete Zeichnung 1.1 zur Abbildung an der dünnen Linse verwendet dieselben Bezeichnungen und kann zur Veranschaulichung herangezogen werden.

1.5.2 Das Modell der dünnen Linse

Ein einfaches physikalisches Modell der Bildgewinnung ist die in Grafik 1.1 skizzierte Abbildung der dünnen Linse. Alle von einem Punkt P auf die Linse auftreffende Lichtstrahlen werden in der Hauptebene H der Linse so gebrochen, dass sie in einem Punkt P' der Bildebene E zusammenlaufen, welche sich im Abstand b zur Hauptebene befindet. Der Betrag der Strecke b wird als Bildweite bezeichnet. Der Abstand zwischen P

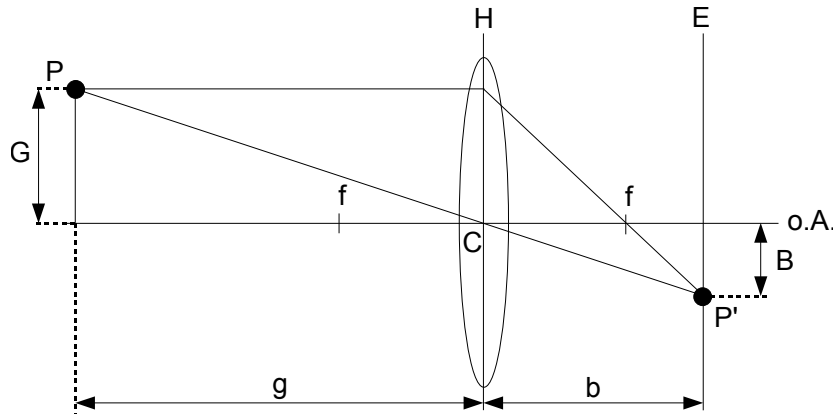


Abbildung 1.1: Abbildung eines Punktes P in P' mit Hilfe einer dünnen Linse

und der Hauptebene H ist die Gegenstandsweite g . Der Abstand zwischen der optischen Achse (in der Grafik als „o.A.“ abgekürzt) und P wird als Gegenstandshöhe G bezeichnet. Als Bildhöhe B bezeichnet man den Abstand zwischen dem Bild des Punktes P' und der optischen Achse. Die Linse besitzt zwei Brennpunkte, welche auf beiden Seiten der Hauptebene H , jeweils im Abstand f zu ihr und auf der optischen Achse liegen. Die optische Achse durchstößt die Hauptebene H senkrecht. Der Durchstoßpunkt soll in Analogie zum Modell der zentralprojektiven Abbildung als C bezeichnet werden.

Aus dem Strahlensatz ergibt sich die Abbildungsgleichung:

$$\frac{B}{G} = \frac{b}{g} \quad (1.1)$$

Eine scharfe Abbildung, also das Zusammenlaufen aller von P ausgehenden Strahlen in P' erfordert, dass b , g und f die Newtonsche Linsengleichung erfüllen:

$$\frac{1}{f} = \frac{1}{b} + \frac{1}{g} \quad (1.2)$$

1.5.3 Koordinatensysteme

Für die Analyse und Beschreibung der vorgestellten Verfahren und Algorithmen ist es zweckmäßig, verschiedene Koordinatensysteme zu definieren. Das Kamerakoordinatensystem (KKS) ist ein rechtshändiges kartesisches Koordinatensystem, welches fest mit der Kamera verbunden ist. Seine Achsen werden mit x , y und z bezeichnet.

Parallel zur x - y -Ebene des Kamerakoordinatensystems verläuft durch $z = -b$ ($b =$ Bildweite der Abbildung) die Projektionsebene E . In ihr liegt das zweidimensionale Bildkoordinatensystem (BKS). Seine Koordinaten werden mit ξ und η bezeichnet. Die ξ -Achse verläuft antiparallel zur x -Achse des KKS und die η -Achse parallel zur y -Achse des KKS (siehe Abbildung 1.2). Die optische Achse durchstößt den Punkt (ξ_0, η_0) des BKS und den Ursprung des KKS. Der Punkt C und der Ursprung des KKS fallen also zusammen. Die optische Achse und die z -Achse des KKS liegen nicht zwingend ineinander,

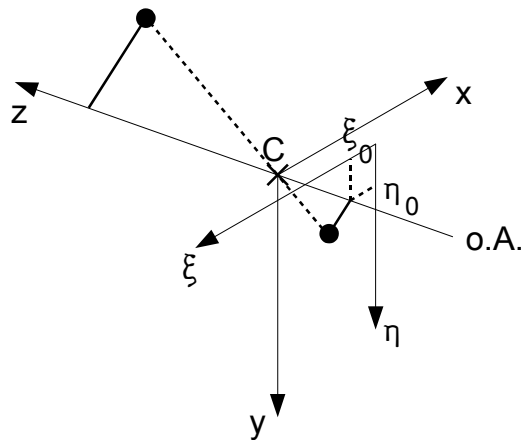


Abbildung 1.2: Kamerakoordinatensystem und Bildkoordinatensystem

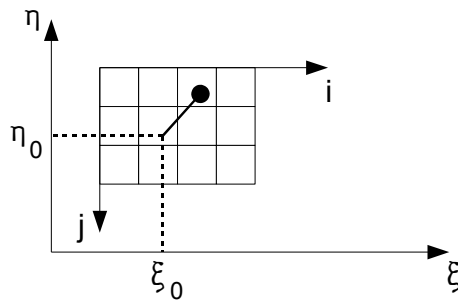


Abbildung 1.3: Bildkoordinatensystem und digitales Koordinatensystem. Die Entfernung des Durchstoßpunktes der optischen Achse durch die Bildebene (ξ_0, η_0) von dem Zentrum des Sensors wurde hier stark übertrieben dargestellt

da Fehler in der optischen Abbildung eine minimale Verkippung beider Geraden gegeneinander bewirken können. Es wird im Folgenden angenommen, dass die sich daraus ergebenden Auswirkungen auf die Entfernungsbestimmung zu vernachlässigen sind.

Innerhalb des Bildkoordinatensystems liegt das digitale Koordinatensystem (DKS). Mit Hilfe der digitalen Koordinaten i und j lassen sich die Positionen einzelner Pixel beschreiben. Die i - und ξ -Achse verlaufen parallel, die j - und η -Achse antiparallel. In Abbildung 1.3 sind beide Koordinatensysteme dargestellt.

Das Kamerakoordinatensystem ist eingebettet in das Maschinenkoordinatensystem (MKS).

Da die verwendeten Abbildungsmodelle Rotationssymmetrie aufweisen, ist für die Beschreibung von Orten in der Bildebene die Verwendung von Polarkoordinaten sinnvoll. Als Ursprung wird der Durchstoßpunkt (ξ_0, η_0) der optischen Achse durch die Bildebene verwendet:

$$\rho = \sqrt{(\xi - \xi_0)^2 + (\eta - \eta_0)^2}$$

$$\varphi = \arctan \frac{(\eta - \eta_0)}{(\xi - \xi_0)}$$

Bei einer (scharfen) Abbildung eines Punktes P in P' gilt also gerade, dass $B = \rho(P')$ ist. Für die spätere Analyse der Effekte von Abbildungsfehlern wird es notwendig sein, ein maximale Bildhöhe ρ_{max} definiert zu haben. Diese ergibt sich aus den Abmessungen des Kamera-Sensors und ist gleich dem maximalen Abstand eines lichtempfindlichen Pixels vom Durchstoßpunkt der optischen Achse durch den Sensor. Außerdem werden die Projektionen von ρ_{max} auf die ξ - und η -Achsen des Bildkoordinatensystems benötigt werden. Sie werden mit $\rho_{\xi_{max}}$ und $\rho_{\eta_{max}}$ bezeichnet¹.

1.5.4 Kamerakalibrierung

Das Modell der Zentralprojektion ist für die Beschreibung realer Kameras nicht ausreichend. Um die Durchführung präziser Messungen zu ermöglichen, ist eine exakte Kenntnis der geometrischen Verhältnisse innerhalb der Kamera notwendig. Auch die durch bei realen Linsensystemen immer vorhandenen Abbildungsfehler verursachten Verzerrungen müssen von dem erweiterten Modell beschrieben werden. Die Parameter dieses erweiterten Modells zur werden in der Photogrammetrie als *Parameter der inneren Orientierung* bezeichnet.

Für die Zuordnung von Bild- zu Raumpunkten muss die Lage der Kamera im Raum bekannt sein. Hierfür werden sechs unabhängige Parameter benötigt, welche als *Parameter der äußeren Orientierung* bezeichnet werden.

Die Problematik der *Kamerakalibrierung*, also die Gewinnung der äußeren und inneren Parameter, ist ein gut untersuchtes Problem, zu dem eine reichhaltige Literatur existiert. Deshalb sollen die hierfür verwendeten Methoden an dieser Stelle nicht weiter vertieft

¹Falls die optische Achse den Sensor der Kamera etwa in dessen Mitte durchstößt entsprechen $\rho_{\xi_{max}}$ und $\rho_{\eta_{max}}$ also ca. der halben Chip-Breite bzw. -Höhe

werden und es wird im Folgenden angenommen, dass die Kalibrierungsparameter der verwendeten Kameras bekannt sind.

Für ein Verfahren, welches die Bewegung einer Kamera innerhalb des Arbeitsraums einer Werkzeugmaschine mit Hilfe der Achsen dieser Maschine ausführt, ist es sicherlich sinnvoll, die Parameter der äußeren Orientierung einmalig durch eine Kalibrierung bei einer bestimmten Kameraposition zu ermitteln, und diese später für alle anderen Kamerapositionen mit Hilfe von Positionsdaten der Maschinensteuerung zu aktualisieren.

2 Stereoskopie bei minimaler Basis

2.1 Einleitung: Probleme Stereoskopischer Verfahren

Im vorangegangenen Kapitel wurden bereits passive triangulationsbasierte Verfahren zur Tiefendatenbestimmung erwähnt. Diese werden auch als stereoskopische Verfahren bezeichnet und sollen nun für den Fall, dass zwei Kameras zur Verfügung stehen, genauer diskutiert werden.

Bei stereoskopischen Verfahren werden die Bilder des zu untersuchenden Objektes von unterschiedlichen Positionen aus aufgenommen. Unter Kenntnis der Lage der Kamerapositionen zueinander und der inneren Kameraparameter lässt sich die Position eines Raumpunktes, der in beiden Bildern identifiziert wurde, bestimmen. Um ein Tiefenbild, also eine Matrix von Abstandsdaten der sichtbaren Objektpunkte zu erhalten, ist es notwendig, korrespondierende Bildregionen zu finden. Das setzt voraus, dass die untersuchten Objekte überhaupt Merkmale aufweisen, z. B. Kanten¹, die sich in beiden Bildern identifizieren lassen. Die Entfernungsmessung homogen gefärbter Flächen lässt sich mit stereoskopischen Verfahren also nur realisieren, indem man diese künstlich mit Merkmalen versieht. Ein Weg hierfür wäre zum Beispiel die Projektion von Mustern auf die zu vermessenden Objekte². Im Folgenden soll vorausgesetzt werden, dass die Oberfläche der zu vermessenden Objekte über günstige Merkmale in geeigneter Dichte verfügt. In Kapitel 3, Abschnitt 3.2.4, wird diese Bedingung genauer spezifiziert. Auf die Besonderheiten, die sich bei der Nutzung eines Muster projizierenden Gerätes ergeben, wird kurz in Kapitel 5 eingegangen.

Die Suche nach korrespondierenden Punkten lässt sich unter gewissen Bedingungen eingrenzen, z. B. ist bekannt, dass Geraden im Raum als Geraden im Bild abgebildet werden. Da dies für alle Aufnahmen gilt, muss ein Objektpunkt, dessen Abbildung auf einer bestimmten Gerade des einen Bildes erscheint, auf einer bestimmten - bei Kenntnis der Kamerapositionen bereits bekannten - Gerade des anderen Bildes abgebildet werden. Es muss also nicht jeweils das gesamte Bild nach einem korrespondierenden Punkt abgesucht werden, sondern nur ein schmaler Streifen. (Stichwort *Epipolarlinien*, *Epipolargeometrie*, z.B. in [VNS95])

Durch die Vorgabe eines erwarteten minimalen und maximalen Abstands der Objektpunkte lässt sich der Suchbereich weiter einschränken: Statt des gesamten Streifens wird nur ein Ausschnitt untersucht, sodass für jedes Pixel eines Bildes mehrere hundert Punkte

¹Kanten sind Diskontinuitäten der räumlichen Grauwertfunktion des Bildes.

²Im Gegensatz zu den aktiven triangulationsbasierten Verfahren können die hier verwendeten Muster jedoch einfacher sein (z. B. statistisch), da sie keine Informationen zur Zuordnung einzelner Sehstrahlen liefern müssen.

des anderen Bildes in Betracht kommen. Die Suche kann weiter beschleunigt werden, indem man iterative Verfahren verwendet und die Bilder erst in einer stark herabgesetzten Auflösung vergleicht, welche dann schrittweise wieder erhöht wird. In jedem Schritt werden die gefundenen korrespondierenden Regionen angepasst und verfeinert. Bisher sind die verwendeten Algorithmen aber noch so rechenintensiv, dass Echtzeit-3D-Erkennung (bei etwa einer halben Mio. Bildpunkte) nur auf spezieller, parallel rechnender Hardware durchgeführt werden kann (ein entsprechendes System wurde zum Beispiel in [TKK⁺06] beschrieben).

Die Lösung dieses Korrespondenzproblems ist also nur mit sehr hohem Rechenaufwand möglich und stellt bei der schnellen Berechnung der Entfernungsdaten das begrenzte Element dar. Falls die zu vermessenden Objekte sehr regelmäßige Oberflächen aufweisen, ist außerdem unter zusätzlichem Aufwand die Vermeidung fehlerhaft erkannter Korrespondenzen sicherzustellen. Der Wunsch, diese Fehlerkennungen grundsätzlich zu vermeiden und die benötigte Rechenzeit drastisch zu minimieren, führte den Autor zu der im folgenden beschriebenen Überlegung und dem darauf basierenden Verfahren zur schnellen Gewinnung von Tiefenbildern.

2.2 Die Idee der Stereoskopie bei minimaler Basis

Rechnerische Vereinfachungen bei der Bestimmung der Entfernungsdaten ergeben sich, wenn die Kameras so ausgerichtet werden, dass ihre Bildebenen in einer gemeinsamen Ebene liegen (Stichwort „rektifizierte Bilder“ bzw. „rectified images“, z. B. in [FP03]). Die optischen Achsen der Aufnahmesysteme sind nun parallel ausgerichtet. Wird der Abstand der optischen Achsen verringert, verkleinert sich auch der nach korrespondierenden Regionen abzusuchende Raum, bis irgendwann der Fall eintritt, dass dieser nur noch ein Pixel umfasst. Die kostenintensive Suche nach korrespondierenden Bildpunkten ist nun nicht mehr notwendig, denn die Information über die Entfernung des Objektes steckt allein in dem Intensitätswert des Pixels (und in denen seiner direkten Nachbarn).

Die zugrundeliegende Idee lässt sich so beschreiben: Bei Verschiebung der Kamera senkrecht zur optischen Achse scheinen sich (aus Sicht der Kamera) nahe liegende Objekte schneller zu bewegen als weiter entfernte. Diese entfernungsabhängige Verschiebung s einzelner Bildregionen zeigt sich bei minimaler Variation der Kameraposition in der Änderung einzelner Intensitätswerte. Wird die Kamera so bewegt, dass Abbildung der Umgebung parallel zu den Pixelreihen oder -Spalten des optischen Sensors verschoben wird, lässt sich die Entfernung einzelner Bildregionen sehr einfach (und schnell) berechnen. Im Folgenden wird angenommen, dass die Kamera in Richtung der x -Achse des Kamerakoordinatensystems um die Strecke x_a verschoben wurde. Da die Verschiebung senkrecht zu den optischen Achsen stattfindet, ist der Abstand der optischen Achsen beider Aufnahmen ebenfalls x_a . Dies ist in Abbildung 2.1 skizziert worden.

Mit Hilfe zweier Grauwertaufnahmen kann s abgeschätzt werden.

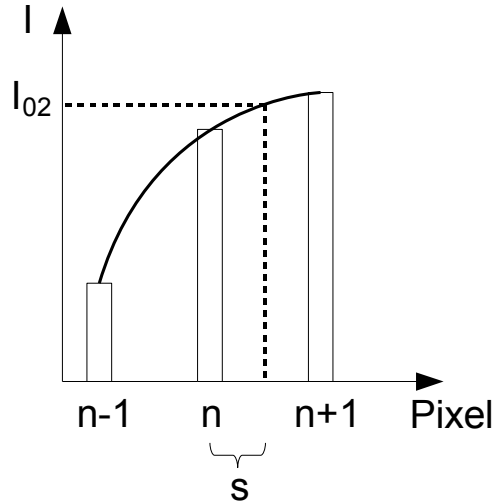


Abbildung 2.2: Zur Bestimmung der Verschiebung s

zu

$$s_{1,2} = -\frac{b}{2a} \pm \sqrt{\frac{b^2}{4a^2} - \frac{c - I_{02}}{a}} \quad (2.4)$$

. Hierbei ist I_{02} die Intensität des betrachteten Pixels in der zweiten Aufnahme. Falls die Wurzel in (2.4) ungleich Null ist, muss eine Lösung von $s_{1,2}$ verworfen werden. Hierzu werden beide Lösungen s_1 und s_2 berechnet und anschließend entschieden: Da nur Werte für s im Intervall $s \in (0, 1)$ sinnvoll sind, wird diejenige Lösung verwendet, die in diesem Intervall liegt. Sollten beide Lösungen die Bedingung erfüllen, wird keine genutzt, da nicht zu bestimmen ist, welche die richtige ist. Ist die Wurzel in (2.4) imaginär, sind die gemessenen Intensitätswerte wahrscheinlich mit einem starken Rauschen behaftet und beide Lösungen werden verworfen.

Die Größe s gibt also die sichtbare Verschiebung in der Einheit „Pixel“ an. Der Raumbereich, den ein Pixel abbildet, hat die Gestalt einer (im Allgemeinen schiefen) Pyramide, deren Spitze im Zentrum der Zentralprojektion C liegt. Abbildung 2.3 zeigt die Projektion dieser Pyramide auf die x - z -Ebene. Für das Pixel ist nur die Grundfläche der Pyramide sichtbar - die Strecke x_B ist also die Breite des von ihm abgebildeten Bereichs. Die für das Pixel sichtbare Verschiebung s ist das Resultat einer räumlichen Verschiebung des abgebildeten Objektpunktes entlang der Strecke x_B . Diese räumliche Verschiebung ist aber gleich dem Abstand der optischen Achsen x_a und somit gilt

$$x_B \cdot s = x_a \quad (2.5)$$

. Die Strecke x_B ist aufgrund des Strahlensatzes proportional zur Höhe der Pyramide z :

$$z \propto x_B \quad (2.6)$$

Also ist

$$z \propto \frac{x_a}{s} \quad (2.7)$$

Werte von s enthält soll im Folgenden als *Verschiebungsbild* bezeichnet werden. Bei der Matrix der z -Werte handelt es sich um das *Tiefenbild*.

2.4 Verwendung einer Marke

Die Strecke x_a ist mit typischen Längen von einigen Hundert μm sehr klein und ihre Messung gestaltet sich deshalb schwierig. Da sich nach (2.7) der relative Fehler von x_a direkt auf die Bestimmung von z fortpflanzen würde, sollte versucht werden, die direkte Messung von x_a zu umgehen. Eine Möglichkeit besteht darin, eine *Marke* im Messraum anzubringen, deren Entfernung z_M von der Kamera bekannt ist.

Ist die Position der Marke im Grauwertbild bekannt, kann sie auch leicht im Verschiebungsbild ausgemacht werden. Die Mittelung der s -Werte über den Bereich der Marke ergibt eine Größe, die als s_M bezeichnet werden soll. Mit ihrer Hilfe ergibt sich für die Entfernungswerte z die folgende Relation:

$$\frac{z}{z_M} = \frac{s_M}{s} \quad (2.9)$$

Bei räumlich hoch auflösenden Kameras erreicht man schon mit kleinen Marken (Fläche einige cm^2) leicht Mittelungen über einige Hundert s -Werte und somit einen relativ geringen Fehler Δs_M . Dazu sollte die Marke eine kontrastreiche Oberfläche aufweisen, die die Gewinnung einer Vielzahl möglichst präziser s -Werte ermöglicht. Günstig scheint zum Beispiel die Verwendung schwarz-weiß gestreifter oder schachbrettartig bedruckter Ebenen. Die Ausrichtung senkrecht zur optischen Achse hat zur Folge, dass alle gemessenen Verschiebungen s des Bereichs der Marke theoretisch identisch sind (weil alle Pixel dieselbe Entfernung z zur Kamera besitzen) und verbessert die Genauigkeit von s_M .

Sind die Position der Marke, die Lage der Kamera im Raum und die inneren und äußeren Kameraparameter hinreichend genau bekannt (und das sollten sie, um das Tiefenbild später in eine exakte 3D-Punktwolke umrechnen zu können), lässt sich die Lage der Marke im Grauwertbild direkt berechnen.

Grundsätzlich besteht die Gefahr, dass die Marke durch Objekte verdeckt wird und im Grauwertbild überhaupt nicht sichtbar ist. Ein Lösungsansatz wäre hier, mehrere Marken im Arbeitsraum zu platzieren, sodass mit hoher Wahrscheinlichkeit immer mindestens eine sichtbar ist. Mit Algorithmen der Mustererkennung könnte überprüft werden, ob sich in den Bereichen der Abbildung, in denen die Marke vermutet wird, auch tatsächlich eine befindet.

Die Nutzung von Kenntnissen über nicht-variable Teile des zu vermessenden Arbeitsraumes ist ebenfalls denkbar. So könnte z. B. die gesamte Aufspanfläche als Tiefenreferenz dienen, falls diese mit einer ausreichend kontrastreichen Oberfläche versehen wurde.

3 Möglichkeiten und Grenzen

Ziel der Tiefenmessung ist die schnelle und präzise Entfernungsbestimmung. Die absolute Genauigkeit Δz des vorgestellten Verfahrens hängt entscheidend von der Genauigkeit der Intensitätsmessung des Lichtes durch die CCD-Kamera ab. Darum werden zuerst die Ursachen für das Rauschen ΔI der Intensitätswerte untersucht und unter der Annahme einer idealen Optik und optimalen Messumgebung bestimmt, in welcher Art die Fehler Δs und demzufolge auch Δz von ΔI abhängen. Hieraus ergeben sich die optimalen Kameraparameter (Blende, Fokus etc.), die Δz minimieren.

Anschließend wird untersucht, welche Auswirkungen äußere Faktoren wie Schwankungen der Beleuchtungsstärke und die mechanische Stabilität des Systems auf die Bestimmung der Entfernung einzelner Punkte haben.

Am Ende des Kapitels soll auf die Auswirkungen von Abbildungsfehlern des Objektivs und Möglichkeiten der Korrektur dieser eingegangen werden.

3.1 Der Fehler der Intensitätsmessung

3.1.1 Sensorrauschen: ΔI

Bei der Erfassung von Grauwertbildern durch die CCD-Kamera führt jedes Pixel des Kamera-Sensors eine Messung der Intensität des auftreffenden Lichtes durch. Die einfallenden Photonen erzeugen durch den inneren Photoeffekt Elektron-Loch-Paare, deren Anzahl proportional zur Zahl der Photonen ist. Die Elektronen werden zum Rand der lichtempfindlichen Fläche transportiert, analog verstärkt und digitalisiert.

Die Standardabweichung der gemessenen Intensitätswerte wird als *Rauschen* bezeichnet und setzt sich aus verschiedenen Komponenten zusammen:

Das *Photonenrauschen* (engl. *shot noise*) ist die Folge statistischer Schwankungen der Zahl der von der Lichtquelle emittierten Photonen und eine elementare Eigenschaft aller Lichtquellen. Die Frage, wie viele Photonen innerhalb einer bestimmten Zeit auf den Detektor auftreffen werden, lässt sich nur wahrscheinlichkeitstheoretisch behandeln. Mit Hilfe der Poisson-Verteilung errechnet sich die durch das Photonenrauschen verursachte Standardabweichung des detektierten Signals zu

$$\sigma_p = \sqrt{N} \tag{3.1}$$

, wobei mit N die Anzahl der detektierten Photonen bezeichnet wurde.

Durch thermische Schwingungen des Gitters werden Elektronen generiert, die den sog. Dunkelstrom erzeugen. Dieser besteht aus einem zeitlich konstanten und einem zufälligen

Anteil, dem *Dunkelstromrauschen*. Durch Subtraktion eines Dunkelbildes¹ vom Grauwertbild lässt sich der durch den konstanten Anteil verursachte Fehler nahezu eliminieren. Das Dunkelstromrauschen σ_d trägt zum Gesamtrauschen bei und ist stark temperaturabhängig. Eine Reduktion der Temperatur um 9°K halbiert das Dunkelstromrauschen, weshalb für bestimmte Anwendungen hauptsächlich gekühlte Kameras eingesetzt werden². Eine Formel zur Abschätzung des Dunkelstromrauschens wird in [DH04] gegeben.

Das *Ausleserauschen* σ_a wird bei der Umwandlung der Ladungsträger in digitale Einheiten erzeugt. Es umfasst das Verstärker- und Quantisierungsrauschen ebenso wie die zufälligen Schwankungen, die durch Verlust von Ladungsträgern beim Verschieben zur Ausleseeinheit und beim Zurücksetzen des Potentials dieser (engl. „reset noise“), entstehen.

Da es sich bei den Komponenten um unkorreliertes weißes Rauschen handelt, errechnet sich das Gesamtrauschen σ_I zu:

$$\sigma_I = \sqrt{\sigma_p^2 + \sigma_d^2 + \sigma_a^2} \quad (3.2)$$

3.1.2 Das Signal-Rausch-Verhältnis

Für die Beurteilung der Qualität einer Messung ist das Verhältnis zwischen Signal und Rauschen (engl. *signal-to-noise ratio*, *SNR*) entscheidend. Wird bei konstantem Lichteinfall mehrmals von einem Pixel der Kamera die Intensität erfasst, sollte der Mittelwert sehr viel größer als die Schwankung dieser Werte sein, also $SNR \gg 1$ gelten.

$$SNR = \frac{I}{\Delta I} \quad (3.3)$$

Bei hohen Intensitäten wird (3.2) durch das Photonenrauschen dominiert, sodass näherungsweise

$$\Delta I = \sigma_I \approx \sigma_p \quad (3.4)$$

gilt. Das Signal-Rausch-Verhältnis lässt sich in diesem Fall nach (3.1) einfach durch

$$SNR_I = \frac{N}{\sqrt{N}} = \sqrt{N}$$

abschätzen. Jedes Pixel der CCD-Kamera akkumuliert während des Belichtungsvorgangs Elektronen, wobei deren Anzahl gleich der Zahl der detektierten Photonen ist. Die im Englischen als *full well capacity* bezeichnete durchschnittliche („Elektronen“-)Kapazität eines einzelnen Pixels entspricht also der maximalen Anzahl detektierbarer Photonen. Somit ergibt sich die folgende Abschätzung:

$$SNR_I \leq \sqrt{\text{full well capacity}} \quad (3.5)$$

¹Als Dunkelbild bezeichnet man eine unter vollständiger Abdunklung der Kamera erzeugte Aufnahme.

²Z. B. in der Fluoreszenzmikroskopie, wo es auf die Detektion sehr kleiner Lichtmengen ankommt.

3.2 Kameraparameter bei idealer Optik und Umgebung

3.2.1 Der Fehler der sichtbaren Verschiebung Δs

In Abschnitt 2.3 aus Kapitel 2 wurde Gleichung (2.4) zur Berechnung der sichtbaren Verschiebung s vorgestellt. Zur besseren Übersicht soll sie hier noch einmal abgedruckt werden:

$$s_{1,2} = -\frac{b}{2a} \pm \sqrt{\frac{b^2}{4a^2} - \frac{c - I_{02}}{a}} \quad (3.6)$$

Mit Hilfe des Gauß'schen Fehlerfortpflanzungsgesetzes lässt sich aus (3.6) der Fehler Δs der s -Werte errechnen.

$$\Delta s = \left| \frac{\partial s}{\partial a} \right| \Delta a + \left| \frac{\partial s}{\partial b} \right| \Delta b + \left| \frac{\partial s}{\partial c} \right| \Delta c$$

wobei

$$\begin{aligned} \frac{\partial s}{\partial a} &= \frac{b}{2 \cdot a^2} \pm \frac{\frac{c - I_{02}}{a} - \frac{b^2}{2 \cdot a^2}}{2 \cdot \sqrt{\frac{b^2}{4} - a(c - I_{02})}} \\ \frac{\partial s}{\partial b} &= \frac{1}{4 \cdot a} \cdot \left(-2 \pm \frac{b}{2 \cdot \sqrt{\frac{b^2}{4} - a(c - I_{02})}} \right) \\ \frac{\partial s}{\partial c} &= \frac{-1}{2 \cdot \sqrt{\frac{b^2}{4} - a(c - I_{02})}} \end{aligned}$$

. Ob in den partiellen Ableitungen von s nach a und b der letzte Bruch jeweils addiert oder subtrahiert wird, richtet sich danach, welche Teillösung von s gewählt wurde.

Die Parameter a , b und c errechnen sich nach (2.1) - (2.3) (Kapitel 2, Abschnitt 2.3) aus den Intensitätswerten dreier benachbarter Pixel, die mit I_l , I_0 und I_r bezeichnet wurden. Das Gauß'sche Fehlerfortpflanzungsgesetz ermöglicht auch hier die einfache Abschätzung der Fehler Δa , Δb und Δc , die sich unter der Annahme, dass die Standardabweichungen der Intensitätswerte für alle Pixel gleich ΔI sind, weiter vereinfachen lassen:

$$\begin{aligned} \Delta a &= \frac{1}{2}(\Delta I_l + \Delta I_r) + \Delta I_0 = 2\Delta I \\ \Delta b &= \frac{1}{2}(\Delta I_l + \Delta I_r) = 2\Delta I \\ \Delta c &= \Delta I_0 = \Delta I \end{aligned}$$

Es ist also möglich, aus den Intensitäten des Grauwertbildes zu jedem einzelnen s -Wert einen Fehler Δs abzuschätzen. Die Informationen, die die Δs -Werte liefern, lassen sich zur Erhöhung der Genauigkeit und Geschwindigkeit der Tiefenbestimmung nutzen. Beispielsweise können vor der Berechnung des Tiefenbildes alle s -Werte, deren Fehler einen gewissen Grenzwert Δs_{krit} überschreitet, verworfen werden. So wird die Zahl der Berechnungen der z -Werte verringert und das Tiefenbild enthält weniger Werte zu kontrastarmen Bildregionen, womit die Genauigkeit der Tiefeninformation der verbleibenden Punkte erhöht wird.

3.2.2 Abhängigkeit des Fehlers der Tiefenbestimmung Δz von Δs

Die Entfernung z eines Pixels errechnet sich nach der in Kapitel 2, Abschnitt 2.4 vorgestellten Gleichung (2.9) zu

$$z = \frac{s_M \cdot z_M}{s} \quad (3.7)$$

Mit Hilfe der Gauß'schen Fehlerrechnung ergibt sich der Fehler Δz der errechneten Entfernung zu

$$\Delta z = \frac{z_M}{s} \cdot \Delta s_M + \frac{s_M}{s} \cdot \Delta z_M + \frac{s_M \cdot z_M}{s^2} \cdot \Delta s \quad (3.8)$$

Die rechte Seite von (3.8) lässt sich stark vereinfachen: Da die Marke aufgrund ihrer großen und speziell präparierten Oberfläche sehr viele Referenzwerte $s_{m,i}$ erzeugt, ist der gemittelte Wert für s_M mit einem verhältnismäßig geringen Fehler behaftet. Daher kann $\Delta s_M = 0$ angenommen werden. Ebenso lässt sich Δz_M vernachlässigen, da beim Einrichten der Maschine die Position der Marke und somit auch z_M relativ genau bestimmt wurden. Δz wird also durch den dritten Summanden dominiert, sodass es sich vereinfacht als (3.9) schreiben lässt.

$$\Delta z = \frac{s_M \cdot z_M}{s^2} \cdot \Delta s \quad (3.9)$$

Durch Einsetzen von (3.7) in (3.9) erhält man den Fehler der Tiefe Δz abhängig von der gemessenen Entfernung z :

$$\Delta z = \frac{z^2 \cdot \Delta s}{s_M \cdot z_M} \quad (3.10)$$

Da nach (3.10) z quadratisch in Δz eingeht, wird der Fehler Δz am Ende des Messbereichs maximal. Um eine möglichst genaue Tiefenwerte zu erhalten, sollten Δs minimal und s_M sowie z_M so groß wie möglich sein. Die Positionierung der Marke sollte also am Ende des Messbereichs erfolgen (z_M maximal).

3.2.3 Wahl von x_a und s_M

Wie bei der Diskussion von (3.10) erwähnt wurde, sollte s_M zur Minimierung des Fehlers der Tiefenbestimmung Δz möglichst groß sein. Da $s_M \propto x_a$ gilt (siehe (2.7) in Kapitel 2, Abschnitt 2.3), muss also auch x_a möglichst groß sein. Eine Erhöhung von x_a bewirkt aber eine Erhöhung der geringsten messbaren Entfernung z_{min} und schränkt somit den Messbereich ein (siehe (2.8) in Kapitel 2, 2.3).

Abhängig von der minimal zu erwartenden Entfernung einzelner Objekte sollte die Verschiebung x_a also so groß wie möglich, aber gleichzeitig so gering gewählt werden, dass alle Objekte erfasst werden.

3.2.4 Abbildungsgröße minimaler Strukturen und Herleitung der Tiefenschärfe einer Kante

Für eine möglichst genaue Bestimmung der sichtbaren Verschiebung s darf die Abbildung der Kamera nicht zu scharf sein. Die Kameraparameter müssen so eingestellt werden, dass der theoretisch unendlich dünne Farbübergang einer Kante auf einen Bereich $B_d \geq B_{d,min}$ abgebildet wird. $B_{d,min}$ ist etwa so groß wie der Abstand der äußersten, zum Schätzen von s verwendeten Pixel. Wird diese Differenz (in Pixeln) mit n_p und der Abstand zweier benachbarter Pixel (in Richtung der ξ -Achse) mit δ_p bezeichnet, gilt für die Schätzung von s aus drei Intensitätswerten also

$$B_d \geq B_{d,min} = n_p \cdot \delta_p = 2 \cdot \delta_p \quad (3.11)$$

. Gleichung (3.11) wird als „Unschärfebedingung“ bezeichnet.

Im Folgenden soll hergeleitet werden, wie B_d von den Kameraparametern abhängt. Dazu wurde der Strahlengang in Abbildung 3.1 skizziert. Mittig ist das Objektiv mit

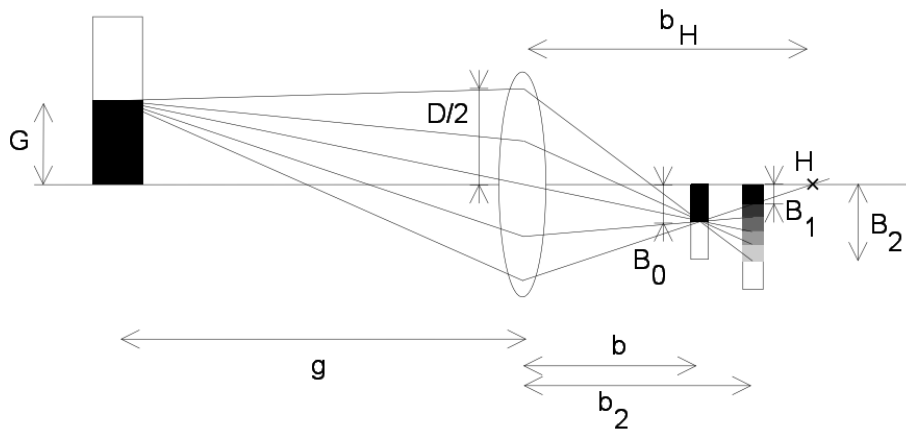


Abbildung 3.1: Skizze zur Herleitung der Breite B_d einer abgebildeten Kante

dem Öffnungsdurchmesser D dargestellt. Zur Vereinfachung wird davon ausgegangen, dass das Objektiv nur eine Hauptebene besitzt, sich also wie eine dünne Linse verhält. Da die gesuchte Größe B_d sehr viel größer als die Auflösung des Objektivs ist, ist diese Näherung gerechtfertigt³. Links vom Objektiv befindet sich der Gegenstandsraum; rechts der Bildraum. Die Kante, die sich im Abstand G zur optischen Achse und im Abstand g zur Linse befindet, wird im Bildraum im Abstand b scharf abgebildet. Die

³Die hier vorgestellte rein strahlengeometrische Herleitung liefert nur einen grobe Schätzwert für die Schärfe der Kante, der jedoch für die weitere Verwendung ausreichend ist. In Abschnitt 3.4 wird kurz auf zusätzliche Effekte wie die Beugung des Lichtes eingegangen. Eine exakte Beschreibung erfordert die Herleitung der Optischen Transferfunktion (OTF) des Objektivs, welche sich nur mit erheblichem numerischen Aufwand und unter genauer Kenntnis aller Kalibrierparameter des Objektivs bestimmen ließe. Außerdem wäre die Kenntnis der spektralen Eigenschaften aller beteiligten optischen Komponenten (Transmission des Objektivs, Emission der Lichtquelle, Reflektanz der zu vermessenden Objekte, ...) erforderlich.

Größe dieser scharfen Abbildung wird mit B_0 bezeichnet. Betrachtet man die Abbildung im Abstand b_2 vom Objektiv, ist diese unscharf. Interessant ist nun die Breite B_d dieser unscharf abgebildeten Kante, die sich als Differenz der Größen B_1 und B_2 ergibt, welche die Grenzen des unscharfen Bereichs darstellen. Der Hilfspunkt H ergibt sich als Schnittpunkt der optischen Achse mit der Verlängerung des Strahls, der durch die Grenze B_1 der unscharfen Abbildung verläuft. Dieser Schnittpunkt befindet sich im Abstand b_H vom Objektiv. Es gelten die bereits in Kapitel 1, speziell Abschnitt 1.5.2 erwähnte Newtonsche Linsengleichung und die Abbildungsgleichung, welche zur besseren Übersicht hier nocheinmal abgedruckt und mit (3.12) und (3.13) bezeichnet werden:

$$\frac{1}{f} = \frac{1}{b} + \frac{1}{g} \quad (3.12)$$

$$\frac{B}{G} = \frac{b}{g} \quad (3.13)$$

Hierbei ist f die Brennweite des Objektivs. b und g bezeichnen die Bild- und Gegenstandsweite, also die jeweilige Entfernung zur Hauptebene. Die Bild- und Gegenstandshöhe wurden mit B und G bezeichnet. Mit Hilfe des Strahlensatzes lassen sich die Relationen

$$\frac{b_H}{D/2} = \frac{b_H - b}{B_0} = \frac{b_H - b_2}{B_1} \quad (3.14)$$

aufstellen. Woraus sich ergibt

$$\begin{aligned} b_H &= \frac{D b_H}{2 B_0} - \frac{D b}{2 B_0} = \frac{D}{2 \cdot B_0} (b_H - b) \\ &\Leftrightarrow b_H = \frac{-b \frac{D}{2 B_0}}{1 - \frac{D}{2 B_0}} \end{aligned} \quad (3.15)$$

Außerdem folgt aus (3.14):

$$B_1 = \frac{D}{2} \left(1 - \frac{b_2}{b_H} \right) \quad (3.16)$$

Setzt man (3.15) in (3.16) ein, erhält man

$$B_1 = \frac{b_2}{b} \cdot B_0 + \frac{D}{2} \cdot \left(1 - \frac{b_2}{b} \right) \quad (3.17)$$

Weiterhin gilt die Relation

$$\frac{B_0 + \frac{D}{2}}{b} = \frac{B_2 + \frac{D}{2}}{b_2}$$

woraus sich für B_2 ergibt:

$$B_2 = \frac{(B_0 + \frac{D}{2}) \cdot b_2}{b} - \frac{D}{2} = \frac{b_2}{b} \cdot B_0 - \frac{D}{2} \cdot \left(1 - \frac{b_2}{b} \right) \quad (3.18)$$

Aus (3.17) und (3.18) folgt also:

$$B_d = |B_2 - B_1| = \left| D \left(\frac{b_2}{b} - 1 \right) \right| = D \left(\left| \frac{b_2}{b} \right| - 1 \right)$$

Aus (3.12) folgt

$$b = \frac{f \cdot g}{g - f}$$

$$b_2 = \frac{f \cdot g_2}{g_2 - f}$$

. Es ist also g die am Objektiv eingestellte Entfernung (Fokus) und die Größe g_2 entspricht der Entfernung eines beobachteten Objektes, weshalb $g_2 \equiv z$ gilt. Beachtet man nun, dass die sich der Objektivdurchmesser als Quotient von Brennweite und Blendenzahl des Objektivs schreiben lässt, also

$$D \equiv \frac{f}{k}$$

, erhält man

$$B_d = \frac{f}{k} \left(\frac{g \cdot z - z \cdot f}{g \cdot z - g \cdot f} - 1 \right) = \frac{f^2}{k} \left(\frac{g - z}{g \cdot (z - f)} \right) \quad (3.19)$$

Für eine Fläche, deren Oberfläche eine unendlich dünne Kante aufweist, die so verläuft, dass ihre Abbildung senkrecht zur ξ -Achse der Bildebene orientiert ist, lässt sich Δs abschätzen. Dazu wird angenommen, dass die Differenz der von den benachbarten Flächen zur Kamera reflektierten Intensitäten das l -fache der minimal auflösbaren Intensitätsdifferenz ist. Mit Hilfe der Breite der abgebildeten Kante B_d und des Abstands der Pixel δ_p erhält man:

$$\Delta s = \frac{B_d}{l \cdot \delta_p} \quad (3.20)$$

Die Größe l ist also die Anzahl der an der Kante auflösbaren Graustufen. Das theoretische Maximum für l entspricht gerade dem Signal-Rausch-Verhältnis der Intensitätsmessung SNR_I . Bei einer realen Abbildung wird dieser Wert jedoch nicht erreicht werden können, da aufgrund von Streulicht selbst sehr dunkle Flächen nicht vollkommen schwarz erscheinen. Außerdem wird man die Belichtungsparameter so wählen, dass keine Übersteuerung auftritt, weshalb die höchsten Intensitätswerte einige Prozent unter dem für Vollaussteuerung erreichbaren Wert liegen werden. l kann somit unter experimentellen Bedingungen zu

$$l \approx 0,90 \cdot SNR_I$$

abgeschätzt werden.

Mit den Objektivparametern in (3.19) ergibt sich aus (3.20) und (3.10):

$$\Delta z = \frac{z^2 \cdot f^2 \cdot (g - z)}{s_M \cdot z_M \cdot l \cdot \delta_p \cdot k \cdot g \cdot (z - f)} \quad (3.21)$$

Durch die Fokussierung der Kamera auf eine Entfernung g , die hinter dem Messbereich liegt ($g > z_{max}$), lässt sich erreichen, dass Δz relativ konstant über den gesamten Messbereich anwächst. Die Belichtungszeit der Kamera und die Blendenzahl k müssen so aufeinander abgestimmt werden, dass die Bedingung $B_d \geq B_{d,min}$ für alle messbaren Werte von z nicht verletzt wird. Da für $g > z_{max}$ die Schärfe für $z = z_{max}$ maximal (also B_d minimal) wird, reicht es sicherzustellen, dass $B_d(z_{max}) \geq B_{d,min}$ erfüllt ist.

Abbildung 3.2 zeigt den nach (3.21) simulierten absoluten Fehler Δz abhängig von der Entfernung z unter folgenden Annahmen: Messbereich $z = 300 \dots 1000\text{mm}$, $z_m = 1000\text{mm}$, $s_m = 0,3\text{Pixel}$, $l = 100$, $f = 8\text{mm}$, $g = 2000\text{mm}$ und $k = 3,2$.

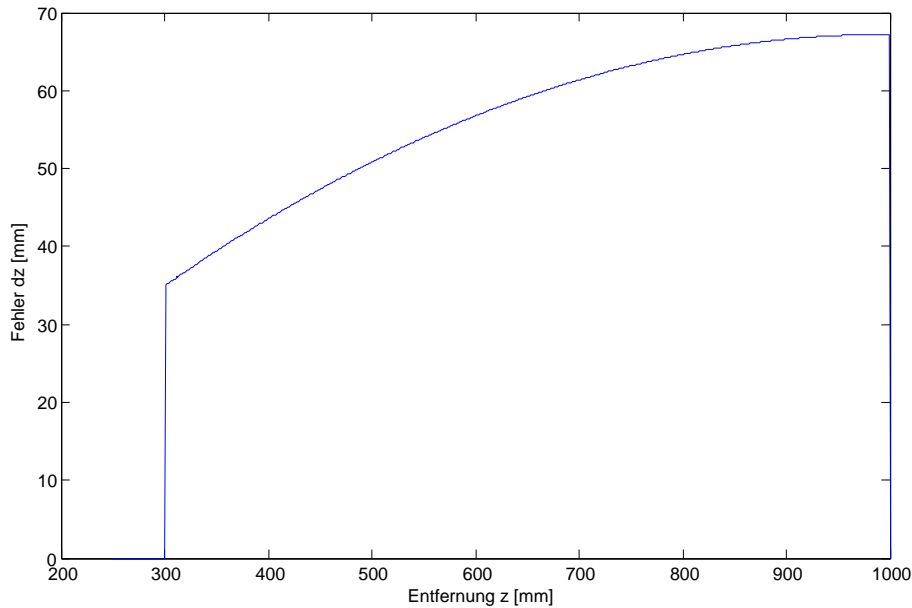


Abbildung 3.2: Errechneter Fehler Δz der Entfernungsbestimmung abhängig von der Entfernung z

Der Fehler Δz der Entfernungsbestimmung liegt nach diesen Berechnungen im gesamten Messbereich zwischen 35 und 68mm.

3.3 Umgebungseinflüsse

3.3.1 Schwankungen der Helligkeit der Lichtquelle

Die gemessene Intensität einer abgebildeten Region des Messraums ist das Produkt der Reflektanz R ihrer Oberfläche und der Gesamthelligkeit, mit der die Region beleuchtet wird:

$$I(\xi, \eta) = R(x, y, z) \cdot L_{ges}$$

Die Gesamthelligkeit L_{ges} ergibt sich durch die Aufsummierung der Intensitäten aller Lichtquellen des Systems L_k , jeweils multipliziert mit einem Korrekturfaktor G_k , der

z. B. die räumliche und spektrale Abstrahlcharakteristik beinhaltet.

$$L_{ges} = \sum_k L_k \cdot G_k$$

Eine genaue Entfernungsbestimmung erfordert, dass die gemessene Intensität $I(\xi, \eta)$ (bei unbewegter Kamera) zeitlich möglichst konstant ist. Die Reflektanz der abgebildeten Region $R(x, y, z)$ kann als unveränderlich angenommen werden. Dass sich die Gesamthelligkeit des Systems nicht ändert, lässt sich hingegen nicht voraussetzen. Eventuelle Schwankungen lassen sich aber korrigieren, falls eine der folgenden Vereinfachungen zutrifft:

- es existiert nur eine einzige Lichtquelle im System - oder -
- jede Lichtquelle beleuchtet die zu vermessenden Oberflächen homogen

Ist mindestens eine dieser Bedingungen erfüllt, lässt sich die Gesamthelligkeit, die während der Aufnahme eines Grauwertbildes wirkte, abschätzen, indem die Intensitätswerte über eine geeignete Fläche des Bildes aufsummiert werden. Eine Division dieses Wertes mit dem entsprechenden Wert des zweiten Bildes ergibt einen Faktor zur Helligkeitskorrektur h_{korrr} . Die Multiplikation aller Intensitätswerte des zweiten Bildes mit diesem Faktor ergeben ein Grauwertbild, das scheinbar unter denselben Beleuchtungsbedingungen wie das erste Bild erzeugt wurde.

3.3.2 Mechanische Stabilität des Systems

Die Positionierung der Kamera im Raum ist nicht beliebig genau. Eine zufällige Verschiebung oder Verkipfung der Kamera bewirkt einen Fehler in der Abbildung, der zu einem zusätzlichen Fehler des erfassten Intensitätswerts führt. Da das vorgestellte Verfahren zur Tiefenbestimmung auf einer definierten Bewegung der Kamera im Raum (d. h. im MKS) und Analyse der Variation der erfassten Intensitätswerte beruht, ist eine möglichst genaue Kontrolle dieser Positions- und Orientierungsänderung erforderlich. Die folgenden Ausführungen beziehen sich also auf die für die Tiefenbestimmung relevante relative Bewegung zwischen den zur Erfassung der Grauwertbilder eingenommenen Positionen der Kamera, und nicht auf die absolute Position und Orientierung dieser im MKS.

Die räumliche Intensitätsfunktion in der Bildebene lässt sich nach dem Satz von Taylor in einer Reihe entwickeln. Mit dieser Reihe (3.22) lässt sich der „positionsbedingte“ Fehler der Intensitätsmessung (3.23) abschätzen:

$$I(\xi + \delta\xi, \eta + \delta\eta) = I(\xi, \eta) + \frac{\partial I(\xi, \eta)}{\partial \xi} \delta\xi + \frac{\partial I(\xi, \eta)}{\partial \eta} \delta\eta + \dots \quad (3.22)$$

$$\Delta I_{Position} = |I(\xi, \eta) - I(\xi + \delta\xi, \eta + \delta\eta)| \approx \left| \frac{\partial I(\xi, \eta)}{\partial \xi} \right| \Delta\xi + \left| \frac{\partial I(\xi, \eta)}{\partial \eta} \right| \Delta\eta \quad (3.23)$$

wobei mit $\Delta\xi$ und $\Delta\eta$ die Beträge des Positions-Fehlers der Abbildung $\delta\xi$ und $\delta\eta$ bezeichnet wurden.

Im Folgenden soll die fehlerhafte Position aus dem Bezugssystem der Kamera heraus untersucht werden. Die zufällige Verschiebung eines Punktes $P = (x, y, z)$ im Messraum wird durch eine Translation entlang der Achsen des KKS beschrieben:

$$P = (x + \delta x, y + \delta y, z + \delta z)$$

Verkippungen sollen durch eine Rotation von P um das Zentrum der Zentralprojektion C modelliert werden. Zur Beschreibung werden die Winkel θ_x , θ_y und θ_z verwendet, welche eine Drehung um die x -, y - und z -Achse des KKS bedeuten.

Da alle Störgrößen klein sind, können ihre Effekte getrennt voneinander betrachtet werden.

Lagefehler

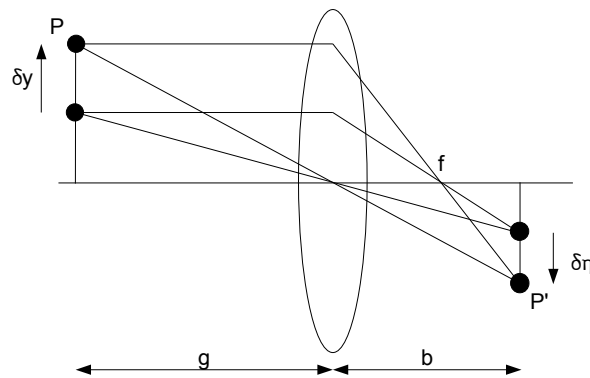


Abbildung 3.3: Auswirkung der Veränderung der Lage eines Punktes in Richtung der y -Achse auf das Bild des Punktes bei scharfer Abbildung

Eine Veränderung der Lage eines Objektpunktes P in Richtung der y -Achse um die Strecke δy bewirkt eine veränderte Position der Abbildung P' in Richtung der η -Achse um die Strecke $\delta \eta$ des BKS. In Abbildung 3.3 ist dies skizziert. Das Verhältnis von der Bild- zur Gegenstandshöhe wird als *Abbildungsmaßstab* M definiert:

$$M = \frac{B}{G}$$

Die Größen δy und $\delta \eta$ sind also proportional zueinander. Der Proportionalitätsfaktor M ist abhängig von der Entfernung z des abgebildeten Objekts und steigt für $z \rightarrow 0$ monoton an. Im Messbereich ergibt sich M_{max} also für den minimalen Messabstand z_{min} . Mit Hilfe der Abbildungsgleichung (3.13) lässt sich M_{max} folgendermaßen abschätzen⁴:

$$M = \frac{B}{G} = \frac{b}{g} \leq \frac{b}{z_{min}} = M_{max} \quad (3.24)$$

⁴Die Abbildungsgleichung (3.13) gilt streng genommen nur für eine scharfe Abbildung. Da die für die Tiefenbestimmung genutzten Grauwertbilder nur eine leicht defokussierte Kamera voraussetzen, entsprechen die geometrischen Verhältnisse in der Kamera fast denen einer scharfen Abbildung. Die Abschätzung (3.24) ist somit gerechtfertigt.

und somit

$$\Delta\eta_y = |M \cdot \delta y| \leq M_{max} \cdot |\delta y| = M_{max} \cdot \Delta y \quad (3.25)$$

Analog ergibt sich für Veränderungen der Lage in Richtung der x -Achse⁵:

$$\Delta\xi_x \leq |-M_{max} \cdot \delta x| = M_{max} \cdot \Delta x \quad (3.26)$$

Da aber bei Verwendung einer Marke die Position der Kamera in Richtung der x -Achse nicht direkt gemessen, sondern aus der sichtbaren Verschiebung der Marke s_M errechnet wird, trägt $\Delta\xi_x$ nicht zum Fehler Δs der Verschiebung s bei. Dies wird modelliert, indem

$$\Delta\xi_x = 0 \quad (3.27)$$

gesetzt wird.

Aufgrund der Orthogonalität der y - und ξ -Achsen sowie der x - und η -Achsen gilt:

$$\Delta\xi_y = \Delta\eta_x = 0$$

Für die Herleitung der Effekte von Positionierungsfehlern in Richtung der z -Achse (siehe Abbildung 3.4) wird die Abbildungsgleichung (3.13) so umgeschrieben, dass sich die Bildhöhe als Funktion von der Entfernung $g = z$ des abzubildenden Punktes ergibt:

$$B(z) = \frac{b \cdot G}{z}$$

Außerdem wird angenommen, dass die Bildweite b für die Abbildungen der Punkte mit den Entfernungen z und $z + \delta z$ vom Zentrum der Zentralprojektion gleich sind. Diese Vereinfachung ist möglich, da $z \gg b$ gilt. Der Δz -bedingte Fehler der Lage der Abbildung

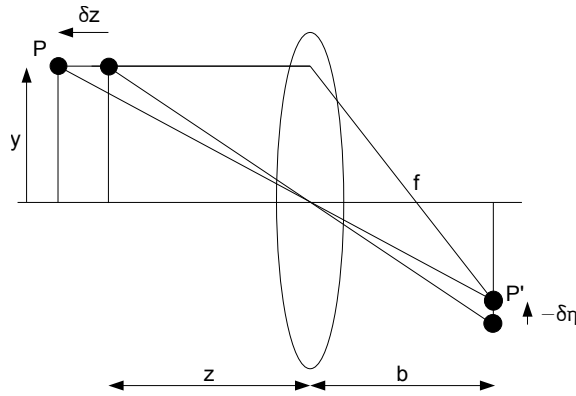


Abbildung 3.4: Auswirkung der Veränderung der Lage eines Punktes in Richtung der z -Achse auf das Bild des Punktes bei scharfer Abbildung

⁵Das negative Vorzeichen folgt aus der antiparallelen Lage der ξ und x -Achsen, während die η und y -Achsen parallel verlaufen.

in Richtung der η -Achse ergibt sich dann zu

$$\delta\eta = \eta - (\eta - \delta\eta) = B(z) - B(z + \delta z) = -\frac{\partial B}{\partial z} \delta z = \frac{b \cdot G}{z^2} \delta z = \frac{B}{z} \delta z \quad (3.28)$$

Die Größe $\delta\eta$ wird nach (3.28) für das kleinstmögliche $z = z_{min}$ und das größtmögliche B maximal. B_{max} entspricht gerade einer Abbildung am Rand des CCD-Sensors, also für $\eta = \rho_{\eta_{max}}$. Das Betragsmaximum $\Delta\eta_z$ ergibt sich somit zu:

$$\Delta\eta_z \leq \frac{\rho_{\eta_{max}}}{z_{min}} \cdot \Delta z \quad (3.29)$$

, wobei hier mit Δz der Betrag von δz bezeichnet wurde.

Analog ergibt sich für die Positionsänderung der Abbildung entlang der ξ -Achse:

$$\Delta\xi_z \leq \frac{\rho_{\xi_{max}}}{z_{min}} \cdot \Delta z \quad (3.30)$$

Im Gegensatz zu dem durch Verschiebungen entlang der x -Achse verursachten Fehlers $\Delta\xi_x$, der homogen auf das gesamte Tiefenbild wirkt und deshalb bei Verwendung einer Marke ignoriert werden und gleich Null gesetzt werden konnte (siehe (3.27)), ist $\Delta\xi_z$ abhängig von der Bildhöhe (und somit der x - und y -Koordinate des Raumpunktes). Er kann deshalb nicht vernachlässigt werden.

Kippfehler

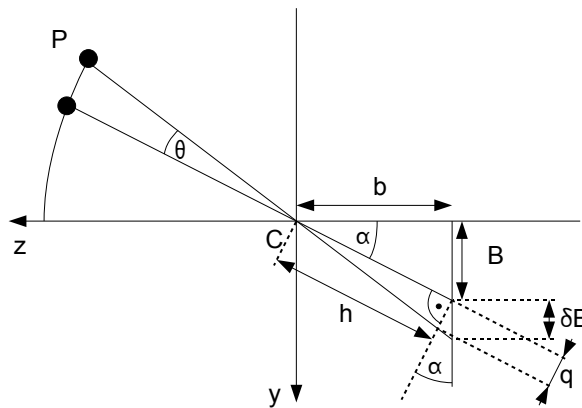


Abbildung 3.5: Veränderung des Ortes der Abbildung aufgrund der Verkippung der Kamera um die x -Achse

In Abbildung 3.5 ist die durch eine Verkippung der Kamera um die x -Achse verursachte Fehler der Bildhöhe δB skizziert. Es gilt

$$\frac{q}{h} = \tan \theta$$

$$h^2 = b^2 + B^2$$

. Da die beiden um den Winkel θ zueinander verkippeten Sehstrahlen fast parallel verlaufen, ist

$$\cos \alpha \approx \frac{q}{\delta B}$$

und deshalb lässt sich der Fehler der Bildhöhe zu

$$\delta B \approx \frac{\tan \theta \cdot h}{\cos \alpha} = \frac{\tan \theta \cdot (b^2 + B^2)}{b}$$

abschätzen.

Für eine Verkippung der Kamera um die x - oder y -Achse wird der Fehler der Position der Abbildung in Richtung der ξ - oder η -Achse also für den größten Abstand der Abbildung von der optischen Achse maximal. Daraus ergeben sich die folgenden Abschätzungen:

$$\Delta \eta_{\theta_x} \leq \frac{\tan \theta \cdot (b^2 + \rho_{\eta_{max}}^2)}{b} \quad (3.31)$$

$$\Delta \xi_{\theta_y} \leq \frac{\tan \theta \cdot (b^2 + \rho_{\xi_{max}}^2)}{b} \quad (3.32)$$

Aufgrund der Orthogonalität der y - und ξ -Achsen sowie der x - und η -Achsen gilt:

$$\Delta \xi_{\theta_x} = 0$$

$$\Delta \eta_{\theta_y} = 0$$

Durch eine Verkippung der Kamera um die z -Achse ergibt sich ebenfalls eine Verschiebung der Position der Abbildung in der Bildebene. Diese beträgt maximal:

$$\Delta \eta_{\theta_z} \leq \cos \varphi \cdot \rho_{max} \cdot \tan \theta_z = \rho_{\xi_{max}} \cdot \tan \theta_z$$

$$\Delta \xi_{\theta_z} \leq \sin \varphi \cdot \rho_{max} \cdot \tan \theta_z = \rho_{\eta_{max}} \cdot \tan \theta_z$$

Zusammenfassung der Positionierungsfehler

Um ein Maximum für die durch eine fehlerbehaftete Kameralage- und -orientierung verursachten Fehler des Ortes der Abbildung angeben zu können, müssen alle ermittelten Fehler addiert werden:

$$\Delta \eta_{Position} = \Delta \eta_y + \Delta \eta_z + \Delta \eta_{\theta_x} + \Delta \eta_{\theta_z} \quad (3.33)$$

$$\Delta \xi_{Position} = \Delta \xi_x + \Delta \xi_z + \Delta \xi_{\theta_y} + \Delta \xi_{\theta_z} \quad (3.34)$$

Nach (3.23) lässt sich aus (3.33) und (3.34) ein positionierungsbedingter Fehler der erfassten Intensität $\Delta I_{Position}$ errechnen:

$$\begin{aligned} \Delta I_{Position} &= \Delta I_{\xi, Position} + \Delta I_{\eta, Position} \\ &= \left| \frac{\partial I(\xi, \eta)}{\partial \xi} \right| \Delta \xi_{Position} + \left| \frac{\partial I(\xi, \eta)}{\partial \eta} \right| \Delta \eta_{Position} \end{aligned} \quad (3.35)$$

Die gemessenen Intensitätswerte sind also zusätzlich mit dem Fehler $\Delta I_{Position}$ behaftet. Falls dieser eine Größenordnung erreicht, die eine Vernachlässigung unmöglich macht, muss die Abschätzung des Fehlers der gemessenen Intensitätswerte (3.4) um den Term (3.35) additiv ergänzt werden.

Umgekehrt lässt sich aus einer geforderten Tiefenauflösung Δz die maximalen Ungenauigkeit der Intensitätsmessung bestimmen und daraus die Toleranzen für die Positioniergenauigkeit errechnen. Dazu ist es erforderlich, eine geeignete Gewichtung festzulegen, welche die Aufteilung der gegebenen Summe $\Delta \eta_{Position}$ bzw. $\Delta \xi_{Position}$ auf die einzelnen Summanden reguliert.

Sollte aufgrund der mechanischen Eigenschaften des Systems der Anteil von $\Delta I_{\eta, Position}$ stark dominieren, ist es denkbar, den daraus resultierenden Fehler für die Tiefenbestimmung zu minimieren, indem vor Beginn der Verarbeitung der Grauwertbilder diejenigen Regionen aus dem Grauwertbild entfernt werden, die eine starke Verfälschung der Intensitätsmessung erfahren. Nach (3.35) lässt sich dies realisieren, indem Bildbereiche, in welchen

$$\left| \frac{\partial I(\xi, \eta)}{\partial \eta} \right| \geq \frac{\Delta I_{\eta, \text{krit}}}{\Delta \eta_{Position}} \quad (3.36)$$

gilt, durch einen entsprechenden Filter aus den Grauwertbildern gelöscht oder im daraus errechneten Tiefenbild als stark fehlerbehaftet gekennzeichnet werden. Mit $\Delta I_{\eta, \text{krit}}$ wurde der maximal vertretbare, durch Lageversetzungen der Abbildung parallel zur η -Achse hervorgerufene Fehler der Intensitätsmessung bezeichnet.

3.4 Einflüsse einer nicht-idealen Optik: Abbildungsfehler des Objektivs

In diesem Abschnitt sollen die Auswirkungen der bei realen Optiken auftretenden Abbildungsfehler auf die Erzeugung von Tiefendaten nach dem vorgestellten Verfahren untersucht werden. Es werden also die sich aus den Abweichungen realer Objektivs zum Modell der Zentralprojektion bzw. dem der dünnen Linse ergebenden Effekte genauer betrachtet.

Jede reale Optik produziert Abbildungen, die nicht frei von Fehlern sind. Nach ihrer Wirkung lassen sich die Abbildungsfehler in verschiedene Klassen einteilen. Da das vorgestellte Verfahren der Tiefendatengewinnung mit Schwarz-Weiß-Kameras arbeitet, sind die *Farbfehler* von eher geringerer Bedeutung⁶.

Schärfefehler, wie z. B. Sphärische Aberration, Koma oder Astigmatismus, bewirken eine lokale Reduktion der Schärfe der Bildpunkte. Auch die häufig den Lagefehlern zugeordnete Bildfeldwölbung bewirkt im Wesentlichen eine Abhängigkeit der Schärfe vom

⁶Farbfehler bewirken eine wellenlängenabhängige Abbildung und somit z. B. Versetzung der farbigen „Teilbilder“ gegeneinander. Bei Verwendung einer Schwarz-Weiß-Kamera, also einer Kamera, die einen großen Teil des sichtbaren Lichtes in einem einzigen Kanal aufzeichnet, werden die Teilbilder gemischt und die Farbfehler bewirken, ähnlich den Schärfefehlern, eine Reduktion der Bildschärfe.

Ort der Abbildung. Solange die Anforderung an die minimale Unschärfe der abgebildeten Kante (3.11) nicht verletzt wird, sind auch die Schärfefehler zu vernachlässigen.

Ein zum Rand hin auftretender Abfall der erfassten Intensität wird als *Vignettierung* bezeichnet. Er tritt besonders bei kurzbrennweitigen Objektiven auf, ist aber bei modernen Optiken meist gut korrigiert. Weist das verwendete Objektiv starke Vignettierung auf, kann der Effekt leicht durch die Multiplikation der gemessenen Intensitäten mit einem ortsabhängigen Faktor korrigiert werden.

Obwohl es sich bei den Effekten der *Beugung* nicht um Fehler des Objektivs handelt, sollen sie hier im Sinne der Abweichung vom einfachen mathematischen Modell der Zentralprojektion kurz betrachtet werden: Aufgrund seiner Wellennatur wird das Licht an der Öffnung der Blende des Objektivs gebeugt. Eine sehr weit (d.h. mehrere hundert Meter) entfernte punktförmige Lichtquelle, die mit der Wellenlänge λ emittiert, wird auf ein sogenanntes Beugungsscheibchen mit dem Radius $r = 1,22 \cdot \lambda \cdot k$ abgebildet, wobei k die Blendenzahl des Objektivs bezeichnet. Deutliche Beugungsscheibchen werden also bei großen Blendenzahlen, weit entfernten Objekten und monochromatischer, scharfer Abbildung sichtbar. Da bei dem verwendeten Tiefenmessverfahren nahe Objekte abgebildet werden, ein großer Wellenlängenbereich genutzt wird und möglichst kleine Blenden verwendet werden, sind Beugungseffekte nicht direkt zu beobachten. Vielmehr kommt es zu einer beugungsbedingten zusätzlichen Unschärfe der Abbildung, welche aber wie auch die Schärfefehler vernachlässigt werden kann, solange die Unschärfebedingung (3.11) erfüllt ist.

Für die beschriebene Methode der Tiefenbestimmung sind besonders die *Lagefehler*, und unter diesen speziell die *Verzeichnung* (Distorsion) relevant. Sie tritt bei asymmetrischer Anordnung der Linsen um die Blende des Objektivs tritt auf und bewirkt, dass der Maßstab in dem ein Objekt abgebildet wird von seiner Entfernung zur optischen Achse abhängt. Die Bildhöhe B ist also in nichtlinearer Weise von der Höhe des entsprechenden Objektpunkts G abhängig. Dies äußert sich zum Beispiel in der gekrümmten Abbildung gerader Linien.

Die Tiefendatenbestimmung hat bisher eine gleichmäßige Verschiebung der Kamera in eine Richtung vorausgesetzt. Im Bezugssystem der Kamera wurde also jeder Punkt des Gegenstandsraum um eine geringe Distanz bewegt. Aufgrund der Distorsion ist diese scheinbare Distanz aber vom Abstand des betrachteten Punktes von der optischen Achse abhängig. Für eine korrekte Entfernungsberechnung muss dieser Effekt also näher betrachtet und wenn möglich korrigiert werden.

Die Distorsion wirkt rotationssymmetrisch um die optische Achse. Um sie zu beschreiben, werden die bereits in 1.5.3 eingeführten Polarkoordinaten verwendet:

$$\rho = \sqrt{(\xi - \xi_0)^2 + (\eta - \eta_0)^2}$$

$$\varphi = \arctan \frac{(\eta - \eta_0)}{(\xi - \xi_0)}$$

Die Verzeichnung des Objektivs lässt sich dann mit Hilfe einer Funktion modellieren, die zu der idealen Bildhöhe ρ die verzerrte Bildhöhe ρ_v liefert. Diese Funktion kann als

Potenzreihe entwickelt werden:

$$\rho_v = f(\rho) = t_1 \cdot \rho + t_2 \cdot \rho^2 + t_3 \cdot \rho^3 \dots \quad (3.37)$$

Da für die folgende Fehlerbetrachtung nur der nichtlineare Teil von (3.37) interessant ist, kann $t_1 = 1$ gesetzt werden⁷. Außerdem wird angenommen, dass die höheren Fehlerterme so schnell abnehmen, dass alle Koeffizienten ab dem dritten Glied gleich Null gesetzt werden können: $t_3 = t_4 = \dots = 0$. Es ergibt sich also die Darstellung

$$\rho_v = f(\rho) = t_1 \cdot \rho + t_2 \cdot \rho^2 \quad (3.38)$$

wobei die Umkehrfunktion

$$\rho_e(\rho_v) = \rho = -\frac{t_1}{t_2} - \sqrt{\frac{t_1^2}{4t_2^2} + \frac{\rho_v}{t_2}}$$

lautet⁸. Die Distorsion wirkt nicht auf den Winkel φ ; es gilt also $\varphi = \varphi_v = \varphi_e$. Typische Werte sind $t_1 = 1$ und $t_2 = -10^{-4}$.

Bisher wurde die Entfernung z eines betrachteten Pixels unter Verwendung einer Marke im Abstand z_M nach (2.9) aus Kapitel 2, Abschnitt 2.4, zu

$$z = \frac{z_M \cdot s_M}{s}$$

errechnet. Aufgrund der Objektivverzeichnung gilt dies aber nur in unmittelbarer Nähe der Marke; die für die Kamera sichtbare Translation x_s ist nicht mehr gleich dem Abstand der optischen Achsen x_a , sondern hängt vom Ort der Abbildung und damit den Bildkoordinaten ξ und η ab:

$$x_s = f(x_a, \xi, \eta)$$

Zur Herleitung denke man sich zwei Punkte des Gegenstandsraums, die so liegen, dass sie unter Annahme einer idealen Optik auf zwei Punkte in der Bildebene abgebildet werden, die sich nur in der ξ -Koordinate geringfügig unterscheiden. Die Punkte in der Bildebene sollen durch ihre Ortsvektoren $\overline{dr_{i1}}$ und $\overline{dr_{i2}}$ beschrieben werden, es gelte:

$$\begin{aligned} r_{i1} &= \begin{pmatrix} \xi_{i1} \\ \eta_{i1} \end{pmatrix} && \text{und} \\ \overline{dr_{i1}} - \overline{dr_{i2}} &= \overline{dr_i} = \begin{pmatrix} \tau \\ 0 \end{pmatrix} && \text{also} \\ r_{i2} &= \begin{pmatrix} \xi_{i2} \\ \eta_{i2} \end{pmatrix} = \begin{pmatrix} \xi_{i1} - \tau \\ \eta_{i1} \end{pmatrix} \end{aligned}$$

⁷Eine Variation von t_1 führt zu einer Änderung des Abbildungsmaßstabs im gesamten Bild. Da die zur Kalibrierung der Tiefenbestimmung benutzte Bildregion der Marke mit demselben Maßstab wie das restliche Bild abgebildet wird, entfällt der Beitrag von t_1 zu den ermittelten Tiefenwerten bei Verwendung einer Marke.

⁸Die Indizes bedeuten (v)erzerzt bzw. (e)ntzerzt.

Aufgrund der Verzeichnung des Objektivs werden sie aber auf die Punkte r_{v1} bzw. r_{v2} abgebildet, welche sich nach (3.38) errechnen:

$$\begin{aligned}
\rho_{i1} &= \sqrt{(\xi_{i1} - \xi_0)^2 + (\eta_{i1} - \eta_0)^2} \\
\rho_{i2} &= \sqrt{(\xi_{i1} - \tau - \xi_0)^2 + (\eta_{i1} - \eta_0)^2} \\
\varphi_{i1} &= \arctan \frac{(\eta_{i1} - \eta_0)}{(\xi_{i1} - \xi_0)} \\
\varphi_{i2} &= \arctan \frac{(\eta_{i1} - \eta_0)}{(\xi_{i1} - \tau - \xi_0)} \\
\rho_{v1} &= t_1 \cdot \rho_{i1} + t_2 \cdot \rho_{i1}^2 \\
\rho_{v2} &= t_1 \cdot \rho_{i2} + t_2 \cdot \rho_{i2}^2 \\
\varphi_{v1} &= \varphi_{i1} \\
\varphi_{v2} &= \varphi_{i2}
\end{aligned}$$

also

$$\begin{aligned}
\overline{r_{v1}} &= \begin{pmatrix} \rho_{v1} \cdot \cos \varphi_{v1} \\ \rho_{v1} \cdot \sin \varphi_{v1} \end{pmatrix} \\
\overline{r_{v2}} &= \begin{pmatrix} \rho_{v2} \cdot \cos \varphi_{v2} \\ \rho_{v2} \cdot \sin \varphi_{v2} \end{pmatrix}
\end{aligned}$$

und somit ist

$$\overline{r_{v1}} - \overline{r_{v2}} = \overline{dr_v} = \begin{pmatrix} \rho_{v1} \cdot \cos \varphi_{v1} - \rho_{v2} \cdot \cos \varphi_{v2} \\ \rho_{v1} \cdot \sin \varphi_{v1} - \rho_{v2} \cdot \sin \varphi_{v2} \end{pmatrix} = \begin{pmatrix} v_\xi \\ v_\eta \end{pmatrix}$$

Das bedeutet, dass zwei Punkte, die bei einer fehlerfreien Optik entlang der ξ -Achse im Abstand τ abgebildet werden würden, sich aufgrund der Distorsion im Allgemeinen nun auch in ihrer η -Koordinate unterscheiden und (bei tonnenförmiger Verzerrung) in einem geringeren Abstand zueinander befinden. Da die vorgestellte Methode der Tiefenbestimmung auf der Schätzung der Kamerabewegung beruht, müssen die verfälschenden Eigenschaften des Objektivs beachtet werden.

Mithilfe numerischer Tests wurde ermittelt, dass für typische Werte von $\tau_0 = 0,2$ Pixel und $\tau \in (0, 1)$ gilt:

$$v_\xi(\tau) \approx v_\xi(\tau_0) \cdot \frac{\tau}{\tau_0}$$

und

$$v_\eta(\tau) \approx v_\eta(\tau_0) \cdot \frac{\tau}{\tau_0}$$

, wobei der relative Fehler jeweils weniger als 10^{-5} betrug. Dies eröffnet die Möglichkeit, schon während der Kamerakalibrierung die Werte von $v_\xi(\tau_0)$ und $v_\eta(\tau_0)$ für alle Pixelkoordinaten zu berechnen und abzuspeichern. Zur Laufzeit der Tiefenberechnung kann

dann auf die im voraus berechneten v -Matrizen zugegriffen werden, was die Distorsionskorrektur stark beschleunigt.

Die ξ -Komponente der sichtbaren Verschiebung lässt sich also schreiben als

$$x_{s,\xi} = x_a \cdot \frac{v_\xi(\xi, \eta, \tau)}{v_\xi(\xi_0, \eta_0, \tau)} \approx x_a \cdot \frac{v_\xi(\xi, \eta, \tau_0)}{v_\xi(\xi_0, \eta_0, \tau_0)}$$

womit sich für z ergibt

$$z \propto \frac{x_a \cdot v_\xi(\xi, \eta, \tau_0)}{v_\xi(\xi_0, \eta_0, \tau_0) \cdot s}$$

was bei Verwendung der Informationen einer Marke auf folgendes führt:

$$z = \frac{v_\xi(\xi, \eta, \tau_0)}{s} \cdot z_M \cdot \underbrace{\frac{s_M}{v_\xi(\xi_M, \eta_M, \tau_0)}}_{\text{arithm.mitteln}} \quad (3.39)$$

Die objektivische Verzeichnung bewirkt, dass die Kamerabewegung scheinbar auch entlang der η -Achse stattfindet. Da dies bei einem starken Kontrast in η -Richtung zu einer Verfälschung der errechneten z -Werte führen würde, ist es sinnvoll, die Information aus v_η für die Einstellung eines Filters zur Aufbereitung der Grauwertbilder analog zu (3.36) aus Abschnitt 3.3.2 zu verwenden, und alle Bereiche aus den Grauwertbildern zu entfernen, die die folgende Bedingung erfüllen:

$$\left| \frac{\partial I(\xi, \eta)}{\partial \eta} \right| \geq \frac{\Delta I_{\eta, \text{krit}}}{\Delta \eta_{\text{Position}} + |v_\eta(\tau_{\text{max}})|} \quad (3.40)$$

Die Lagedifferenz τ der betrachteten Punkte ist für eine Abschätzung im Kontext der Tiefengewinnung nach dem vorgestellten Verfahren nur für Werte für $|\tau| \leq 1$ interessant. Der Wert τ_{max} kann also gleich Eins gesetzt werden.

4 Messungen

4.1 Das Bildaufnahmesystem

Gegenstand dieses Kapitels ist die praktische Durchführung des beschriebenen Verfahrens zur Tiefendatengewinnung. Wie die vorangegangenen Überlegungen zeigten, wird die Tiefenauflösung durch die Grauwertauflösung der Kamera begrenzt. Um eine besonders genaue Tiefenbestimmung durchzuführen ist es daher notwendig eine Kamera zu verwenden, die sehr viele Graustufen mit einem möglichst hohen Signal-Rausch-Verhältnis SNR_I auflösen kann. Bei Verwendung einer Schwarz-Weiß-Kamera sollte dies deutlich höher als bei einer Farbkamera sein, da hier ein größerer Anteil der Sensorfläche zur Intensitätsmessung eingesetzt wird. Nach Abschnitt 3.1.2 aus Kapitel 3 muss die Kamera außerdem über eine möglichst hohe full-well-capacity verfügen, da diese das Photonenrauschen und damit das gesamte Signal-Rausch-Verhältnis maßgeblich beeinflusst. Um eine ausreichende latererale Auflösung sicherzustellen sollte der Sensor der verwendeten Kamera mindestens eine Million Pixel bereit stellen. In Hinblick auf den späteren Einsatz waren weiterhin der Preis mit maximal 5000€ sowie eine kompakte Bauform der Kamera entscheidend.

Während der Suche nach einer geeigneten Kamera erklärten sich freundlicherweise zwei Hersteller bereit, leihweise Kameras zur Durchführung von Experimenten zur Verfügung zu stellen. So konnten mit der Mikroskopkamera PROGRES MF von JENOPTIK sowie dem Modell PIXELFLY HIRES von der PCO AG Versuche durchgeführt werden.

Die Kamera PROGRES MF¹ lief unter der FirmWare Version 2.6. Angesteuert wurde die Kamera über den Treiber FIRECAMJ in Version 1.25.0.0 mit der Software Capture Pro, Version 2.1. Sie wurde über die FireWire-Schnittstelle mit Energie versorgt. Da der Anschluss des verwendeten Notebooks nicht in der Lage war, eine ausreichende Stromversorgung sicherzustellen, kam zwischen Kamera und Laptop zusätzlich ein FireWire-Hub zum Einsatz. Dieser *6-Port Hub*² der Firma BELKIN verfügte über ein eigenes Netzteil. Die Kamera verwendet den CCD-Sensor ICX285AL von SONY, welcher eine Diagonale von 11mm (dies entspricht $2/3''$) aufweist und über 1360×1024 Pixel mit einer full-well-capacity von $18000e^-$ verfügt. Die Pixelgröße³ beträgt $6,45\mu\text{m} \times 6,45\mu\text{m}$.

Die Kamera PIXELFLY HIRES⁴ war über ein Netzkabel und die Framegrabber-Karte 540SC⁵ mit einem PC verbunden und wurde darüber mit Energie versorgt.

¹Seriennummer 2206-27-0114

²Seriennummer C600245172

³Hier ist nicht die optisch sensitive Fläche, sondern die Pixelgröße inkl. der Totbereiche gemeint.

⁴Seriennummer 270XD10002

⁵Seriennummer 2502002058

Zum Ansteuern wurde die Software CAMWARE in Version 2.19.0141 verwendet. Die Kamera erfasst 1392×1024 Pixel mit einer full-well-capacity von $13000e^-$. Der CCD-Sensor hat eine Diagonale von 7,68mm (welches $1/2''$ entspricht), was eine Pixelgröße von $4,65\mu\text{m} \times 4,65\mu\text{m}$ ergibt.

Beide Kameras verfügen über einen 12Bit AD-Wandler und werden nicht aktiv gekühlt. Für alle Aufnahmen wurde die Sensitivität des Sensors minimal gewählt (also eine ggf. verfügbare zusätzliche analoge Signalverstärkung abgeschaltet) und die Gammakorrektur durch die Wahl von $\Gamma = 1$ deaktiviert. Vor den Messungen wurde bei abgedunkelter Kamera ein Korrekturbild aufgenommen, welches von jeder Aufnahme subtrahiert wurde (Dunkelstromkorrektur). Außerdem wurde der Schwarzpunkt manuell auf die Intensität eines dunklen Punktes im Bild gesetzt. Als Referenzmarke für das Schwarz diente das offene Ende eines einseitig geschlossenen, hohlen Zylinders, der mit seiner Rotationsachse auf das Objektiv der Kamera ausgerichtet wurde.

Bei der verwendeten Optik handelte es sich um das Objektiv M0814-MP der Firma COMPUTAR mit einer Brennweite von $f = 8\text{mm}$. Die Blendenzahl konnte zwischen $k = 1,4$ und $k = \infty$ variiert werden. Da das Objektiv über keine Skala verfügt, ließ sich die aktuelle Blendenzahl nicht abgelesen.

Die Verarbeitung der Bilddaten wurde in MATLAB, Version 7.0.0.19920 (R14), durchgeführt. Dazu wurde die Möglichkeit des Exports der Rohbilddaten als 16Bit TIFF-Datei genutzt, die von der Software beider Kameras geboten wurde. Der Import der TIFF-Dateien in Matlab erfolgte über die interne Funktion `imread`. Obwohl beide Kameras die AD-Wandlung nur mit einer Auflösung von 12Bit durchführen, nutzen die exportierten Bilddaten den vollen Dynamikbereich der TIFF-Dateien von 16Bit. Es findet also während der Verarbeitung durch die Kamera-Software eine Streckung der Intensitätswerte vom Intervall $(0, 2^{12} - 1)$ auf $(0, 2^{16} - 1)$ statt.

4.2 Voruntersuchungen

4.2.1 Bestimmung des maximalen Signal-Rausch-Verhältnisses

Um das Signal-Rausch-Verhältnis zu ermitteln, wurden Abbildungen von homogen beleuchtete Flächen aufgenommen und ausgewertet. Die Abbildung erfolgte dabei möglichst unscharf, sodass sich der eventuell vorhandene tatsächliche Kontrast der betrachteten Fläche vernachlässigen ließ. Ausgewertet wurde jeweils nur ein rechteckiger Ausschnitt der Abbildung, der eine Fläche von ca. 5000 Pixeln aufwies. Als Lichtquelle wurde ausschließlich mit Sonnenlicht gearbeitet, um die Messergebnisse nicht durch eine zusätzlich rauschende Lichtquelle zu verfälschen.

Als abgebildete Objekte kamen unterschiedliche Dinge zum Einsatz: Ein weißes Blatt Papier (Beleuchtung mit Auf- oder Durchlicht), die Oberfläche einer Milch-Wasser-Lösung, eine weiße Moosgummimatte und stark bewölkter sowie klarer Himmel.

Die Intensitätswerte der in dem betrachteten Ausschnitt liegenden Pixel wurden arithmetisch gemittelt (Mittelwert I_0) und es wurde die Standardabweichung σ berechnet. Als Signal-Rausch-Verhältnis ergab sich so $SNR_I = I_0/\sigma$. Bei dieser Methode wurde

vorausgesetzt, dass sich die unterschiedlichen Pixel in ihrer Charakteristik sehr stark ähneln und sich die zeitlichen Schwankungen der Intensitätswerte eines Pixels durch die Schwankungen der Intensitätswerte vieler Pixel, die identisch beleuchtet und gleichzeitig ausgelesen wurden, beschreiben lässt.

Die besten Werte für das Signal-Rausch-Verhältnis wurden für die Abbildung der Moosgummimatte ermittelt. Es betrug bei der Kamera PROGRES MF $SNR_{ProgResMF} = 95,3$ und bei der PIXELFLY $SNR_{pixelfly} = 115,3$.

4.2.2 Ermittlung der Parameter zur Verzeichnungskorrektur

Um die in Kapitel 3, Abschnitt 3.4 beschriebenen Auswirkungen einer nicht-idealen Optik korrigieren zu können, mussten die Verzeichnungsparameter des verwendeten Objektivs bestimmt werden. Aus den dort beschriebenen Gründen ist allein für die Tiefendatenbestimmung (unter Verwendung einer Marke) die Kenntnis des Parameters t_2 (siehe (3.38)) bereits ausreichend.

Um ihn zu bestimmen wurden mit der Kamera PROGRES MF unter Verwendung des zu korrigierenden Objektivs Aufnahmen von geraden Objekten im Raum gemacht und die Eigenschaft ausgenutzt, dass diese aufgrund der Distorsion im Bild gekrümmt erscheinen. Als Objekte dienten Holzstäbe, die in etwa einem Meter Abstand zur Kamera, parallel zur Bildebene und am Rand des abbildbaren Raumbereiches positioniert wurden.

Die erzeugten Grauwertbilder wurden in Matlab geladen und mithilfe der Umkehrfunktion von (3.38) wurde manuell der Wert $t_{2,Test}$ gesucht, der die Abbildungen der Stäbe wieder gerade erscheinen ließ. Als Durchstoßpunkt der optischen Achse wurde der Mittelpunkt des Bildes angenommen. Es wurde so

$$t_{2,ProgResMF} = (-1,8 \pm 0,1) \cdot 10^{-5} \cdot \frac{1}{\text{Pixel}^2}$$

ermittelt. Da die Pixel-Abmessungen beider verwendeten Kameras bekannt waren, ließ sich daraus der entsprechende Korrekturparameter für die PIXELFLY-Kamera berechnen:

$$t_{2,pixelfly} = t_{2,ProgResMF} \cdot \frac{4,65^2 \mu\text{m}^2}{6,45^2 \mu\text{m}^2} \cdot \frac{1}{\text{Pixel}^2} = (9,355 \pm 0,516) \cdot 10^{-6} \cdot \frac{1}{\text{Pixel}^2}$$

4.3 Aufbau zur Aufnahme geeigneter Grauwertbilder

Zur Durchführung des beschriebenen Verfahrens der Tiefendatengewinnung mussten mindestens zwei Bilder aus unterschiedlichen Raumpositionen aufgenommen werden. Dazu wurde die Kamera auf einen Translationsschlitten montiert, welcher auf einen stabilen Tisch geschraubt wurde. Der Translationsschlitten ermöglicht durch Drehung einer Stellschraube die Verschiebung der Kamera in Richtung der x -Achse des Kamerakoordinatensystems, sodass also die Abbildung antiparallel zur ξ -Achse des Bildkoordinatensystems verschoben wurde. Die Stellschraube verfügte über eine Skala zum Ablesen der Position. Ein Skalenteil entsprach einer Verschiebung der Kamera um $4\mu\text{m}$ - die Position der Kamera konnte also sehr genau manipuliert werden. Um Hystereseeffekte zu vermeiden, wurde bei den Messungen der Schlitten stets nur in eine Richtung verfahren.

4.3.1 Karten

Zur Untersuchung der Tiefenmessung wurden quadratische Karten mit einer Kantenlänge von 80mm verwendet, deren Oberfläche mit einem kontrastreichen Muster bedruckt war. Das Muster bestand aus schwarzen und weißen Quadraten mit einer Kantenlänge von 5mm, die in der Form eines Schachbretts angeordnet waren. Diese Karten wurden parallel zur Bildebene ausgerichtet und - abhängig vom jeweiligen Experiment - in verschiedenen Abständen z zur Kamera aufgestellt. Die letzte Karte diente jeweils als Marke und befand sich im Abstand $z_M = 1000\text{mm}$. Die Verschiebung der Kamera wurde bei den durgeführten Experimenten so gewählt, dass die minimale Messdistanz $z_{min} \approx 25\text{cm}$ betrug.

Da beide Kameras nicht kalibriert wurden und das Zentrum der Zentralprojektion somit nicht genau bekannt war, wurde dieses in einer Ebene ca. 10mm hinter der Seite des Kameragehäuses angenommen, in welcher sich die Objektivaufhängung befand. Die letzte Karte war also 990mm von der Vorderseite des Gehäuses entfernt.

4.4 Datenauswertung

Die aufgenommenen Grauwertbilder wurden in MATLAB durch die Funktion `imread` geladen und in einer Matrix hinterlegt. Mithilfe einer weißen Fläche, die in beiden Aufnahmen manuell markiert wurde, konnte die Veränderung der Helligkeit der Lichtquelle abgeschätzt werden. Der so ermittelte Korrekturfaktor h_{korr} ergab sich als Quotient der aufsummierten Intensitätswerte des markierten Bereichs beider Bildern. Durch Multiplikation der zweiten Aufnahme mit diesem Faktor standen nun zwei Bilder zur Verfügung, die bei scheinbar gleicher Helligkeit der Lichtquelle aufgenommen wurden.

Die Verarbeitung der Daten ist schematisch in Abbildung 4.1 dargestellt. Zur besseren Übersicht wurden alle Matrizen mit einem rechteckigen Rahmen versehen. Funktionen sind durch ein Oval gekennzeichnet worden und skalare Datentypen wurden nicht eingfasst. Der Datenfluss ist durch Pfeile gekennzeichnet worden.

Aus den korrigierten Grauwertbildern I wurde die sichtbare Verschiebung s und der zugehörige Fehler Δs errechnet. Die Berechnung von Δs erfordert die Festlegung der zu erwartenden Fehler der Intensitätsmessung. Um die spätere Analyse zu vereinfachen wurde hierfür ein unrealistisch kleiner Wert, nämlich eine Intensitätseinheit gewählt: $\Delta I = 1$. Da sich unter den experimentellen Bedingungen die äußeren Einflüsse recht gut kontrollieren ließen, wurden eventuelle zufällige Positionsänderungen oder Verkippen der Kamera als vernachlässigbar angenommen, also der in Kapitel 3, Abschnitt 3.3.2, diskutierte positionsbedingte Fehler der Intensitätsmessung $\Delta I_{Position} = 0$ gesetzt.

Abhängig von dem angenommenen Fehler der Intensitätserfassung ΔI und dem Kontrast des Grauwertbildes lässt sich eine Schwelle des maximalen Fehlers der sichtbaren Verschiebung Δs_{krit} definieren. Die Matrix der gefilterten Verschiebungswerte s_{flt} ergibt sich dann als Kopie der Matrix s , wobei alle Werte gelöscht werden, für deren Fehler $\Delta s \geq \Delta s_{krit}$ gilt. Die in Kapitel 3, Abschnitt 3.4 vorgeschlagene Distorsionskorrektur lässt sich an dieser Stelle einfach implementieren, indem elementweise durch die nume-

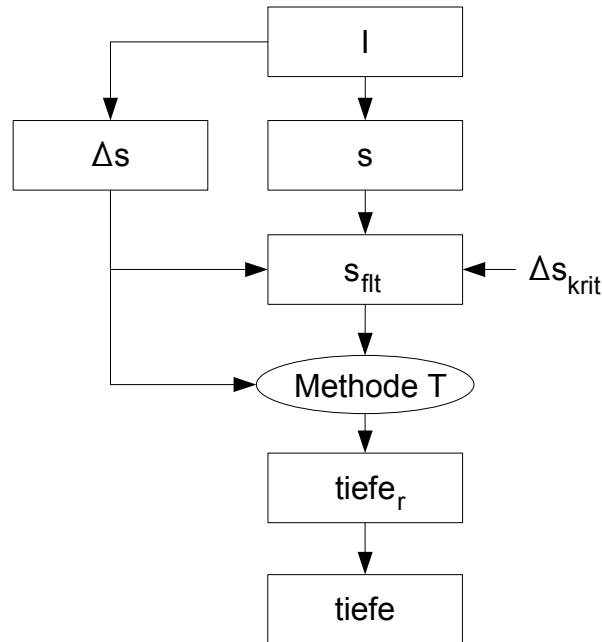


Abbildung 4.1: Schematische Darstellung der Datenverarbeitung

risch ermittelte Matrix v_ξ dividiert wird⁶ (siehe dazu Gleichung (3.39)). Ebenso lässt sich der im selben Abschnitt vorgeschlagene Filter zum Ausschluss kritischer, parallel zu ξ -Achse verlaufender Kanten (siehe (3.40)) günstig an dieser Stelle einsetzen. Bei den durchgeführten Experimenten zeigte sich allerdings, dass dieser Filter einen nicht messbaren Einfluss auf das generierte Tiefenbild ausübt, da die Verzerrung der Optik in Richtung der η -Achse selbst bei stark ausgeprägte Kanten einen sehr viel geringeren Fehler der Intensitätsmessung bewirkt, als die durch das Sensor- und Photonenrauschen verursachten Schwankungen.

In MATLAB-Notation ergibt sich für die Erzeugung von s_{flt} also folgender Code⁷:

```
shift_flt=(shift.*(dshift<dshift_krit))./v_xi.*~etaFilter;
```

Da die aufgenommenen Grauwertbilder in der Regel nicht über kontrastreiche Merkmale in ausreichender Dichte verfügen, ist eine räumliche Mittelung erforderlich. Hierzu wurde die Matrix in rechteckige „Kacheln“ unterteilt, aus deren Elementen die **Methode T** die Werte einer Matrix mit reduzierten Abmessungen generierte. Die Kantenlänge der Kacheln (in der Einheit „Pixel“) wurden mit $KachelX$ und $KachelY$ bezeichnet. Als **Methode T** wurden verschiedene Varianten zur Schätzung eines geeigneten Kachelwertes untersucht, z.B. die Bildung von Mittelwerten (arithmetisch, geometrisch, harmonisch), die Verwendung des Medians, gewichtete Methoden, die die Fehlermatrix Δs verwenden, sowie Schätzer, die den Mittelwert der am dichtest beieinander liegenden s -Werte wählen. Als genaueste Methode erwies sich die folgende: Die s -Werte einer Kachel werden ihrer

⁶Der zur Erzeugung von v_ξ verwendete Wert τ_0 sollte im Intervall $(0, 1)$ liegen.

⁷Im Quellcode wurden s und Δs zur besseren Lesbarkeit als „shift“ bzw. „dshift“ ausgeschrieben.

Größe nach sortiert. Nun werden die höchsten und niedrigsten 25% der Werte verworfen. Der arithmetische Mittelwert der verbleibenden Werte wird als Wert der betrachteten Kachel festgesetzt. Diese Operation lässt sich mit der MATLAB-internen Funktion `trimmean` durchführen.

Der Bereich der in den Grauwertbildern abgebildeten, als Entfernungsreferenz verwendeten Marke wurde bei den durchgeführten Versuchen manuell markiert. Nach (2.9) ergibt sich die Tiefe, also Entfernung z der Bildpunkte von der Kamera durch Division des Produktes von s_M und z_M durch die (gefilterte) sichtbare Verschiebung s_{flt} . Die direkte Ermittlung von s_M aus den Werten der Matrix s des markierten Bereichs erwies sich als ungünstig. Sehr viel einfacher zu realisieren und numerisch stabiler ist die Bestimmung der Tiefendaten mithilfe einer zusätzlichen Matrix $tiefe_r$, die unter der Annahme $s_M = 1$ und $z_M = 1$ analog zur tatsächlich gewünschten Tiefenmatrix aus s_{flt} erzeugt wurde.

```
tiefe_r=1./(double(glatt(shift_flt,KachelX,KachelY,'trimmean',50)));
```

Aus dem Verhältnis des arithmetischen Mittelwerts $Messung_{Marke}.mean$ dieser „Tiefen-Roh-Daten“ aus dem Bereich der Tiefenmarke zu dem Tiefen-Soll-Wert des Bereichs z_M kann so die Matrix der Tiefendaten $tiefe$ errechnet werden. Pixel, deren Tiefe als größer oder gleich 1500mm errechnet wurde, wurden als fehlerhaft angenommen und herausgefiltert:

```
tiefe=tiefe_r./Messung_Marke.mean.*z_M;
tiefe(tiefe > 1500)=0;
```

Aufgrund der räumlichen Mittelung umfasst das so erzeugte Tiefenbild weniger „Pixel“ als die verwendeten Grauwertbilder. Die Breite des Tiefenbildes errechnet sich aus der Breite eines Grauwertbildes, dividiert durch die Breite der Kacheln $KachelX$. Analog ergibt sich die Höhe des Tiefenbildes. Sollten bei dem Prozess der Kachelung aufgrund eines ungünstigen Verhältnisses zwischen Kachel- und Grauwertbildgröße am Rand nicht mehr genügend Pixel zur Verfügung stehen, um eine Kachel komplett zu füllen, so werden diese verworfen.

Die Erhöhung der Kachelgröße vergrößert die Genauigkeit der Tiefenbestimmung, wobei jedoch die laterale Auflösung reduziert wird. Die Auflösungsgrenze in Richtung der x - und y -Achse errechnen sich zu

$$\begin{aligned}\Delta x &= z \cdot KachelX \cdot \alpha_P \\ \Delta y &= z \cdot KachelY \cdot \alpha_P\end{aligned}$$

, wobei z die Entfernung der abgebildeten Region und α_P den Öffnungswinkel des auf ein Pixel, welches nahe bei der optischen Achse liegt, abgebildeten Raumbereiches darstellt.

Zur Durchführung einheitlicher Messungen wurden bei allen Versuchen die folgenden Parameter verwendet:

$$\begin{aligned}KachelX &= 10 \\ KachelY &= 10\end{aligned}$$

Dies entspricht in einem Abstand von $z = z_M = 1\text{m}$ einer lateralen Auflösung von $\Delta x = \Delta y = 6,6\text{mm}$.

Zur Beschreibung von „Pixel“-Positionen im erzeugten Tiefenbild sollen an dieser Stelle die *skalierten Pixelkoordinaten* i' und j' eingeführt werden. Das skalierte und nicht-skalierte (digitale) Koordinatensystem haben denselben Ursprung: Diejenige Kachel, in der das Pixel $(i, j) = (1, 1)$ liegt, besitzt die skalierten Koordinaten $(i', j') = (1, 1)$. Die Achsen des skalierten Koordinatensystems sind zu denen des nicht-skalierten Systems parallel.

4.5 Analyse der Genauigkeit der ermittelten Tiefenwerte

4.5.1 Untersuchung der Unabhängigkeit der Tiefenwerte von der Position der abgebildeten Objekte im Messraum

In einem ersten Test sollte die Frage geklärt werden, ob das untersuchte Verfahren tatsächlich Tiefenwerte liefert, die unabhängig von der Distanz des abgebildeten Objektpunktes zur optischen Achse, d.h. seiner Gegenstandshöhe und damit auch seiner Bildhöhe sind. Dazu wurde die Kamera PROGRES MF eingesetzt und die Messkarten wie in der Abbildung 4.2 gezeigt, aufgestellt: Die als Marke verwendete Karte befand sich in der Entfernung $z_M = 1000\text{mm}$ zur x - y -Ebene des Kamerakoordinatensystems und etwa in der Mitte der späteren Abbildung. Im Abstand $z = 500\text{mm}$ wurden weitere Messkarten so aufgestellt, dass auf einer Geraden parallel zur x -Achse das gesamte Sichtfeld der Kamera abgedeckt wurde. Alle Karten waren parallel zur x - y -Ebene ausgerichtet.

Bei schwach geöffneter Blende wurde das Objektiv auf ein etwa 4m entferntes Objekt fokussiert. Nun wurde die Blende soweit geöffnet, bis die Kanten der Tiefen-Marke auf etwas mehr als ein Pixel abgebildet wurden. Die Belichtungszeit wurde dann schrittweise soweit erhöht, bis die interessierenden Bildbereiche gerade noch nicht überbelichtet wurden.

Nun wurden zwei Grauwertbilder im Abstand von $x_a \approx 400\mu\text{m}$ bei einer Belichtungszeit von jeweils $t_{bel} = 32\text{ms}$ erzeugt und in MATLAB geladen. Ein solches Grauwertbild ist in Abbildung 4.3 gezeigt.

Die Erzeugung der Tiefendaten erfolgte auf die in Abschnitt 4.4 beschriebene Weise, wobei der dort erwähnte Parameter Δs_{krit} solange manuell verringert wurde, wie die Dichte der erzeugten Tiefendaten noch ausreichend hoch für die weitere Datenverarbeitung war. So wurde der Wert zu $\Delta s_{krit} = 0,0006$ festgesetzt.

Eine dreidimensionale Darstellung des errechneten Tiefenbildes ist in Abbildung 4.4 wiedergegeben. Hier wurde die Entfernung z (in mm) über den skalierten Pixelpositionen i' und j' aufgetragen. Neben den als Ebenen erscheinenden Messkarten sind diverse Spitzen zu erkennen, welche das Resultat einzelner Kanten in den erzeugten Grauwertaufnahmen sind (z.B. der im oberen Teil des Bildes 4.3 zu erkennenden Büroklammern).

Zur Beurteilung der Unabhängigkeit der bestimmten Tiefenwerte z von der Gegenstandshöhe G wurden die Tiefendaten des im Grauwertbild 4.3 den rechten Messkar-



Abbildung 4.2: Aufbau des Distorsionstests: Im Vordergrund ist die Kamera auf dem Translationsschlitten zu sehen. Dahinter sind die Messkarten, die weiße Fläche zur Helligkeitskorrektur und der als Schwarzreferenz dienende Zylinder zu erkennen.

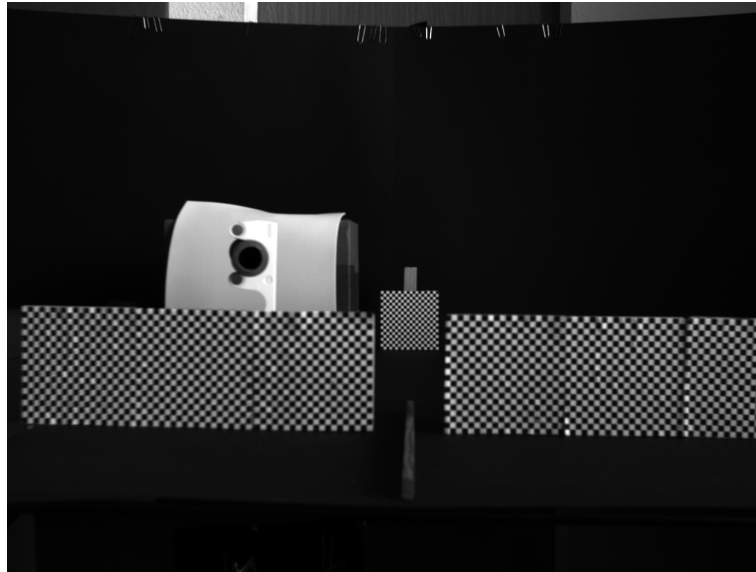


Abbildung 4.3: Grauwertbild zur Untersuchung der Unabhängigkeit der Tiefe z von der Gegenstandshöhe G .

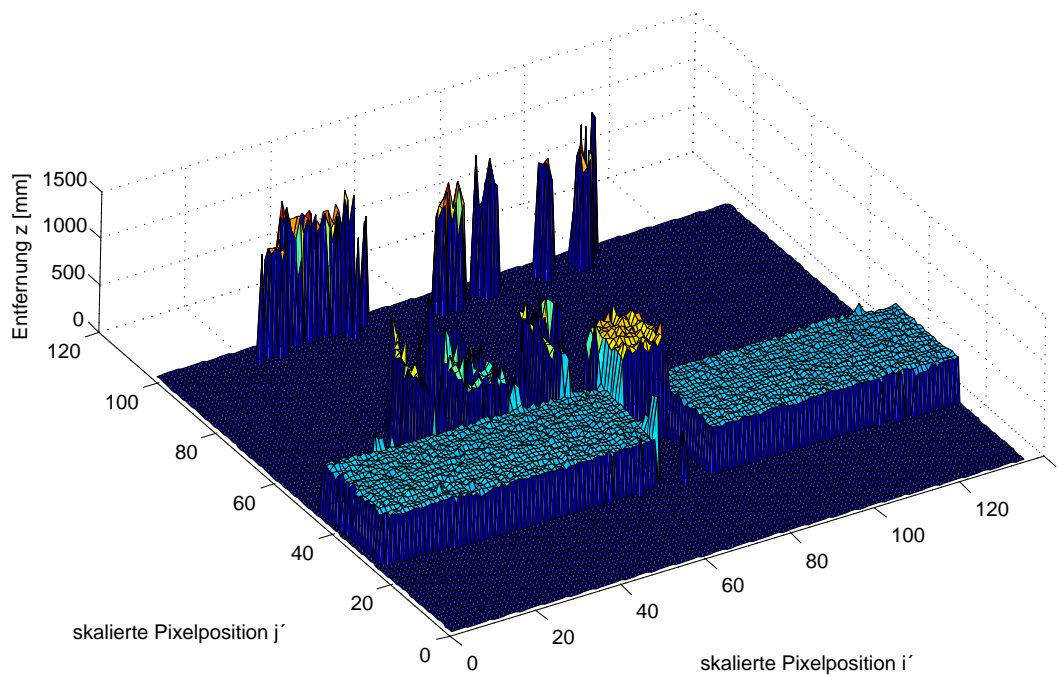


Abbildung 4.4: Dreidimensionale Darstellung der ermittelten Tiefenwerte z bei aktivierter Distorsionskorrektur. Deutlich zu erkennen sind die in einer Ebene liegenden Punkte der vermessenen Oberfläche, die die Bildregion der Tiefenmarke einschließen (Erhöhung im Zentrum der Abbildung). Die darüber erscheinenden Spitzen entsprechen vereinzelt Regionen im Grauwertbild, deren hoher lokaler Kontrast eine Tiefenbestimmung ermöglichte.

ten entsprechenden Bereichs isoliert betrachtet⁸. Diese Tiefendaten wurden entlang der j' -Achse arithmetisch gemittelt und über der i' -Position des betrachteten Ausschnitts aufgetragen. Die Berechnung erfolgte einmal mit und einmal ohne Distorsionskorrektur. Die so erzeugten Datensätze wurden durch eine lineare Funktion mit den Parametern $j' = p_1 \cdot i' + p_2$ angenähert und in einem gemeinsamen Diagramm (Abbildung 4.5) aufgetragen. Die Regressionsparameter und die Wurzel des mittlere Fehlerquadrat σ als Indikator für die Güte der Regression lauteten:

$$\begin{aligned} \text{Fit1: } & p_1 = -0,01272 \quad p_2 = 472,2 \quad \sigma = 2,439 \\ \text{Fit2: } & p_1 = 0,15410 \quad p_2 = 473,3 \quad \sigma = 2,450 \end{aligned}$$

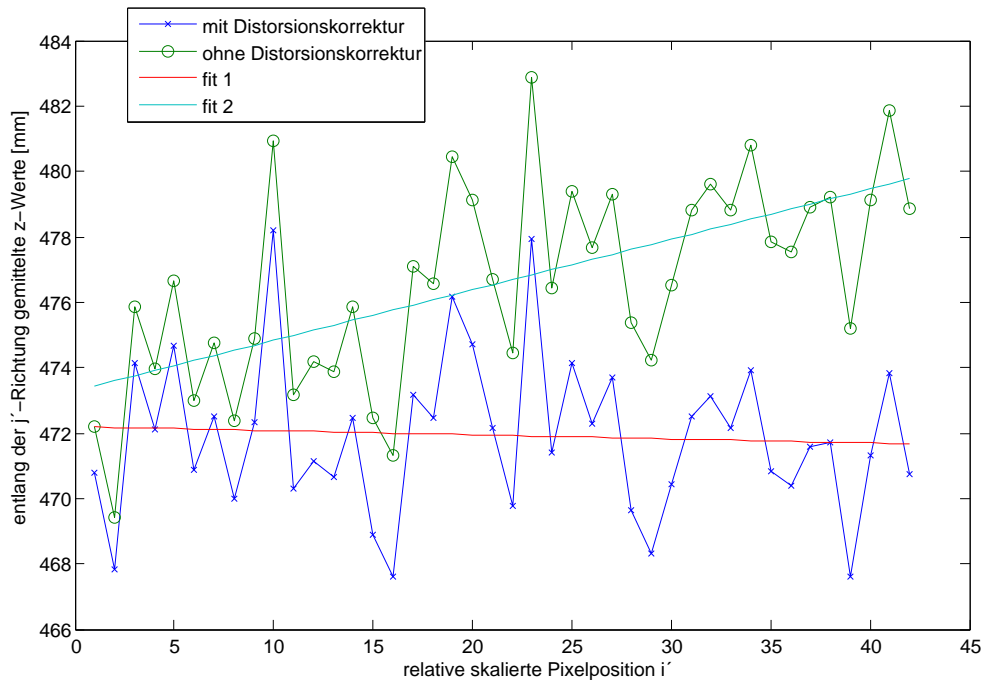


Abbildung 4.5: Entlang der j' -Achse gemittelte z -Werte aufgetragen über den skalierten Pixelpositionen i' des ausgewählten Ausschnitts und lineare Regressionen

Obwohl die Güte beider Näherungen nicht sehr hoch ist, lässt sich doch vergleichend aus dem geringeren Betrag des Wertes p_1 schließen, dass die Distorsionskorrektur zu einer Minderung der Abhängigkeit der ermittelten Tiefenwerte von der Gegenstandshöhe führt und somit eine Verbesserung bewirkt.

Der Effekt der Distorsionskorrektur lässt sich direkt beurteilen, indem man die Differenz der (j' -Richtung) gemittelten Tiefenwerte über der relativen skalierten Pixelposition i' aufträgt, wie dies in Abbildung 4.6 gezeigt wurde. Unabhängig von dem großen Fehler der Tiefenbestimmung lässt sich hier erkennen, dass die Distorsionskorrektur zu einer Verringerung des systematischen Fehlers der Tiefenbestimmung beiträgt, der etwa linear

⁸Die skalierten Koordinaten des Bereichs betragen: $i' = 80 \dots 122$ und $j' = 58 \dots 77$

mit der Bild- bzw. Gegenstandshöhe des abgebildeten Raumpunktes anwächst und am Rand des von der Kamera erfassten Bereichs ca. 8mm beträgt. An dieser Stelle muss betont werden, dass dieses Resultat nur für einen Abstand der zu messenden Objekte von $z = 500\text{mm}$ gültig ist. Eine Verallgemeinerung des systematischen Fehlers der Tiefenbestimmung bei deaktivierter Distorsionskorrektur ist auf Basis der aufgenommenen Messdaten nicht möglich. Es ist jedoch naheliegend, dass eine Berücksichtigung der objektivischen Verzeichnung eine Korrektur bewirkt, die im gesamten Messbereich in einer ähnlichen Größenordnung liegt.

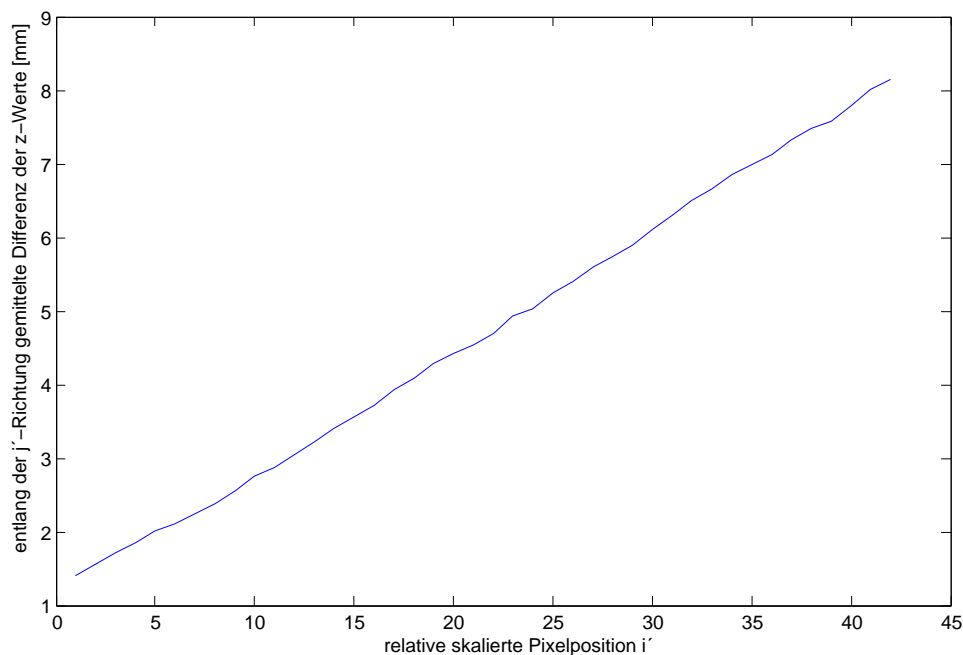


Abbildung 4.6: Differenz zwischen distorsionsunkorrigierten und -korrigierten, in j' -Richtung gemittelten Tiefenwerten z , abhängig von der relativen skalierten Pixelposition i'

4.5.2 Bestimmung der Tiefenauflösung Δz abhängig von z

Gegenstand dieses Versuchs war die Bestimmung der tatsächlich erzielbaren Tiefenauflösung Δz in Abhängigkeit von der Entfernung z des zu vermessenden Objektes. Dazu wurden die Messkarten parallel zur x - y -Ebene ausgerichtet und so angeordnet, dass sich die Entfernung z der Karten um ganzzahlige Vielfache von 100mm unterschied. Die letzte Karte wurde im Abstand $z = 1000\text{mm}$ aufgestellt und diente als Tiefenmarke. Sie stellte das Ende des Messbereichs dar: $z_{max} = z_M = 1000\text{mm}$. Die erste Karte befand sich einer Entfernung von 400mm zum Zentrum der Zentralprojektion. In Abbildung 4.7 ist die Aufstellung der Karten wiedergegeben.



Abbildung 4.7: Blick von oben auf den Aufbau der Messkarten für die Bestimmung der Tiefenauflösung Δz

Die Fokussierung des Objektivs und die Einstellung der Blende wurden wie bei dem in Abschnitt 4.5.1 beschriebenen Versuch durchgeführt: Es wurde auf ein ca. 4m entferntes Objekt fokussiert und die Blende soweit geöffnet, bis die Kanten der Tiefen-Marke auf etwas mehr als ein Pixel abgebildet wurden. Anschließend wurde die Belichtungszeit schrittweise erhöht, bis die erzeugten Aufnahmen die gewünschte Helligkeit aufwiesen.

Es wurden zwei Grauwertbilder im Abstand von $x_a \approx 400\mu\text{m}$ bei einer Belichtungszeit von jeweils $t_{bel} = 32\text{ms}$ aufgenommen und in MATLAB importiert. Anschließend wurden die Tiefendaten erneut nach der in Abschnitt 4.4 beschriebenen Methode generiert. Der Parameter Δs_{krit} wurde auch hier auf $\Delta s_{krit} = 0,0006$ festgesetzt.

Das so erzeugte Tiefenbild wurde in Abbildung 4.8 wiedergegeben. Es wurden die Bereiche der Messkarten markiert und die arithmetischen Mittelwerte sowie die Stan-

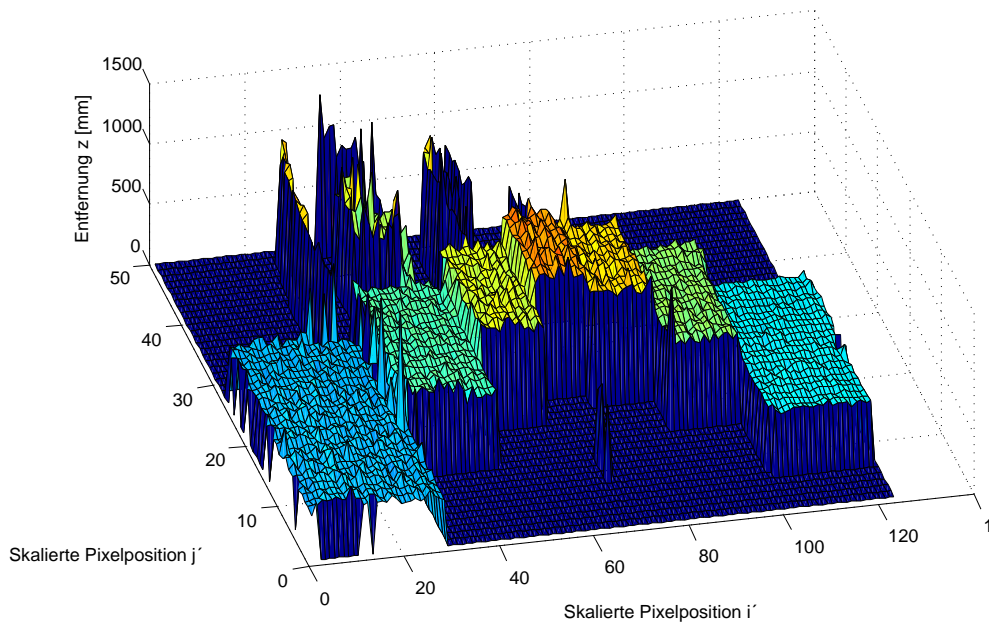


Abbildung 4.8: Dreidimensionale Darstellung des erzeugten Tiefenbildes. Die Entfernung der Bildregionen, die aufgrund unzureichenden Kontrastes keine Tiefenbestimmung ermöglichten, wurden auf $z = 0$ gesetzt.

Standardabweichung der Tiefendaten untersucht. Die ermittelte Standardabweichung Δz_{ist} kann als Maß für den zufälligen Fehler der Tiefenbestimmung verstanden werden, während die über die Messkarten räumlich gemittelten Tiefenwerte z_{ist} eine Abschätzung des systematischen Fehlers des Verfahrens ermöglichen. Die errechneten Standardabweichungen der Messwerte wurden zusammen mit den berechneten Mittelwerten in Abbildung 4.9 über dem Abstand z_{soll} der Messkarten aufgetragen. Diese Werte wurden in Tabelle 4.1 zusammengefasst. Zusätzlich wurde hier die Differenz zwischen den Ist- und Soll-Entfernungen ausgerechnet.

z_{soll} [mm]	1000,00	900,00	800,00	700,00	600,00	500,00	400,00
z_{ist} [mm]	1000,00	900,80	811,54	697,33	605,25	489,51	424,00
Δz_{ist} [mm]	44,11	32,66	32,70	19,62	17,94	11,42	24,71
$(z_{ist} - z_{soll})$ [mm]	0,00	0,80	11,54	-2,67	5,25	-10,49	24,00

Tabelle 4.1: Ermittelte Tiefenwerte z und Fehler Δz .

Bei Betrachtung der Resultate aus Tabelle 4.1 lässt sich feststellen, dass Abweichung der ermittelten Tiefenwerte von den Soll-Werten $z_{ist} - z_{soll}$ im gesamten Messbereich unter 24mm liegt. Im Bereich von $z = 1000 \dots 500$ mm ist diese Abweichung sogar geringer als 1,2% (bezogen auf die maximale Messdistanz z_{max}). Das Verschwinden der Abweichung gegen Ende des Messbereichs ist eine Folge der Marken-basierten Datenverarbeitung.

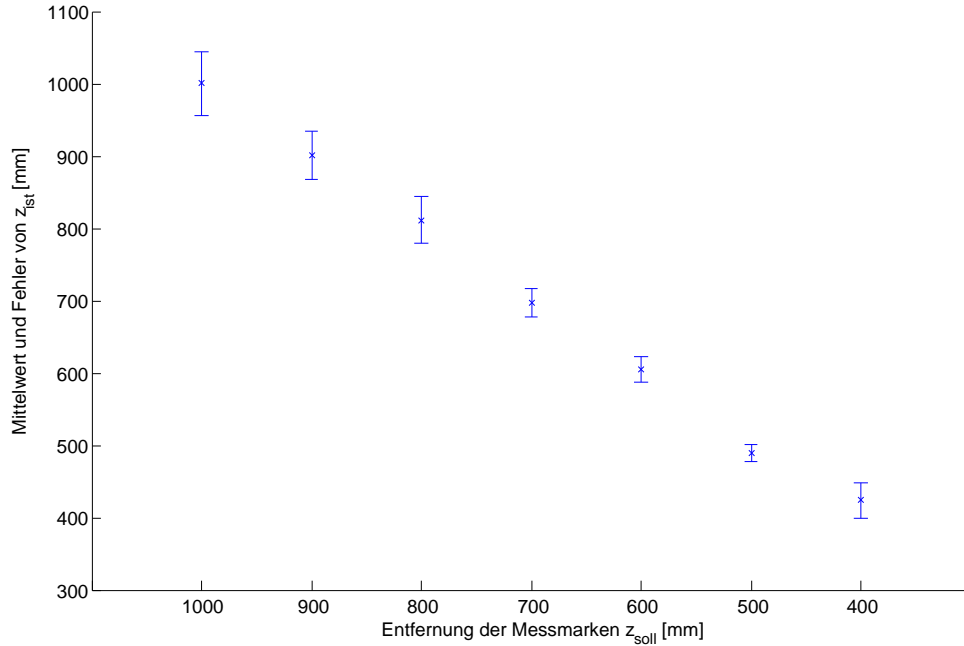


Abbildung 4.9: Mittelwert und Fehler der Tiefenwerte z_{ist} zu den im Abstand z_{soll} positionierten Messkarten.

Die zufälligen Fehler Δz_{ist} nehmen erwartungsgemäß mit Verringerung der Messdistanz z_{soll} ab und liegen im gesamten Messbereich unter $\Delta z < 44,11\text{mm}$. Wie in Abschnitt 4.4 erläutert, wurden die Parameter zur räumlichen Mittelung dabei so gewählt, dass die laterale Auflösung im gesamten Messbereich mindestens $\Delta x = \Delta y = 6,6\text{mm}$ betrug.

Der Anstieg des zufälligen Fehlers Δz_{ist} sowie der systematischen Abweichungen $z_{ist} - z_{soll}$ für die am dichtesten bei der Kamera positionierten Messkarte ist wahrscheinlich durch eine ungünstige, d. h. inhomogene Beleuchtung der Karten zu begründen. Allerdings ist auch ein Zusammenhang mit der bei Verringerung der Messentfernung stark sinkenden Schärfe der abgebildeten Kanten nicht ausgeschlossen.

4.5.3 Untersuchung der Abhängigkeit des Fehlers der Tiefenbestimmung Δz vom Signal-Rausch-Verhältnis der Kamera SNR_I

In Kapitel 3, Abschnitt 3.2.4, wurde durch (3.21) postuliert, dass sich der Fehler der Tiefenbestimmung Δz antiproportional zum Signal-Rausch-Verhältnis der Kamera SNR_I verhält:

$$\Delta z \propto \frac{1}{l} \propto \frac{1}{SNR_I} \quad (4.1)$$

Das Signal-Rausch-Verhältnis der Kamera lässt sich in der Regel nicht direkt manipulieren. Unter der Annahme, dass die Schwankungen der Intensitätsmessung durch das

Photonenrauschen dominiert werden (siehe Abschnitt 3.1.2 desselben Kapitels), und dieses somit zu

$$SNR_I = \sqrt{\text{full well capacity}} \quad (4.2)$$

abschätzen lässt, kommen aber Möglichkeiten in Betracht, die auf einer „künstlichen Erhöhung“ der full-well-capacity beruhen. Denkbar wäre hier z. B. das Zusammenfassen mehrerer Pixel zu größeren Gruppen und der Analyse der über diese Gruppen aufsummierten Intensitäten. Dieser als *binning* bezeichnete Arbeitsmodus der Kamera war für die geplanten Experimente zur Untersuchung der Genauigkeit der Tiefenbestimmung nicht geeignet, da auch bei der Nutzung größerer „Pixel“ eine ausreichend unscharfe Abbildung sichergestellt sein musste (siehe Kapitel 3, Sektion 3.2.4, Unschärfebedingung (3.11)). Hier wurden die Grenzen der möglichen Defokussierung des verwendeten Objektivs erreicht.

Eine Alternative zu dieser räumlichen Mittelung der Grauwertbilder stellt die zeitliche Mittelung dar. Hierbei werden n Aufnahmen nacheinander erzeugt und aufaddiert. Die Summe dieser Grauwertbilder verhält sich ähnlich einer Aufnahme, die bei n -fach höherer full-well-capacity generiert wurde. Da jeder einzelne Auslesevorgang Ausleserauschen erzeugt und jede Einzelmessung mit Dunkelstromrauschen behaftet ist, steigt auch deren absoluter Beitrag zum Intensitätsrauschen. Für das aufsummierte Grauwertbild gilt jedoch weiterhin, dass es durch das Photonенrauschen dominiert und deshalb durch den Faktor $n \cdot \text{full well capacity}$ maßgeblich beeinflusst wird. Für die zeitlich gemittelten Messungen lässt sich so ein „synthetisches Signal-Rausch-Verhältnis“ von

$$SNR_{I, syn} = \sqrt{n \cdot \text{full well capacity}}$$

definieren.

Die Kamera PIXELFLY bietet die Möglichkeit, intern und damit sehr schnell solche aufsummierten Intensitätsbilder zu erzeugen. Deshalb wurde die PIXELFLY-Kamera analog zu der in Abschnitt 4.5.2 beschriebenen Messung positioniert und eingestellt. Die um jeweils 10cm versetzte Aufstellung der Messkarten entsprach ebenfalls dem dort beschriebenen Aufbau.

Nun wurden zeitlich gemittelte Grauwertbilder im Abstand von jeweils etwa $x_a \approx 400\mu\text{m}$ erzeugt und mit MATLAB ausgewertet. Die Belichtungszeit für ein einzelnes Bild betrug 16ms - die Aufnahmezeit der gemittelten Bilder ergab sich also etwa als das n -fache diesen Wertes:

$$t_{bel} = n \cdot 16\text{ms}$$

Die Erzeugung der Tiefendaten erfolgte auf die in Abschnitt 4.4 beschriebene Weise, wobei wieder $\Delta s_{krit} = 0,0006$ gewählt wurde.

In den so generierten Tiefenbild wurden die zu den einzelnen Messkarten gehörenden Bereiche erneut manuell markiert und deren Tiefenwerte analysiert. Die Standardabweichungen der ermittelten Tiefendaten als Maß für den zufälligen Fehler der Tiefenbestimmung wurden in Tabelle 4.2 abhängig von der Anzahl der durchgeführten Grauwertmittlungen n sowie der Entfernung der jeweiligen Messkarte z_{soll} aufgetragen.

Anzahl der Mittelungen n	z_{soll} [mm]						
	1000,00	900,00	800,00	700,00	600,00	500,00	400,00
1	39,46	34,21	35,90	21,42	21,76	12,21	31,42
4	31,55	31,32	26,30	11,97	11,15	6,02	8,33
16	31,43	27,73	21,98	10,56	7,12	4,89	3,75
64	29,09	32,06	26,35	9,38	9,14	9,95	4,09

Tabelle 4.2: Fehler der Tiefenbestimmung Δz abh. von der Anzahl der erfolgten Grauwert-Mittelungen n und der Entfernung der abgebildeten Messkarte von der Kamera z_{soll}

Die erwartete Abnahme des zufälligen Fehlers Δz_{soll} mit zunehmender Anzahl der Mittelungen n sowie abnehmender Entfernung z_{ist} der Messkarten von der Kamera wird von verschiedenen Störeffekten überlagert: Die Abbildung der als Messmarke verwendete Karte im Abstand $z_M = 1000\text{mm}$ ist möglicherweise zu scharf, sodass die Anforderung an die minimale Unschärfe der abgebildeten Kante (siehe Kapitel 3, Abschnitt 3.2.4, (3.11)) nicht erfüllt wurde. Die Messungen der im Abstand $z_{ist} = 400\text{mm}$ positionierten Karte waren stets mit einem sehr hohen Fehler belastet, der die in Abschnitt 4.5.2 erwähnte Vermutung, dies könne im Zusammenhang mit der stark defokussierten Abbildung der Kanten stehen, bestärkt. Die starke Defokussierung bewirkt einen verminderten Kontrast der abgebildeten Oberfläche, welcher sich in einer Verringerung der Tiefenaufösung auswirkt.

Um den erwarteten antiproportionalen Zusammenhang zwischen dem aufgrund der zeitlichen Intensitätsmessung erhöhten synthetische Signal-Rausch-Verhältnisses $SNR_{I, syn}$ und dem zufälligen Fehler der Tiefenmessung (siehe (4.1)) zu visualisieren, wurden in Abbildung 4.10 die Fehler Δz_{ist} der Tiefenwerte jeder Karte über dem reziproken Signal-Rausch-Verhältnis $1/SNR_{I, syn}$ aufgetragen. In dieser Darstellung sollten sich der antiproportionale Zusammenhang der beiden Größen dadurch zeigen, dass die eingezeichneten Punkte auf Geraden liegen. Bei Betrachtung der entsprechenden Abbildung ist dieser lineare Zusammenhang erkennbar. Auffällig ist, dass alle Punkte, die über $1/SNR_{I, syn} = 4,385 \cdot 10^{-3}$ aufgetragen wurden, die jeweiligen linearen Verläufe deutlich unterbrechen. Diese Messpunkte entsprechen der maximalen Mittelung von $n = 64$. Zeitlich inhomogene Eigenschaften der Lichtquelle (vor der Sonne vorbeiziehende Wolken), die aufgrund der deutlich erhöhten Aufnahmedauer der verwendeten Grauwertbilder hier stärker ins Gewicht fielen, scheinen die Ursache für diese Unregelmäßigkeiten zu sein.

4.6 Geschwindigkeit

Ziel des vorgestellten Verfahrens war neben einer möglichst hohen Auflösung in allen drei Dimensionen die Gewinnung von 3D-Daten in sehr kurzer Zeit. Zur Untersuchung der Geschwindigkeit der verwendeten MATLAB-Implementierung wurde der in Abschnitt 4.5.2 beschriebene Versuch wiederholt und die Laufzeit jeder der an der Tiefenberechnung beteiligten Komponenten bestimmt. Hierzu wurde die MATLAB-interne Funktion

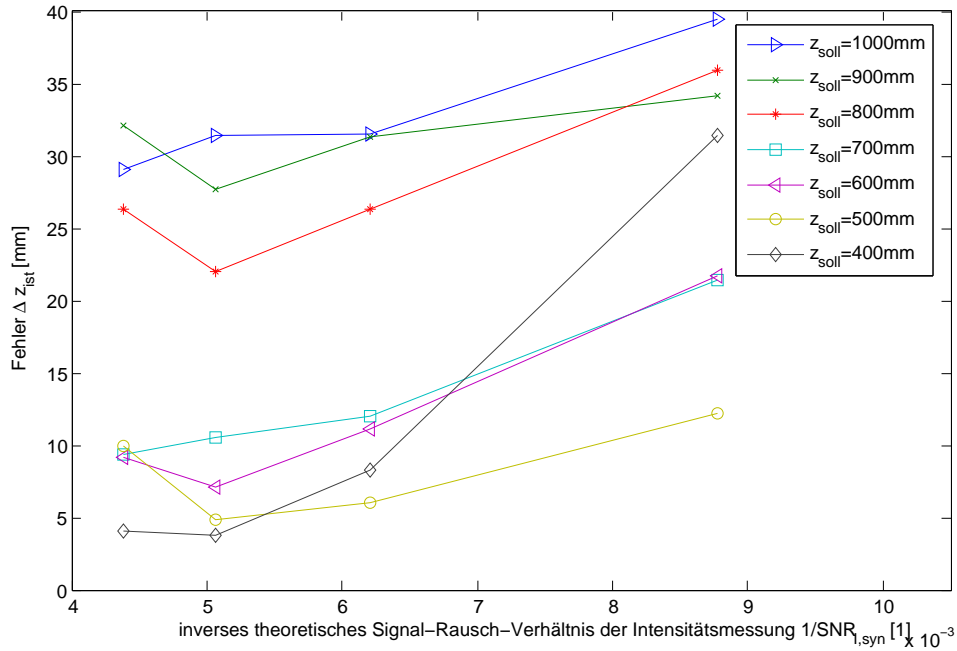


Abbildung 4.10: Fehler der Tiefenwerte Δz_{ist} abh. vom inversen Signal-Rausch-Verhältnis der Kamera für die in verschiedenen Abständen z_{soll} positionierten Messkarten

cputime zum Auslesen der Systemzeit verwendet. Der verwendete Computer war ein Laptop mit einem 1,4GHz INTEL PENTIUM M-Prozessor und 992MB Arbeitsspeicher, der unter WINDOWS XP lief.

Bei dem in Abschnitt 4.5.2 beschriebenen Versuch wurde zur Berechnung der Tiefenmatrix nur ein Ausschnitt der Grauwertbilder verwendet, der $503 \times 1243\text{Pixel}$ umfasste. Die beiden verwendeten Rohbilder hatten also jeweils eine Fläche von $n_{Pixel} = 625229\text{Pixel}^2$.

Vor Durchführung der Tiefenberechnung mussten die Matrizen v_ξ und v_η berechnet werden. Diese Initialisierung ist nur einmal vor Beginn der Tiefenbestimmung notwendig. Da die Resultate unabhängig von den zu verarbeitenden Grauwertbildern sind und somit zur Laufzeit bereits im Speicher hinterlegt sein können, wurde diese Zeit zum Initialisieren nicht zur die Gesamtarbeitszeit der Tiefenbestimmung hinzugezählt. Die Dauer zum Berechnung der Korrekturmatrizen betrug $t_{Distorsion} = 14,560938\text{s}$.

Die Gesamtzeit der Tiefenbestimmung betrug $t_{Gesamt} = 8,311952\text{s}$. Sie setzte sich aus den Laufzeiten ihrer Komponenten zusammen, die in Tabelle 4.3 wiedergegeben wurden.

Alle Rechenschritte besitzen lineare Komplexität - eine Verdoppelung der zu verarbeitenden Grauwertpixel benötigt somit die doppelte Zeit. Aus der Gesamtzeit und der Anzahl der verarbeiteten Grauwerte lässt sich die Verarbeitungsgeschwindigkeit zu 75221Pixel/s errechnen. Die Geschwindigkeit zum Initialisieren der Matrizen zur Distorsionskorrektur betrug 42941Pixel/s .

Da der für die Berechnungen verwendete Quellcode in Hinblick auf eine gute Lesbarkeit und Flexibilität geschrieben wurde, kann man die gemessenen Laufzeiten keinesfalls

Komponente	Laufzeit [s]
Laden der Grauwertbilder (von der Festplatte)	1,602304
Durchführen der Helligkeitskorrektur	0,070101
Berechnung $shift$ und $dshift$	4,226077
Erstellen der Matrix s_{fit} und Durchführung der Distorsionskorrektur	0,140202
Räumliche Mittelung und Errechnung des Tiefenbildes $tiefe$	2,273269

Tabelle 4.3: Laufzeit der Komponenten der Tiefenbestimmung

als Geschwindigkeitsoptimum verstehen. Ein erheblicher Geschwindigkeitszuwachs ließe sich zunächst durch die Implementierung der Algorithmen in einer effizienteren Programmiersprache erzielen. Aber auch die folgenden Optimierungsoptionen wurden bisher nicht genutzt:

- Laden der Grauwertbilder über Schnittstellen direkt von der Kamera
- Einführen von Abbruchbedingungen - bisher wird die Berechnung der Verschiebung s für alle Pixel durchgeführt, eine frühzeitige Einschränkung auf die kontrastreichen Regionen der Grauwertbilder wäre sinnvoll.
- Nutzung spezieller Prozessor-Befehle zum Durchführen mehrerer Einzeloperationen pro Takt
- Parallelisierung für mehrere Prozessorkerne usw.

Die Berechnung des Tiefenbildes umfasst ohne Berücksichtigung der Initialisierung der Distorsionskorrekturmatrizen sowie des Ladens der Bilddaten von der Festplatte etwa 120 elementare Operationen pro Pixel. Theoretisch lassen sich somit bereits auf einem 1GHz-System bei Verwendung einer 1MP-Kamera etwa 8,3 Bilder pro Sekunde prozessieren. In Kombination mit einer geeigneten Optik, die die parallele Erfassung der Grauwertbilder ermöglicht, wäre mit dieser Methode also die Erfassung von Tiefenbildern in Echtzeit auf Standard-Hardware möglich.

5 Zusammenfassung und Vergleich

Es konnte gezeigt werden, dass das vorgestellte Verfahren die Gewinnung von Tiefenbildern unter sehr geringem Rechenaufwand ermöglicht. Da im Gegensatz zu konventionellen stereoskopischen Verfahren kein Korrespondenzproblem auftritt und das Verfahren somit nur lineare Komplexität besitzt, skaliert es stets besser als alle konventionellen stereoskopischen Ansätze. Weil die Informationen zur Entfernungsbestimmung allerdings direkt aus den gemessenen Intensitäten errechnet werden, ist die Verwendung einer Kamera mit hoher Intensitätsauflösung und das Vorhandensein kontrastreicher Oberflächen erforderlich. Für die Vermessung homogener Flächen ist die Projektion von geeigneten Mustern notwendig. Da die Aufgabe des Musters nur darin besteht, den Kontrast der abgebildeten Oberflächen zu erhöhen und nicht - wie bei den aktiven triangulatorischen Verfahren mit einer Kamera - eine Zuordnung von Projektorstrahlen zu den Sehstrahlen des Aufnahmesystems erforderlich ist, ist bereits ein einziges hochfrequentes Muster ausreichend. Streifen etwa, die senkrecht zur Verschiebungsrichtung der Kamera abgebildet werden, würden ein geeignetes Muster darstellen.

Die derzeit von kommerziell erhältlichen flugzeitbasierten Erfassungssystemen für Tiefenbilder (TOF-Systeme) erreichten räumlichen Auflösungen liegen in einer mit der des vorgestellten Verfahrens vergleichbarer Größenordnung. Im Gegensatz zu den TOF-Systemen reagiert die vorgestellte Methode jedoch viel robuster auf Fremdlicht. Weiterhin sollte sich die Geschwindigkeit der TOF-Systeme von derzeit 25fps bereits mit aktuellen Desktop-Prozessoren¹ übertreffen lassen (siehe [MES08] oder [OLK⁺04]).

Dennoch muss betont werden, dass das vorgestellte Messverfahren noch keine praktikable Lösung für den Einsatz in Werkzeugmaschinen mit dem Ziel der Kollisionsvermeidung darstellt. Das Versehen der zu vermessenden Objekte mit kontrastreichen Oberflächen wird sich in einer realen Umgebung kaum umsetzen lassen. Daher kommen nur projektionsbasierte Varianten ernsthaft in Frage, die sicherlich nur in Ausnahmefällen eine optimale Ausleuchtung mit idealen Oberflächenkontrasten erreichen werden. Im praktischen Einsatz ist somit davon auszugehen, dass sich die Tiefenauflösung gegenüber der experimentell ermittelten etwa um den Faktor 2 verringert. Zur Gewinnung von Tiefenbildern mit höherer Auflösung sollten Kameras mit höherer full-well-capacity eingesetzt werden. Außerdem ist bei Sicherstellung einer ausreichenden Ausleuchtung des Messraums die in Kapitel 4, Abschnitt 4.5.3, vorgeschlagene künstliche Erhöhung der full-well-capacity durch zeitliches Mitteln mehrerer Grauwertbilder denkbar.

Nach (3.21) aus Kapitel 3, Abschnitt 3.2.4, und (2.5) sowie (2.8) aus Kapitel 2, Abschnitt 2.3, verringert sich der Fehler der Tiefenbestimmung Δz bei Vergrößerung der

¹Bei Verwendung eines Prozessors mit 3GHz Taktfrequenz und 4 Prozessorkernen sollte sich die in Kapitel 4, Abschnitt 4.6, abgeschätzte Geschwindigkeit von 8,3fps um den Faktor 12 auf ≈ 100 fps steigern lassen.

Kameraverschiebung x_a , wobei sich jedoch auch der Messbereich einschränkt, da z_{min} vergrößert wird. Falls eine längere Zeit zur Aufnahme der Daten und Prozessierung tolerierbar ist, kann auch eine mehrstufige Bilddatenerfassung in Erwägung gezogen werden. Hier würden drei oder mehr Grauwertbilder in unterschiedlichen Positionen erfasst, die Tiefendaten aus jeweils zwei Bildern errechnet und anschließend kombiniert werden.

Die vorgestellte Tiefenbildgewinnung durch minimale Verschiebung einer Kamera liefert nur korrekte Daten, falls sich die zu vermessende Szenerie im Zeitraum der Bilddatenerfassung nicht ändert. Bewegliche Objekte oder auch Schatten von in der Nähe befindlichen Personen oder Körpern können die Messung massiv stören. Für den Einsatz des Verfahrens in Werkzeugmaschinen sollte daher eine optische Isolation zur Umgebung sichergestellt werden. Viele Werkzeugmaschinen verwenden ein Schutzgehäuse zur mechanischen Abgrenzung des Arbeitsraums, welches bereits eine geeignete Abgrenzung darstellen würde. Falls das Gehäuse über ein Fenster zur einfacheren Kontrolle der Vorgänge innerhalb der Maschine durch das Personal verfügt, bestünde ein möglicher Weg die optische Isolation dennoch sicherzustellen darin, mit der optischen Messung in einen anderen Wellenlängenbereich auszuweichen (z. B. nahes Infrarot) und das Fenster des Arbeitsbereichs z. B. durch einen Folienfilter für diesen Wellenlängenbereich unpassierbar zu machen.

Zusammenfassend lässt sich also festhalten, dass das vorgestellte Verfahren das Potenzial besitzt, zur optischen Raumdatenerfassung in Werkzeugmaschinen eingesetzt werden zu können. Die stetige Steigerung der verfügbaren Kameras bezüglich Empfindlichkeit und Auflösungsvermögen wird das Erreichen der angestrebten räumlichen Genauigkeitsgrenzen nach der vorgestellten Methode in naher Zukunft ermöglichen.

Literaturverzeichnis

- [Bla03] BLAIS, F.: Review of 20 years of range sensor development. In: EL-HAKIM, S. F. (Hrsg.) ; GRUEN, A. (Hrsg.) ; WALTON, J. S. (Hrsg.): *Videometrics VII. Edited by El-Hakim, Sabry F.; Gruen, Armin; Walton, James S. Proceedings of the SPIE, Volume 5013, pp. 62-76 (2003)*. Bd. 5013, 2003 (Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference), S. 62–76
- [DH04] DUSSAULT, D. ; HOESS, P.: Noise performance comparison of ICCD with CCD and EMCCD cameras. In: DERENIAK, E. L. (Hrsg.) ; SAMPSON, R. E. (Hrsg.) ; JOHNSON, C. B. (Hrsg.): *Image Reconstruction from Incomplete Data III. Edited by Bones, Philip J.; Fiddy, Michael A.; Millane, Rick P. Proceedings of the SPIE, Volume 5563, pp. 195-204 (2004)*. Bd. 5563, 2004 (Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference), S. 195–204
- [FP03] FORSYTH ; PONCE: *Computer Vision - A modern approach*. U.S.A. : Pearson Education, 2003
- [GYB04] GOKTURK, S. B. ; YALCIN, Hakan ; BAMJI, Cyrus: A Time-Of-Flight Depth Sensor - System Description, Issues and Solutions. In: *CVPRW '04: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Volume 3*. Washington, DC, USA : IEEE Computer Society, 2004. – ISBN 0-7695-2158-4, S. 35
- [MES08] MESA IMAGING AG: *SR-3000 Data Sheet*. Version: Jan 2008, 2008. http://www.mesa-imaging.ch/pdf/SR3000_Flyer_Jan07.pdf, Abruf: 28.01.2008
- [MK98] MARTINEZ, Patrick ; KLOTZ, Alain: *A practical guide to ccd astronomy*. Cambridge : Cambridge University Press, 1998
- [OLK⁺04] OGGIER, T. ; LEHMANN, M. ; KAUFMANN, R. ; SCHWEIZER, M. ; RICHTER, M. ; METZLER, P. ; LANG, G. ; LUSTENBERGER, F. ; BLANC, N.: An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger). In: MAZURAY, L. (Hrsg.) ; ROGERS, P. J. (Hrsg.) ; WARTMANN, R. (Hrsg.): *Optical Design and Engineering. Edited by Mazuray, Laurent; Rogers, Philip J.; Wartmann, Rolf. Proceedings of the SPIE, Volume 5249, pp. 534-545 (2004)*. Bd. 5249, 2004 (Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference), S. 534–545

- [TKK⁺06] TORNOW, M. ; KASZUBIAK, J. ; KUHN, R. W. ; MICHAELIS, B. ; KRELL, G.: *Stereophotogrammetric Real-Time 3D Machine Vision*. Otto-von-Guericke University Magdeburg, 2006. <http://dx.doi.org/10.1134/S1054661806010317>
- [VNS95] VOSS, Klaus ; NEUBAUER, Ralf ; SCHUBERT, Michael: *Monokulare Rekonstruktion für Robotvision*. Aachen : Verlag Shaker, 1995

Abbildungsverzeichnis

1.1	Abbildung eines Punktes P in P' mit Hilfe einer dünnen Linse	7
1.2	Kamerakoordinatensystem und Bildkoordinatensystem	8
1.3	Bildkoordinatensystem und digitales Koordinatensystem.	8
2.1	Schematische Darstellung des Messaufbaus.	13
2.2	Zur Bestimmung der Verschiebung s	14
2.3	Zur Berechnung der Entfernung z aus der sichtbaren Verschiebung s . . .	15
3.1	Skizze zur Herleitung der Breite B_d einer abgebildeten Kante	21
3.2	Errechner Fehler Δz der Entfernungsbestimmung abhängig von der Entfernung z	24
3.3	Auswirkung der Veränderung der Lage eines Punktes in Richtung der y -Achse auf das Bild des Punktes bei scharfer Abbildung	26
3.4	Auswirkung der Veränderung der Lage eines Punktes in Richtung der z -Achse auf das Bild des Punktes bei scharfer Abbildung	27
3.5	Veränderung des Ortes der Abbildung aufgrund der Verkippung der Kamera um die x -Achse	28
4.1	Schematische Darstellung der Datenverarbeitung	39
4.2	Aufbau des Distorsionstests	42
4.3	Grauwertbild zur Untersuchung der Unabhängigkeit der Tiefe z von der Gegenstandshöhe G	43
4.4	Dreidimensionale Darstellung der ermittelten Tiefenwerte z bei aktivierter Distorsionskorrektur.	43
4.5	Entlang der j' -Achse gemittelte z -Werte aufgetragen über den skalierten Pixelpositionen i' des ausgewählten Ausschnitts und lineare Regressionen	44
4.6	Differenz zwischen distorsionsunkorrigierten und -korrigierten Tiefenwerten z	45
4.7	Blick von oben auf den Aufbau der Messkarten für die Bestimmung der Tiefenauflösung Δz	46
4.8	Dreidimensionale Darstellung des erzeugten Tiefenbildes.	47
4.9	Mittelwert und Fehler der Tiefenwerte z_{ist} zu den im Abstand z_{soll} positionierten Messkarten.	48
4.10	Fehler der Tiefenwerte Δz_{ist} abh. vom inversen Signal-Rausch-Verhältnis der Kamera für die in verschiedenen Abständen z_{soll} positionierten Messkarten	51

Tabellenverzeichnis

4.1	Ermittelte Tiefenwerte z und Fehler Δz	47
4.2	Fehler der Tiefenbestimmung Δz abh. von der Anzahl der erfolgten Grauwert-Mittelungen n	50
4.3	Laufzeit der Komponenten der Tiefenbestimmung	52

Erklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

.....
Ort, Abgabedatum

.....
Unterschrift des Verfassers

Seitens des Verfassers/Verfasserin bestehen keine Einwände, die vorliegende Diplomarbeit für die öffentliche Nutzung in der Thüringer Universitäts- und Landesbibliothek zur Verfügung zu stellen.

Jena, den 29. Oktober 2008

.....
Unterschrift