

**Jan Böhm**

**Modellbasierte Segmentierung und Objekterkennung  
aus Distanzbildern**

**München 2005**

---

Verlag der Bayerischen Akademie der Wissenschaften  
in Kommission beim Verlag C. H. Beck



Modellbasierte Segmentierung und Objekterkennung  
aus Distanzbildern

Von der Fakultät für Luft- und Raumfahrttechnik und Geodäsie  
der Universität Stuttgart  
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)  
genehmigten Dissertation

vorgelegt von

Dipl.-Inform. M.Sc. Jan Böhm

aus Stuttgart

München 2005

---

Verlag der Bayerischen Akademie der Wissenschaften  
in Kommission beim Verlag C. H. Beck

Adresse der Deutschen Geodätischen Kommission:

**Deutsche Geodätische Kommission**

Marstallplatz 8 • D – 80 539 München

Telefon +49 - (0)89 - 23 031 -0 / -1113 • Telefax +49 - (0)89 - 23 031 -1283 / -1100

E-mail [hornik@dgfi.badw.de](mailto:hornik@dgfi.badw.de) • <http://dgk.badw.de>

Hauptberichter: Prof. Dr.-Ing.habil. Dieter Fritsch

Mitberichter: Prof. Dr.rer.nat.habil. Paul Levi

Tag der Einreichung: 22.01.2004

Tag der mündlichen Prüfung: 22.12.2004

---

© 2005 Deutsche Geodätische Kommission, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,  
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen

---

## Kurzfassung

Die Verfügbarkeit immer leistungsfähigerer dreidimensionaler Sensorsysteme erfordert die Entwicklung entsprechender Systeme zur automatisierten Verarbeitung der aufgenommenen 3D-Daten. Im Gegensatz zu den herkömmlichen Intensitätsbildern, die nur die Helligkeitsinformation der Oberfläche eines Objektes enthalten, sind in einem Distanzbild die Abstände der einzelnen Punkte einer Objektoberfläche zum Sensor enthalten. Für die Verarbeitung dieser Daten auf hoher Abstraktionsebene müssen neue Konzepte entwickelt werden, die über die Verfahren der zweidimensionalen Bildanalyse hinausgehen. Zur Entwicklung automatisierter Systeme im dreidimensionalen maschinellen Sehen stehen für diese Arbeit zwei Aspekte im Vordergrund, die Zerlegung der Distanzbilder und die Herstellung der Korrespondenz zu einem Modell.

Die vorliegende Arbeit hat zum Ziel ein modellbasiertes System zur Segmentierung und Objekterkennung von Objekten mit gekrümmten Flächen basierend auf Distanzbildern zu entwerfen. Die Repräsentation des Modells beruht auf einem CAD-Modell, das eine mathematisch exakte und zuverlässige Beschreibung beliebiger Körper erlaubt.

Im Verlauf der Arbeit wird ein neu entwickeltes modellbasiertes Segmentierungsverfahren vorgestellt, das Krümmungen als invariante Merkmale einsetzt. Durch die Integration der Modellinformation wird der Vorgang der Segmentierung gestützt und es kann eine Aufteilung in Regionen erfolgen, die der im CAD-Modell entspricht.

Die in der Segmentierung aus dem Distanzbild abgeleiteten Merkmale werden den Merkmalen des Modells durch ein Baumsuchverfahren zugeordnet. Die Anzahl der zur Zuordnung in Betracht kommenden Merkmale reduziert sich durch die Berücksichtigung der Modelldaten bei der Segmentierung. Nach einer erfolgten Zuordnung wird die räumliche Transformation des Modells zum Objekt in der Szene berechnet, womit die Objekterkennung abschließt.

Die Arbeit zeigt auf, wie Objekte mit gekrümmten Flächen in Distanzbildern segmentiert, aus beliebiger Lage erkannt und zu einem CAD-Modell transformiert werden können. Es wird damit ein Beitrag geliefert zur Entwicklung von automatisierten Systemen in der Inspektion, der Fertigungstechnik und der Robotik.

## Abstract

The availability of highly capable three-dimensional sensor systems creates the need for adequate systems for automated processing of three-dimensional data. Contrary to conventional intensity-images, which contain information on the intensity of an objects surface, range images store the distance of each point on a surface to the sensor. In order to process these data on a high level of abstraction, new concepts need to be developed, exceeding two-dimensional image analysis. In the context of developing automated three-dimensional computer vision systems, this thesis brings two aspects into focus: partitioning the image data and establishing a correspondence in-between image- and model-data.

This work aims at developing a model-based system for the segmentation and object recognition of three-dimensional objects with curved surfaces using range images. The model data is represented using a CAD-model, providing a mathematical precise and reliable description of arbitrary shapes.

Within this thesis a newly developed method for model-based range image segmentation is presented, using curvature as invariant features. By integrating model information into the segmentation stage, the segmentation process is guided to provide a partitioning corresponding to that of the CAD-model.

The features extracted during segmentation stage are brought into correspondence with the features of the model using a tree-search strategy. The number of possible pairings is reduced due to the fact that model information has been utilized during segmentation. After successful pairing, the three-dimensional transformation of the model onto the object in the scene is computed, completing the object recognition process.

This work provides a way to compute the segmentation of objects in range images containing curved surfaces, detect objects in arbitrary positions and derive the transformation onto a CAD-model. Thereby it contributes to the development of automated systems in the areas of inspection, manufacturing and robotics.

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>7</b>
1.1	Motivation . . . . .	7
1.2	Ziele der vorliegenden Arbeit . . . . .	9
1.3	Gliederung der Arbeit . . . . .	9
<b>2</b>	<b>Grundlagen</b>	<b>11</b>
2.1	3D Sensoren . . . . .	11
2.1.1	Triangulierende Verfahren . . . . .	12
2.1.2	Laufzeitverfahren . . . . .	17
2.1.3	Datenrepräsentation . . . . .	20
2.2	Geometrisches Modellieren . . . . .	21
2.2.1	Solid Modeling . . . . .	21
2.2.2	Randbeschreibung . . . . .	22
2.3	Objekterkennung . . . . .	24
2.3.1	Repräsentationsformen zur Objekterkennung . . . . .	24
2.3.2	Systeme zur Objekterkennung . . . . .	25
<b>3</b>	<b>Geometrie der Fläche</b>	<b>27</b>
3.1	Einleitung . . . . .	27
3.2	Differential Geometrie . . . . .	28
3.2.1	Fundamentalformen . . . . .	28
3.2.2	Krümmungen . . . . .	29
3.2.3	Abgeleitete Größen . . . . .	29
3.2.4	Lokale Flächenklassifikation . . . . .	30
3.2.5	Formklassifikation . . . . .	30
3.2.6	Ausgezeichnete Punkte . . . . .	31
3.2.7	Spezielle Flächentypen . . . . .	31
3.3	Stückweise Flächenrückführung . . . . .	32
3.3.1	Problembeschreibung . . . . .	32
3.3.2	Flächenbeschreibung . . . . .	32
3.3.3	Parametrisierte Flächen . . . . .	32
3.3.4	Implizite Flächen . . . . .	33
3.3.5	Explizite Flächen . . . . .	36
3.3.6	Spezielle Flächentypen aus der Gruppe der Quadriken . . . . .	36
<b>4</b>	<b>Modellbasierte Segmentierung</b>	<b>38</b>
4.1	Einleitung . . . . .	38
4.1.1	Spalten und Verschmelzen . . . . .	39
4.1.2	Bereichswachstum . . . . .	40

4.1.3	Segmentierung durch Clustering . . . . .	42
4.1.4	Krümmungsbasierte Segmentierung . . . . .	42
4.2	Motivation für ein modellbasiertes Verfahren . . . . .	43
4.3	Krümmungseigenschaften in CAD Daten . . . . .	43
4.3.1	Flächen mit konstanter Krümmung . . . . .	43
4.3.2	Rotationsflächen mit variabler Krümmung . . . . .	46
4.3.3	Regelflächen mit variabler Krümmung . . . . .	47
4.3.4	Freiformflächen mit variabler Krümmung . . . . .	48
4.3.5	Details der Implementierung . . . . .	49
4.4	Krümmungsextraktion auf Distanzbildern . . . . .	50
4.4.1	Krümmungen aus der Normalenänderung . . . . .	51
4.4.2	Krümmungsschätzung aus Ableitungen . . . . .	52
4.4.3	Methoden zur analytischen Krümmungsschätzung . . . . .	52
4.4.4	Problematik der Krümmungsschätzung . . . . .	55
4.5	Integration der Modell-Information . . . . .	56
4.6	Experimente . . . . .	58
4.6.1	Testszene mit Grundkörpern . . . . .	58
4.6.2	Beispiele aus der SAMPL Datenbank . . . . .	58
4.6.3	Prüfkörper des Sonderforschungsbereichs “Aktive Exploration” . . . . .	62
<b>5</b>	<b>Zuordnungsanalyse</b>	<b>64</b>
5.1	Einleitung . . . . .	64
5.2	Modellbasierte Objekterkennung . . . . .	64
5.2.1	Zuordnungsanalyse durch Baumsuche . . . . .	65
5.3	Zuordnungsmengen aus der Segmentierung . . . . .	65
5.4	Transformationsbestimmung . . . . .	71
5.4.1	Transformation aus Punktkorrespondenzen . . . . .	71
5.4.2	Ableitung der Punktkorrespondenzen . . . . .	73
5.4.3	Transformationsergebnisse . . . . .	73
<b>6</b>	<b>Zusammenfassung</b>	<b>83</b>
<b>A</b>	<b>Eigenwert Verfahren</b>	<b>84</b>
<b>B</b>	<b>Alternative Herleitung der Krümmungen</b>	<b>85</b>
B.1	Erste Ableitung . . . . .	85
B.1.1	Normalkoordinaten erster Ordnung . . . . .	85
B.2	Zweite Ableitung . . . . .	85
B.2.1	Hauptkrümmungen und Hauptkrümmungsrichtungen . . . . .	86
B.2.2	Normalkoordinaten zweiter Ordnung . . . . .	86
	<b>Literaturverzeichnis</b>	<b>89</b>



# 1 Einführung

## 1.1 Motivation

In den letzten Jahren hat es einen enormen Fortschritt in der Entwicklung dreidimensionaler bildgebender optischer Sensoren gegeben. Eine Vielzahl kommerzieller Systeme ist heute am Markt erhältlich, mit denen nahezu alle Anforderungen nach Messvolumen, Genauigkeit und Geschwindigkeit erfüllt werden können. Im Vergleich zur Entwicklung der Sensoren befindet sich die Entwicklung entsprechender Systeme zur Verarbeitung der Daten in deutlichem Rückstand. Im Vergleich zu den Systemen des zweidimensionalen maschinellen Sehens befinden sich die Systeme des dreidimensionalen maschinellen Sehens sogar erst am Anfang der Entwicklung (Blais, 2003). Heute realisierte Anwendungen erfordern typischerweise ein hohes Maß an Interaktivität vom Benutzer. Aufgabe der Forschung muss es sein den Automatisierungsgrad zu steigern, um idealerweise vollautomatische Systeme zu erreichen.

Die Anwendungen dreidimensionalen maschinellen Sehens sind äußerst vielfältig und liegen in den Bereichen der Robotik, der Fertigungstechnik, der Konstruktion, der Inspektion und Qualitätskontrolle, um nur einige zu nennen. Dreidimensionale bildgebende optische Sensoren liefern im Vergleich zu den einfachen zweidimensionalen Sensoren ein wesentlich höheres Maß an Information. Sie erlauben nicht nur bisherige Aufgaben aus dem maschinellen Sehen mit höherer Präzision und besserer Zuverlässigkeit durchzuführen, sondern beziehen ihren besonderen Wert aus der Tatsache, dass sie völlig neue Anwendungsfelder erst ermöglichen.

Die Daten dreidimensionaler bildgebender optischer Sensoren, also die Distanzbilder, teilen viele Eigenschaften zweidimensionaler Bilder und daher lassen sich viele bestehende Methoden übernehmen oder übertragen. Die in den Distanzbildern direkt enthaltene räumliche Information macht es jedoch möglich den Bezug der Sensordaten zur wahren Geometrie eines Objektes viel direkter herzustellen, als dies bei zweidimensionalen Bildern möglich ist, aufgrund der dort vorliegenden perspektivischen Abbildung. Um diesen Bezug automatisiert herzustellen, lassen sich zwei wesentliche Schritte identifizieren, die Segmentierung und die Objekterkennung.

Als Segmentierung bezeichnet man den Vorgang, ein Bild in sinnvolle Bestandteile zu zerlegen. Das Problem das Objekt zu detektieren wird als Objekterkennung bezeichnet. Eine allgemein gültige Definition des Begriffs Objekterkennung ist nicht bekannt, jedoch wird üblicherweise von dem Ziel ausgegangen, ein bestimmtes Objekt in einer Szene zu identifizieren und dessen genaue Position und Orientierung zu bestimmen.

Ein ideales Objekterkennungssystem sollte in der Lage sein das Objekt in der Szene zu lokalisieren, unabhängig von ungünstigen Randbedingungen. Hierzu zählen Objekte mit komplexen beliebig gekrümmten Oberflächen, die Tatsache, dass das Objekt eine beliebige Lage im Raum einnehmen kann oder dass das Objekt teilweise verdeckt sein kann. Die Anwendungen für ein solches System sind äußerst vielfältig. Häufig anzutreffende Beispiele sind die Greifplanung bei Robotern, die Navigation autonomer Robotersysteme, industrielle Fertigung und Inspektion. Die hier vorliegende Arbeit zielt vor allem auf Anwendungen der letztgenannten Kategorie ab.

Ein System des künstlichen Sehens wird als modellbasiert bezeichnet, wenn es ein Modell des Objektes verwendet. Um ein solches System zu realisieren, müssen bei seiner Entwicklung folgende Fragestellungen gelöst werden: die einzusetzenden Sensoren, die Methoden zur Modellbildung, die Repräsentation des Modells und der Sensordaten und die Methoden zur Zuordnung der gewonnenen Sensordaten zu den Modelldaten. Der Sensor bestimmt die Art der von der Szene gewonnenen Daten, und ihre Auflösung. Das Modell stellt die Informationen über das Objekt bereit, auf dem die folgenden Verarbeitungsschritte aufbauen. Um eine Zuordnung zu ermöglichen, müssen die Modelldaten und die Sensordaten in eine adäquate Repräsentationsform gebracht werden. Die Repräsentation nimmt eine Schlüsselrolle in der Objekterkennung ein. Aus ihr werden die notwendigen Merkmale abgeleitet und sie bestimmt maßgeblich die Methoden zur Zuordnungsanalyse.

Für die Fragestellung nach den einzusetzenden Sensoren, müssen vorhandene Sensorsysteme betrachtet werden. Moderne Entwicklungen in der Sensortechnologie ermöglichen es heute problemlos direkt räumliche Informationen über ein Objekt in einem dichten Raster zu erfassen, so genannte Distanzbilder. Weitere mögliche Sensoreingaben sind Wärmebilder, aufgenommen mit speziellen thermischen IR-Kameras, oder volumetrische Daten, die mit Computertomographen erfasst wurden. Die Wahl der Sensoren hängt von der angestrebten Anwendung und ihrem Kontext ab. Im Rahmen dieser Arbeit steht die Erkennung kleinerer Objekte mit industriell gefertigten

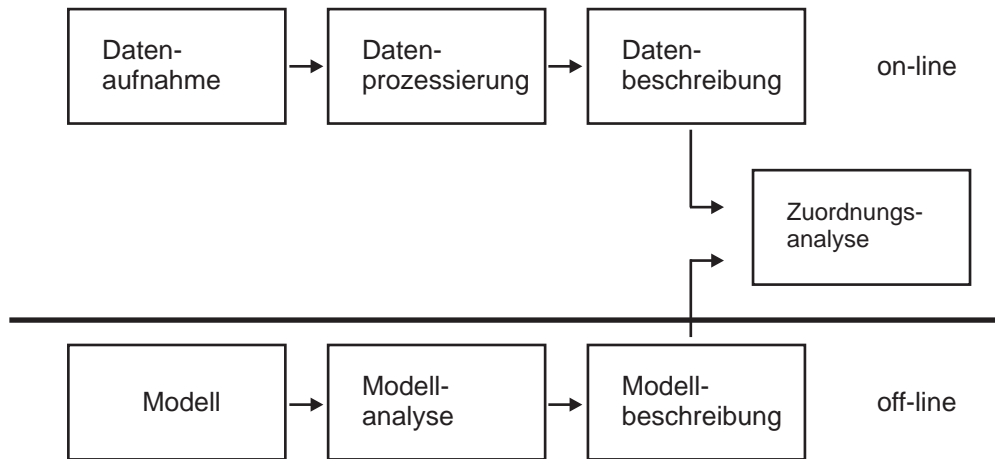


Abbildung 1.1: Ein allgemeines Schema des modellbasierten maschinellen Sehens in Anlehnung an Arman und Aggarwal (1993).

gekrümmten Oberflächen im Vordergrund. Die hierfür betrachteten Sensoren stammen aus der Kategorie der distanzmessenden Systeme.

Wie Grimson (1990) darlegt, gibt es zwei wesentliche Richtungen unter den Objekterkennungssystemen, die sich in zwei Varianten des Suchraums äußern. Zum einen kann der Suchraum aus allen möglichen Kombinationen der Parameter bestehen, die die Lage des zu erkennenden Objektes im Raum beschreiben, also zum Beispiel die Translation- und Rotationsparameter. Zum anderen kann der Suchraum aus allen möglichen Zuordnungen zwischen den Merkmalen eines Modells und den Merkmalen der Szene bestehen. Diese beiden unterschiedlichen Systemtypen lassen sich auch als ansichtsbasiert beziehungsweise merkmalsbasiert bezeichnen. Für ein beliebig positionierbares Objekt ist der Suchraum über alle Translation- und Rotationsparameter sechsdimensional. Dabei muss noch zusätzlich eine Quantisierung der Parameter eingeführt werden. Bei einem Ansatz mit Merkmalszuordnung beträgt die Größe des Suchraums bei  $m$  Merkmalen des Modells und  $n$  Merkmalen der Szene  $m^n$ . In beiden Fällen ist der Suchraum in der Regel zu groß, um ihn vollständig zu durchsuchen und er muss daher mit geeigneten Methoden eingeschränkt werden. Neben der Wahl des Suchraums ist die gewählte Repräsentationsform ein weiteres wichtiges Kriterium, um Objekterkennungssysteme zu klassifizieren. Ein allgemeines Schema eines modellbasierten Systems zur Objekterkennung ist in Abbildung 1.1 gegeben.

Viele der CAD-basierten maschinellen Sehsysteme nutzen experimentelle CAD-Systeme, die speziell für diese Aufgaben entwickelt wurden. Der Einsatz dieser experimentellen Systeme ist jedoch nur beschränkt und nur im Rahmen von Forschungsarbeiten akzeptabel.

Ein großer Hinderungsgrund kommerzielle CAD-Systeme einzusetzen besteht darin, dass diese Systeme in erster Linie zur Konstruktion und Herstellung von Objekten entwickelt wurden und nicht zu ihrer Erkennung. Daher sind viele der kommerziellen Systeme nicht in der Lage die für die Objekterkennung relevanten Informationen zu liefern. Herkömmliche CAD-Systeme legen vielmehr ihr Augenmerk auf die komfortable interaktive Manipulation der Objekte bis hin zu ihrer photorealistischen Darstellung.

Zudem ist die Einbindung vieler kommerziellen CAD-Systemen in die Anwendung eines Objekterkennungssystems schwierig, da die Systeme geschlossen sind, das heißt nicht die benötigten Schnittstellen bereitstellen. Wie Arman und Aggarwal (1993) bemerken, ist die Fähigkeit kommerzielle CAD-Systeme einzubinden jedoch entscheidend für den Erfolg von Objekterkennungssystemen in der Praxis.

Ein Ausweg aus dem Mangel an direkten Schnittstellen, den zahlreiche Systeme nutzen, ist die Möglichkeit Informationen über standardisierte Dateiformate auszutauschen. Hierzu wurde in vielen Arbeiten der IGES-Standard verwendet (Flynn und Jain, 1991). Durch die Ablage in standardisierte Dateiformate gehen aber zumeist wichtige Informationen des CAD-Systems verloren. Speziell bei IGES geht zum Beispiel die Information zur Topologie der Flächen vollständig verloren und muss, soweit benötigt, nachträglich rekonstruiert werden.

## 1.2 Ziele der vorliegenden Arbeit

Die hier vorliegende Arbeit hat zum Ziel ein modellbasiertes System zur Segmentierung und Objekterkennung basierend auf Distanzbildern zu entwerfen. Die Repräsentation des Modells beruht dabei auf einem CAD-Modell, dessen Handhabung vollständig in einem standardisierten, kommerziellen CAD-System erfolgt. Die Entscheidung ein CAD-System zur Modellbildung einzusetzen, wurde unter Betrachtung einer Reihe vorteilhafter Gesichtspunkte getroffen. Zunächst sind in dem hier betrachteten industriellen Kontext zumeist bereits CAD-Modelle der zu untersuchenden Objekte vorhanden. Dies bedeutet, dass der zusätzliche Aufwand, ein Modell für die Objekterkennung zu erstellen, entfällt. Auch wenn ein Modell neu erstellt werden muss, wird dies vorzugsweise in einem etablierten CAD-System geschehen, da es in hohem Maße automatisiert arbeitet und zahlreiche Operationen und Modellprimitive zur Verfügung stellt, die die Erstellung erleichtern. Zudem ist der Vorgang an sich standardisiert, wodurch eine effektive Qualitätskontrolle der Modelle möglich wird. Insgesamt bieten also CAD-Modelle eine vollständige, mathematisch exakte und zuverlässige Beschreibung, womit sie vielen anderen Methoden zur Modellbildung überlegen sind.

Die manuelle Erstellung von Modellen in spezialisierten Repräsentationsformen erfordert einen ebenso spezialisierten Anwender und generell einen hohen Zeitaufwand. Ebenfalls einen hohen Zeitaufwand benötigen Methoden zur Modellbildung, die ein Modell mit Verfahren des maschinellen Lernens aus Sensordaten ableiten.

Eine Modellbildung, welche auf einem standardisierten CAD-System beruht, kann auf erweiterte Information des Modells zugreifen, zum Beispiel auf Toleranzen und Materialdichten. Zudem kann das Objekterkennungssystem in idealer Weise mit anderen Komponenten, die sich um das CAD-System gruppieren, zu einem Gesamtsystem integriert werden.

Da innerhalb eines CAD-Systems nahezu alle Objekt- und Flächentypen bearbeitet werden können, sollten auch die übrigen Komponenten eines Objekterkennungssystems diesem universellen Anspruch nach Möglichkeit genügen. Diese Arbeit stellt daher ein neu entwickeltes Verfahren zur Segmentierung von Distanzbildern vor. Das Verfahren wurde mit dem Ziel entwickelt möglichst viele Flächentypen, insbesondere gekrümmte Flächen, verarbeiten zu können.

Das Segmentierungsverfahren arbeitet modellbasiert, das heißt, es greift auf die im CAD-Modell enthaltene Information zu, um eine Aufteilung der Daten zu erzielen. Durch die Integration der Modellinformation wird der Vorgang der Segmentierung gestützt und es kann eine Aufteilung in Regionen erfolgen, die der im CAD-Modell entspricht. Um Information aus dem CAD-Modell in die Verarbeitung von Distanzbildern einbeziehen zu können, werden, basierend auf den Krümmungen, invariante Merkmale aus den CAD-Daten extrahiert, die in einen gemeinsamen Merkmalsraum eingebracht werden können.

Der finale Schritt der Objekterkennung, also die Zuordnung der Merkmale aus Szene und Modell, erfolgt im Rahmen einer Baumsuche. Aufgrund der frühzeitigen Einbindung der Modelldaten bei der Segmentierung, kann die Anzahl der zur Zuordnung in Betracht kommenden Merkmale reduziert werden. Somit bleibt die exponentielle Komplexität beherrschbar.

## 1.3 Gliederung der Arbeit

Kapitel 2 legt wichtige Grundlagen, auf denen die folgenden Kapitel aufbauen. Zu Beginn wird eine Übersicht über gängige Verfahren zur dichten Oberflächenerfassung gegeben. Es werden die für diese Arbeit wichtigsten Verfahren und die sich daraus ableitenden Sensortypen dargelegt und die Repräsentationsformen der gewonnenen Tiefeninformation bei den einzelnen Systemen beschrieben. Es folgt eine Einführung in das Geometrische Modellieren, speziell des Solid Modeling. Zum Ende werden die in der Literatur bereits vorgestellten Systeme zur Objekterkennung basierend auf Distanzbildern vorgestellt und diskutiert.

Die Methoden zur Beschreibung von Flächen und zur Ableitung von charakteristischen Eigenschaften von Flächen werden in Kapitel 3 definiert. Die wichtigsten Methoden zur Charakterisierung von Flächen liefert dabei die Differentialgeometrie. Zudem werden Methoden zur lokalen Flächenrückführung vorgestellt.

Kapitel 4 beschreibt zunächst grundsätzlich das Problem der Segmentierung. Danach werden detailliert bestehende Ansätze zur Segmentierung von Distanzbildern vorgestellt. Darauf aufbauend wird ein eigenständiges modellbasiertes Verfahren entwickelt. Hierzu werden Methoden zur Krümmungsschätzung sowohl aus dem CAD-Modell als auch aus den Sensordaten benötigt. Das entwickelte Verfahren wird dann anhand praktischer Beispiele diskutiert.

Das ausgewählte Verfahren zur Zuordnungsanalyse wird in Kapitel 5 erläutert. Basierend auf den Ergebnissen aus der Segmentierung werden die Ergebnisse der Zuordnung dargelegt. Hierzu wird die räumliche Transformation des Modells zum Objekt in der Szene berechnet, womit die Objekterkennung abschließt.

## 2 Grundlagen

### 2.1 3D Sensoren

Die Anzahl der Sensoren zur dichten Erfassung von Oberflächen ist in den letzten Jahren sprunghaft gestiegen. Während R. A. Jarvis (1983) Mitte der achtziger Jahre in einer umfassenden Studie noch etwa 20 Systeme aus Forschungs- und Entwicklungslabors aufzählt, die verschiedenste physikalische Prinzipien zur Distanzgewinnung verwenden, recherchierte F. Blais (2003) eine Liste von über achtzig Firmen, die oftmals mehrere Sensoren zur dichten Distanzmessung anbieten, allein aus dem Bereich der laser- und projektionsbasierten Systeme. Es ist also nicht nur die Anzahl der verfügbaren Sensoren stark gestiegen, gleichzeitig hat sich deren Entwicklung aus der Forschung in die kommerzielle Anwendung verschoben.

Im nun Folgenden soll nicht versucht werden eine vollständige Aufzählung der möglichen Verfahren zur dichten Oberflächenerfassung zu erstellen. Stattdessen sollen die für diese Arbeit wichtigsten Verfahren und die sich daraus ableitenden Sensortypen untersucht werden. Dies geschieht insbesondere im Hinblick auf die Repräsentationsformen der gewonnenen Tiefeninformation bei den einzelnen Systemen. Diese Charakteristik der Daten ist ein entscheidendes Kriterium bei der Wahl der Operatoren in den anschließenden Schritten der Datenverarbeitung. Durch eine geeignete Kombination von Datenrepräsentation und Algorithmus lassen sich die einzelnen Verarbeitungsschritte trotz hohen Datenaufkommens beschleunigen.

Neben den eigenen Erfahrungen im Umgang mit verschiedenen Sensorsystemen und dem Austausch mit Herstellern solcher Systeme stützen sich die folgenden Beschreibungen auf eine ganze Reihe von Publikationen, die jeweils mit unterschiedlichen Schwerpunkten und Ansätzen veröffentlicht wurden. R. A. Jarvis (1983) klassifiziert die Methoden zur Distanzmessung in strukturierte Beleuchtung, einige monookulare Verfahren wie Überdeckung, Texturgradient, Schattierung und Fokus sowie Stereobildzuordnung, optischer Fluss, Moiré und Laufzeitverfahren entweder auf Basis von Ultraschall oder Laser.

P. J. Besl (1988a) unterscheidet nach bildgebenden Radarsystemen, also laufzeitmessende Verfahren wie gepulsten Lasern, amplitudenmodulierte Laser oder frequenzmodulierte Laser, aktive Triangulation, dazu gehören Punktriangulation, Lichtschnitt, kodierte Binärmuster und Zufallsmuster, Moiré Techniken, holografische Interferometrie und Fokusverfahren. Er entwickelt eine Kenngröße, die alle Systeme unabhängig vom Verfahren in ihrer Leistungsfähigkeit charakterisieren soll. Bezeichnet  $T$  die Zeitspanne für die Entfernungsmessung eines einzelnen Punktes,  $L_x, L_y, L_z$  das maximale zu vermessende Volumen und  $\sigma_x, \sigma_y, \sigma_z$  die Messgenauigkeit in der Koordinatenrichtung, so ergibt sich die Kenngröße  $M$  zu

$$M = \frac{1}{\sqrt{T}} \left( \frac{L_x L_y L_z}{\sigma_x \sigma_y \sigma_z} \right)^{\frac{1}{3}}$$

Diese Kenngröße hat sich aus mehreren Gründen nicht durchgesetzt. Vor allem scheitert die Bestimmung oftmals schon an der Festlegung der Genauigkeiten  $\sigma_x, \sigma_y, \sigma_z$ . Diese Größen sind zum einen extrem abhängig von den Umgebungsbedingungen, wie zum Beispiel Beleuchtungsverhältnisse, Oberflächeneigenschaften und Reflektionsverhalten. Zum anderen sind bereits die Vorschriften zur experimentellen Bestimmung der Kenngrößen völlig offen. Erst in jüngster Zeit beginnen Vertreter von Herstellern, Anwendern und Wissenschaftlern derartige Richtlinien zu entwickeln (VDI, 2002). Die Diskussionen in diesem Gremium zeigen, wie schwierig die Einigung auf allgemein-gültige Verfahrensweisen ist. Die von P. J. Besl vorgeschlagene Kenngröße ist dennoch interessant, da sie die Möglichkeit bietet, den Entwicklungsfortschritt über einen größeren Zeitraum zu dokumentieren. Das in seiner Veröffentlichung beste System nach dem Prinzip der kodierten Streifenprojektion erreichte eine Kennzahl von 100.000. Ein aktuelles System eines kommerziellen Anbieters nach ähnlichem Verfahren erreicht eine Kennzahl von 1.000.000.

B. Breuckmann (1993) führt den Begriff der Topometrie ein. Er unterscheidet grundsätzlich zwischen Methoden, in denen die Bildinformation durch Scannen des Messbereichs erfasst wird, und direkten bildgebenden Methoden. Zu den topometrischen 3D-Messtechniken zählt er die Lasertriangulation, das Lichtschnittverfahren, die bildhafte Triangulation, die Streifenprojektion und Moiré Techniken. Die ersten beiden gehören zur Kategorie der scannenden Verfahren, die übrigen zu den bildgebenden Verfahren.

S. F. El-Hakim (1995) vergleicht in seiner Veröffentlichung die Leistungsfähigkeit von passiver Triangulation, aktiver Triangulation mittels punktförmiger Laserprojektion und strukturierter Beleuchtung mittels Laserlinien. Insbesondere untersuchte er die Aspekte der Kalibrierung und der Leistungsfähigkeit der Systeme.

H. J. Tiziani (1997) beschreibt speziell Verfahren zur optischen Vermessung industriell gefertigter Oberflächen. Hierzu zählt er Triangulation, Streifenprojektion und verschiedene Varianten der Interferometrie. Er zeigt die unterschiedlichen Genauigkeitspotenziale der Systeme auf, die insgesamt einen Genauigkeitsbereich von  $10^{-6}$  Millimetern bis zu mehreren Millimetern abdecken.

G. Häusler (1999) begründet detailliert die physikalischen Grenzen triangulierender Verfahren. Die Genauigkeit triangulierender Verfahren ist nach seiner Darstellung vor allem durch die Speckle Problematik begrenzt. Den Fehler für die Distanzmessung, der durch Speckle verursacht wird, gibt er an mit

$$\delta_{z_m} = \frac{c\lambda}{2\Pi \sin u_{obs} \sin \theta}$$

Wobei  $\theta$  der Triangulationswinkel,  $\sin u_{obs}$  die Apertur der beobachtenden Optik,  $\lambda$  die Wellenlänge der Beleuchtungsquelle und  $c$  der Speckle-Kontrast ist. Er betont, dass nicht die Monochromatizität der Grund des Rauschens ist, sondern die räumliche Kohärenz und daher die Probleme nicht nur bei Laser sondern auch bei Weißlicht auftreten können, wenn auch nur in abgeschwächter Form. Dies ist immer dann der Fall, wenn die Apertur der beleuchtenden Optik klein ist im Verhältnis zu der beobachtenden Optik. Insbesondere aus Gründen der Schärfentiefe kann auch bei der Beleuchtung zum Beispiel eines Streifenprojektionssystems keine beliebig große Blende gewählt werden. Die oben angegebene physikalische Grenze der Messgenauigkeit wird jedoch in der Praxis kaum von einem System erreicht werden können, da bei realen Objekten die Auswirkungen der Mikrostruktur der Oberfläche einen größeren Fehler verursachen. So werden zum Beispiel bei gefrästen Metalloberflächen die Bearbeitungsspuren die Struktur der Beleuchtung beeinflussen und somit zu wesentlich größeren Fehlern beitragen.

F. Chen et. al. (2000) führen in einer neueren Übersicht zum Thema mehrere Verfahren zur dreidimensionalen Messtechnik mit einer extensiven Literaturliste auf. Zu den aufgeführten Verfahren gehören Punkt-Lasertriangulation, Moiré, Laser Speckle Pattern Sectioning, Interferometrie, Photogrammetrie und strukturierte Beleuchtung.

F. Blais (Blais, 2003) betrachtet in seiner Veröffentlichung die Entwicklung von Distanzsensoren speziell aus Sicht der industriellen Anwendung im Laufe der letzten 20 Jahre. Ausgehend von einfacher Laserpunkt Triangulation zeigt er die Entwicklung der synchronisierten Scansysteme auf. Ihr Vorteil liegt in der hohen Ortsauflösung, die durch einen geringeren Öffnungswinkel der Optik erzeugt wird. Um dennoch ein großes Gesichtsfeld des Sensors zu erzielen, werden Beleuchtung und Empfänger synchron abgelenkt. Als weitere Entwicklungsschritte in der triangulierenden Distanzdatengewinnung zählt er die Laserlichtschnitt Systeme auf sowie Musterprojektion und Moiré. Die betrachteten Lasersysteme haben gegenüber den musterprojizierenden Systemen den Vorteil eines größeren Tiefenbereiches, da die Tiefenschärfe der Projektion durch optische Gesetze begrenzt ist. Für Objekte mit relativ großer Tiefenausdehnung werden von ihm daher Laufzeitverfahren bevorzugt, die entweder mit gepulstem oder amplitudenmoduliertem Laser operieren.

## 2.1.1 Triangulierende Verfahren

### 2.1.1.1 Stereobildzuordnung

Das photogrammetrische Verfahren der Stereobildzuordnung (Abbildung 2.1) kann wohl als das klassische Verfahren zur dichten Distanzdatenerfassung bezeichnet werden. Basierend entweder auf einfacher Korrelation oder auf einer kleinsten Quadrate Zuordnung werden in einem Rasterverfahren korrespondierende Bildausschnitte gesucht. Erfolgt diese Suche in einem dichten Raster im Ausgangsbild so kann für jede gefundene Korrespondenz eine dreidimensionale Koordinate bestimmt werden und somit ein ebenso dichter Distanzdatensatz erzeugt werden. Bei bekannter oder rechnerisch bestimmter äußerer Orientierung der beiden Bilder sowie einer kalibrierten inneren Orientierung des Aufnahmesystems ergibt sich der Zusammenhang zwischen den Bildkoordinaten der korrespondierenden Punkte und den räumlichen Koordinaten durch die Kolinearitätsgleichung:

$$x = -c \frac{r_{11}(X-X_0)+r_{21}(Y-Y_0)+r_{31}(Z-Z_0)}{r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)}$$

$$y = -c \frac{r_{12}(X-X_0)+r_{22}(Y-Y_0)+r_{32}(Z-Z_0)}{r_{13}(X-X_0)+r_{23}(Y-Y_0)+r_{33}(Z-Z_0)}$$

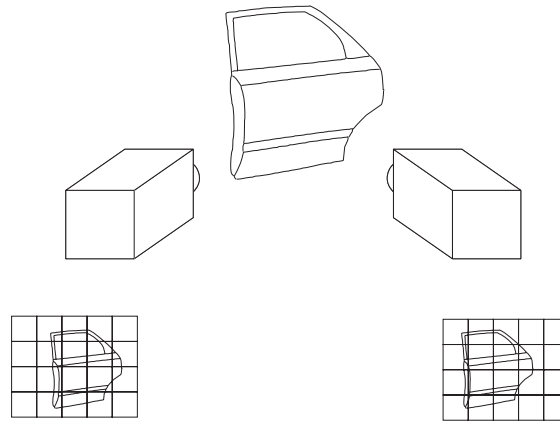


Abbildung 2.1: Die Stereo-Anordnung von zwei Kameras zum Objekt.

Neben der Berechnung der räumlichen Koordinaten  $X$ ,  $Y$  und  $Z$  erhält man noch eine weitere wichtige Information. Unter der Annahme der Stetigkeit und Geschlossenheit der beobachteten Fläche sowie ihrer Beschränkung auf eine Graph-Fläche, kann man aus der Nachbarschaft der Bildpunkte auf die Nachbarschaft der räumlichen Punkte schließen. Somit ergibt sich neben der Geometrie der gemessenen Punktwolke auch eine Ordnung dieser Punkte, also ihre Topologie. Durch die rasterförmige Suche in den Bildern ergibt sich eine rasterförmige Nachbarschaft der Punkte im Raum. Diese kann sehr einfach in einer Matrix festgehalten werden.

Das Verfahren der Bildzuordnung kann auf mehrere Bilder ausgeweitet werden. Wird die Zuordnung von einem Ankerbild aus auf mehrere Bilder berechnet, bleiben die oben genannten Eigenschaften bezüglich der Nachbarschaft erhalten. Werden statt dessen aus mehreren Bildern jeweils paarweise Zuordnungen bestimmt, ergibt sich nur noch eine paarweise Ordnung der Objektpunkte. Die globale Ordnung der Punktwolke muss nachträglich berechnet werden. Dieser nicht-triviale Schritt kann abhängig von der Geometrie der Fläche sehr aufwändig und fehleranfällig sein und bedarf oft auch des manuellen Eingreifens (siehe zum Beispiel (Gruen et al., 2001)).

### 2.1.1.2 Projektionssysteme mit Zufallsmuster

Das eben beschriebene Verfahren der Bildzuordnung verwendet die natürliche Bildinformation, um eine Zuordnung einzelner Bildbereiche zu ermöglichen. Dies gelingt immer dann, wenn die Objektoberfläche eine hinreichende Textur besitzt. Im Bereich der industriell gefertigten Flächen ist dies aber eher die Ausnahme. Um das Verfahren dennoch einsetzen zu können, kann künstlich eine Textur auf die Oberfläche aufgebracht werden. Um die Oberfläche dabei nicht dauerhaft zu verändern, zum Beispiel durch Aufbringung einer Prägung oder einer Beschichtung, kann die Textur aufprojiziert werden.

Bereits sehr früh zu Zeiten der analogen Nahbereichsphotogrammetrie wurden solche Systeme entwickelt. Die Firma ZEISS brachte das System INDUSURF auf den Markt, das im Automobilbereich zur Karosserievermessung eingesetzt wurde. Die Firma INPHO betrieb später eine Weiterentwicklung und kombinierte das Projektionssystem mit modernen hoch auflösenden Digitalkameras. Heute befindet sich noch die Firma COGNITENS mit ihrem OPTIGON Sensor (siehe Abbildung 2.2) am Markt. Dieser Sensor verfügt über eine Projektionseinheit in Kombination mit drei Kameras. Die Mathematik des Trifokaltensors ermöglicht die rasche Zuordnung aus den drei Bildern.

Der grundsätzliche Vorteil dieser Systeme, die auf dem Prinzip der Bildzuordnung und der Projektion eines einzelnen Musters basieren, liegt in ihrer Fähigkeit aus einem einzelnen Bildpaar oder Bildtupel die gesamte räumliche Information zu berechnen. Somit entfällt die Notwendigkeit, Bildsequenzen aufzunehmen. Dadurch sind die Systeme relativ unempfindlich gegenüber Schwingungen oder sonstigen kurzzeitig auftretenden Deformationen am Objekt. Nachteilig wirkt sich die vergleichsweise hohe Rechenzeit aus, die benötigt wird, um die Zuordnungen zu berechnen.

### 2.1.1.3 Punkt Triangulation

Zu den einfachsten aktiven Verfahren, die auf der Triangulation beruhen, gehört die Punkt Triangulation. Typischerweise mit einem Laser wird ein einzelner Punkt auf die zu messende Oberfläche projiziert. Das auf



Abbildung 2.2: Sensorsystem mit Musterprojektion und drei Kameras der Firma CogniTens.

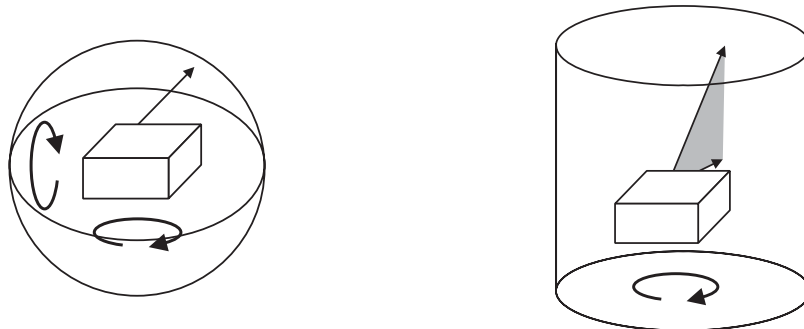


Abbildung 2.3: Mechanische Scanbewegungen für punkt- oder linienhafte Sensoren.

der Oberfläche reflektierte Licht wird durch eine Kamera beobachtet. Bei bekanntem Aufbau der Komponenten, das heißt bei bekannten äußeren und inneren Orientierungen, kann aus dem Schnitt des Beobachtungsstrahls mit dem Projektionsstrahl der räumliche Punkt bestimmt werden.

Durch eine Bewegung des Laserstrahls über die Oberfläche kann ein Objekt durch dieses Prinzip auch flächenhaft vermessen werden. Hierzu kann zum einen nur der Laserstrahl durch eine zweidimensional Ablenkeinheit bewegt werden oder der gesamte Sensor zweidimensional verschwenkt werden (siehe Abbildung 2.3). Wird nur der Laserstrahl abgelenkt, bleibt das Gesichtsfeld des Sensors auf das Gesichtsfeld der Kamera beschränkt. Dadurch ergibt sich eine Abwägung aus der Ortsauflösung der Kamera, die wiederum die Genauigkeit begrenzt, mit der der reflektierte Punkt bestimmt werden kann und somit der Messgenauigkeit und dem Gesichtsfeld des Sensors. Wird der gesamte Sensorkopf geschwenkt, begrenzt das Gesichtsfeld der Kamera lediglich die Tiefenauflösung des Systems. Das Gesichtsfeld wird durch die Mechanik bestimmt.

Einen Mittelweg aus beiden Verfahren bietet das so genannte synchronisierte Scannen (Rioux, 1984). Dabei werde sowohl der projizierende als auch der beobachtende Strahl abgelenkt. Da diese opto-mechanische Ablenkung synchron erfolgt, sind die Winkelbeziehungen der Strahlen bekannt. Der Vorteil des Verfahrens besteht darin, dass der gleiche Vorteil eines großen Gesichtsfeldes erzielt werden kann, aber nur geringe Massen mechanisch bewegt werden müssen.



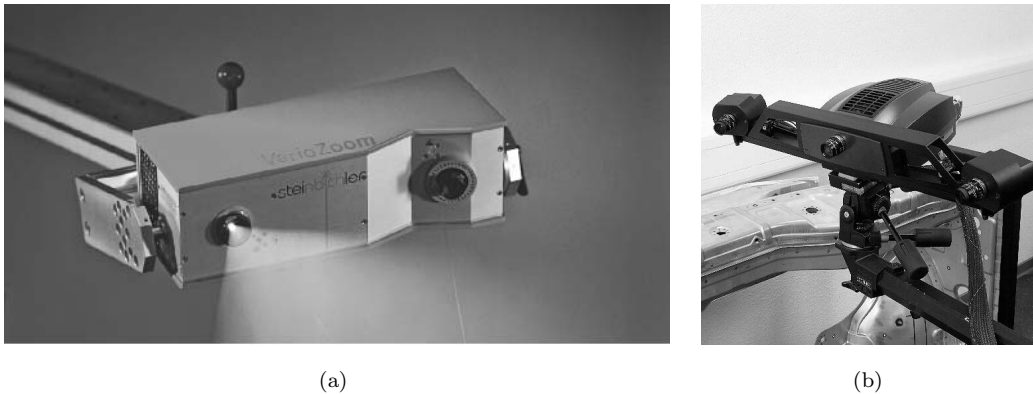


Abbildung 2.4: Kommerzielle Streifenprojektionssysteme. (a) COMET-400 Vario Zoom der Firma STEINBICHLER. (b) ATOS IIe der Firma GOM.

Da bei dem Verfahren der Punkttriangulation pro elementarem Messvorgang immer nur ein Punkt gemessen wird, ist das Verfahren relativ langsam. Auch durch den Einsatz von schnellen Kameras oder anderen Detektoren können typischerweise nur wenige tausend Punkte pro Sekunde erfasst werden. Durch die Notwendigkeit mechanischer Bewegungen kommen zudem schwer kalibrierbare Einflüsse auf die Messung hinzu.

#### 2.1.1.4 Lichtschnittverfahren

Eine logische Erweiterung der Punkttriangulation ist das Lichtschnittverfahren (Abbildung 2.5). Anstelle eines eindimensionalen Punktes wird eine zweidimensionale Struktur, eine Linie, auf das Objekt projiziert. Eine Linie kann aus einem punktförmigen Laserstrahl sehr kostengünstig durch eine Zylinderlinse erzeugt werden. Die Reflektion der projizierten Linie wird wiederum mit einer Kamera beobachtet. Aus der Deformation der Linie kann die dreidimensionale Form des Objektes entlang der projizierten Struktur errechnet werden.

Um die Messung auf die Fläche auszudehnen, muss wiederum eine mechanische Bewegung erfolgen. Aber im Gegensatz zur Punkttriangulation muss die Bewegung nur in einer Dimension erfolgen, zum Beispiel um eine Achse (siehe Abbildung 2.3). Ansonsten treffen auch hier dieselben Abwägungen und Möglichkeiten der Scanbewegung zu. Eindeutig können pro elementarem Messvorgang mehr Punkte bestimmt werden als bei der punktweisen Triangulation.

#### 2.1.1.5 Streifenprojektions-Systeme

Streifenprojektions-Systeme gehören sicherlich zu den populärsten Sensoren zur dichten Oberflächenerfassung im industriellen Kontext. Streifenprojektion wird hierbei als ein Überbegriff über zahlreiche verwandte Verfahren gebraucht, denen allen die Projektion eines oder mehrerer streifenförmiger Muster gemeinsam ist. Zahlreiche kommerzielle Realisierungen des prinzipiellen Verfahrens sind heute auf dem Markt, zum Beispiel von ABW, BREUCKMANN, GOM, STEINBICHLER oder IMETRIC (siehe auch Abbildung 2.4).

Zunächst kann die Streifenprojektion als eine natürliche Erweiterung des Lichtschnittverfahrens angesehen werden. Statt eine einzelne Linie mehrmals auf das Objekt zu projizieren und dabei zu verschieben, werden gleichzeitig mehrere Linien versetzt zueinander projiziert. Jeder einzelne der Streifen liefert nun die Information über die Form des Objektes entlang der Linie. Die eigentliche Schwierigkeit des Verfahrens besteht nun aber darin, die globale Zuordnung der im Sensor detektierten Streifen zu den projizierten Streifen herzustellen.

Ohne auf die Details einzelner Implementierungen eingehen zu können, sollen im Folgenden die Grundzüge der Streifenprojektionstechnik nach dem codierten Lichtansatz kurz dargestellt werden (Altschuler et al., 1979). Um eine Oberfläche dicht zu erfassen, muss eine Vielzahl von Linien in dichten Abständen projiziert werden. Um dennoch die Eindeutigkeit zu gewährleisten, werden die Linien nicht gleichzeitig sondern nach einem bestimmten Verfahren in zeitlicher Folge gestaffelt als binäres Muster projiziert. Das hierfür von vielen Systemen verwendete Verfahren ist die Gray-Code-Sequenz. Mit der Gray-Code-Sequenz müssen für  $n$  Linien lediglich  $\lg n$  Muster auf das Objekt projiziert werden (siehe Abbildung 2.6).

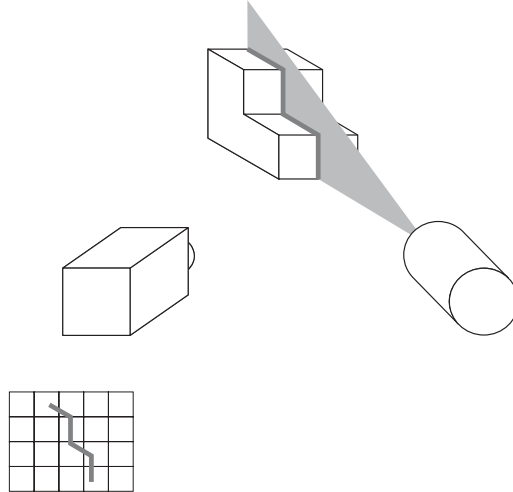


Abbildung 2.5: Das Prinzip des Lichtschnitts mit einer Kamera und einem Linienprojektor.

Die Gray-Code-Sequenz löst die Korrespondenz zwischen Projektorkoordinaten und Kamerakoordinaten jedoch nur grob auf Pixelebene. Um Subpixel-Genauigkeit zu erreichen, muss eine weitere Mustersequenz projiziert werden. Viele Systeme verwenden hierfür den Phasen-Shift (siehe Abbildung 2.6). Zwei Nachteile des Phasen-Shift-Verfahrens sind anzumerken. Zum einen ist die Annahme einer sinusförmigen Abbildung des Streifenmusters auf dem Sensor bei einem rasterbasierten Projektionssystem zunächst nicht erfüllt. Dies führt zu deutlich nachweisbaren Messfehlern. Die Sinusform des Streifenmusters muss erst erzeugt werden. Dies kann zum einen durch ein gezieltes 'Unschärf'-Stellen der Projektionseinheit erfolgen. Der Nachteil hierbei ist, dass diese Unschärfe nur in einem bestimmten Abstand des Objektes vom Projektor das gewünschte Ergebnis erzielt. Variiert der Abstand der Objektoberfläche stark in Aufnahmerichtung, dann wird das Streifenmuster nur in manchen Bereichen des Objektes sinusförmig abgebildet, in anderen nicht. Eine zweite Möglichkeit zur Erzeugung sinusförmiger Muster aus Raster-Projektoren ist die Anbringung optischer (Tiefpass-) Filter vor dem Projektor. Dies ist jedoch nur in einer Streifenrichtung möglich, sodass keine Kreuzprojektion mehr erfolgen kann.

Ein weiterer Nachteil des Phasen-Shift-Verfahrens betrifft die Auswertung der Korrespondenz eines Kamerapixels zu einem Projektorstreifen. Beim Phasen-Shift wird die ganzzahlige Pixelkoordinate der Kamera zu einer Subpixelkoordinate des Projektors in Beziehung gesetzt. Beobachten mehrere Kameras dieselbe Streifensequenz, so kann keine direkte Pixel-zu-Pixel-Korrespondenz zwischen den beteiligten Kameras aufgebaut werden.

Um diesen Nachteilen des Phasen-Shifts zu begegnen, wurde am Institut für Photogrammetrie von Gühring et al. (2000) ein neues Verfahren für die Durchführung und Auswertung von Messungen nach dem codierten Lichtansatz entwickelt. Das neue Line-Shift-Verfahren basiert nicht mehr auf der Annahme einer sinusförmigen Abbildung des Streifenmusters. Stattdessen wird getrennt für jede projizierte Linie das Zentrum im Bild bestimmt (siehe Abbildung 2.6). Durch das neue Verfahren werden ganzzahlige Projektorkoordinaten mit Subpixel-Koordinaten der Kameras verknüpft. Dies ermöglicht die direkte Verknüpfung von Bildmessungen mehrerer Kameras, die dieselbe Streifensequenz beobachten. Sind mehrere Kameras an der Messung beteiligt führt dies zu einer Steigerung der Redundanz und somit zu zuverlässigeren und genaueren Resultaten. Ein weiterer Vorteil der neuen Methode ist die verbesserte Kompensation von Reflektivitätsunterschieden in der Objektoberfläche.

Sowohl der einfache Phasen-Shift als auch der Line-Shift dienen dazu, das Gray-Code-Verfahren um eine höhere Messgenauigkeit zu ergänzen. Die Auflösung der Mehrdeutigkeit wird durch die Gray-Code-Sequenz sichergestellt. Eine Alternative zu dieser Vorgehensweise bietet das heterodyne Phasen-Shift-Verfahren. Hierbei werden mehrere Phasen-Shifts mit unterschiedlicher Wellenlänge hintereinander durchgeführt. Werden zwei Streifenmuster mit den unterschiedlichen Wellenlängen  $\lambda_1$  und  $\lambda_2$  eingesetzt, so ergibt sich eine größere effektive Wellenlänge

$$\Lambda = \frac{\lambda_1 \lambda_2}{|\lambda_1 - \lambda_2|}$$

und somit auch ein wesentlich größeres Eindeutigkeitsintervall.

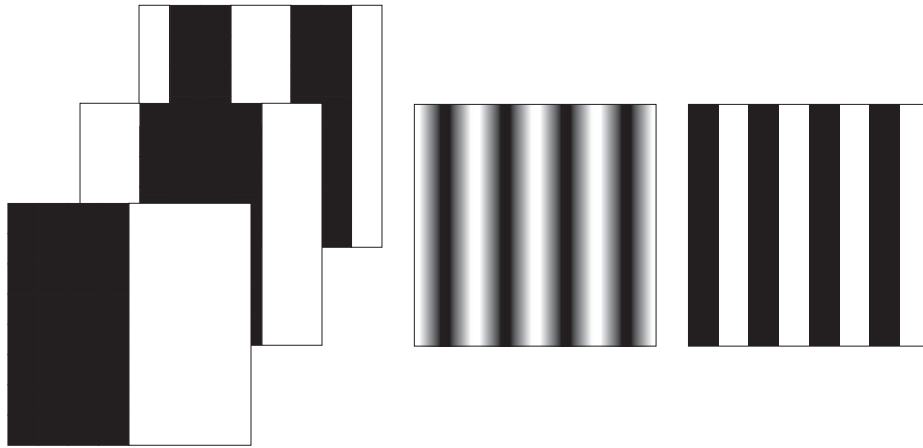


Abbildung 2.6: Die Projektionssequenzen für den Gray-Code und die beiden Alternativen zur Subpixel Bestimmung, Phasen-Shift und Line-Shift.

### 2.1.1.6 Photogrammetrische Prozessierung

Besitzt der Projektor die Möglichkeit das Streifenmuster sowohl horizontal als auch vertikal zu projizieren, so kann der Projektor als inverse Kamera modelliert werden. Das heißt, die in der Photogrammetrie üblichen Verfahren zur Kalibrierung und Berechnung von räumlichen Koordinaten können analog auf den Projektor angewandt werden. Für den Projektor werden die Parameter der inneren und äußeren Orientierung bestimmt, außerdem werden noch die zusätzlichen Parameter wie zum Beispiel die Radialverzeichnung bestimmt (Brenner et al., 1999). Dieses Vorgehen liefert exzellente Resultate, wenn ein geometrisch stabil aufgebauter Projektor in Kombination mit einer Kamera eingesetzt wird.

Für geometrisch weniger stabile Projektionseinheiten ist eine Kombination mit mehreren Kameras vorzuziehen, bei der die Koordinaten des Projektors nicht in die Berechnung der räumlichen Koordinaten einbezogen werden. Daher muss der Projektor nicht kalibriert werden. Stattdessen wird die Streifensequenz nur dazu verwendet, die Korrespondenz zwischen den beteiligten Kameras zu lösen.

Die photogrammetrische Kalibrierung und Prozessierung hat gegenüber der häufig anzutreffenden Methode der direkten Kalibrierung den Vorteil, dass keine hoch genaue Kalibriereinheit, zum Beispiel eine Referenzebene und ein Verschiebetisch, benötigt wird. Zudem liegen die Parameter der äußeren und inneren Orientierung explizit vor und behalten ihre Gültigkeit auch außerhalb des ursprünglichen Kalibriervolumens. Die Residuen und die Kovarianz Matrizen, die auch bei der Auswertung anfallen, erlauben eine gute Selbstdiagnose des Systems.

Die oben beschriebenen Algorithmen und Verfahren wurden in mehreren Aufbauten realisiert und getestet. In einem Aufbau mit einem ABW 640 Projektor und einer hoch auflösenden Digitalkamera (Abbildung 2.7 Links) kann auf einer Fläche von  $200 \times 200 \text{ mm}^2$  ein Raster von 640 auf 640 Punkten erzeugt werden. Dies entspricht in etwa einer Auflösung von  $0,3 \text{ mm}$  am Objekt. In einem Ebenheitstest an einer Referenzebene, wie er in ähnlicher Weise auch für die Abnahme- und Überwachungsrichtlinien optischer 3D-Sensoren vorgeschlagen wurde, erzielt das System eine Ebenheit von  $0,031 \text{ mm}$  bei einer maximalen Abweichung von  $0,174 \text{ mm}$ .

Ein weiteres Beispiel ist die Integration eines Streifenprojektions-Systems in eine Multi-Sensor-Messmaschine (Böhm et al., 2001). Die Maschine, konzipiert für den automatisierten Messvorgang industrieller Teile, trägt mehrere Sensorplattformen, darunter eine Stereokamera mit einer Basis von ca.  $300 \text{ mm}$  (Abbildung 2.7 Mitte). Diese Stereo-Kamera wurde mit einem LCD-Projektor zu einem dicht messenden 3D-Sensor kombiniert. Durch die Aktorik der Maschine kann die Stereo-Kamera automatisiert und optimal zum Objekt positioniert werden.

## 2.1.2 Laufzeitverfahren

Die oben beschriebenen Verfahren haben aufgrund ihres Triangulationsprinzips einen wesentlichen Nachteil. Durch die beiden unterschiedlichen Beobachtungswinkel kann es bei Oberflächen mit größeren Höhenunterschieden zu Abschattung kommen. Bei Laufzeitverfahren hingegen ist die Richtung der Energiequelle und die Beobachtungsrichtung identisch, daher sind Abschattungen ausgeschlossen. Einen weiteren Vorteil bieten die Laufzeitverfahren hinsichtlich ihrer Messgenauigkeit im gesamten Messvolumen. Bei der Triangulation nimmt

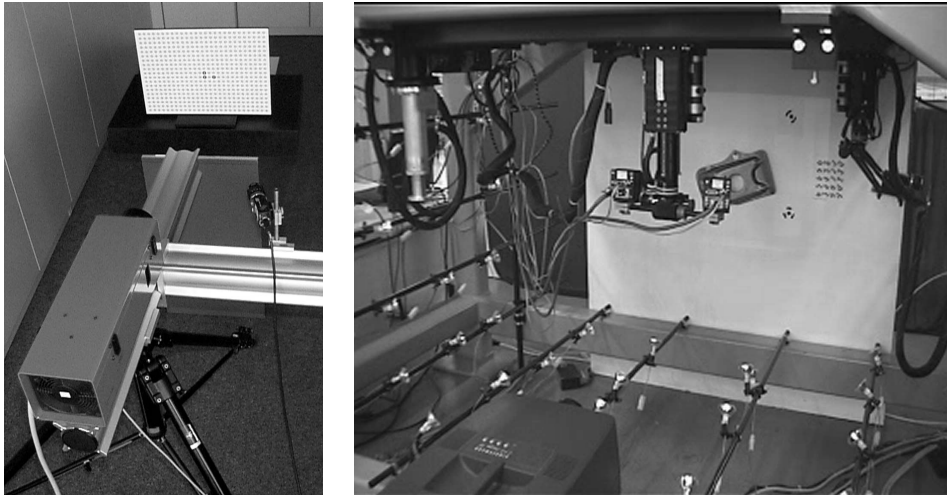


Abbildung 2.7: Aufbau und Anwendung von Streifenprojektionssystemen. Links ein Labor Aufbau, rechts ein Streifenprojektionssystem integriert in einer Multi-Sensor-Messmaschine.



Abbildung 2.8: Ein Datensatz, wie er mit der Multi-Sensor-Messmaschine aus Abbildung 2.7 erzeugt wurde.



Abbildung 2.9: Leica Laser Radar System LR 200 für mittlere Volumina in industriellen Anwendungen mit  $100\ \mu\text{m}$  Messgenauigkeit.

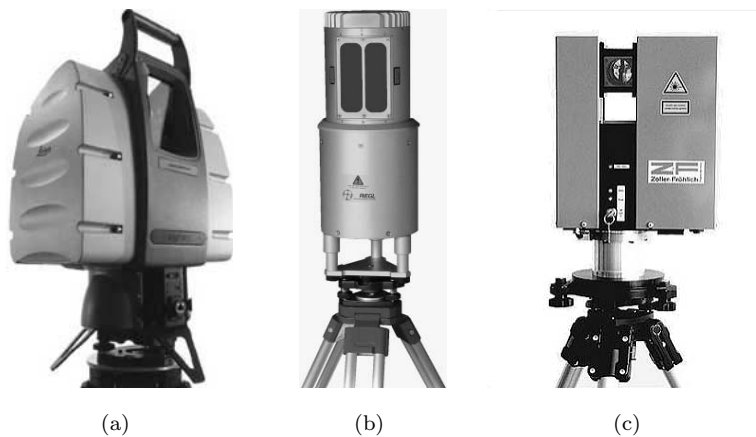


Abbildung 2.10: Typische terrestrische Laser Scanner für geodätische Anwendungen. (a) Leica HDS 3000. (b) Riegl LMS-Z420. (c) Zoller + Fröhlich Imager 5003.

die Tiefengenauigkeit bekanntermaßen quadratisch zum Abstand ab. Bei Laufzeitverfahren bleibt sie hingegen meist relativ konstant. Als Energiequellen sind vor allem zwei physikalische Prinzipien zum Einsatz gekommen, zum einen Ultraschall zum anderen Laser. Im Kontext der industriellen Messtechnik ist Ultraschall nicht relevant. Laser-Laufzeitverfahren gehören jedoch zu den sehr verbreiteten Messtechniken.

Ein wichtiger Vertreter der Laser-Laufzeitmessung ist die Puls-Laufzeitmessung. Hierbei wird die Zeit gemessen, die vergeht, bis ein ausgesandter Laserpuls von der Oberfläche zurückreflektiert wird. Die hierzu notwendige Elektronik muss in der Lage sein das Zeitintervall im Pico-Sekunden-Bereich aufzulösen. Eine weitere Methode ist die Phasenmessung. Hierzu wird der Laser amplitudenmoduliert. Die Laufzeit wird bestimmt aus der Differenz der Phasenlage des ausgesendeten zum reflektierten Signal. Die Phasenmessung bringt typischerweise eine bessere Auflösung und damit eine bessere Messgenauigkeit. Jedoch ergeben sich zusätzliche Probleme mit dem Eindeutigkeitsbereich und der Messung an Kanten. Die Genauigkeit der Distanzmessung hängt offensichtlich von der Wellenlänge des aufmodulierten periodischen Signals ab. Die Entwicklung von immer kürzeren Wellenlängen zur Steigerung dieser Genauigkeit bildet den Übergang zu den interferometrischen Distanzmessern. In jüngster Zeit wurden Geräte entwickelt die in Kombination mit aufmodulierten Signalen auch die Wellenlänge des Lichts selbst zur Phasenmessung verwenden.

Um aus einer einzelnen Distanzmessung ein flächenhaft messendes System zu bilden, muss der ausgesandte Laserstrahl über das zu vermessende Objekt bewegt werden. Hierzu werden optische Ablenkeinheiten eingesetzt. Entweder in zwei Dimensionen oder in nur einer Dimension in Kombination mit einer Verschwenkung des gesamten Sensorkopfes.

Laser-Laufzeitmessung hat sich vor allem bei der Vermessung großer Volumina etabliert, so bei der Vermessung von Bauwerken, Brücken oder Ähnlichem. Mittlerweile gibt es eine große Anzahl von Anbietern solcher Geräte,

zum Beispiel CALLIDUS, CYRA/LEICA, MENSI, RIEGL und OPTECH (siehe auch Abbildung 2.10). Aber auch zur Messung kleinerer Objekte sind Systeme nach diesem Verfahren entwickelt worden (siehe Abbildung 2.10).

### 2.1.3 Datenrepräsentation

In seiner Studie zählt J.P. Besl (1988a) eine Reihe von Namen auf, die man für Repräsentationsformen von Distanzdaten kennt: Distanzkarte, Tiefenkarte, Tiefenbild, Distanzbild, 3D-Bild,  $2\frac{1}{2}$ D-Bild, digitales Oberflächenmodell, topographische Karte, Oberflächenprofile, XYZ-Punktliste und andere. Es ist zu beachten, dass diese Begriffe nicht alle äquivalenten Repräsentationsformen beschreiben. So enthält zum Beispiel eine XYZ-Punktliste keine Information über die Nachbarschaft der Punkte, im Gegensatz zum Distanzbild, das diese Information liefert. Auch ist die Mächtigkeit der Repräsentationsformen bezüglich der mit ihr repräsentierbaren Oberflächentypen nicht immer die gleiche.

Die Repräsentation der gemessenen Daten wird in erster Linie vom Sensortyp und von der Anwendung bestimmt. Alle in den vorherigen Abschnitten aufgezählten Sensoren liefern Nachbarschaftsinformation zu den gemessenen Punkten. Es ist daher sinnvoll eine Repräsentationsform zu wählen, die diese Information speichert. Liegen die gemessenen Punkte in einem kartesischen Koordinatensystem, so spricht man von der XYZ-Form der Daten. Misst der Sensor die Entfernung entlang eines räumlichen Vektors  $r$ , der durch zwei Ganzzahlen  $i$  und  $j$  indiziert werden kann, so spricht man von der  $r_{ij}$  Form der Daten. Ist das Abtastintervall am Objekt äquidistant, so kann das gesamte Distanzbild in einer einzigen Matrix abgelegt werden. Die Matrixelemente  $(i, j)$  enthalten dann die Entfernungskomponente dieser Punkte. Wird diese Distanz entlang der  $z$ -Achse gemessen, wird die Matrix als  $Z$ -Matrix bezeichnet, oder bei einer Darstellung als Bild als  $Z$ -Bild. Die Nachbarschaft der Punkte ergibt sich über die Nachbarschaft im Raster. Die Äquidistanz der Abtastung kann aber von den meisten Sensoren nicht sichergestellt werden. Daher empfiehlt sich die Repräsentation durch drei getrennte Matrizen, jeweils für die  $X$ -,  $Y$ - und  $Z$ -Komponente der Punkte.

Die Umrechnung einer  $Z$ -Matrix eines äquidistant abgetasteten Distanzbildes in drei getrennte Matrizen ist trivial. Mit den Abtastintervallen  $s_x$  und  $s_y$  und dem Verschiebungsvektor  $\mathbf{a}$  ergibt sich für die Matrix  $\mathbf{m}$  an den Rasterelementen mit den Indices  $i$  und  $j$

$$x = a_x + s_x i \quad y = a_y + s_y j \quad z = a_z + s_z m_{ij}$$

Der umgekehrte Weg indes ist schwieriger und unter Umständen nicht verlustfrei durchzuführen. Zum einen bedarf es der Interpolation von Werten, was zu Ungenauigkeiten führen kann. Zum anderen ist es bei einer Darstellung in der Form einer einzigen  $Z$ -Matrix unbedingt erforderlich, dass pro  $(X, Y)$  Tupel nur ein einziger  $Z$ -Wert abgelegt ist, also nur eine Graph-Fläche vorliegt. Dies ist bei der Darstellung in drei getrennten Matrizen nicht unbedingt erforderlich. Diese Darstellungsform ist also mächtiger, da sie eine größere Vielfalt von Flächentypen darstellen kann. Im folgenden wird für den Begriff Distanzbild immer diese Repräsentationsform in drei getrennten Matrizen angenommen.

Das Prinzip einiger Sensoren bringt es mit sich, dass sie die Punkte in einem sphärischen Koordinatensystem messen. Es ist ebenfalls offensichtlich möglich, diese Daten in ein kartesisches Koordinatensystem umzurechnen. Mit dem Abtastintervall  $s_\phi$  des Elevationswinkels, dem Abtastintervall  $s_\theta$  des Azimuthwinkels, dem Verschiebungsvektor  $\mathbf{a}$  und der Matrix der gemessenen Distanzen  $\mathbf{m}$  ergibt sich

$$\begin{aligned} x &= a_x + m_{ij} \cos(is_\phi) \sin(js_\theta) \\ y &= a_y + m_{ij} \sin(is_\phi) \\ z &= a_z + m_{ij} \cos(is_\phi) \cos(js_\theta) \end{aligned}$$

Ebenso lassen sich zylindrische Koordinaten mancher Sensortypen umrechnen. Komplexere Messergebnisse, insbesondere von Systemen, die Resultate aus mehreren Einzelmessungen fusionieren, lassen sich im Allgemeinen nur noch als Punktwolke oder in triangulierter Form repräsentieren.

Die Repräsentation als Matrix, und damit als Bild, hat Vorteile bei der Bearbeitung der Daten. Zum einen lassen sich direkt bestimmte Operatoren aus der klassischen Bildverarbeitung anwenden, insbesondere Filteroperatoren. Zum anderen kann in der Matrixrepräsentation die Nachbarschaft der Punkte in der Regel schneller erfasst werden. Dies kann durch einfache Maskenoperatoren erfolgen. In der Triangulation werden nur die direkten Nachbarn eines Punktes erfasst. Indirekte Nachbarschaften müssen erst abgeleitet werden. Hierdurch ergibt sich ein klarer Geschwindigkeitsvorteil der Matrizen, immer dann, wenn häufig die Umgebung eines Punktes betrachtet werden muss.

## 2.2 Geometrisches Modellieren

Der Begriff geometrisches Modellieren, oder der englischsprachige Fachbegriff *Geometric Modeling*, stammt aus den Fachbereichen der Computergrafik, dem computergestützten Entwurf (CAD) und der Produktionstechnologie. Der Begriff umfasst dabei eine Gruppe von Verfahren, welche die exakte Beschreibung der Form und anderer geometrischer Merkmale eines Körpers ermöglichen. Er schließt somit auch die Bereiche der *Computational Geometry* und des *Solid Modeling* mit ein. Das geometrische Modellieren bedient sich dabei einer ganzen Reihe von Techniken aus unterschiedlichen Bereichen, wie zum Beispiel der Geometrie, der Topologie, der Numerik und der linearen Algebra, um nur einige zu nennen. Aus dieser Kombination hat sich mittlerweile ein eigenständiges Forschungsgebiet etabliert.

Waren die Computergrafik, das CAD und die computergestützte Produktion (CAM) die ursprünglichen treibenden Kräfte in der Entwicklung des geometrischen Modellierens, wächst neben dem erfolgreichen Einsatz in diesen Kerngebieten auch die Anzahl der Anwendungen in anderen Gebieten. Hier sind insbesondere die Robotik, das maschinelle Sehen und allgemein die künstliche Intelligenz zu nennen. Drei wesentliche Aspekte des geometrischen Modellierens lassen sich dabei identifizieren: die Repräsentation, also die mathematische Beschreibung der Geometrie, der Entwurf, die möglichst intuitive Erzeugung von Geometrie und das Rendering, die Erzeugung eines Abbildes der Geometrie.

In der Computergrafik steht die Generierung realistischer zweidimensionaler Abbilder geometrisch modellierter Objekte im Vordergrund. Durch die Integration von Textur und Oberflächeneigenschaften wie Reflexion und die Generierung hoch detaillierter Modelle, können Bilder auch in Echtzeit erzeugt werden, die von der Realität kaum zu unterscheiden sind.

Ziel im CAD ist die Weiterentwicklung und Verbesserung der Entwurfssysteme. Neben dem bloßen Ersatz der Zeichenbretter tritt hier die parallel zum Entwurf mögliche Analyse des Modells in den Vordergrund. Dies betrifft nicht nur die Analyse der geometrischen Eigenschaften, also zum Beispiel die Aufdeckung topologischer Fehler, sondern auch die Analyse physikalischer Eigenschaften, die in modernen hoch entwickelten CAD-Systemen angeboten wird. Sogar die Analyse dynamischer Eigenschaften eines Objektes ist möglich. Somit können Entwurfsfehler bereits frühzeitig aufgedeckt werden.

Neue Möglichkeiten bietet das geometrische Modellieren auch im Bereich der Produktionstechnik. Durch die Verwendung eines exakten Modells ist die vollautomatische Planung ganzer Fertigungsabläufe möglich. Dabei wird im Modell nicht nur die geometrische Form des Objektes hinterlegt, sondern auch weiterreichende geometrische Eigenschaften wie Toleranzen und nicht-geometrische Eigenschaften wie Materialeigenschaften. Die Weiterentwicklung dieser Möglichkeiten führt zum Konzept der virtuellen Fabrik, in der alle Abläufe simuliert werden können und zur vollautomatischen Inspektion aller Teile.

Für den Bereich der Robotik ist der Aspekt der Repräsentation im geometrischen Modellieren wichtig. Bestehende Modelle der räumlichen Umgebung eines Roboters können dazu verwendet werden automatische Bahnplanungen durchzuführen, Kollisionen zu vermeiden oder mit Objekten zu interagieren. Genauso können durch Sensoren gewonnene geometrische Informationen in Modellen abgelegt werden oder bestehende Modelle ergänzt werden.

Ebenso wird bei Anwendungen im maschinellen Sehen vorgegangen. Aus den Sensoren extrahierte und durch Algorithmen aufbereitete geometrische Information wird in geeigneten geometrischen Modellen abgelegt. Bestehende geometrische Modelle können dabei den Extraktionsprozess stützen, Ergebnisse verifizieren oder eine Interpretation der Ergebnisse auf höherer Abstraktionsebene ermöglichen. Bekannte und wichtige Arbeiten aus dem Bereich des künstlichen Sehens, die auf geometrischen Modellen beruhen werden in Abschnitt 2.3.1 noch einmal aufgegriffen.

### 2.2.1 Solid Modeling

Die Entwicklung des geometrischen Modellierens begann mit einfachen zweidimensionalen Vektorzeichnungen. Diese wurden bald zu dreidimensionalen Systemen erweitert. Repräsentationen für gewölbte Kurven und Flächen wurden entwickelt und in die Systeme integriert, wie zum Beispiel *Coons Patches*, Bezierflächen und NURBS. Eine wichtige Entwicklungsstufe des geometrischen Modellierens ist das so genannte Solid Modeling, also die Modellierung eines festen Körpers.

Ein Solid Model ist die eindeutige und vollständige mathematische Beschreibung eines physikalischen Körpers, in einer Art und Weise, die einem Computer zugänglich ist (Mortenson, 1997). Es werden also nicht Teilaspekte, Teilflächen oder Teilansichten repräsentiert, sondern die geschlossene Beschreibung eines Körpers. Aus den

Methoden zur Konstruktion eines Solid Models lassen sich sechs Hauptkategorien unterscheiden: parametrisierte Körper, Zellzerlegung, Rotationskörper oder allgemeiner *Sweep Representation*, CSG (*Constructive Solid Geometry*), Randbeschreibung, oder mit dem englischsprachigen Fachbegriff *Boundary Representation*, und Drahtgitter-Modelle. Zu den wichtigsten Vertretern im CAD gehören heute CSG und die Randbeschreibung.

Für diese Arbeit ist vor allem die Randbeschreibung von Bedeutung. Dies lässt sich zum einen durch die im Abschnitt 2.1 beschriebenen Sensoren motivieren, die eben genau die Randfläche eines Körpers zu seiner Umgebung erfassen. Zum anderen ergibt sich keine Einschränkung in der Wahl der CAD-Technologie, da auch moderne CSG-Systeme in der Lage sind, aus ihrer internen Repräsentation eine Randbeschreibung abzuleiten.

## 2.2.2 Randbeschreibung

Die Randfläche eines Körpers trennt die Punkte, die innerhalb des Körpers liegen von den außerhalb liegenden ab. Die Randfläche definiert wesentliche Eigenschaften eines Körpers. Die Eigenschaften wie Rauheit, Oberflächentextur und Streuverhalten bestimmen seine optische Erscheinung. Typische Bearbeitungsprozesse eines Werkstücks finden an dessen Randfläche statt.

Nach Requicha (1980) gelten für die Randfläche folgende Bedingungen:

1. Der Körper wird durch eine endliche Anzahl von Flächen begrenzt.
2. Jede Fläche eines Körpers ist ein Teil seiner Randfläche.
3. Die Vereinigungsmenge aller Flächen eines Körpers bildet seine Randfläche.
4. Jede Fläche ist selbst wiederum Teil einer größeren Fläche.
5. Jede Fläche besitzt einen endlichen Flächeninhalt und ihre Dimension entspricht der der Randfläche.

Jedes Flächenstück sollte für sich geschlossen, durch eine Randkurve begrenzt, orientierbar, nicht selbst-überschneidend und zusammenhängend sein. Man kann zum Beispiel einen einfachen Würfel in seine natürlichen sechs Teilflächen untergliedern. Alle sechs Flächen zusammen bilden die Randfläche des Würfels. Jede der Teilflächen ist durch eine Randkurve in der Form eines Quadrats begrenzt und jeder Punkt auf einer solchen Teilfläche kann mit einem anderen Punkt dieser Teilfläche durch eine Kurve verbunden werden, ohne diese Teilfläche zu verlassen. Jede der Teilflächen kann bei Bedarf wiederum in mehrere Flächen untergliedert werden.

### 2.2.2.1 Mehrdeutigkeit der Zerlegung einer Randbeschreibung

Jede Randbeschreibung beschreibt eindeutig einen Körper. Dies bedeutet, dass auch die Vereinigung jeder Zerlegung einer Randbeschreibung wiederum eindeutig den Körper beschreibt. Jedoch gibt es zu ein und demselben Körper mehrere Möglichkeiten seine Randbeschreibung zu zerlegen. Damit gibt es mehrere Möglichkeiten der Randbeschreibung desselben Körpers.

Bei bestimmten Gruppen von Körpern, zum Beispiel bei Polyedern (siehe Abbildung 2.11) kann eine eindeutige Zerlegung in maximal zusammenhängende Teilflächen erzielt werden. Jedoch stehen bei der weiteren Unterteilung dieser Teilflächen wiederum mehrere Möglichkeiten offen.

Besonders bei Körpern, die überwiegend aus gekrümmten Flächen bestehen, ist die Mehrdeutigkeit der Zerlegung offensichtlich. In Abbildung 2.12 ist als Beispiel für einen solchen Körper ein Glas gegeben. Die Randfläche des Körpers kann nach funktionalen Gesichtspunkten, also in Standfläche, Stiel und das eigentliche Behältnis unterteilt werden (links). Die Randfläche kann aber auch nach konstruktiven Gesichtspunkten in konkave und konvexe Teilflächen unterteilt werden (rechts).

Diese Fragen bei der Zerlegung einer Randfläche sind die wesentliche Ursache für die Schwierigkeiten bei der Segmentierung von Distanzbildern die Objekte mit gekrümmten Flächen enthalten. Hierauf wird in Kapitel 4 nochmals verwiesen.



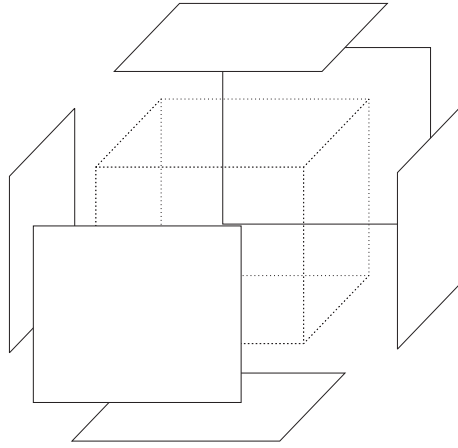


Abbildung 2.11: Eindeutige Zerlegung eines Würfels in seine sechs Teilflächen.

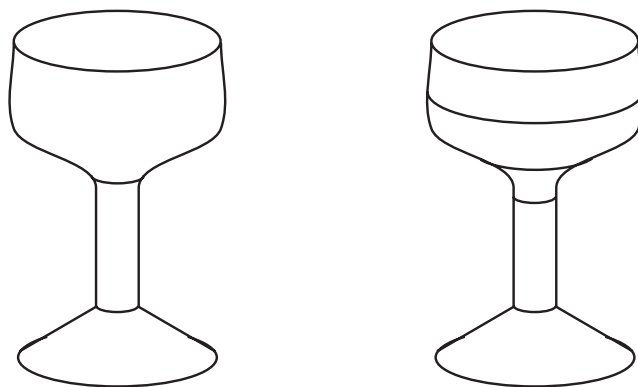


Abbildung 2.12: Mehrere gültige Unterteilungen der Randfläche für denselben Körper.

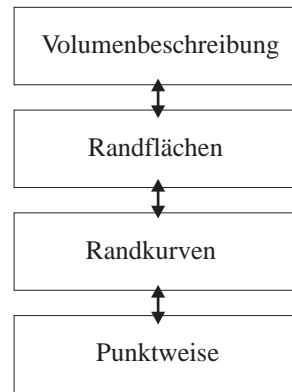


Abbildung 2.13: Ordnung der Repräsentationsformen nach ihrer Dimension.

## 2.3 Objekterkennung

Die Objekterkennung beschäftigt die Wissenschaft bereits seit langem. Entsprechend groß ist die Anzahl der hierzu veröffentlichten Arbeiten. Daher soll an dieser Stelle nur ausschnittsweise ein Einblick in alternative Systeme und deren Verfahrensweisen gegeben werden. Die Auswahl beschränkt sich dabei auf Systeme, die mit räumlichen Daten arbeiten. Hiervon werden einige besonders typische und wichtige Vertreter vorgestellt. Ausführliche Studien wurden von Besl und Jain (1985) und Arman und Aggarwal (1993) veröffentlicht.

### 2.3.1 Repräsentationsformen zur Objekterkennung

Um die dreidimensionale Form eines Objektes im Computer zu beschreiben, muss eine bestimmte Repräsentation gewählt werden. Eine geeignete Repräsentationsform sollte eine Reihe von Kriterien erfüllen. Zum einen sollte die Repräsentation möglichst *vollständig* sein, das heißt, es muss möglich sein aus der Repräsentation die Ausgangsform des Objektes hinreichend genau rekonstruieren zu können. Die Repräsentation sollte außerdem *stabil* sein. Diese Anforderung soll sicherstellen, dass eine kleine Änderung der Daten, etwa hervorgerufen durch Sensorrauschen, nicht zu einer stark veränderten Modellbeschreibung führen. Dies würde den späteren Erkennungsprozess unnötig erschweren. Es ist außerdem erstrebenswert eine möglichst *natürliche* Repräsentationsform zu wählen. Dies erleichtert den Zugang zu den Modellen und ermöglicht damit auch eine Benutzerinteraktion im Fehlerfall. Wie für jedes EDV-System ist auch für ein Objekterkennungssystem die Effizienz wichtig. Dies bedeutet, dass die Repräsentation eines Objektes nur einen begrenzten Rechen- und Speicherbedarf entwickeln sollte. Versucht man alle Kriterien gleichzeitig zu berücksichtigen, erkennt man schnell, dass sich die Anforderungen teilweise widersprechen. So wird etwa ein System, das eine besonders vollständige Repräsentation besitzt automatisch einen höheren Speicher- und Rechenbedarf aufweisen.

Abbildung 2.3.1 ordnet die Repräsentationsformen in Stufen nach ihrer Dimension. Während die Repräsentationsformen auf niedriger Stufe besonders einfach zu berechnen sind, sind sie meist weder vollständig noch besonders stabil. Systeme welche eine geometrisch motivierte Repräsentationsform nutzen, basieren oft auf den Methoden des Solid Modeling, wie sie in Abschnitt 2.2.1 beschrieben wurden. Wobei das Solid Modeling dabei ein Überbegriff für Repräsentationen auf verschiedenen Stufen ist. Darüber hinaus wurden jedoch für Objekterkennungssysteme eine Reihe spezieller Verfahren entwickelt, die unabhängig vom Bereich des CAD existieren.

Auf der höchsten Stufe stehen die Systeme, die auf einer Volumenbeschreibung aufbauen. Parametrisierte Körper und CSG werden dort eingesetzt, wo sich die Erkennung auf eine eingeschränkte Menge von Formen konzentriert, zum Beispiel bei der Erkennung von Gebäuden aus luftgestützten Daten, wie von Brenner (2000) beschrieben.

Methoden der Zellzerlegung, die ebenfalls zu den volumetrischen Methoden zählen, sind die Voxelrepräsentation, also eine einfache Zerlegung des Volumens in diskreten Schritten entlang der Koordinatenachsen, *Octtrees* und KD-Bäume. Die letzteren beiden bieten eine effizientere Speicherung des Volumens. Da die Zerlegung jedoch nicht entlang oder an bestimmten Merkmalen erfolgt, sondern in hierarchisch gewählten Schritten, können solche Repräsentation nur schwer für eine merkmalsbasierte Zuordnung verwendet werden. Sie eignen sich vielmehr für eine Suche über alle möglichen Positionen.

Von Binford (1971) wurden die generalisierten Zylinder (*generalized cylinders*) eingeführt. Eine Methode, bei der eine ebene und geschlossene Kurve entlang einer Bahn durch den Raum gezogen wird. Die erzeugende Kurve kann

als Querschnitt der entstehenden Oberfläche verstanden werden, wobei zu beachten ist, dass in der allgemeinen Form die Kurve entlang der Bahn beliebigen Veränderungen unterworfen werden kann. Generalisierte Zylinder aus Sensordaten zu extrahieren ist eine schwierige Aufgabe. Generell eignet sich diese Repräsentationsform zur Darstellung lang gestreckter Objekte, aber ist unpassend für allgemeine Körper.

Barr (1981) führte Superquadriken (*superquadrics*) in das maschinelle Sehen ein, eine mathematische Beschreibung einer geschlossenen Randfläche. Durch zahlreiche Parameter kann die Fläche in ihrer Ausdehnung, Rundheit und Stauchung verändert werden. Hierbei können jedoch nur schwer lokale Änderungen eingebracht werden. Zudem ist die Flächenrückführung kein triviales Problem, und in der Regel nur iterativ mit guten Startwerten möglich. Da eine Superquadrik mangels lokaler Repräsentationsmöglichkeit ein Objekt nur approximieren kann, ist die Repräsentation nicht eindeutig. Das heißt, ein und dasselbe Objekt kann von mehreren Superquadriken mit verschiedenen Parametern approximiert werden. Daher werden Superquadriken vor allem zur Modellierung von Formen aber kaum zur Objekterkennung eingesetzt.

Auf der nächsten Stufe der Beschreibungshierarchie stehen Flächenbeschreibungen. Hierbei unterscheidet man zwischen globalen und segmentierten Flächenbeschreibungen. Globale Beschreibungen werden in einem Schritt für das gesamte Objekt aus den Sensordaten extrahiert. Segmentierte Beschreibungen erfordern zuerst eine Segmentierung, das heißt eine Zerlegung der Eingangsdaten.

Zu den globalen Flächenbeschreibungen zählt die Gauß-Abbildung (*gaussian sphere*). Hierbei wird die Orientierung des Normalenvektors an einem Punkt auf einen korrespondierenden Punkt auf der Einheitskugel abgebildet (do Carmo, 1976). Die Repräsentation ist rotationsinvariant. Sie ist jedoch nur für konvexe Objekte eindeutig und zudem werden die Skalierung und die Translation des Objektes nicht repräsentiert. Es existieren erweiterte Varianten (*extended gaussian image*), die diese Einschränkungen zum Teil aufheben (Horn, 1984). Grundsätzlich ist zu erwarten, dass globale Flächenbeschreibungen empfindlich gegenüber Abschattung und Überlappung sind.

Segmentierte Flächenbeschreibungen hingegen arbeiten mit lokaler Information und können somit lokale Defekte besser abfangen. Zahlreiche Verfahren zur Segmentierung von Distanzdaten wurden von verschiedenen Forschergruppen vorgeschlagen. Die wichtigsten Verfahren werden in Kapitel 4 beschrieben.

Kurven- und punktbasierte Repräsentationen speichern die Information entlang den Diskontinuitäten der Flächen. Speziell für Randkurven einer Fläche führt dies zu Problemen, wenn zwei benachbarte Flächen einen stetigen Übergang besitzen. Während die Kante im Modell repräsentierbar ist, kann im allgemeinen nicht erwartet werden, dass die so genannten glatten Kanten aus Distanzdaten extrahiert werden können. Thirion (1993b) entwickelte die *Extremal Meshes*, eine Repräsentation für dreidimensionale Flächen aus Punkten und Kanten. Die Punkte sind dabei Punkte der Fläche, an denen beide Krümmungen (siehe Kapitel 3) Extremwerte annehmen. Das Netz aus diesen Extrempunkten wird vervollständigt durch die so genannten Extremkurven, Kurven auf der Fläche, an denen eine der beiden Krümmungen einen Extremwert annimmt (Thirion, 1993a). Die Repräsentation wurde entwickelt zur Registrierung von medizinischen dreidimensionalen Bilddaten.

### 2.3.2 Systeme zur Objekterkennung

In der Vergangenheit wurden bereits eine Reihe von Objekterkennungssystemen für Distanzdaten entwickelt. Zu meist handelt es sich hierbei um experimentelle Systeme, die aus der Forschung stammen. Eine Auswahl aus diesen Systemen soll nun vorgestellt werden.

Bolles und Horaud (1986) entwickelten mit 3DPO ein System zur Greifplanung von räumlichen Objekten. Die Modellbildung erfolgt in einem attributierten CAD-Modell, in dem besonders markante Merkmale gekennzeichnet werden. Das System ist stellvertretend für eine typische Klasse von Systemen, bei denen ein explizites geometrisches Modell verwendet wird. Die daraus extrahierten Merkmale werden den aus der Szene extrahierten Merkmalen durch ein Suchverfahren zugeordnet. Die Merkmale, die für 3DPO Verwendung finden sind die Kanten des Objektes. Hierbei wird zwischen geraden und zylindrischen Kanten unterschieden. Somit wird die Menge der behandelbaren Objekte deutlich eingeschränkt.

Greenspan und Boulanger (1999) präsentierten ein System, das eine endliche Anzahl von Ansichten eines Objektes in ein Voxelraaster abbildet. Diese Voxelrepräsentation wird als Template für die Erkennung verwendet, indem das Template in verschiedenen Positionen über die Distanzdaten gelegt wird. Das vorgestellte System arbeitet bei der Erkennung also im Suchraum, der aus allen möglichen Positionen und Orientierungen des Objektes besteht. Das herausragende Merkmal der Arbeit besteht in der effizienten Speicherung der Voxeltemplates in einem Suchbaum, der den Aufwand für die Suche nach der korrekten Lage minimiert. Dennoch bleibt der Suchraum hinreichend groß, sodass die Autoren planten, die Methode auf einem Rechnernetzwerk mit mehreren Prozessoren zu implementieren.

Von Fan (1989) wurde ein System vorgeschlagen, das auf dem Regionen-Adjazenz-Graphen (RAG) beruht. Es stellt einen Kompromiss aus einer vollständigen dreidimensionalen Repräsentation und einem ansichtsbasierten System dar. Ein Objekt wird durch mehrere Ansichten repräsentiert, wobei in jeder Ansicht das Objekt in seine Teilflächen zerlegt wird. Die Teilflächen werden in einem Adjazenzgraphen gespeichert. Die Distanzbilder der Szene werden segmentiert und die gewonnenen Regionen mit den Teilflächen aus dem RAG zur Zuordnung gebracht. Das System wurde auf Szenen mit mehreren Objekten und Verdeckungen erfolgreich getestet. Wobei die Anzahl der gespeicherten Ansichten mit vier bis sechs sehr gering erscheint für Objekte, deren Erscheinung sich unter unterschiedlichem Blickwinkel stark ändert.

Ein System, welches sich weitestgehend von geometrischen Vorstellungen löst, wurde von Campbell und Flynn (1999) vorgestellt. Das System arbeitet ansichtsbasiert, das heißt, in einer Trainingsphase werden mehrere Ansichten eines Objektes abgespeichert. Die erzeugten Distanzbilder werden als ein Merkmalsvektor aufgefasst. Der Merkmalsraum wird anschließend durch eine Eigenvektorzerlegung verkleinert. Die Autoren bezeichnen das Ergebnis eines solchermaßen transformierten Distanzbildes als *Eigenshape*. In der Erkennungsphase wird das eingehende Distanzbild gleichermaßen transformiert und im Merkmalsraum in einer Nächste-Nachbar-Klassifikation dem Modell mit dem geringsten Abstand zugewiesen. Die Ansichten werden in einem zweidimensionalen Parameterraum der Rotationen bestimmt. Um einen Abstand von zum Beispiel  $10^0$  einzuhalten, müssen 362 Ansichten generiert werden. Das System zeigt die Fähigkeit, mit größeren Objektdatenbanken zu arbeiten. Ein ebenfalls ansichtsbasiertes System wurde von Hetzel et al. (2001) vorgestellt. Das System arbeitet auf einer Datenbank von 30 Objekten und speichert dabei insgesamt nahezu 2000 Ansichten, bei etwa  $25^0$  Abstand in den Winkeln. Jede Ansicht wird durch einen Merkmalsvektor aus Intensitäts-, Normalen- und Krümmungshistogramm repräsentiert. Das System arbeitet selbst bei einer Vielzahl von Objekten noch sehr effizient. Die maßgebliche Motivation für ansichtsbasierte Systeme, liegt darin, dass so eine Segmentierung der Eingangsdaten vermieden werden kann. Eines der Hauptprobleme liegt in der Frage, welche Ansichten beziehungsweise wie viele Ansichten für ein Objekt gespeichert werden müssen, um das Objekt vollständig zu repräsentieren.

## 3 Geometrie der Fläche

### 3.1 Einleitung

In diesem Kapitel werden die Grundlagen zur Behandlung von Flächen gelegt. Die Definition und Beschreibung von Flächen werden ebenso erläutert wie charakteristische Eigenschaften von Flächen. Wichtige Methoden zur Charakterisierung von Flächen liefert die Differentialgeometrie, ein äußerst komplexes Themengebiet, aus dem hier nur sehr grundlegende Ausschnitte dargestellt werden können.

Während die Differentialgeometrie Hinweise zur Behandlung von Flächen liefert, die in einer mathematischen Beschreibung vorliegen, ist für das räumliche maschinelle Sehen die Behandlung von rohen Punktdaten ebenso wichtig. Der Abschnitt über die lokale Flächenrückführung liefert den Übergang von rohen Punktdaten zur mathematischen Beschreibung. Er ermöglicht es somit, die aus dem ersten Teil gewonnenen Analysemethoden auf Punktdaten zu übertragen.

Flächen im Raum, die uns in der praktischen Erfahrung vor allem als Oberflächen eines Körpers oder Objektes begegnen, lassen sich sowohl als zusammenhängende Menge von Punkten auffassen, aber auch als eine Abbildung eines zweidimensionalen Parameterraums in den dreidimensionalen Raum. Aus diesen Grundgedanken ergeben sich in der Mathematik mehrerer Möglichkeiten, Flächen mathematisch zu beschreiben. Als erstes wird die Beschreibung als Abbildung erläutert.

Eine Fläche  $S$  im dreidimensionalen Raum ist gegeben durch die Abbildung einer offenen Menge  $U$  aus  $\mathbb{R}^2$  in den dreidimensionalen Raum  $\mathbb{R}^3$  mit

$$X : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3 \quad (3.1)$$

Die Funktion dieser Abbildung lässt sich als Vektor aufschreiben in der Form

$$\mathbf{X}(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix}, \quad (u, v) \in U \quad (3.2)$$

Die Funktion  $X$  ist differenzierbar und für alle  $q \in U$  ist das Differential  $dX_q$  injektiv (do Carmo, 1976). Durch die Beschreibung der Fläche in Abhängigkeit des Parameterraums  $U$  wird diese Darstellung als parametrisierte Darstellung bezeichnet. Ohne Beschränkung der Allgemeinheit kann das Intervall der Parameter  $(u, v)$  auf  $[0, 1] \times [0, 1]$  beschränkt werden.

Hält man einen der beiden Parameter auf einem bestimmten Wert fest, erhält man eine Kurve auf der Fläche  $S$  parametrisiert nach dem jeweils anderen Parameter. Wiederholt man diesen Vorgang mehrere Male zunächst für den einen und dann für den anderen Parameter, erhält man ein Netz von Kurven auf der Fläche. Ebenso einfach lässt sich ein Gitter von Punkten aus der Darstellung ableiten.

Ein einfaches Beispiel für eine parametrisierte Flächendarstellung lässt sich für die Kugel aufstellen. Eine Kugel mit dem Radius  $r$  und dem Mittelpunkt  $(x_0, y_0, z_0)$  wird dargestellt durch

$$\begin{aligned} x(u, v) &= x_0 + r \cos u \cos v \\ y(u, v) &= y_0 + r \cos u \sin v \\ z(u, v) &= z_0 + r \sin u \end{aligned} \quad u \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], v \in [0, 2\pi] \quad (3.3)$$

Hält man den Parameter  $v$  fest, ergeben sich für freies  $u$  die Kreise der Breitengrade auf der Kugel und entsprechend die Längengrade, wenn man  $u$  festhält. Die Abbildung 3.1 zeigt eine Kugel mit den entsprechenden Kurven.

Bei der Betrachtungsweise einer Fläche als Menge von Punkten, kann eine Fläche definiert werden als die Menge der Nullstellen eines impliziten Ausdrucks mit

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid F(x, y, z) = 0\} \quad (3.4)$$

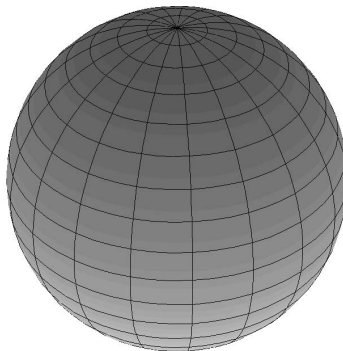


Abbildung 3.1: Darstellung einer nach Gleichung 3.3 parametrisierten Kugel mit den Kreiskurven entlang der Parameter.

Diese Darstellungsform wird als implizite Darstellung bezeichnet. Jede Fläche ist in dieser Form darstellbar.

Wird der implizite Ausdruck nach einer der drei Variablen aufgelöst, zum Beispiel nach  $z$  zu  $z = F(x, y)$ , erhält man die Darstellung einer Fläche als Menge von Punkten, die einen expliziten Ausdruck erfüllen mit

$$S := \{(x, y, z) \in \mathbb{R}^3 \mid z = F(x, y)\} \quad (3.5)$$

Diese Darstellung wird analog als explizite Darstellung bezeichnet. Die resultierenden Flächen sind auch als Graphflächen bekannt.

## 3.2 Differential Geometrie

Eine Fläche  $X$  im dreidimensionalen Raum sei gegeben durch ihre parametrisierte Darstellung über dem zweidimensionalen Parameterraum  $U \subset \mathbb{R}^2$  durch die Vektorfunktion

$$\mathbf{X}(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix}, \quad (u, v) \in U \quad (3.6)$$

Die partiellen Ableitungen  $\frac{\partial \mathbf{X}}{\partial u}$  und  $\frac{\partial \mathbf{X}}{\partial v}$  werden abgekürzt mit  $\mathbf{X}_u$  und  $\mathbf{X}_v$ . Mit der Vektorfunktion  $\mathbf{X}$  sind diese partiellen Ableitungen einfach darzustellen als

$$\mathbf{X}_u = \left( \frac{\partial x(u, v)}{\partial u}, \frac{\partial y(u, v)}{\partial u}, \frac{\partial z(u, v)}{\partial u} \right), \quad \mathbf{X}_v = \left( \frac{\partial x(u, v)}{\partial v}, \frac{\partial y(u, v)}{\partial v}, \frac{\partial z(u, v)}{\partial v} \right) \quad (3.7)$$

Eine gegebene Fläche in ihrer expliziten Form lässt sich leicht durch

$$\mathbf{X} = (u, v, F(u, v)) \quad (3.8)$$

in eine Vektorfunktion umformen. Aus dieser Umformung erhält man für die partiellen Ableitungen

$$\mathbf{X}_u = \left( 1, 0, \frac{\partial F(u, v)}{\partial u} \right), \quad \mathbf{X}_v = \left( 0, 1, \frac{\partial F(u, v)}{\partial v} \right) \quad (3.9)$$

Es ist wichtig anzumerken, dass jede Fläche lokal durch eine explizite Funktion beschrieben werden kann.

### 3.2.1 Fundamentalformen

Für die differenzierbare Fläche  $X$  soll außerdem noch die Bedingung  $\mathbf{X}_u \times \mathbf{X}_v \neq 0$  erfüllt sein. Der Einheitsnormalenvektor  $\mathbf{N}$  ist dann gegeben durch

$$\mathbf{N} = \frac{\mathbf{X}_u \times \mathbf{X}_v}{\|\mathbf{X}_u \times \mathbf{X}_v\|} \quad (3.10)$$

Der Ausdruck

$$\begin{aligned} I &= d\mathbf{X} \cdot d\mathbf{X} \\ &= (\mathbf{X}_u du + \mathbf{X}_v dv) \cdot (\mathbf{X}_u du + \mathbf{X}_v dv) \\ &= Edu^2 + 2Fdudv + Gdv^2 \end{aligned} \quad (3.11)$$

mit

$$E = \mathbf{X}_u \cdot \mathbf{X}_u, \quad F = \mathbf{X}_u \cdot \mathbf{X}_v, \quad G = \mathbf{X}_v \cdot \mathbf{X}_v \quad (3.12)$$

heißt erste Fundamentalform.

Der Ausdruck

$$\begin{aligned} II &= -d\mathbf{X} \cdot d\mathbf{N} \\ &= -(\mathbf{X}_u du + \mathbf{X}_v dv) \cdot (\mathbf{N}_u du + \mathbf{N}_v dv) \\ &= Ldu^2 + 2Mdudv + Ndv^2 \end{aligned} \quad (3.13)$$

mit

$$L = -\mathbf{X}_u \cdot \mathbf{N}_u, \quad M = -\frac{1}{2}(\mathbf{X}_u \cdot \mathbf{N}_v + \mathbf{X}_v \cdot \mathbf{N}_u), \quad N = -\mathbf{X}_v \cdot \mathbf{N}_v \quad (3.14)$$

heißt zweite Fundamentalform. Es gilt alternativ

$$L = \mathbf{X}_{uu} \cdot \bar{\mathbf{N}}, \quad M = \mathbf{X}_{uv} \cdot \mathbf{N}, \quad N = \mathbf{X}_{vv} \cdot \mathbf{N} \quad (3.15)$$

### 3.2.2 Krümmungen

Aus den beiden Fundamentalformen lassen sich die beiden Hauptkrümmungen ableiten:

- die maximale Krümmung  $k_1$  und
- die minimale Krümmung  $k_2$

$$k_{1,2} = H \pm \sqrt{H^2 - K} \quad (3.16)$$

mit

$$\begin{aligned} K &= \frac{LN - M^2}{EG - F^2} \\ H &= \frac{EN + GL - 2FM}{2(EG - F^2)} \end{aligned} \quad (3.17)$$

Sie sind translations- und rotationsinvariant. Die zu den Hauptkrümmungen gehörenden Hauptkrümmungsrichtungen stehen immer senkrecht zueinander.

### 3.2.3 Abgeleitete Größen

Alle Größen, die sich aus Parametern ableiten lassen, die invariant gegen Translation und Rotation sind, sind wiederum invariant.

#### 3.2.3.1 Gaußsche Krümmung

Das Produkt von  $k_1$  und  $k_2$

$$K = k_1 \cdot k_2 \quad (3.18)$$

wird als Gaußsche Krümmung bezeichnet und wird mit  $K$  abgekürzt.

### 3.2.3.2 Mittlere Krümmung

Der Mittelwert von  $k_1$  und  $k_2$

$$H = \frac{k_1 + k_2}{2} \quad (3.19)$$

wird als mittlere Krümmung bezeichnet und mit  $H$  abgekürzt.

### 3.2.3.3 Abweichung von der Ebenheit

Ein Kriterium, um die Ebenheit einer Fläche abzuschätzen, ist die Abweichung der Ebenheit mit

$$k_1^2 + k_2^2$$

Sie charakterisiert die lokale Unebenheit und ist geeignet ebene Flächenstücke zu identifizieren. Darüber hinaus liefert der Term eine Größe, um über die Stärke der Gekrümmtheit eines Flächenstücks zu entscheiden, im Gegensatz zu den einzelnen Hauptkrümmungen, bei denen immer beide Größen gemeinsam betrachtet werden müssen.

## 3.2.4 Lokale Flächenklassifikation

Anhand der Hauptkrümmungen kann folgendes Flächenklassifikationsschema entwickelt werden:

	$k_1 < 0$	$k_1 = 0$	$k_1 > 0$
$k_2 < 0$	negativ elliptisch	negativ zylindrisch	Sattelfläche
$k_2 = 0$	-	Ebene	positiv zylindrisch
$k_2 > 0$	-	-	positiv elliptisch

Verwendet man stattdessen  $H$  und  $K$  erhält man folgendes Schema (Besl, 1988b):

	$K < 0$	$K = 0$	$K > 0$
$H < 0$	negative Sattelfläche	negativ zylindrisch	negativ elliptisch
$H = 0$	Minimalfläche	Ebene	-
$H > 0$	positive Sattelfläche	positiv zylindrisch	positiv elliptisch

Der Unterschied zum vorherigen Schema ist die weitere Unterteilung des Typs Sattelfläche in die drei Typen negative Sattelfläche, Minimalfläche und positive Sattelfläche. Dieses Schema wird auch als HK-Signmap bezeichnet.

## 3.2.5 Formklassifikation

Jan Koenderink entwickelte ein Formklassifikationsschema (Koenderink und van Doorn, 1992), bei dem jeder Punkt in den zweidimensionalen kartesischen Raum  $k_1 \times k_2$  projiziert wird. Es ist jedoch üblich, diesen Raum in Polarkoordinaten darzustellen:

$$\begin{aligned} S &= \tan^{-1} \left( \frac{k_2}{k_1} \right) \\ C &= \sqrt{k_1^2 + k_2^2} \end{aligned} \quad (3.20)$$

wobei  $S$  die Form, und  $C$  die Stärke der Krümmung erfasst. ( $C$  entspricht der Wurzel aus der Abweichung von der Ebenheit.) Punkte mit demselben Wert für  $S$  aber unterschiedlichem  $C$ , können als Punkte mit gleicher Form unter stärkerer Krümmung betrachtet werden.

Der erste Quadrant dieses Raumes ist konvex, der dritte Quadrant ist konkav, der zweite und vierte Quadrant sind hyperbolisch. Ein wesentlicher Unterschied zur HK-Signmap ist die Einteilung von Ebenen mit nur einem Parameter  $C < \tau$ .



### 3.2.6 Ausgezeichnete Punkte

#### 3.2.6.1 Parabolischer Punkt

Ein parabolischer Punkt ist ein Punkt, an dem die Gaußsche Krümmung 0 ist, die mittlere Krümmung jedoch ungleich 0, also  $K = 0$ ,  $H \neq 0$ . Alle Punkte eines Zylinders sind parabolische Punkte.

#### 3.2.6.2 Nabelpunkt

Ein Nabelpunkt ist ein Punkt, an dem die Hauptkrümmungen identisch sind, also  $k_1 = k_2$ . An einem solchen Punkt sind die Eigenvektoren nicht eindeutig definiert, das heißt, der Eigenraum ist die gesamte Tangentialebene. Mit anderen Worten, an diesem Punkt ist die Krümmung in alle Richtungen identisch. Insbesondere sind alle Punkte einer Ebene und alle Punkte einer Kugel Nabelpunkte.

#### 3.2.6.3 Minimaler Punkt

Analog zum Nabelpunkt kann der minimale Punkt definiert werden als der Punkt an dem die Hauptkrümmungen den gleichen Betrag jedoch unterschiedliche Vorzeichen haben, also  $k_1 = -k_2$ . Dies korrespondiert zur Definition der minimalen Fläche, siehe oben.

### 3.2.7 Spezielle Flächentypen

Für viele praktische Anwendungen sind spezielle Flächentypen der Form  $z = f(x, y)$  von großer Bedeutung. Diese Flächen werden als Graph-Flächen bezeichnet. Im Folgenden werden die fundamentalen Größen dieser speziellen Flächen hergeleitet.

Wie oben angedeutet lässt sich eine Vektorfunktion erstellen der Form

$$\bar{X} = \begin{bmatrix} x \\ y \\ f \end{bmatrix} \quad (3.21)$$

und man erhält für die partiellen Ableitungen

$$\bar{X}_x = \begin{bmatrix} 1 \\ 0 \\ f_x \end{bmatrix}, \bar{X}_y = \begin{bmatrix} 0 \\ 1 \\ f_y \end{bmatrix}, \bar{X}_{xx} = \begin{bmatrix} 0 \\ 0 \\ f_{xx} \end{bmatrix}, \bar{X}_{xy} = \begin{bmatrix} 0 \\ 0 \\ f_{xy} \end{bmatrix}, \bar{X}_{yy} = \begin{bmatrix} 0 \\ 0 \\ f_{yy} \end{bmatrix} \quad (3.22)$$

Der Normalenvektor ergibt sich zu

$$\bar{N} = \frac{1}{\sqrt{f_x^2 + f_y^2 + 1}} \begin{bmatrix} -f_x \\ -f_y \\ 1 \end{bmatrix} \quad (3.23)$$

die Parameter der ersten Fundamentalform zu

$$\begin{aligned} E &= 1 + f_x^2 \\ F &= f_x f_y \\ G &= 1 + f_y^2 \end{aligned} \quad (3.24)$$

die Parameter der zweiten Fundamentalform zu

$$\begin{aligned} L &= \frac{f_{xx}}{\sqrt{f_x^2 + f_y^2 + 1}} \\ M &= \frac{f_{xy}}{\sqrt{f_x^2 + f_y^2 + 1}} \\ N &= \frac{f_{yy}}{\sqrt{f_x^2 + f_y^2 + 1}} \end{aligned} \quad (3.25)$$

	lokale Form	lokale Veränderbarkeit	leicht Abzutasten	leicht rückführbar
Parametrisch	nein	ja	ja	nein
Implizit	ja	nein	nein	ja
Explizit	ja	nein	ja	ja
Superquadrik	ja	nein	ja	ja
Generalisierte Zylinder	ja	nein	ja	nein

Tabelle 3.1: Eigenschaften verschiedener Flächenbeschreibungen.

die Gaußsche Krümmung zu

$$K = \frac{f_{xx}f_{yy} - f_{xy}^2}{(f_x^2 + f_y^2 + 1)^2} \quad (3.26)$$

die mittlere Krümmung zu

$$H = \frac{(1 + f_y^2)f_{xx} + (1 + f_x^2)f_{yy} - 2f_xf_yf_{xy}}{2(f_x^2 + f_y^2 + 1)^{\frac{3}{2}}} \quad (3.27)$$

Eine alternative Herleitung der Krümmungseigenschaften eines Flächenstück über lokale Koordinaten wird in Anhang B gegeben.

### 3.3 Stückweise Flächenrückführung

#### 3.3.1 Problembeschreibung

Es sei gegeben eine Punktwolke  $S$ , wobei  $S \subset \mathbb{R}^3$ . Ziel der Flächenrückführung ist es aus der Punktwolke  $S$  die Parameter einer mathematischen Flächenbeschreibung zu bestimmen. Es werden hier vor allem die Fälle betrachtet, in denen die Anzahl der Punkte groß ist im Verhältnis zur Anzahl der zu bestimmenden Parameter, es wird also nach der Lösung eines überbestimmten Gleichungssystems gesucht. Die Methoden der Flächenrückführung unterscheiden sich nach

- der Art der mathematischen Flächenbeschreibung, also dem Modell und
- dem Minimierungskriterium und dem davon abhängigen Lösungsverfahren.

#### 3.3.2 Flächenbeschreibung

Die Mathematik bietet eine Vielzahl von Möglichkeiten, eine Fläche im Raum darzustellen. Bereits in der Einleitung dieses Kapitels wurden drei grundlegende Methoden beschrieben, für die tatsächliche Berechnung einer mathematischen Beschreibung müssen aber konkretere Formeln entwickelt werden. Neben den parametrisierten, impliziten und expliziten Flächen werden im maschinellen Sehen noch Superquadriken und generalisierte Zylinder eingesetzt. Die Tabelle 3.1 gibt eine Übersicht über die wichtigsten Eigenschaften der verschiedenen Darstellungsformen. Dies sind die Möglichkeiten die lokale Flächenform abzuleiten, die Fläche lokal zu verändern, die Fläche abzutasten, also in eine Punktwolke zu überführen und die Berechenbarkeit aus einer gegebenen Punktwolke.

#### 3.3.3 Parametrisierte Flächen

Für die parametrisierte Darstellung einer Fläche

$$\mathbf{x}(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix} \quad (3.28)$$

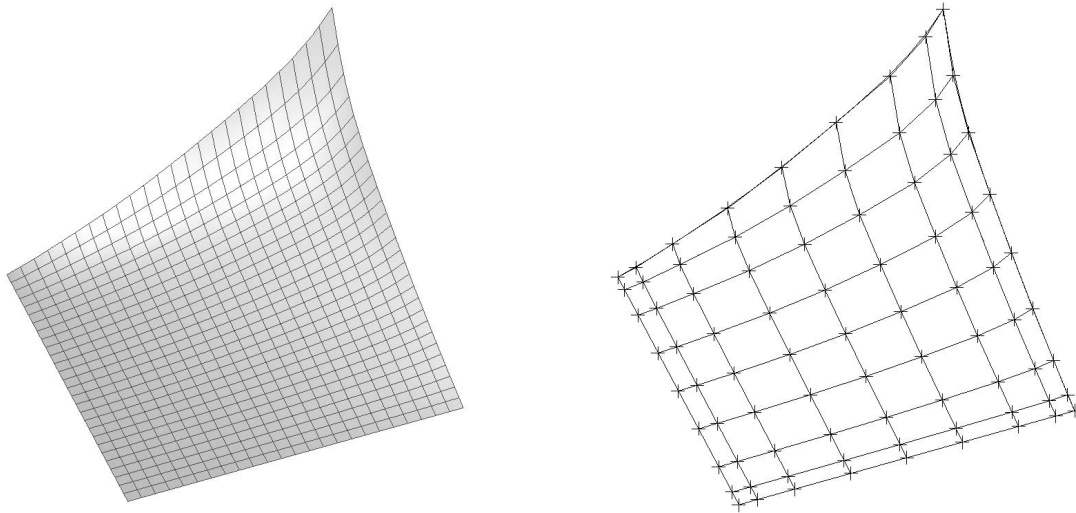


Abbildung 3.2: Beispiel einer NURBS-Freiformfläche. Die Fläche ist über einem  $9 \times 9$  Gitter von Kontrollpunkten definiert. Links eine schattierte Darstellung der Fläche. Rechts das definierende Gitter von Kontrollpunkten.

müssen die drei Funktionen  $x(u, v)$ ,  $y(u, v)$  und  $z(u, v)$  definiert werden. Bekannt vor allem aus dem Bereich CAD ist die Verwendung von nicht-uniformen rationalen B-Splines (NURBS), definiert durch

$$x(u, v) = \sum_i \sum_j B_{i,j}^h N_{i,j}(u) M_{i,l}(v)$$

wobei  $N_{i,j}(u)$  und  $M_{i,l}(v)$  die B-Spline-Basisfunktionen der Ordnung  $k$  und  $l$  sind, und  $B_{i,j}^h$  die homogenen Koordinaten der Stützpunkte enthält. Die Methode eine Fläche über ein Netz von Stützpunkten zu bestimmen ist für den Entwurf hervorragend geeignet, da die Fläche über ihre Stützpunkte gezielt und sehr intuitiv verändert werden kann. Die Abbildung 3.2 zeigt eine NURBS-Fläche, die interaktiv auf einem  $9 \times 9$  Gitter von Stützpunkten definiert wurde.

Im maschinellen Sehen hat sich diese Repräsentationsform jedoch wenig durchsetzen können. Dies liegt unter anderem daran, dass eine automatische Ableitung dieser Beschreibung aus einer Punktwolke schwierig ist. Die Definition der Stützstellen ist nicht trivial und erfordert im Allgemeinen manuelle Interaktion (siehe auch Besl, 1990).

### 3.3.4 Implizite Flächen

Bei einer Repräsentation als implizite Funktion

$$F(x, y, z) = 0 \tag{3.29}$$

wählt man für die Funktion  $F$  in der Regel ein Polynom vom Grad  $n$ . Somit erhält man

$$F(x, y, z) = \sum_{i=0}^n \sum_{j=0}^n \sum_{k=0}^n a_{ijk} x^i y^j z^k = 0 \tag{3.30}$$

Bei der Wahl des Grades stehen zwei Aspekte im Vordergrund, die Differenzierbarkeit und die numerische Stabilität. Die Differenzierbarkeit ist immer dann relevant, wenn die analytischen Ableitungen der Funktion gebildet werden sollen, wie zum Beispiel bei der Bestimmung der Krümmungen. Hierfür wird maximal die zweite Ableitung benötigt, ein Polynom zweiten Grades ist daher ausreichend. Andererseits findet man in der Literatur Berichte, nach denen Polynome bereits ab dem dritten Grad dazu neigen “überzuschwingen” (Fan, 1989), andere Quellen sehen die Grenze bei  $n = 4$  (Campbell und Flynn, 2002).

Aus diesen Überlegungen heraus und aus den Erfahrungen der zu erwartenden Flächen konzentrieren sich die Betrachtungen zur Flächenrückführung meist auf die so genannten Quadriken, also implizite Funktionen als Polynome zweiten Grades

$$K : F(x, y, z) = \begin{aligned} & a_1x^2 + a_2y^2 + a_3z^2 + a_4xy + a_5xz + a_6yz \\ & + a_7x + a_8y + a_9z + a_{10} \end{aligned} = 0 \quad (3.31)$$

Die Gruppe der darstellbaren Flächen umfasst unter anderem die Kugel, den Zylinder und den Kegel aber auch den zweischaligen Hyperboloid. Die wichtigsten Flächen aus der Gruppe der Quadriken werden in einem späteren Abschnitt detailliert beschrieben.

Eine äquivalente mathematische Darstellung ergibt sich durch die Matrixschreibweise

$$F(x, y, z) = \begin{bmatrix} x & y & z & 1 \end{bmatrix} \begin{bmatrix} a_1 & a_4' & a_5' & a_7' \\ a_4' & a_2 & a_6' & a_8' \\ a_5' & a_6' & a_3 & a_9' \\ a_7' & a_8' & a_9' & a_{10} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.32)$$

wobei  $2a_4' = a_4$  gilt, und die übrigen gestrichelten Koeffizienten analog umgerechnet werden können.

Bezeichnen wir die  $4 \times 4$  Matrix aus Gleichung 3.32 mit  $\mathbf{Q}$  und die obere  $3 \times 3$  Teilmatrix mit  $\mathbf{Q}_u$ . Aus dem Rang der beiden Matrizen, dem Vorzeichen der Determinante von  $\mathbf{Q}$ , sowie der Frage, ob alle Eigenwerte von  $\mathbf{Q}$  beziehungsweise  $\mathbf{Q}_u$ , die verschieden von Null sind, dasselbe Vorzeichen besitzen, lassen sich detaillierte Aussagen über die Quadrik treffen (Feddema und Little, 1997), die in Tabelle 3.2 zusammengefasst sind.

Wird eine Quadrik in der Darstellung von Gleichung 3.31 einer Hauptachsentransformation unterzogen, so entfallen die gemischten Terme, und man gelangt zu der normalisierten Darstellung

$$a_1x^2 + a_2y^2 + a_3z^2 + a_7x + a_8y + a_9z + a_{10} = 0 \quad (3.33)$$

Durch einen direkten Koeffizientenvergleich lassen sich bereits Aussagen über die Gestalt der Fläche ableiten, die in der Tabelle 3.3 nach Fan (1989) angegeben sind.

Zur Rückführung von impliziten Flächen und speziell von Quadriken wurden zahlreiche Vorschläge gemacht (Ahn et al. (2002); Faber und Fisher (2001); Bricault und Monga (1997); Taubin (1991); Guy und Medioni (1997); Lukacs et al. (1998); Feddema und Little (1997); Werghi et al. (1999)). S. Petitjean (2002) veröffentlichte eine Studie zum Vergleich der verschiedenen Verfahren speziell für triangulierte Daten, die sich jedoch auf Distanzbilder übertragen lässt.

### 3.3.4.1 Lösungsverfahren

Die Lösung des gestellten Problems entspricht einer Minimierung der Summe der quadratischen Fehler pro Punkt  $\sum \varepsilon_{x,y,z}^2 \rightarrow \min$ . Um aus einer Punktwolke die Parameter eines bestimmten Flächenmodells zu schätzen, muss also zunächst die zu minimierende Fehlerfunktion bestimmt werden. Idealerweise wählt man den geometrischen (euklidischen) Abstand eines Punktes  $P$  zur Fläche. Es zeigt sich, dass dieser Ansatz zu einem nicht-linearen Gleichungssystem führt (Redding, 2000). Dieses Problem ist zwar durchaus lösbar, ist aber mit einem erhöhten Rechenaufwand verbunden.

Bei den verschiedenen Flächenmodellen bieten sich unterschiedliche Alternativen zum geometrischen Abstand an. Bei den impliziten Flächen bietet sich der so genannte algebraische Abstand an.

$$F(x_i, y_i, z_i) = v_i \quad (3.34)$$

wiederum mit  $\sum_{i=1}^n v_i^2 \rightarrow \min$ . Diese Formulierung des Fehlers birgt jedoch für den klassischen Ansatz zur kleinsten Quadrate Minimierung ein Problem. So sind von den zehn Parametern der quadratischen Gleichung nur neun wirklich unabhängig. Die gesamte Gleichung ist frei skalierbar. Multipliziert man den Lösungsvektor mit einer beliebigen Zahl, erhält man wieder eine gültige Lösung. Dies ist insbesondere für Null erfüllt. Diese Eigenheit macht es notwendig zusätzlich Bedingungen einzuführen. Dies lässt sich an einer kurzen Betrachtung veranschaulichen. Für  $v_i = 0$  ist der Fehler minimal. Bei der Wahl eines Polynoms als Flächenfunktion führt dies zum Beispiel für die oben beschriebene konische Funktion zur trivialen Lösung  $a_i = 0$  für  $i = 1, \dots, 10$ . Um

Quadrik Typ	$Rang(Q_u)$	$Rang(Q)$	$\det(Q)$	Vorzeichen der Eigenwerte von $Q_u$	Vorzeichen der Eigenwerte von $Q$
reeller Ellipsoid	3	4	$< 0$	gleich	
imaginärer Ellipsoid	3	4	$> 0$	gleich	
einschaliger Hyperboloid	3	4	$> 0$	verschieden	
zweischaliger Hyperboloid	3	4	$< 0$	verschieden	
reeller Kegel	3	3		verschieden	
imaginärer Kegel	3	3		gleich	
elliptischer Paraboloid	2	4	$< 0$	gleich	
hyperbolischer Paraboloid	2	4	$> 0$	verschieden	
reeller elliptischer Zylinder	2	3		gleich	verschieden
imaginärer elliptischer Zylinder	2	3		gleich	gleich
hyperbolischer Zylinder	2	3		verschieden	
zwei reelle Ebenen	2	2		verschieden	
zwei imaginäre Ebenen	2	2		gleich	
parabolischer Zylinder	1	3			
imaginäre parallele Ebenen	1	2			verschieden
deckungsgleiche Ebenen	1	2			gleich
reelle parallele Ebenen	1	1			

Tabelle 3.2: Typisierung einer Quadrik.

$a_1$	$a_2$	$a_3$	$a_7$	$a_8$	$a_9$	$a_{10}$	
0	0	0					Ebene
+	+	+				-	Ellipsoid
+	+	-				-	einschaliger Hyperboloid
+	+	0			$\neq 0$		elliptischer Paraboloid
+	-	0			$\neq 0$		hyperbolischer Paraboloid
+	+	0			0	-	Zylinder oder elliptischer Zylinder
+	-	0			0	-	hyperbolischer Zylinder
+	0	0	$a_8$ oder $a_9 \neq 0$			-	parabolischer Zylinder
+	+	-				0	elliptischer Kegel

Tabelle 3.3: Klassifikation von Flächenpatches einer Quadrik in normalisierter Darstellung.

Ebene	$ax + by + cz = 0$
Kugel	$(x^2 + y^2 + z^2) - a = 0$
Ellipsoid	$ax^2 + by^2 + cz^2 - d = 0$
Zylinder	$x^2 + y^2 - a = 0$
Kegel	$x^2 + y^2 - az^2 = 0$

Tabelle 3.4: Übersicht über einige typische Flächen aus der Gruppe der Quadriken in ihrer Normallage.

diesen entarteten Fall zu umgehen, müssen bei diesem Ausgleichungsansatz zusätzliche Bedingungen eingeführt werden (siehe Bookstein, 1979 und Fitzgibbon et al., 1999). Im einfachsten Fall wird der konstante Term auf einen bestimmten Wert, zum Beispiel 1 gesetzt. Eine weitere Möglichkeit besteht darin den gesamten Lösungsvektor zu normieren, also  $|[a_1, \dots, a_n]| = 1$ . Andere Bedingungen ergeben sich abhängig vom genauen Typ der erwarteten Fläche.

Durch ein anderes Lösungsverfahren lässt sich die triviale Lösung ganz ohne Nebenbedingungen vermeiden. Hierzu wählt man das Verfahren der Eigenwert- beziehungsweise Eigenvektorbestimmung. Dieses Verfahren wurde im Bereich der Computer Vision erstmals zur Bestimmung einer ausgleichenden Ebene verwendet (Duda und Hart, 1973), es eignet sich jedoch ebenso für Polynome höherer Ordnung. Dieses Verfahren wird in Anhang A motiviert.

### 3.3.5 Explizite Flächen

Für explizite Funktion ergibt sich bei der Wahl von Polynomen eine Darstellung analog zu den impliziten Flächen.

$$F(x, y) = \sum_{i=0}^n \sum_{j=0}^n a_{ijk} x^i y^j \quad (3.35)$$

Die Repräsentation einer Fläche als explizite Funktion ist jedoch mit Einschränkungen im Umfang der darstellbaren Flächen verbunden. Für eine lokale Beschreibung, zum Beispiel zur Extraktion von Krümmungseigenschaften, reicht diese Darstellungsform jedoch aus. Hierbei kann man sich ebenso wie bei den Quadriken auf ein Polynom zweiten Grades beschränken.

$$F(x, y) = a_1 x^2 + a_2 y^2 + a_3 xy + a_4 x + a_5 y + a_6 \quad (3.36)$$

#### 3.3.5.1 Lösungsverfahren

Bei den expliziten Funktionen bietet sich der Abstand entlang der ausgesuchten Achse als Kriterium an. Dies führt zu einer Beobachtungsgleichung der folgenden Form

$$F(x_i, y_i) = z_i + v_i \quad (3.37)$$

mit  $\sum_{i=1}^n v_i^2 \rightarrow \min$ .

Es ist anzumerken, dass das dargestellte Fehlerkriterium nicht die Realität der Sensorsysteme widerspiegelt. Die Annahme, dass die Fehler entlang der  $x$  und  $y$  Koordinatenachsen zu vernachlässigen sind und alleine der Fehler in  $z$  Richtung zu minimieren ist, ist grundsätzlich falsch. Zu Gunsten des einfachen Rechenschemas wird dies jedoch oft bewusst in Kauf genommen.

### 3.3.6 Spezielle Flächentypen aus der Gruppe der Quadriken

Im Folgenden wird die stückweise Rückführung einer Auswahl von häufig anzutreffenden Flächentypen im Detail betrachtet. Die Gleichungen für die einzelnen Funktionen werden zunächst in Anlehnung an die oben gegebene konische Gleichung dargestellt.

### 3.3.6.1 Ebene

Eine Variante der konischen Gleichung, bei der keine quadratischen oder gemischten Terme existieren:  $K \mid a_1, \dots, a_6 = 0$ .

$$a_7x + a_8y + a_9z + a_{10} = 0 \quad (3.38)$$

Mögliche Nebenbedingungen  $a_{10} = \text{const.}$ ,  $|a_7^2 + a_8^2 + a_9^2 + a_{10}^2| = \text{const.}$ ,  $|a_7^2 + a_8^2 + a_9^2| = 1$ . Nur bei der letzten Nebenbedingung ergibt sich die gewohnte geometrische Interpretation der Parameter, nach der  $a_7, a_8, a_9$  den Normalenvektor der Ebene und  $a_{10}$  den Abstand der Ebene zum Ursprung enthalten. Für alle anderen Darstellungen lassen sich die geometrischen Parameter wie folgt herleiten:

$$r = |a_7^2 + a_8^2 + a_9^2| \quad (3.39)$$

$$d = \frac{a_{10}}{r} \quad (3.40)$$

$$[e_1, e_2, e_3] = \frac{1}{r} [a_7, a_8, a_9] \quad (3.41)$$

### 3.3.6.2 Kugel

Alle gemischten Terme der quadratischen Gleichung verschwinden:  $K \mid a_4, \dots, a_6 = 0$ .

$$a_1x^2 + a_2y^2 + a_3z^2 + a_7x + a_8y + a_9z + a_{10} = 0 \quad (3.42)$$

Mit der notwendigen Bedingung  $a_1 = a_2 = a_3$  ergibt sich eine alternative Darstellung in fünf Parametern.

$$a_1(x^2 + y^2 + z^2) + a_7x + a_8y + a_9z + a_{10} = 0 \quad (3.43)$$

Mögliche Nebenbedingungen  $a_{10} = \text{const.}$  Die geometrischen Parameter (Mittelpunkt, Radius) lassen sich wie folgt aus den algebraischen Parametern ableiten:

$$x_0 = -\frac{a_4}{2a_1}, y_0 = -\frac{a_5}{2a_1}, z_0 = -\frac{a_6}{2a_1}$$

$$r = \sqrt{a_1(x_0^2 + y_0^2 + z_0^2) - a_{10}}$$

### 3.3.6.3 Achsparalleles Ellipsoid

Die gemischten Terme zu null:  $K \mid a_4, \dots, a_6 = 0$ .

$$a_1x^2 + a_2y^2 + a_3z^2 + a_7x + a_8y + a_9z + a_{10} = 0 \quad (3.44)$$

Mögliche Nebenbedingungen  $a_{10} = \text{const.}$  Die geometrischen Parameter (Mittelpunkt, Halbachsen) lassen sich wie folgt aus den algebraischen Parametern ableiten:

$$x_0 = -\frac{a_4}{2a_1}, y_0 = -\frac{a_5}{2a_2}, z_0 = -\frac{a_6}{2a_3},$$

$$r = \sqrt{a_1(x_0^2 + y_0^2 + z_0^2) - a_{10}},$$

$$a = \sqrt{\frac{r}{a_1}}, b = \sqrt{\frac{r}{a_2}}, c = \sqrt{\frac{r}{a_3}}.$$

## 4 Modellbasierte Segmentierung

### 4.1 Einleitung

Als Segmentierung bezeichnet man den Vorgang, ein Bild in sinnvolle Bestandteile zu zerlegen (Jiang und Bunke, 1997). Welche Zerlegung sinnvoll ist, hängt von der konkreten Aufgabenstellung, dem vorliegenden Datenmaterial und dem Systemdesign ab. Im Kontext der Objekterkennung dient die Segmentierung zur Merkmalsextraktion. Sie sollte daher robust und zuverlässig eine dem Modell möglichst genau entsprechende Zerlegung erzielen. Bei Distanzbildern ist es üblich eine Zerlegung in Regionen zu erzeugen, bei der jede Bildregion einer Fläche des Objektes entspricht.

Zwei grundsätzliche Verfahrensweisen stehen zur Verfügung, eine solche Zerlegung zu erzielen, die kantenorientierte und die regionenorientierte Vorgehensweise. Bei der kantenorientierten Methode versucht man, die Grenzlinien zwischen den gesuchten Bildregionen zu ermitteln. Hierzu müssen verschiedenen Kantentypen betrachtet werden: Sprungkanten, die durch den Versatz zweier Flächen in Aufnahme­richtung entstehen, Schnittkanten, die durch den Schnitt zweier Flächen mit unstetigem Übergang entstehen und so genannte glatte Kanten, die durch den stetigen Übergang zweier Flächen ineinander entstehen. Soll ein Segmentierungsverfahren möglichst viele Objekttypen behandeln können und damit möglichst viele Flächentypen, also insbesondere beliebig gekrümmte Flächen, bereiten vor allem die glatten Kanten bei der kantenorientierten Segmentierung Probleme. Sie sind in der Regel nur unzuverlässig detektierbar. Daher ist ein regionenorientiertes Vorgehen für die Ziele dieser Arbeit zu bevorzugen.

Bei der regionenorientierten Segmentierung versucht man, den Pixeln eines Distanzbildes eine Markierung (*label*) zuzuordnen, wobei alle Pixel, die zur selben Fläche gehören, dieselbe Markierung tragen. Eine formale Beschreibung dieser Forderung wurde von Horowitz und Pavlidis (1974) aufgestellt und auch von zahlreichen anderen Autoren in zum Teil leicht abgewandelter Form angewandt (Ballard und Brown, 1982; Gonzalez und Woods, 1992; Haralick und Shapiro, 1992). Alle Punkte eines Distanzbildes bilden die Gesamtregion  $R$ , welche in die Regionen  $R_i$  zerlegt werden soll. Die formalen Bedingungen für eine gültige Segmentierung lauten dann:

1.  $\bigcup_{i=1}^n R_i = R$ .
2.  $R_i$ , mit  $i = 1, 2, \dots, n$ , ist zusammenhängend.
3.  $R_i \cap R_j = \emptyset$  für  $i \neq j$ .
4.  $H(R_i)$  ist WAHR für  $i = 1, 2, \dots, n$ .
5.  $H(R_i \cup R_j)$  ist FALSCH für  $i \neq j$  und falls  $R_i$  und  $R_j$  benachbart sind.

Die erste Bedingung fordert die Vollständigkeit der Segmentierung, das heißt, alle Pixel des Distanzbildes müssen einer Region zugeordnet werden. Dies ist allerdings nicht immer möglich. So treten zum Beispiel bei triangulierenden Messverfahren Abschattungen auf, sodass das Distanzbild an sich bereits Lücken aufweist. Somit muss der Segmentierung erlaubt werden, manche Pixel als nicht zu einer Fläche gehörig zu kennzeichnen.

Das Prädikat  $H(R_i)$  drückt die Homogenität der Region  $R_i$  aus. Dieses Kriterium hängt wiederum von der konkreten Anwendung ab. Häufig anzutreffende Homogenitätskriterien sind:

- die gleiche Ausrichtung<sup>1</sup> des Normalenvektors, hiermit erreicht man eine Segmentierung in ebene Regionen,
- die Bedingung, dass alle Punkte einer Region einem gemeinsamen mathematischen Flächenmodell folgen, dies ermöglicht eine Segmentierung zum Beispiel in gekrümmte Regionen,
- die gleiche Ausprägung der Krümmungseigenschaften aller Punkte einer Region.

<sup>1</sup>Hier sowie in den folgenden Punkten ist Gleichheit nicht mathematisch exakt zu verstehen, sondern im Sinne von gleich innerhalb einer gewissen Toleranz.



Das Problem der Segmentierung sowohl im Allgemeinen, als auch im speziellen Fall der regionenorientierten Segmentierung von Distanzbildern ist ein seit langer Zeit bekanntes und intensiv behandeltes Thema, zu dem zahlreiche Lösungsvorschläge für verschiedenste Anwendungen existieren. Eine wichtige Studie von Hoover et al. (1996) (siehe auch Hoover et al., 1995; Jiang et al., 2000) mit dem Ziel ein Verfahren zur Evaluation und zum Vergleich von Segmentierungsverfahren auf Distanzbildern zu entwickeln zählt siebzehn verschiedene Verfahren allein zur Segmentierung in ebene Regionen auf. Doch auch diese Aufzählung ist noch unvollständig. Die Studie selbst evaluiert vier bekannte Verfahren. Eine Folgepublikation von Powell et al. (1998) unternahm den Versuch die Methoden zum Vergleich von Segmentierungsverfahren auf Verfahren zu erweitern, die gekrümmte Regionen segmentieren. Nur zwei Verfahren konnten in dieser Studie eingesetzt werden. Die Resultate der Segmentierungen wurden von den Autoren selbst als zum Teil unzureichend bezeichnet. Es ist daher davon auszugehen, dass das Problem der Segmentierung trotz der intensiven Forschung, immer noch ein sehr schweres Problem darstellt, für das keine allgemeine Lösung existiert. Um das im Rahmen dieser Arbeit entwickelte Verfahren zu motivieren und die Zusammenhänge und Unterschiede mit bereits bestehenden Verfahren hervorzuheben, sollen im Folgenden kurz die wesentlichen Ansätze zur regionenorientierten Segmentierung dargestellt werden.

### 4.1.1 Spalten und Verschmelzen

Verfahren, welche zu dieser Gruppe gehören, gehen bei der Segmentierung in zwei Schritten vor. Im ersten Schritt wird die Gesamtregion  $R$  durch ein möglichst einfaches Kriterium in viele kleine Teilregionen  $R'_i$  aufgespalten. Somit wird bewusst eine Übersegmentierung erzeugt. Dabei muss bereits im ersten Schritt sichergestellt werden, dass alle Teilregionen  $R'_i$  die Bedingungen einer gültigen Segmentierung erfüllen. Lediglich die Forderung nach der Maximalität einer einzelnen Region (siehe oben Bedingung 5), muss noch nicht erfüllt sein. Im zweiten Schritt werden dann benachbarte Regionen, sofern sie gemeinsam das Homogenitätskriterium erfüllen zu einer Region verschmolzen. Dies wird fortgesetzt, bis keine weiteren Regionen gefunden werden können. Damit wird die letzte Bedingung nach der Maximalität erfüllt.

Um die anfängliche Zerlegung zu erzielen, sollte ein möglichst einfaches Verfahren gewählt werden. Hauptsächlich handelt es sich dabei um hierarchische Ansätze, die die Gesamtregion  $R$  immer weiter unterteilen, bis jede Teilregion das Homogenitätskriterium erfüllt. Ein wichtiger Vertreter solcher Verfahren ist die Aufspaltung durch einen *Quadtree* (siehe zum Beispiel Parvin und Medioni, 1986).

#### 4.1.1.1 Gruppierung der Abtastzeilen

Zu der Gruppe der Verfahren nach dem Ansatz des Spaltens und Verschmelzens, kann auch ein Verfahren gezählt werden, das von Jiang und Bunke (1994) vorgestellt wurde. Die Idee des Ansatzes besteht darin, das Distanzbild zunächst entlang der Abtastzeilen in kurze Liniensegmente aufzuspalten. Danach werden benachbarte Liniensegmente zu zweidimensionalen Regionen verschmolzen.

Das Verfahren wurde ursprünglich für die Segmentierung in ebene Regionen vorgeschlagen. Unter der Annahme, dass ein äquidistant abgetastetes Distanzbild vorliegt (Z-Matrix), kann die allgemeine explizite Ebenengleichung (vergleiche Abschnitt 3.3.5) für einen Abschnitt in der Abtastzeile, für die  $y = y_0 = \text{const.}$  gilt, als Geradenstück in der XZ-Ebene umgeformt werden.

$$z = a_0x + a_1y_0 + a_3 = a_0x + a'_1$$

In der zweiten Phase des Verfahrens werden die Liniensegmente, welche benachbart sind und in einer gemeinsamen Ebene liegen, zu einer Region verschmolzen. Zu den prägnanten Details des Verfahrens zählen eine geeignete Datenstruktur für den effizienten Zugriff auf die Liniensegmente und deren Nachbarschaft, sowie die Formulierung geeigneter Kriterien für die Verschmelzung.

Die Stärken des Verfahrens liegen zum einen in seiner Schnelligkeit. Da das Verfahren auf Liniensegmente als Grundprimitiv zugreift anstatt auf einzelne Pixel, wird das Datenvolumen für den Verschmelzungsprozess deutlich reduziert. Zum anderen bleibt das Kriterium für die Phase der Aufspaltung des Distanzbildes einfach zu berechnen (siehe Duda und Hart, 1973).

Eine Variante des Verfahrens wurde später auf die Segmentierung von Distanzbildern in gekrümmte Flächen erweitert (Jiang und Bunke, 1998). Hierzu wird das Distanzbild anstelle von geraden Liniensegmenten in gekrümmte Kurven aufgespalten.



Abbildung 4.1: Beispiel für eine Segmentierung durch Gruppierung der Abtastzeilen aus Brenner und Fritsch (1997). Es sind linienhafte Artefakte sichtbar, die durch das Verfahren hervorgerufen werden.

Nachteilig an dem Verfahren ist die Eingangsbedingung, dass das Distanzbild äquidistant abgetastet sein sollte. Außerdem weist das Verfahren die typischen Schwächen der Verfahren auf, die eine anfängliche Zerlegung entlang hierarchischer Strukturen (hier die Abtastzeilen) vornehmen. Die Struktur dieser Zerlegung schlägt sich auch in der Struktur der erzeugten Regionen nieder. In Abbildung 4.1 ist ein Beispiel für eine Segmentierung eines Distanzbildes nach dem Verfahren von Jiang und Bunke aus Brenner und Fritsch (1997) gegeben. Deutlich sind die durch das Verfahren bedingten linienhaften Strukturen in den Regionen zu erkennen.

#### 4.1.2 Bereichswachstum

Das Verfahren des Bereichswachstums ähnelt der zweiten Phase beim Spalten und Verschmelzen. Ausgehend von einer ursprünglichen Saatregion wird diese Region solange erweitert, solange das Homogenitätskriterium erfüllt bleibt. Beim Bereichswachstum wird jedoch direkt auf Pixelebene gearbeitet. Es findet also keine anfängliche Unterteilung in größere Grundprimitive statt.

Der Prozess beginnt mit der Auswahl einer geeigneten Saatregion  $R_i^0$ , welche selbstverständlich das Homogenitätskriterium  $H(R_i^0)$  erfüllen muss. Iterativ wird jedes Nachbarpixel  $P_j$  im Verfahrensdurchlauf  $k$  durch  $H(R_i^k \cup P_j)$  auf Homogenität mit der Region getestet und, falls erfüllt, die Region  $R_i^{k+1} = R_i^k \cup P_j$  gebildet. Das Wachstum wird beendet, wenn kein weiteres Nachbarpixel mehr gefunden werden kann, das das Homogenitätskriterium erfüllt. Danach wird mit der nächsten Saatregion fortgefahren, bis keine geeignete Saatregion mehr existiert und alle Pixel abgearbeitet wurden.

Durch die ausschließliche Betrachtung von Pixeln in direkter Nachbarschaft der auszudehnenden Region, sowie die ständige Überprüfung des Homogenitätskriteriums, sind alle Anforderungen an eine gültige Segmentierung erfüllt. Die Qualität der Segmentierung hängt entscheidend von der Wahl der Saatregionen und des Homogenitätskriteriums ab. Bei der Segmentierung von Distanzdaten ist es üblich, die Homogenität einer Region durch die Schätzung einer Flächenfunktion zu realisieren. Bleibt der Schätzfehler unterhalb einer gewissen Schranke, gilt die Region als homogen. Da die Homogenität für jedes Pixel erneut geprüft werden muss, ergibt sich ein erheblicher Aufwand bei der Berechnung. Der Aufwand kann reduziert werden, wenn zusätzliche Kriterien herangezogen werden, die mögliche Kandidaten vor dem Schätzprozess ausschließen. Hierzu kann zum Beispiel die Kompatibilität des Normalenvektors geprüft werden. Dies ist insbesondere bei einer Ebenensegmentierung populär.

Abbildung 4.2 zeigt an einem Beispiel ein Problem bei der Segmentierung eines Distanzdatensatzes durch Bereichswachstum auf. Das Verfahren wurde implementiert, um ein Distanzbild in gekrümmte Flächen zu segmentieren. Die Homogenität wird dabei mit einem Flächenschätzverfahren für ein explizites Polynom zweiten Grades (siehe 3.36) überprüft. Je nachdem wo die Saatregion platziert wird, ergibt sich ein unterschiedliches Ergebnis für die Ausdehnung der einzelnen Region. In beiden dargestellten Fällen wird die Saatregion auf einer ebenen Fläche des Objektes gebildet. Im zweiten Fall jedoch befindet sie sich dicht am Rand der Fläche. In den Ergebnisbildern am Ende des Wachstumsprozesses wird deutlich, wie im zweiten Fall die Region in den gekrümmten Bereich übergelaufen ist. Die Saatregion wäre optimal platziert, wenn sie in der Mitte der Fläche liegen würde. Hierzu müsste man aber bereits vorab die Aufteilung in die Regionen kennen. In Abbildung 4.3 ist das Ergebnis nach dem vollständigen Ablauf des Verfahrens dargestellt.

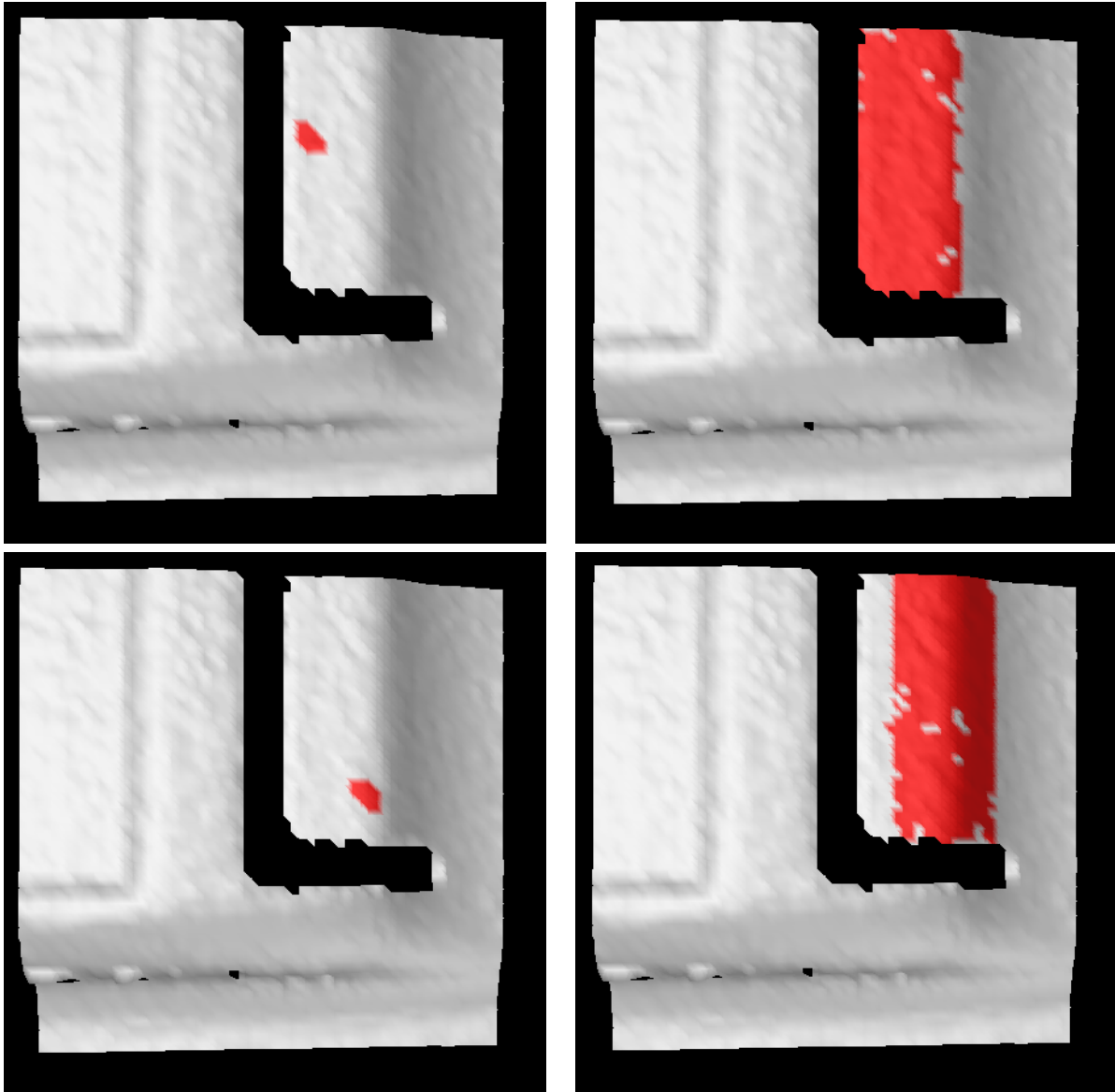


Abbildung 4.2: Segmentierung durch Bereichswachstum. Zweimal wird eine einzelne Region aus demselben Distanzbild mittels eines Bereichswachstumsverfahrens extrahiert. In der linken Spalte ist jeweils die Ausgangssituation mit der Saatregion dargestellt, in der rechten Spalte das Resultat am Ende des Wachstums.

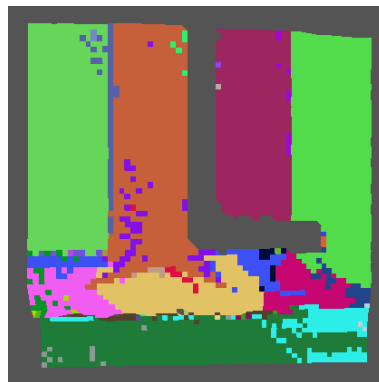


Abbildung 4.3: Beispiel für eine vollständige Segmentierung durch Bereichswachstum in gekrümmte Flächen. Es ist eine grundsätzliche Übersegmentierung und eine Vielzahl kleiner Splitterregionen sichtbar.

Eine detaillierte Beschreibung des Regionenwachstums aus der Sicht des Reverse Engineering zur Rückführung von gekrümmten Flächen wurde von Spapidis und Besl (1995) gegeben.

### 4.1.3 Segmentierung durch Clustering

Mit Clustering bezeichnet man eine Gruppe allgemeiner Verfahren der Mustererkennung. Ziel der Verfahren ist es eine Menge von Muster- oder Merkmalsvektoren nach ihrer Ähnlichkeit zu gruppieren. Jedes Element der Eingangsdaten, beschrieben durch einen Merkmalsvektor, wird so einer Gruppe (*cluster*) zugeordnet, dass die Merkmalsvektoren aller Elemente des Clusters im  $n$ -dimensionalen Merkmalsraum eine Häufung bilden. Eine allgemeine Übersicht über Clustering-Techniken für Aufgaben der Mustererkennung wurde von Jain et al. (1999) gegeben.

Es ist wichtig darauf hinzuweisen, dass es sich bei Clustering-Techniken um unüberwachte Verfahren der Mustererkennung handelt. Sowohl die Lage der Cluster im Merkmalsraum als auch ihre Anzahl sind a priori nicht bekannt. Ist diese Information gegeben, zum Beispiel durch einen Lernprozess mithilfe von Trainingsdaten, handelt es sich um ein überwachtes Verfahren der Mustererkennung. Bei einer bekannten Aufteilung des Merkmalsraums, besteht die Aufgabe darin neue unklassifizierte Merkmalsvektoren den korrekten Gruppen (Klassen) zuzuweisen. Man spricht dann von einer überwachten Klassifikation (Jain et al., 1999).

Der Unterschied des Clustering zur Segmentierung besteht im Betrachtungsraum. Beim Clustering geschieht die Gruppierung im Merkmalsraum. Bei der Segmentierung steht die Gruppierung im Bildraum im Vordergrund. Bei den Segmentierungsverfahren kommt es zu einem Wechselspiel zwischen der zusammenhängenden und überlappungsfreien Gruppierung im Bildraum und der Gruppierung im Merkmalsraum, für den jedoch nicht unbedingt eine überlappungsfreie Gruppierung angestrebt wird (Haralick und Shapiro, 1992). Umgekehrt erzeugt eine überlappungsfreie Gruppierung im Merkmalsraum durch ein Clustering nicht notwendigerweise zusammenhängende Regionen im Bild. Dies kann jedoch in einem nachträglichen Verarbeitungsschritt erzwungen werden. Somit können Clustering-Verfahren zur Segmentierung eingesetzt werden.

Ein wichtiges Beispiel für den Einsatz eines Clustering-Verfahrens zur Segmentierung von Distanzbildern wurde von Hoffman und Jain (1986) gegeben. Hierbei werden die Messungen in einen sechsdimensionalen Merkmalsraum übertragen. Durch ein Clustering-Verfahren wird eine initiale Segmentierung erzeugt, die anschließend verfeinert wird. Für ein einzelnes Pixel  $P_i$  eines Distanzbildes wird der Merkmalsvektor  $v_i$  berechnet, der sich aus den räumlichen Koordinaten und dem Normalenvektor zu  $v_i = (x_i, y_i, z_i, n_i^x, n_i^y, n_i^z)$  ergibt. Zur Gruppierung wird eine Variante des populären  $k$ -Means Algorithmus mit dem Namen CLUSTER eingesetzt (Besl und Jain, 1988). Um zusammenhängende Regionen zu erzeugen, wird eine Analyse der Zusammenhangskomponenten durchgeführt (*connected component labeling*). Zudem werden noch weitere Verarbeitungsschritte angehängt, die eine Übersegmentierung vermeiden sollen.

Weitere Beispiele von Clustering-Verfahren zur Segmentierung von Distanzbildern sind in Frigui und Krishnapuram (1999) und Koster und Spann (2000) gegeben.

### 4.1.4 Krümmungsbasierte Segmentierung

Besl (1988b) leistete mit seiner Arbeit einen wesentlichen Beitrag zur Einführung differenzialgeometrischer Merkmale in der Segmentierung von Distanzbildern. Er nutzt die mittlere und die Gaußsche Krümmungen, um eine initiale Segmentierung zu erzeugen. Dabei werden für jedes Pixel des Distanzbildes nach dem in Abschnitt 4.4.3.1 beschriebenen Verfahren der orthogonalen Polynome die Krümmungen geschätzt. In einem starren Klassifikationsschema nach dem Vorzeichen der Krümmungen, das in Abschnitt 3.2.4 eingeführt wurde, wird dann jedem Pixel einer von acht Flächentypen zugewiesen.

Da die Klassifikation nur in einem groben Schema erfolgt, entsteht eine Untersegmentierung. Benachbarte Pixel, die zu unterschiedlichen Flächen gehören, können den gleichen Flächentyp aufweisen, zum Beispiel den Typ konvexe zylindrische Fläche. Um diese Flächen wieder zu trennen, müssen die Regionen noch weiter zerkleinert werden. Dies geschieht im wesentlichen durch Erosionsverfahren.

Die anfängliche Segmentierung wird einer Reihe von morphologischen Filteroperationen unterzogen, bei denen kleine Splitterregionen gelöscht und große zusammenhängende Regionen auf einen kleinen Kern zusammengesmolzen werden. Dieser Prozess wird fortgesetzt, bis die Regionen auf eine bestimmte Mindestgröße reduziert sind. Diese kleinen Kernregionen bilden die Saatregionen  $R_i^0$  für ein Regionenwachstumsverfahren.

Die von Besl und Jain (1986) vorgestellte Variante eines Regionenwachstumsverfahrens verwendet dabei explizite Polynome in der Form von Gleichung 3.35. Das Besondere an dem Ansatz besteht darin, dass der Grad

des Polynoms flexibel ist und auch während des Expansionsprozesses angepasst wird. Begonnen wird mit der Saatregion  $R_i^0$  und einem Polynom von möglichst niederem Grad. Der angemessene Polynomgrad wird dabei anhand des Approximationsfehlers, der unter einem Schwellwert liegen muss, bestimmt. Während der Expansion im Schritt  $k$  wird beim Übergang von der Region  $R_i^k$  zu  $R_i^{k+1}$  erneut eine Fläche approximiert. Überschreitet der Approximationsfehler die Schranke, wird der Polynomgrad erhöht. Der maximal zulässige Polynomgrad beträgt vier. Im Detail arbeitet das Verfahren deutlich komplexer, so wird zum Beispiel zur Bestimmung des Polynomgrads nicht nur der Approximationsfehler berücksichtigt sondern ein erweitertes Testverfahren angesetzt. Dabei werden die Vorzeichen der punktwisen Approximationsfehler ausgewertet. Der Grund für das erweiterte Verfahren liegt darin, dass aus Effizienzgründen möglichst frühzeitig ein notwendiger höherer Polynomgrad erkannt werden soll.

## 4.2 Motivation für ein modellbasiertes Verfahren

Viele der für die Segmentierung von Distanzbildern vorgeschlagenen Verfahren arbeiten rein datengetrieben, sie benötigen also kein Vorwissen über die in der Szene abgebildeten Flächen. Dieser Ansatz ist vor allem dort anzutreffen, wo die Zerlegung der Randbeschreibung eines Objektes eindeutig ist. So sind zum Beispiel datengetriebene Clustering-Verfahren zur Segmentierung von polyedrischen Objekten sehr erfolgreich. Auch in anderen Anwendungen, in denen die Formenvielfalt eingeschränkt werden kann, können vorwiegend datengetriebene Verfahren eingesetzt werden, so bei Szenen, die nur aus zylindrischen Flächen bestehen, oder bei Objekten, bei denen die gekrümmten Flächen stets durch scharfe Kanten getrennt sind.

Im Abschnitt 2.2.2.1 wurde bereits darauf hingewiesen, dass die Zerlegung der Randbeschreibung eines Objektes, also die Zerlegung in Teilflächen, im Allgemeinen nicht eindeutig ist. Für solche Objekttypen sind rein datengetriebene Ansätze weniger tauglich. Beim Regionenwachstum zum Beispiel hängt die Zerlegung der Oberfläche im Wesentlichen von der Platzierung der Saatregionen ab. Diese lässt sich jedoch nicht eindeutig und optimal aus den Daten bestimmen. Zusätzlich wird beim Regionenwachstum, sofern eine Flächenapproximation eingesetzt wird, der Flächentyp auf einfach approximierbare Flächen (siehe Tabelle 3.1) eingeschränkt, zumeist auf Polynome.

Gerade wenn die Segmentierung zur Merkmalsextraktion in einem merkmalsbasierten Objekterkennungssystem eingesetzt werden soll, ist es wichtig, dass die Zerlegung des Distanzbildes in die Regionen möglichst gut der Zerlegung des Modells in seine Teilflächen entspricht. Die Information, wie diese Zerlegung zu erfolgen hat, ist prinzipiell im CAD-Modell abgelegt. Es ist daher erstrebenswert diese Information bereits im Extraktionsprozess zu berücksichtigen. Dies bedeutet nichts anderes, als bereits bei der Segmentierung die Modellinformation zu verwenden.

Da auf der Stufe der Segmentierung noch keinerlei Beziehung zwischen dem Koordinatensystem des Modells und dem Koordinatensystem des Objektes in der Szene hergestellt wurde, können keine Informationen aus dem Modell verwendet werden, die mit dem Koordinatensystem behaftet sind. Es sind invariante Merkmale zu definieren. Dies führt zur Betrachtung differentialgeometrischer Merkmale, wie sie in Kapitel 3 eingeführt wurden.

In den folgenden Abschnitten wird das vorgeschlagene Verfahren zur modellbasierten Segmentierung in seinen einzelnen Schritten erläutert. Zunächst wird die Extraktion invarianter differentialgeometrischer Merkmale aus dem CAD-Modell betrachtet. In Abschnitt 4.4 wird dann die Berechnung äquivalenter Merkmale aus Distanzbildern erläutert. Die aus dem CAD-Modell extrahierten Merkmale bilden die Grundlage für eine überwachte Klassifikation, wie sie in Abschnitt 4.5 beschrieben wird. In Abbildung 4.4 wird dargestellt, wie sich das grobe Schema des Ablaufs durch diesen frühzeitigen Zugriff auf das Modell ändert.

## 4.3 Bestimmung der Krümmungseigenschaft lokaler Flächenstücke in CAD Daten

### 4.3.1 Flächen mit konstanter Krümmung

Für eine Gruppe von Flächen ist die Bestimmung ihrer Krümmungseigenschaften trivial, da sie konstante Krümmungen über die gesamte Fläche besitzen. Zu dieser Gruppe zählen die Ebene, die Kugel und der Zylinder. Für diese Flächen reichen wenige geometrische Angaben, im Fall einer Ebene sind gar keine Angaben erforderlich, um die Krümmungen berechnen zu können.

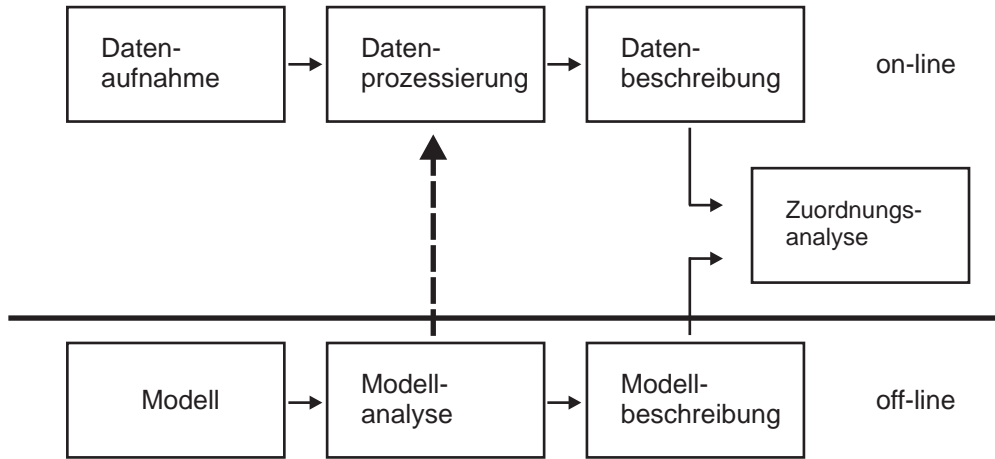


Abbildung 4.4: Veränderter Ablauf eines allgemeinen Schemas des modellbasierten künstlichen Sehens in Anlehnung an Arman und Aggarwal (1993). Bereits in der Stufe der Datenprozessierung wird auf Ergebnisse der Modellanalyse zugegriffen.

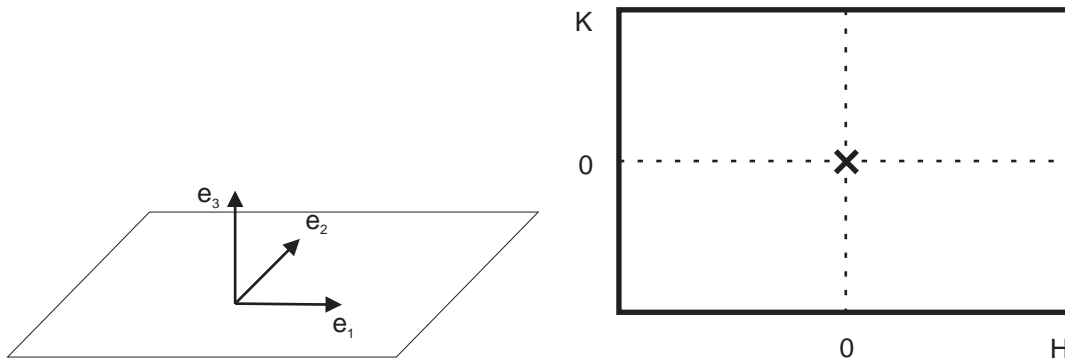


Abbildung 4.5: Definition und Krümmungseigenschaften einer Ebene.

#### 4.3.1.1 Ebene

Eine Ebene lässt sich beschreiben durch zwei senkrecht zueinander stehende Einheitsvektoren  $\vec{e}_1$  und  $\vec{e}_2$ , die in Richtung der Flächenparameter  $u$  und  $v$  liegen, dem Normalenvektor der Ebene  $\vec{e}_3$  und dem Ursprungspunkt der Ebene  $O$ . Aus dieser Beschreibung kann man eine Parametrisierung der Ebene ableiten, wie folgt

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = u * \vec{e}_1 + v * \vec{e}_2 + O$$

Aus dieser Parametrisierung kann wiederum die Krümmungsinformation nach Abschnitt 3.2 hergeleitet werden. Offensichtlich besitzt die Ebene keine Krümmung. Somit erhalten die Hauptkrümmungen  $k_1$  und  $k_2$  beide den Wert null. Ebenso sind  $H$  und  $K$  beide null.

#### 4.3.1.2 Kugel

Eine Kugel kann beschrieben werden durch ihren Ursprung  $O$  und ihren Radius  $r$ . Mit den Parametern  $u$  und  $v$  ergibt sich

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = r * \begin{bmatrix} \cos u \cos v \\ \cos u \sin v \\ \sin u \end{bmatrix} + O \quad u \in \left[-\frac{\Pi}{2}, \frac{\Pi}{2}\right], v \in [0, 2\Pi]$$

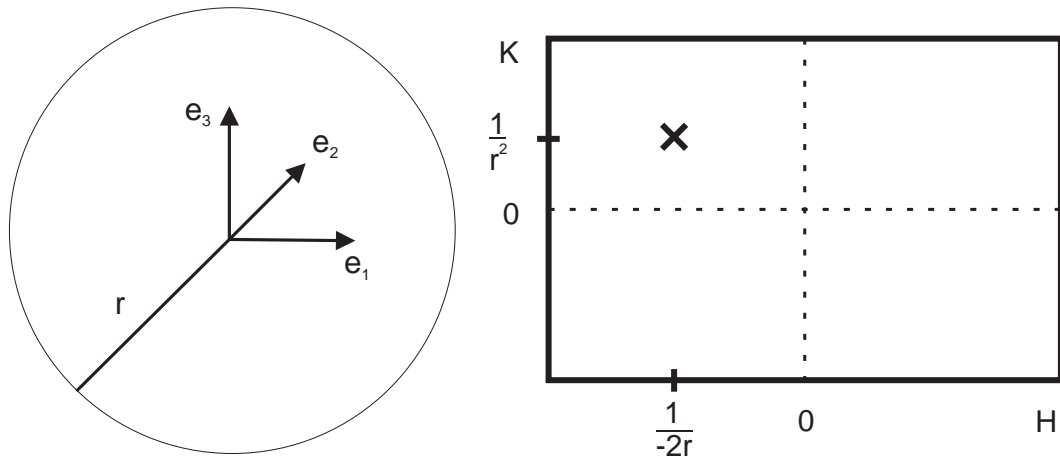


Abbildung 4.6: Definition und Krümmungseigenschaften einer Kugel.

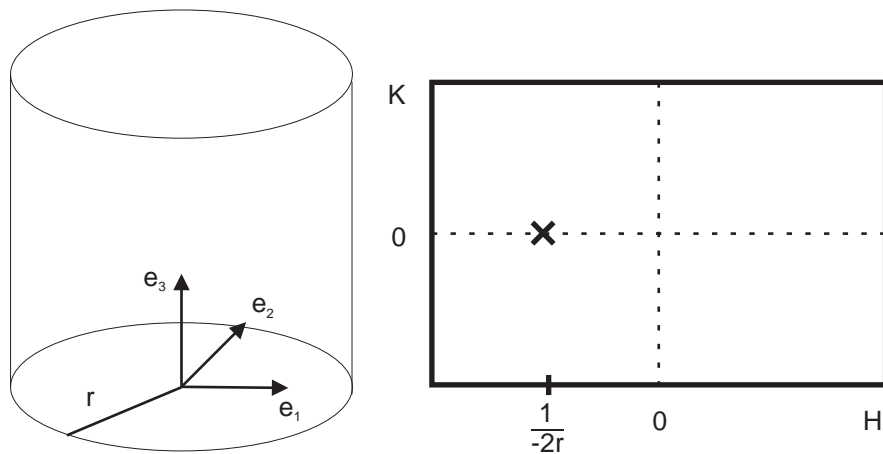


Abbildung 4.7: Definition und Krümmungseigenschaften eines Zylinders.

Die Krümmungen einer Kugel ergeben sich aus der Darstellung als  $k_1 = k_2 = \pm \frac{1}{r}$ . Das Vorzeichen hängt davon ab, ob die Innenfläche oder die Außenfläche einer Kugel betrachtet wird. Dies lässt sich mithilfe des Normalenvektors überprüfen, der entweder in Richtung des Ursprungs zeigt oder davon abgewandt ist. Per Definition wird für eine konvexe Fläche, also für die Außenfläche der Kugel ein negatives Vorzeichen gewählt. Dementsprechend ergibt sich für  $H$  und  $K$ :  $H = k_1$  und  $K = k_1^2$ .

#### 4.3.1.3 Zylinder

Eine Zylinderfläche kann erzeugt werden durch die Rotation einer Geraden um eine parallele Achse. Der Abstand der Geraden zur Rotationsachse ist der Radius  $r$  des Zylinders. Für die Parametrisierung des Zylinders, wie allgemein für Rotationskörper, wird ein lokales Koordinatensystem aus drei senkrechten Einheitsvektoren  $\vec{e}_1$ ,  $\vec{e}_2$  und  $\vec{e}_3$  und dem Ursprungspunkt  $O$  etabliert. Die Rotationsachse ist gegeben durch den Einheitsvektor  $\vec{e}_3$ , die Gerade liegt in der Ebene, die durch  $\vec{e}_1$  und  $\vec{e}_3$  aufgespannt wird. Sie wird von  $\vec{e}_1$  in Richtung  $\vec{e}_2$  gedreht. Die Formel der Parametrisierung ergibt sich zu

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = r * [\cos u * \vec{e}_1 + \sin u * \vec{e}_2] + v * \vec{e}_3 + O$$

Die Hauptkrümmungen ergeben sich zu  $k_1 = 0$  beziehungsweise  $k_2 = \pm \frac{1}{r}$ . Für das Vorzeichen gilt das oben Gesagte analog. Die mittlere und die Gaußsche Krümmung ergeben sich zu  $H = \frac{k_2}{2}$  und  $K = 0$ .

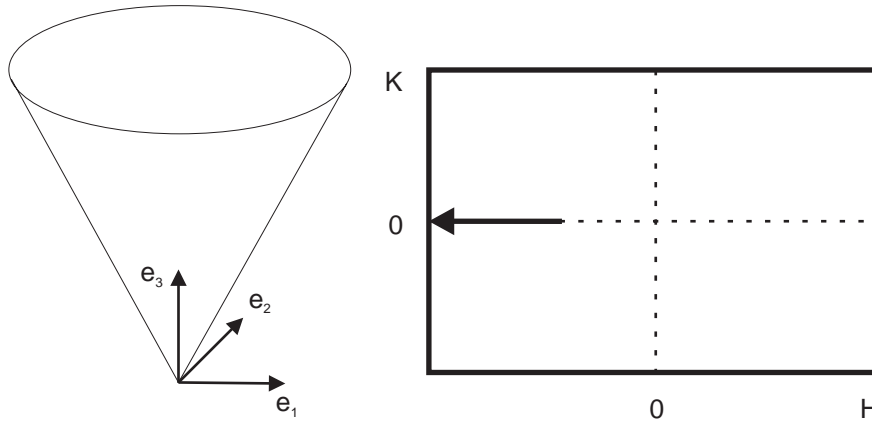


Abbildung 4.8: Definition und Krümmungseigenschaften eines Kegels.

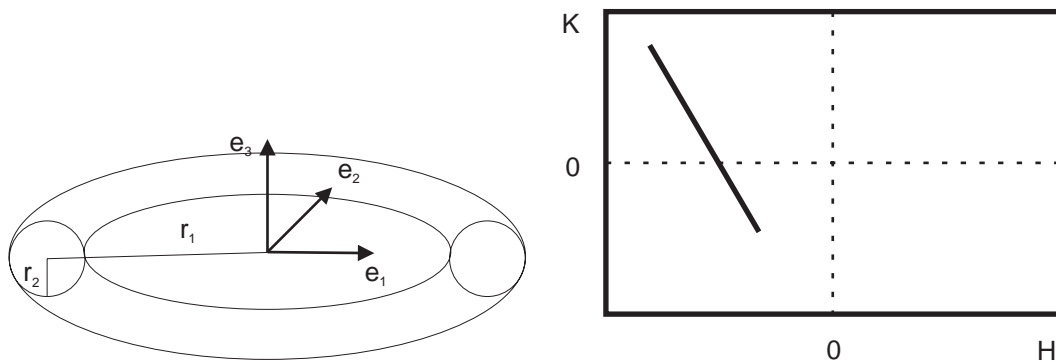


Abbildung 4.9: Definition und Krümmungseigenschaften eines Torus.

## 4.3.2 Rotationsflächen mit variabler Krümmung

Bereits im vorangegangenen Abschnitt wurde eine Rotationsfläche mit konstanter Krümmung, der Zylinder, eingeführt. In diesem Abschnitt werden Rotationsflächen mit variabler, also nicht konstanter Krümmung betrachtet. Für den Kegel und den Torus genügen wiederum geometrische Grundangaben zur Berechnung der Krümmungseigenschaften. Für eine allgemeine Rotationsfläche muss die Krümmung der generierenden Kurve ausgewertet werden.

### 4.3.2.1 Kegel

Im Gegensatz zum Zylinder wird beim Kegel die Gerade in einem Winkel  $\alpha$  zur Rotationsachse gedreht. Der Schnittpunkt der generierenden Geraden mit der Rotationsachse ist der Ursprung des Kegels. Der Parameter  $v$  drückt die Höhe eines Punktes entlang der Achse aus. Der radiale Abstand des Punktes zur Achse ist durch  $v * \tan \alpha$  gegeben. Die Formel der Parametrisierung ergibt sich zu

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = v * \tan \alpha * [\cos u * \vec{e}_1 + \sin u * \vec{e}_2] + v * \vec{e}_3 + O$$

Die Krümmungen variieren mit dem Parameter  $v$ . Eine der beiden Hauptkrümmungen  $k_1$  ist konstant null. Das andere Krümmungsextremum  $k_2$  ergibt sich zu  $\pm \frac{1}{v * \tan \alpha}$ . Die mittlere und die Gaußsche Krümmung ergeben sich zu  $H = \frac{k_2}{2}$  und  $K = 0$ .



### 4.3.2.2 Torus

Ein Torus wird erzeugt durch die Rotation einer Kreiskurve mit dem Radius  $r_2$ . Diese Kurve wird um die Rotationsachse  $\vec{e}_3$  in einem Abstand  $r_1$  gedreht. Der erzeugende Kreis liegt in der Ebene, die durch  $\vec{e}_1$  und  $\vec{e}_2$  aufgespannt wird. Er wird von  $\vec{e}_1$  in Richtung  $\vec{e}_2$  gedreht. Der äußerste Punkt der Fläche liegt in einem Abstand  $r_1 + r_2$  vom Ursprung des Torus. Der radiale Abstand eines beliebigen Punktes ergibt sich aus  $r_1 + r_2 \cos v$ . Die Höhe des Punktes entlang der Rotationsachse ist  $r_2 \sin v$ . Die Formel der Parametrisierung ergibt sich damit zu

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (r_1 + r_2 \cos v) * [\cos u * \vec{e}_1 + \sin u * \vec{e}_2] + r_2 \sin v * \vec{e}_3 + O$$

Die Krümmungen variieren mit dem Parameter  $v$ . Ohne Beschränkung der Allgemeinheit nehmen wir an, dass der Radius  $r_2$  kleiner ist als der Radius  $r_1$ . Eine der beiden Hauptkrümmungen  $k_1$  ist dann konstant  $\pm \frac{1}{r_2}$ . Das andere Krümmungsextremum  $k_2$  variiert zwischen den beiden Extremwerten  $-\frac{1}{r_1+r_2}$  und  $\frac{1}{r_1-r_2}$ , die sich durch den maximalen äußeren Radius und den minimalen inneren Radius des Torus ergeben, wobei sich beim Übergang von der äußeren Ringseite zur inneren Seite noch das Vorzeichen wechselt. Die Gaußsche und die mittlere Krümmung ergeben sich dann entsprechend aus den Formeln. Alternativ kann auch nur die konstante der beiden Hauptkrümmungen aus den geometrischen Parametern abgeleitet werden und die variable Hauptkrümmung entsprechend dem Vorgehen bei Freiformflächen durch regelmäßige Abtastung entlang des Parameters  $v$  ermittelt werden.

### 4.3.2.3 Allgemeine Rotationsfläche

Eine allgemeine Rotationsfläche entsteht durch die Rotation einer Kurve  $s(v)$  um eine Achse. Eine räumliche Kurve liegt typischerweise in einer parametrisierten Form vor, es kann sich zum Beispiel um ein Spline handeln. Die Kurve wird an der Stelle  $v$  ausgewertet. Der so erzeugte Punkt wird um die Achse bewegt. Dies wird für alle Punkte entlang der Kurve wiederholt. Wir verwenden wiederum das in Abschnitt 4.3.1.3 eingeführte lokale Koordinatensystem mit der Rotationsachse  $\vec{e}_3$  und dem Ursprung  $O$ . Die Kurve liegt in der Ebene, die durch  $\vec{e}_1$  und  $\vec{e}_2$  aufgespannt wird. Sie wird von  $\vec{e}_1$  in Richtung  $\vec{e}_2$  gedreht. Mit dem Punkt

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = s(v)$$

auf der erzeugenden Kurve ergibt sich die Parametrisierung zu

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = [c_1 \cos u - c_2 \sin u] * \vec{e}_1 + [c_1 \sin u + c_2 \cos u] * \vec{e}_2 + c_3 * \vec{e}_3 + O$$

Mit dem Abstand  $r$  des Punktes zur Achse ergibt sich eine der beiden Hauptkrümmungen zu  $k_1 = \pm \frac{1}{r}$ . Die andere Krümmung  $k_2$  ergibt sich aus der Krümmung der Kurve im Punkt  $s(v)$ . Daraus lassen sich mit den Formeln 3.18 und 3.19 aus Kapitel 3 die mittlere und die Gaußsche Krümmung berechnen.

## 4.3.3 Regelflächen mit variabler Krümmung

Regelflächen sind ein elementares Werkzeug in der Konstruktion von Flächen. Sie haben alle die Eigenschaft gemein, dass ihre Gaußsche Krümmung null ist. Dies ergibt sich daraus, dass Regelflächen durch die Bewegung einer Geraden im Raum (Abramowski und Müller, 1992) erzeugt werden und somit immer eine der Hauptkrümmungen null ist. Im Folgenden werden die Erzeugung von Regelflächen und die Herleitung ihrer Eigenschaften an zwei Beispielen verdeutlicht.

### 4.3.3.1 Einfache Regelfläche

Eine einfache Regelfläche entsteht durch lineare Interpolation zwischen zwei Kurven im Raum. Zwei Kurven  $c_1(u)$  und  $c_2(u)$  sind gegeben. Beide Kurven werden jeweils an derselben Stelle des Parameters  $u$  ausgewertet.

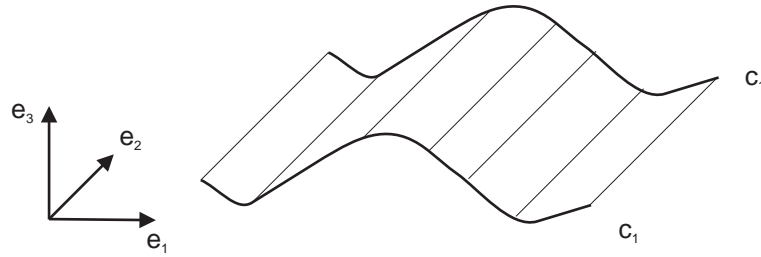


Abbildung 4.10: Eine Regelfläche erzeugt aus zwei identischen Kurven  $c_1$  und  $c_2$ , zwischen denen linear interpoliert wird.

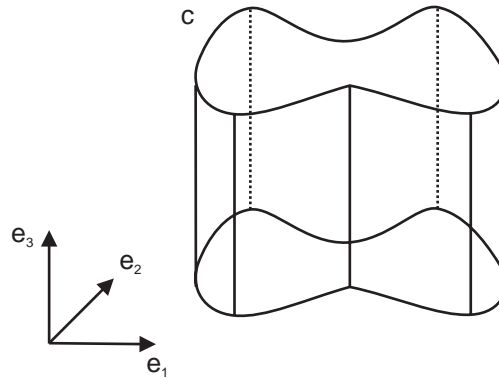


Abbildung 4.11: Eine Regelfläche entstanden aus einem Zugkörper mit der erzeugenden Kurve  $c$ .

Die beiden resultierenden Punkte werden durch eine Gerade verbunden. Der zweite Parameter  $v$  ist der lineare Parameter zwischen den beiden Punkten und läuft entlang der Geraden. Wird dieser Vorgang für alle möglichen Werte von  $u$  wiederholt, entsteht eine Regelfläche (siehe Abbildung 4.10).

Entlang der Geraden ist die Krümmung offensichtlich null. Die zweite der beiden Hauptkrümmungen  $k_2$  hängt ab von den beiden Kurven  $c_1(u)$  und  $c_2(u)$ . Die mittlere Krümmung  $H$  ergibt sich zu  $\frac{k_2}{2}$  und die Gaußsche Krümmung ist wie erwähnt null.

#### 4.3.3.2 Regelfläche aus einem Zugkörper

Eine geschlossene Kurve  $c(u)$  wird entlang einer Geraden durch den Raum bewegt oder “gezogen”. Die resultierende Fläche bezeichnet man auch als Zugkörper. Im speziellen Fall handelt es sich mit dem englischsprachigen Fachwort um einen *tabulated cylinder* (siehe Abbildung 4.11). Ein lokales Koordinatensystem, ähnlich dem der Rotationskörper wird etabliert. Die geschlossene Kurve liegt in der Ebene, die durch  $\vec{e}_1$  und  $\vec{e}_2$  aufgespannt wird. Die Kurve wird entlang der Richtung von  $\vec{e}_3$  mit dem Parameter  $v$  bewegt. Alternativ kann man die Fläche auch beschreiben als lineare Interpolation zwischen zwei identischen axial versetzten geschlossenen Kurven und erhält somit den Zusammenhang mit obiger Herleitung der Regelflächen.

Entlang der Richtung von  $\vec{e}_3$  besitzt die Fläche keine Krümmung. Somit ergeben sich dieselben Krümmungseigenschaften wie oben. Die zweite der beiden Hauptkrümmungen  $k_2$  hängt jedoch nur noch von einer Kurve  $c(u)$  ab.

#### 4.3.4 Freiformflächen mit variabler Krümmung

Die Mächtigkeit und Variabilität moderner CAD-Systeme liegt zum großen Teil in ihrer Fähigkeit, über die oben aufgeführten Grundgeometrien hinaus nahezu jeden denkbaren Oberflächenverlauf zu beschreiben. Diese Fähigkeit wird durch die so genannten Freiformflächen realisiert. Hierzu zählen einfache Abrundungen als Übergänge zwischen zwei oder mehreren Flächen ebenso wie Coons-Patches, Splineflächen, Bezierflächen und Nonuniform Rational B-Spline Surfaces, kurz NURBS.

Derartige Flächen können nicht mehr durch einfache geometrische Kenngrößen, wie Radius oder Winkel beschrieben werden. Stattdessen werden sie durch ein Raster von Kontrollpunkten und Stetigkeitsbedingungen

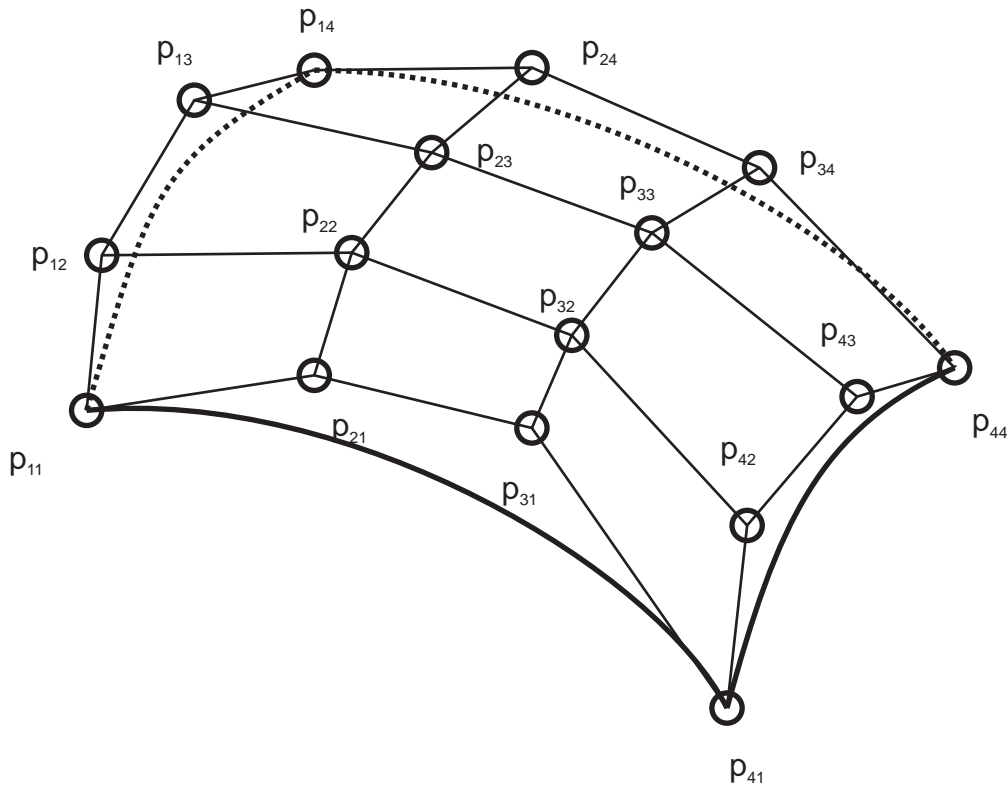


Abbildung 4.12: Eine kubische NURBS-Fläche mit einem  $4 \times 4$  Gitter von Kontrollpunkten.

an diesen Kontrollpunkten beziehungsweise an ihren Randkurven beschrieben. Ein Beispiel für eine kubische NURBS-Fläche und das dazugehörige Gitter der Kontrollpunkte ist in Abbildung 4.12 gegeben.

Daraus folgt, dass auch ihre Krümmungseigenschaften nicht mehr direkt hergeleitet werden können. Stattdessen kann jeder Punkt im Parameterraum  $(u, v)$  mit der jeweiligen Parametrisierung der Fläche ausgewertet werden. Da die Beschreibung der Fläche mathematisch exakt vorliegt, kann auch die erste und zweite Ableitung für jeden Punkt aus der Parametrisierung berechnet werden. Somit können aus den Formeln 3.26 und 3.27 aus Abschnitt 3.2 die Krümmungen berechnet werden. Durch eine regelmäßige Abtastung entlang der Parameter  $u$  und  $v$  ergibt sich ein Raster im Parameterraum. Für jeden Knoten dieses Rasters können die Krümmungen berechnet werden. Die Krümmungseigenschaften einer Teilfläche ergeben sich also nach dieser Methode als die Menge der Krümmungen an den Knotenpunkten.

### 4.3.5 Details der Implementierung

Die Behandlung von CAD-Daten wurde mithilfe eines kommerziellen CAD-Systems realisiert, dem Pro/Engineer CAD-System. Das System bietet eine Programmierschnittstelle, über die alle internen Repräsentationsparameter eines Volumenmodells und seiner Teilflächen abgefragt werden können. Somit war es möglich alle oben in theoretischer Form beschriebenen Operationen zu implementieren. Bei der Implementierung waren jedoch noch einige Details zu beachten.

Zum einen muss im System jede Fläche aufgrund der in Abschnitt 2.2 formulierten Bedingung von *einer* Randkurve begrenzt werden. Dies bedeutet, dass zum Beispiel jede Zylinderfläche in zwei Teilflächen aufgespalten wird. Beide Teilflächen weisen natürlich dieselben Krümmungseigenschaften auf und würden somit doppelt ausgewertet. Diese spezielle Situation muss gesondert behandelt werden. Auch hierzu standen Programmiermöglichkeiten zur Verfügung.

Zum anderen werden in den vorhergehenden Abschnitten immer nur vollständige Flächen ohne Aussparung behandelt. Tatsächlich werden die Flächen durch Randkurven begrenzt. Zum Beispiel bei einer Bohrung oder bei einem Schnitt mit einer anderen Fläche können diese Randkurven die Fläche auch unterbrechen. Daher muss bei der regelmäßigen Rasterung im  $(u, v)$  Parameterraum vor jeder Berechnung immer die Gültigkeit des Punktes untersucht werden. Daher muss getestet werden, ob der Punkt überhaupt auf der Fläche liegt.

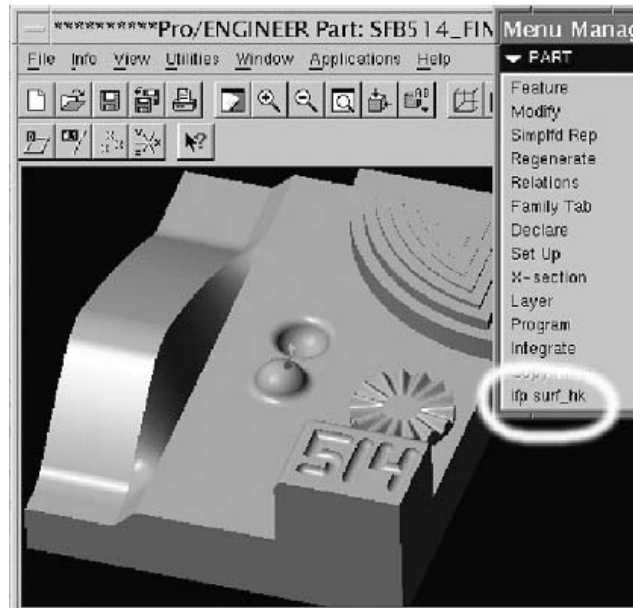


Abbildung 4.13: Einbindung der Krümmungsextraktion in das CAD-System Pro/Engineer.

Die Funktionen zur Berechnung der Krümmungseigenschaften jedes möglichen Flächentyps wurden im CAD-System integriert und können auf beliebige Objekte angewandt werden. Die Ausgabe erfolgt in eine Datei, die für jede Fläche des Objektes ihre Krümmungscharakteristik enthält. Entweder als einzelnes *HK*-Tupel für Flächen mit konstanter Krümmung oder als Vektor von Tupeln für Flächen mit variabler Krümmung. Ein Bildschirmauszug der Benutzeroberfläche des CAD-Systems, das die Einbindung der Funktionalität zeigt, ist in Abbildung 4.13 gegeben.

Die Menge der *HK*-Tupel pro Fläche bilden die Klassen für die im nächsten Abschnitt beschriebene Gruppierung. Um unnötige Mehrdeutigkeiten zu vermeiden, müssen gleichartige Klassen beziehungsweise verschmolzen werden. Identische Krümmungsbeschreibungen für getrennte Flächen sind sowohl aufgrund grundsätzlicher geometrischer Eigenschaften, als auch aufgrund systemspezifischer Eigenheiten nicht auszuschließen. Es können in einem Objekt mehrere Flächen gleicher Krümmung enthalten sein. Dies ist zum Beispiel bei einem Objekt, welches mehr als eine ebene Fläche enthält, immer der Fall. Jede der ebenen Flächen erzeugt ein einzelnes *HK*-Tupel mit den Werten Null. Im implementierten System wird dies aufgrund der gleichartigen Geometrie erkannt und die beiden Klassen werden verschmolzen. Die Information über die Verschmelzung und die zu Grunde liegenden Ausgangsklassen müssen abgespeichert werden. Sie enthalten wesentliche Informationen für die Phase der Zuordnung. Außerdem hat das CAD-System die Eigenheit bestimmte Flächen, wie oben beschrieben, in mehrere Teilflächen aufzutrennen. Dies würde wiederum zu mehreren im Merkmalsraum identischen Klassen führen.

## 4.4 Krümmungsextraktion auf Distanzbildern

Die Krümmungen einer Fläche und andere differentialgeometrische Parameter wie der Normalenvektor und die Hauptkrümmungsrichtungen kommen im dreidimensionalen maschinellen Sehen weit verbreitet zum Einsatz. Zu den Einsatzfeldern gehören neben der Segmentierung und der Flächenklassifikation auch die Flächenrekonstruktion, bei der die differentiellen Parameter den Schätzvorgang stützen können und die Registrierung, bei der die Krümmungen die Korrespondenzfindung stützen, da sie invariant gegen Rotation und Translation sind.

Aufgrund ihrer wichtigen Stellung innerhalb der verschiedenen Anwendungsgebiete sind die Verfahren zur Ableitung von Krümmungseigenschaften aus Distanzbildern intensiv untersucht worden (Abdelmalek, 1990; Tang und Medioni, 1999; Hamann, 1992; Hilton et al., 1995; Koenderink und van Doorn, 1992, 1994; Page et al., 2001; Sacchi et al., 1999; Taubin, 1995; Trucco und Fisher, 1995; Yuen et al., 1999). Mehrere Untersuchungen und Vergleiche von wichtigen Verfahren wurden durchgeführt, sowie Übersichten über bestehende Verfahren erstellt (Stokely und Wu, 1992; Flynn und Jain, 1989; Krsek et al., 1998).

Die Verfahren zur Schätzung der Krümmungswerte aus digitalen Distanzdaten lassen sich grob in zwei Katego-

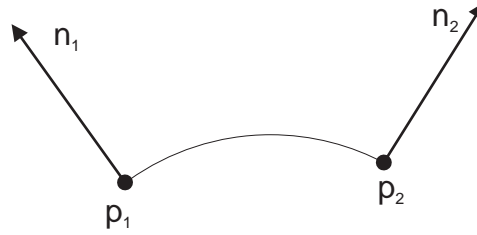


Abbildung 4.14: Die Normalenvektoren  $\mathbf{n}_1$  und  $\mathbf{n}_2$  zweier benachbarter Punkte  $\mathbf{p}_1$  und  $\mathbf{p}_2$ .

rien unterteilen, numerische Verfahren und analytische Verfahren. Zu den numerischen Verfahren gehören die Krümmungsschätzung durch Beobachtung der Veränderungen im Normalenvektor einer Fläche sowie die Berechnung der Krümmung aus den numerischen Ableitungen. Die analytischen Verfahren basieren alle auf einer lokalen Flächenrückführung (siehe Abschnitt 3.3). Die grundsätzliche Strategie dieser Verfahren besteht aus drei Teilschritten

1. Die Bestimmung einer lokalen Flächenfunktion in der Nachbarschaft des zu untersuchenden Punktes.
2. Die analytische Berechnung der partiellen Ableitungen aus den Parametern der lokalen Flächenfunktion.
3. Die Berechnung der Krümmungen aus den partiellen Ableitungen nach den Formeln 3.26 und 3.27 aus Abschnitt 3.2.

#### 4.4.1 Krümmungen aus der Normalenänderung

Um die Krümmungen aus der Beobachtung der Veränderung des Normalenvektors abzuleiten, muss für jeden Punkt zunächst der Normalenvektor bestimmt werden. Dies kann zum Beispiel einfach durch die Bestimmung einer lokalen Ebene erfolgen. Die Ebene kann mit einem einfachen Eigenwert Verfahren bestimmt werden (Duda und Hart, 1973). Details hierzu finden sich auch im Anhang A. Der Normalenvektor der lokalen Ebene wird dann zum Normalenvektor in diesem Punkt.

Ausgehend von einem Punkt  $\mathbf{p}_1$  wird die Veränderung des Normalenvektors im Vergleich zu seinem benachbarten Punkt  $\mathbf{p}_2$  untersucht (Abbildung 4.14). Die Schätzung der Krümmung im Punkt  $\mathbf{p}_1$  in Richtung von  $\mathbf{p}_2$  ergibt sich zu (Hoffman und Jain, 1986)

$$k_{1-2} = \frac{\|\mathbf{n}_1 - \mathbf{n}_2\|}{\|\mathbf{p}_1 - \mathbf{p}_2\|} \quad (4.1)$$

Das Vorzeichen dieser Krümmungsschätzung muss noch durch die Bedingung

$$\|\mathbf{p}_1 - \mathbf{p}_2\| \leq \|(\mathbf{n}_1 + \mathbf{p}_1) - (\mathbf{n}_2 + \mathbf{p}_2)\| \quad (4.2)$$

bestätigt werden, sonst muss der Wert negiert werden.

Die so geschätzte Krümmung ist gültig entlang der Richtung von  $\mathbf{p}_1$  nach  $\mathbf{p}_2$ . Um die Hauptkrümmungen zu erhalten, muss diese Schätzmethode für möglichst viele Nachbarpunkte von  $\mathbf{p}_1$  in möglichst verschiedenen Richtungen wiederholt werden. Das Maximum und das Minimum der geschätzten Werte ergeben dann die Hauptkrümmungen.

Die Verfahren, welche die Krümmungen aus der Beobachtung der Veränderung des Normalenvektors ableiten, werden häufig auf triangulierten Daten angewandt. Eine alternative Formulierung speziell für triangulierte Daten wurde von Sacchi et al. (1999) vorgeschlagen. Mit der Formel

$$k_{1-2} = \frac{\|\mathbf{n}_1 \times \mathbf{n}_2\|}{\|\mathbf{p}_1 - \mathbf{p}_2\|}$$

können auch Punkte  $\mathbf{p}_i$  behandelt werden, die nicht auf der Fläche liegen, etwa wenn sie durch Interpolation aus Nachbarpunkten abgeleitet wurden.

### 4.4.2 Krümmungsschätzung aus Ableitungen

Aus den gerichteten Ableitungen eines Distanzbildes können die Hauptkrümmungen nach folgendem Verfahren ermittelt werden (Fan et al., 1986). Ausgehend von einer bestimmten Richtung  $\psi$  im Distanzbild, werden mittels Filterkernen die ersten und zweiten Ableitungen  $f'_\psi$  und  $f''_\psi$  entlang dieser Richtung durch Faltung berechnet. Die Krümmung in einem Punkt ergibt sich dann aus

$$k_\psi(u, v) = \frac{f''_\psi(u, v)}{\left(1 + f'_\psi(u, v)\right)^{\frac{3}{2}}}$$

Werden für  $\psi$  die Richtungen 0, 45, 90 und 136 Grad gewählt, ergeben sich die Krümmungen  $k_0, k_{45}, k_{90}$  und  $k_{135}$ . Aus diesen können wiederum mit

$$\begin{aligned} H &= \frac{1}{4}(k_0 + k_{45} + k_{90} + k_{135}) \\ K &= k_0 k_{90} + k_{45} k_{135} - H^2 \end{aligned}$$

die mittlere und die Gaußsche Krümmung berechnet werden.

Numerische Methoden zur Krümmungsbestimmung besitzen den Vorteil, dass sie sehr effizient arbeiten und daher eine sehr rasche Berechnung auch auf großen Datensätzen ermöglichen. Sie zeigen jedoch deutliche Mängel bei der Genauigkeit der Krümmungsschätzung. Insbesondere die Ermittlung der zweiten Ableitung bei den ableitungsbasierten Verfahren erweist sich als extrem rauschempfindlich.

### 4.4.3 Methoden zur analytischen Krümmungsschätzung

Die wohl genauesten Methoden zur Berechnung der Krümmung liefern die Verfahren zur analytischen Krümmungsschätzung. Die bekannten Verfahren unterscheiden sich in der Wahl der Funktion  $f$  zur Flächenapproximation. Grundsätzlich kommen alle im Abschnitt Section 3.3 zur lokalen Flächenrückführung vorgestellten Verfahren in Betracht. Tatsächlich haben sich vor allem zwei Verfahren etabliert, die Approximation mittels orthogonaler Polynome und die Approximation durch ein explizites bivariates Polynom.

#### 4.4.3.1 Orthogonale Polynome

Besl und Jain (1986) entwickelten eine Methode zur lokalen Flächenapproximation basierend auf orthogonalen Polynomen. Jedem Punkt  $p_i$  innerhalb einer  $N \times N$  Maske in der Umgebung eines bestimmten Punktes  $p$  aus dem Distanzbild wird ein Tupel  $(u, v)$  der lokalen Parametrisierung zugewiesen. Dabei stammt  $(u, v)$  aus dem Wertebereich  $U \times U$  mit

$$U = \left\{ -\frac{N-1}{2}, \dots, -1, 0, 1, \dots, \frac{N-1}{2} \right\}$$

Die drei orthogonalen Polynome

$$\Phi_0(u) = 1, \quad \Phi_1(u) = u, \quad \Phi_2(u) = \left(u^2 - \frac{M(M+1)}{3}\right) \quad (4.3)$$

mit  $M = \frac{N-1}{2}$  ermöglichen eine biquadratische Approximation. Der zweite Grad eines Polynoms genügt aus dem bereits in Abschnitt 3.3.4 genannten Grund der zweimaligen Differenzierbarkeit.

Zu den Polynomen  $\Phi_i(u)$  wird noch ein korrespondierender Satz von normalisierten Polynomen  $b_i(u)$  berechnet mit den Normalisierungskonstanten  $P_i(M)$  und  $b_i(u) = \frac{\Phi_i(u)}{P_i(M)}$ . Die normierten Basisfunktionen  $b_i(u)$  erfüllen mit den Basisfunktionen  $\Phi_i(u)$  die Orthogonalitätsbedingung

$$\sum_{i+j \leq 2} \Phi_i(u) b_j(u) = \delta_{ij} \quad (4.4)$$

wobei das Kronecker-Symbol  $\delta_{ij} = 1$  für  $i = j$  und  $\delta_{ij} = 0$  sonst.

Eine Flächenfunktion  $z_{uv} = f(u, v)$  kann nun in die lokale Umgebung  $U^2$  eines Punktes eingepasst werden durch

$$f(u, v) = \sum_{i+j \leq 2} a_{ij} \Phi_i(u) \Phi_j(v) \quad (4.5)$$

wobei der quadratische Fehler

$$\epsilon^2 = \sum_{(u,v) \in U^2} (z_{uv} - f(u, v))^2 \quad (4.6)$$

zu minimieren ist. Die gesuchten Koeffizienten  $a_{ij}$  können durch

$$a_{ij} = \sum_{(u,v) \in U^2} z_{uv} b_i(u) b_j(v) \quad (4.7)$$

erhalten werden. Für die ersten und zweiten partiellen Ableitungen an der Stelle  $(u, v) = (0, 0)$  erhält man

$$f_u = a_{10}, \quad f_v = a_{01}, \quad f_{uv} = a_{11}, \quad f_{uu} = 2a_{20}, \quad f_{vv} = 2a_{02} \quad (4.8)$$

Hieraus kann man wiederum nach dem üblichen Schema die Krümmungen berechnen.

Dieses Verfahren ist besonders effizient, da die zur Berechnung der Parameter  $a_{ij}$  benötigten Basisfunktionen  $b_i(u)$  beziehungsweise  $b_j(v)$  für jede feste ungerade Maskengröße  $N$  vorab berechnet werden können. Die Berechnung selbst kann als Faltung des Distanzbildes mit Filtermasken implementiert werden, wobei die separierbaren Filtermasken aus den Basisfunktionen abgeleitet werden. Für eine typische Maskengröße von  $N = 7$  erhält man so zum Beispiel die eindimensionalen Filterkerne

$$d_0 = \frac{1}{7} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T$$

$$d_1 = \frac{1}{28} \begin{bmatrix} -3 & -2 & -1 & 0 & 1 & 2 & 3 \end{bmatrix}^T$$

$$d_2 = \frac{1}{84} \begin{bmatrix} 5 & 0 & -3 & -4 & -3 & 0 & 5 \end{bmatrix}^T$$

Damit lassen sich die zweidimensionalen Filtermasken zur Bestimmung der partiellen Ableitung darstellen zu

$$\begin{aligned} D_u &= d_0 d_1^T & D_v &= d_1 d_0^T \\ D_{uu} &= d_0 d_2^T & D_{vv} &= d_2 d_0^T & D_{uv} &= d_1 d_1^T \end{aligned} \quad (4.9)$$

Dem Vorteil der effizienten Realisierung und dem damit verbundenen Laufzeitvorteil stehen aber auch einige Nachteile gegenüber. So geht die ursprüngliche Notation von einem äquidistanten Raster der Distanzdaten mit einer Rasterweite von 1 aus. Eine abweichende Rasterweite kann durch eine einfache Reparametrisierung eingebunden werden. Die Äquidistanz des Rasters jedoch ist für die Korrektheit Voraussetzung. In der in Abschnitt 2.1.3 eingeführten Bezeichnung ist dies die  $r_{ij}$ -Form eines Distanzbildes. Diese Form lässt sich nur durch eine Reinterpolation aus einer  $XYZ$ -Form ableiten und auch dies ist nicht für alle Objektgeometrien verlustfrei möglich. In einer lokalen Umgebung eines Punktes ist jedoch typischerweise keine gravierende Veränderung der Rasterweite zu erwarten, sodass dieses Verfahren auch auf nicht äquidistanten Distanzbildern eingesetzt werden könnte.

Das Verfahren ist besonders dann sehr schnell, wenn es auf optimierten Faltungsoperatoren aufbaut, wie sie von kommerziellen Bildverarbeitungssystemen angeboten werden. Jedoch behandeln solche Operatoren alle Punkte innerhalb der Maske gleich, das heißt, es kann keine Gewichtung einzelner Beobachtungen vorgenommen werden oder einzelne Punkte ausgeschlossen werden. Da Distanzbilder aber häufig Lücken von einzelnen Pixeln aufweisen, führt dies zu einer starken Verfälschung der Krümmungen. Abbildung 4.15 zeigt die Auswirkungen solcher Fehler bei der Implementierung mittels optimierter Faltungsoperatoren. Die quadratischen Artefakte in dem Krümmungsbild rühren von einzelnen Fehlpixeln her.

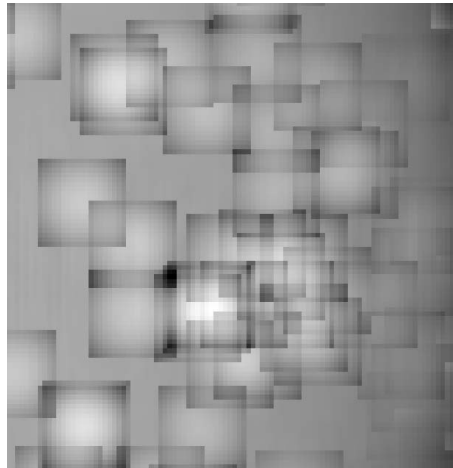


Abbildung 4.15: Artefakte einzelner Fehlpixel bei der Filtermethode.

#### 4.4.3.2 Approximation durch ein explizites bivariates Polynom

Eine weitere Möglichkeit zur analytischen Krümmungsschätzung ergibt sich durch die direkte Schätzung eines expliziten bivariaten Polynoms zweiten Grades, wie es in Gleichung 3.36 beschrieben wurde. Wie auch bei der Methode der orthogonalen Polynome, wird ein Punkt  $p(x_p, y_p, z_p)$  in seiner Umgebung betrachtet, die durch eine  $N \times N$  Maske im Distanzbild gegeben ist. Jeder Punkt  $p_i$  innerhalb der Umgebung wird als Beobachtung in den Schätzprozess eingebracht. Mit der Methode der kleinsten Quadrate werden dann die Parameter  $a_1 \dots a_6$  der Flächenfunktion  $f$  bestimmt. Die partiellen Ableitungen im Punkt  $p$  ergeben sich dann zu

$$\begin{aligned} f_x &= 2a_1x_p + a_2y_p + a_4 \\ f_y &= 2a_1y_p + a_2x_p + a_5 \\ f_{xx} &= 2a_1 \\ f_{xy} &= a_2 \\ f_{yy} &= 2a_3 \end{aligned}$$

Es ist von Vorteil das lokale Koordinatensystem der Beobachtungen in das Zentrum  $p$  der Maske zu verschieben:

$$\bar{p}_i = p_i - p$$

sodass sich die Koordinaten von  $p$  zu null ergeben. Damit werden die obigen Formeln der partiellen Ableitungen nochmals vereinfacht zu

$$\begin{aligned} f_x &= a_4 \\ f_y &= a_5 \\ f_{xx} &= 2a_1 \\ f_{xy} &= a_2 \\ f_{yy} &= 2a_3 \end{aligned}$$

Nachteilig am Verfahren mit Polynomen zweiten Grades ist die Tatsache, dass die Krümmungen am betrachteten Punkt meist überschätzt werden, wie dies auch schon von Krsek et al. (1998) bemerkt wurde. Der Grund liegt darin, dass der Paraboloid zweiten Grades an seinem Scheitelpunkt ein Krümmungsmaximum besitzt, welches die Krümmungen in der Umgebung eines Punktes auf einer Fläche mit konstanter Krümmung, zum Beispiel einer Kugel, nicht exakt widerspiegelt.

Es ist auch nochmals wie in Abschnitt 3.3.5.1 darauf hinzuweisen, dass das zu minimierende Fehlerkriterium nicht die Realität der Sensorsysteme widerspiegelt und vom gewählten Koordinatensystem abhängig ist. Auch Flynn und Jain (1989) bemerken einen systematischen Fehler in der Schätzung, abhängig von der Orientierung des Flächenstücks. Als Korrektur kann eine Neuorientierung der Fläche durch eine lokale Hauptachsentransformation durchgeführt werden, bevor die Beobachtungen in den Schätzprozess eingeführt werden. Dies verursacht jedoch



einen erheblichen Mehraufwand. Zudem sind bei der Betrachtung eines einzelnen Distanzbildes die Daten bereits zum Beobachtungsstandpunkt hin ausgerichtet. In den üblicherweise angesetzten Sensorkoordinatensystemen bedeutet dies bereits eine ungefähre Ausrichtung der Normalen entlang der z-Achse. Somit bringt eine zusätzliche Rotation keinen Verbesserungen bei der Schätzung.

Die Methode der expliziten Polynomschätzung wurde in dieser Arbeit verwendet, da sie eine sehr flexible Verarbeitung der Daten erlaubt. Die Punkte aus der Umgebung können einzeln gewichtet werden, wenn eine Genauigkeitsinformation pro Punkt aus den Sensordaten ableitbar ist. Somit können auch einzelne Fehlmessungen oder Lücken in den Daten bei der Krümmungsbestimmung gezielt ausgeschlossen werden. Dies vermeidet Probleme, wie sind in Abbildung 4.15 angedeutet sind. Allerdings hat die Methode der expliziten Polynome gegenüber der Methode der orthogonalen Polynome einen deutlichen Nachteil bezüglich der Rechenzeit.

#### 4.4.4 Problematik der Krümmungsschätzung

Bereits in einer sehr frühen Studie zu den Verfahren der Krümmungsextraktion bemerkten Flynn und Jain (1989), dass sich alle untersuchten Verfahren extrem empfindlich gegenüber Rauschen verhalten und ihre direkte Anwendung auf Distanzdaten zu völlig unbefriedigenden Ergebnissen führt. Um die Rauschempfindlichkeit zu kompensieren, müssen die Daten mehrfach geglättet werden. Die Autoren berichten von Glättungsverfahren, die aus 10 und bis zu 20 Anwendungen eines Glättungsfilters bestehen. Die Folgen solch extensiver Glättung sind offensichtlich. Kanten und andere Details, die in den Daten enthalten sind, gehen verloren. Flynn und Jain (1989) argumentieren, dass die Methoden ohnehin die Berechnung der Krümmungen innerhalb "glatter" Regionen zum Ziel haben und nicht die Kantendetektion und daher die Auswirkungen der Glättung akzeptabel sind. Dies ist allerdings vor dem Hintergrund zu sehen, dass die meisten der zu diesem Zeitpunkt berücksichtigten Anwendungen lediglich das Vorzeichen der Krümmungen auswerten und nicht deren tatsächlichen Wert verwenden. Der absolute Wert der Krümmung ist wesentlich empfindlicher.

Die Glättung beziehungsweise die Filterung der Daten wird somit zu einem wichtigen Thema. Boulanger et al. (2002) entwickelten spezielle Filterverfahren, die auf einem physikalischen Sensormodell beruhen. Alvertos und D'Cunha (1991) berichten über den Einfluss der Medianfilterung auf Distanzbilder, konnten aber mit der Methode keine überzeugenden Ergebnisse erzielen. Karbacher et al. (2001) berichten, dass Medianfilter ungeeignet sind, um Rauschen zu eliminieren und geben einem einfachen Mittelwertfilter als erstem Glättungsschritt den Vorzug. Daran anschließend werden weitere komplexere Filterschritte unternommen, die einer lokalen Flächenrückführung sehr ähnlich sind und somit im Kontext einer Krümmungsschätzung redundant sind.

Neben den Verfahren zur Filterung spielt die Größe der eingesetzten Masken eine Rolle, sowohl zur Krümmungsbestimmung als auch zur Glättung. Grundsätzlich gilt je größer die eingesetzten Masken sind, desto besser kann das Rauschen unterdrückt werden. Je kleiner die Masken gewählt werden desto ungenauer wird die Krümmungsbestimmung. Der Zusammenhang zwischen dem Rauschen der Messungen und der Genauigkeit der Krümmungsbestimmung wurde von Abdelmalek (1990) und Hilton et al. (1995) untersucht. Dabei weist Abdelmalek (1990) auch statistisch nach, dass der Fehler mit zunehmender Maskengröße abnimmt. Die komplexen Terme zur Abschätzung einer oberen Grenze des Fehlers bei der Bestimmung der mittleren und der Gaußschen Krümmung sind in der referenzierten Literatur nachzulesen. Wie der Autor selbst anmerkt, sind diese Terme zu groß und werden in der Praxis nicht erreicht. Hilton et al. (1995) gehen von einer Krümmungsschätzung nach der Methode der orthogonalen Polynome aus. Sie geben realistischere Terme zur Fehlerabschätzung an und bestätigen diese theoretische Abschätzung an praktischen Experimenten. Beide Abschätzungen benötigen dabei den zu erwartenden Fehler in der Tiefenmessung als Eingabe. Dieser Fehler steht allerdings nicht ohne weiteres zur Verfügung.

Ein zweiter gegenläufiger Aspekt der Maskengröße ist ebenso zu beachten. Bei großen Masken wird die untersuchte Umgebung des Punktes ausgedehnt. Somit steigt die Wahrscheinlichkeit, dass der "glatte" Bereich der Fläche verlassen wird. Durch Diskontinuitäten in der Umgebung eines Punktes werden die Krümmungen erheblich verfälscht. Somit nimmt die Genauigkeit der Krümmungen wiederum ab (Arman und Aggarwal, 1993).

Da das Rauschen der Daten sowohl vom eingesetzten Sensor als auch von der Objektoberfläche und vielen weiteren Umgebungsbedingungen abhängt, und der durch das Rauschen verursachte Fehler gegen die Lokalität der Schätzung der Objektgeometrie abgewägt werden muss, kann es selbstverständlich keine allgemein gültige Maskengröße geben. Stattdessen werden für die hier vorgestellten Ergebnisse in einer vorab durchgeführten Messung die Werte kalibriert. Dabei wurde für das eingesetzte Sensorsystem und eine kalibrierte Sollgeometrie - eine Kugel mit definiertem Durchmesser - eine Messung durchgeführt, die Krümmungswerte bestimmt und mit den Sollwerten verglichen. Diese Vorgehensweise hat deutliche Ähnlichkeit mit den Bestimmungen zur Abnahme

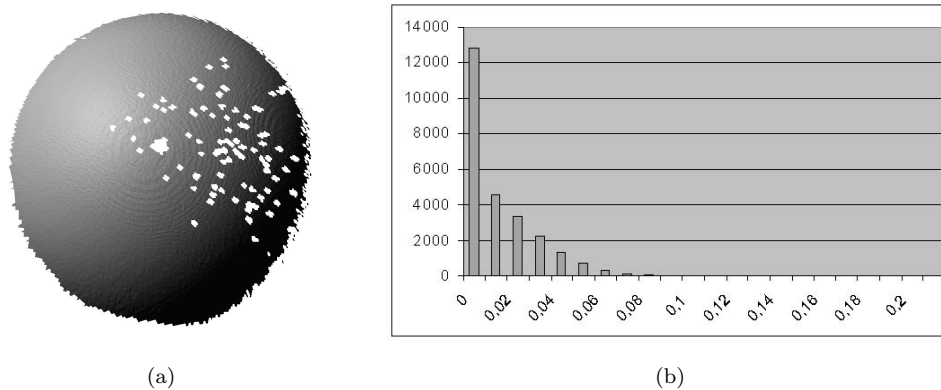


Abbildung 4.16: Untersuchungen zur Genauigkeit eines Streifenprojektionssystems. (a) Schattierte Darstellung eines Datensatzes, der mit einem Streifenprojektor erfasst wurde. Das Objekt ist eine mattierte Stahlkugel mit kalibrierter Rundheit. (b) Abweichung der Messung gegenüber der Sollform. Das Balkendiagramm trägt die Anzahl der Messpunkte über der Abweichung in Millimetern ab.

und Überwachung von optischen 3D-Messsystemen nach VDI/VDE 2634 Blatt 2 (VDI, 2002). Abbildung 4.16 zeigt das Ergebnis einer solchen Probemessung.

## 4.5 Integration der Modell-Information in den Extraktionsprozess

Sind in dem Distanzbild für jeden Bildpunkt die Krümmungen berechnet, kann eine Gruppierung der Pixel nach ihren Krümmungen zu Regionen erfolgen. Dieser Schritt ähnelt denen von klassischen Klassifikationsschemata, bei denen die Pixel in ebene, konvex oder konkav gekrümmte Regionen eingeteilt werden. Insbesondere hat er Gemeinsamkeiten mit dem von Besl (1988b) entwickelten Klassifikationsschema, das zwischen acht verschiedenen Typen unterscheidet, je nach dem Vorzeichen der mittleren und der Gaußschen Krümmung (siehe 3.2.4 auf Seite 30). Jedoch wird bei jenen Ansätzen nach einem starren Klassifikationsschema gearbeitet.

Im Gegensatz dazu soll, wie bereits in Abschnitt 4.2 motiviert, in dem hier entwickelten Verfahren eine modellbasierte Gruppierung erfolgen. Hierzu wird vorab in der so genannten off-line Phase die notwendige Information aus dem CAD-Modell abgeleitet, wie dies in Abschnitt 4.3 dargestellt wurde.

Für jede Teilfläche der Randbeschreibung des Modells wird die mittlere und die Gaußsche Krümmung berechnet. Jede Teilfläche wird dabei durch eine Menge  $C_i$  repräsentiert, die aus den Merkmalsvektoren gebildet wird. Die Merkmalsvektoren  $v_j$  enthalten dabei eben die mittlere und die Gaußsche Krümmung, über denen sich der Merkmalsraum aufspannt. Die Menge  $C_i$  kann dabei aus nur einem Vektor bestehen, wenn die Fläche konstante Krümmungseigenschaften besitzt, oder aus mehreren Vektoren, wenn die Krümmungen über der Fläche variieren und die Merkmalsvektoren im Parameterraum der Fläche abgetastet wurden. Somit ergibt sich für die  $i$ -te Teilfläche

$$C_i = \{v_j \in H \times K \mid j \geq 1\}$$

Die Mengen  $C_i$  bilden die Masterklassen für eine überwachte Klassifikation des Distanzbildes. Für jedes Pixel, das noch unklassifiziert ist, wird die korrekt Klasse in einem Nächste-Nachbar-Verfahren bestimmt.

Abbildung 4.17 zeigt das Beispiel einer Testszene, welche aus drei grundlegenden Flächentypen - Ebene, Kugel und Zylinder - besteht. Das Intensitätsbild der Szene ist dem Krümmungshistogramm des Distanzbildes gegenübergestellt. Das Krümmungshistogramm zeigt deutlich die drei zu erwartenden Häufungspunkte, die den drei Flächentypen entsprechen, die in Tabelle 4.1 aufgeführt sind.

Die Distanzfunktion  $d_n$  für die Nachbarschaftsanalyse ergibt sich aus der Euklidischen Distanz  $d_E$  des Merkmalsvektors  $w_j$  für den  $j$ -ten Bildpunkt zur Menge  $C_i$  durch

$$d_N(w_j, C_i) = \min_{v_k \in C_i} [d_E(w_j, v_k)]$$

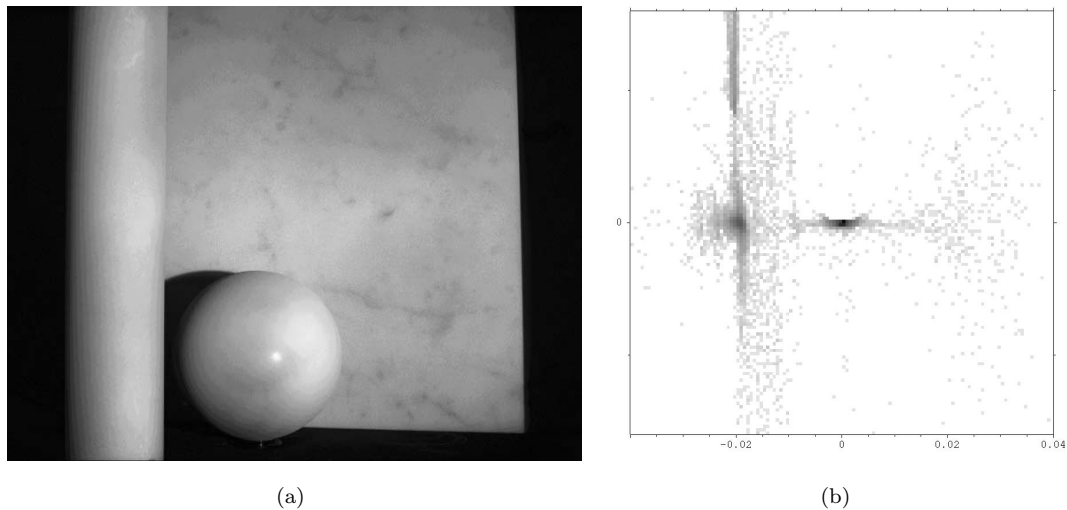


Abbildung 4.17: Untersuchungen an einer Testszene aus verschiedenen Flächentypen. (a) Intensitätsbild der Test Szene aufgenommen mit einem Streifenprojektionssystem. (b) Zweidimensionales Krümmungshistogramm der Gaußschen Krümmung über der mittleren Krümmung.

Flächentyp	H	K
Ebene	-0.0000	0.0000
Zylinder	-0.0200	0.0000
Kugel	-0.0200	0.0004

Tabelle 4.1: Die Tabelle der Krümmungen der drei Flächentypen aus der Testszene.

Die Zuweisung des noch unklassifizierten Merkmalsvektors  $w_j$  erfolgt zur Klasse mit dem geringsten Abstand, also

$$w_j \in C_i \text{ wenn } d_N(w_j, C_i) = \min_l [d_N(w_j, C_l)]$$

Alternativ zur Euklidischen Distanz kann auch eine gewichtete Distanzfunktion angesetzt werden. Die Mahalanobis Distanz ist definiert durch

$$d_M(w_j, v_k) = ((w_j - v_k)^T \Sigma^{-1} (w_j - v_k))^{\frac{1}{2}}$$

mit der Kovarianzmatrix  $\Sigma$ .

Der Einsatz einer gewichteten Distanzfunktion zur krümmungsbasierten Klassifikation motiviert sich durch die Beobachtung, dass die Gaußsche Krümmung eine größere Streuung aufweist als die mittlere Krümmung, wie dies auch aus Abbildung 4.17 ersichtlich ist. Dies kann auch durch Berechnungen nachgewiesen werden, die zum Beispiel von Abdelmalek (1990) durchgeführt wurden. Cantzler und Fisher (2001) verweisen auf eine Formel für die Herleitung des Schwellwerts  $\tau_K$  für die Gaußsche Krümmung aus dem Schwellwert für die mittlere Krümmung  $\tau_H$ , die auch als Formel für die Herleitung einer Gewichtung aufgefasst werden kann:

$$\tau_K = \tau_H * (\tau_H + 2 * \max |H|)$$

In der praktischen Anwendung ist es jedoch oft schwierig das Maximum der mittleren Krümmung  $\max |H|$  in den Bilddaten herzuleiten. In den im Rahmen dieser Arbeit durchgeführten Experimenten konnte bei den untersuchten Datensätzen kein Vorteil einer gewichteten Distanzfunktion festgestellt werden.

Um aus dem Klassifikationsergebnis im Bildraum zusammenhängende Regionen zu erzeugen, wird anschließend eine Analyse der Zusammenhangskomponenten durchgeführt. Hierbei wird auf die Implementierung einer kommerziellen Bildverarbeitungsbibliothek zurückgegriffen.

## 4.6 Experimente

### 4.6.1 Testszene mit Grundkörpern

Erste Untersuchungen wurden auf einer Testszene gemacht, welche aus einfachen Grundkörpern besteht. Sie besteht aus einer Stahlkugel mit bekanntem Radius, die mit Zinkoxydpulver besprüht wurde (ein Verfahren, das sich in der Praxis etabliert hat, um spiegelnde Oberflächen mit optischen Sensoren zu vermessen), einem weißen Kunststoffzylinder mit dem halben Radius der Kugel und einer ebenen Steinplatte, die ebenfalls besprüht wurde, um Reflexionen zu unterdrücken. Ein Intensitätsbild dieser Szene ist in Abbildung 4.17 (a) gegeben. Die Szene wurde mit einem eigenen am Institut entwickelten Streifenprojektionssystem nach dem Line-Shift Prinzip und mit photogrammetrischer Kalibrierung aufgenommen und ausgewertet.

Mit den bekannten Sollwerten der mittleren und der Gaußschen Krümmung wurde eine Klassifikation des Distanzbildes durchgeführt. Das Resultat der Klassifikation ist in Abbildung 4.18 (a) gegeben. Aus der Abbildung wird eine deutliche Anzahl von falsch klassifizierten Pixeln ersichtlich. Diese beruhen auf falschen Schätzungen der Krümmungen im Wesentlichen verursacht durch Sensorrauschen. Systematische Fehler entlang der Kanten des Zylinders, an denen Gruppen von Pixeln in lokal sphärische Flächentypen klassifiziert wurden, resultieren aus einer systematischen Welligkeit der Sensordaten. Dies ist ein typisches Problem bei der Streifenprojektion, bei der sich die Rasterstruktur der Projektionseinheit mit dem Raster des digitalen Bildsensors der Kamera überlagert. Diese Überlagerung verursacht ein Moiré Muster, welches sich in Form von konzentrischen Wellen auf dem Distanzbild niederschlägt.

Das Resultat der Klassifikation kann durch die Anwendung eines morphologischen Filters, welcher kleine einzeln stehende Regionen entfernt, und durch ein Regionenwachstum, welches Lücken innerhalb großer Regionen schließt, verbessert werden. Das Ergebnis dieses Säuberungsschritts ist in Abbildung 4.18 (b) gegeben. Der Schritt des Regionenwachstums kann allerdings nur ausgeführt werden, wenn die in der Szene abgebildeten Flächen einem bestimmten Flächentyp zugehören, der leicht rückführbar ist und sich damit als Homogenitätskriterium eignet. In diesem Fall kann die Homogenität mit einem einfachen Polynom zweiten Grades überprüft werden. Da dies aber nicht generell erfüllt ist, wird in den folgenden Untersuchungen dieser Schritt nicht durchgeführt.

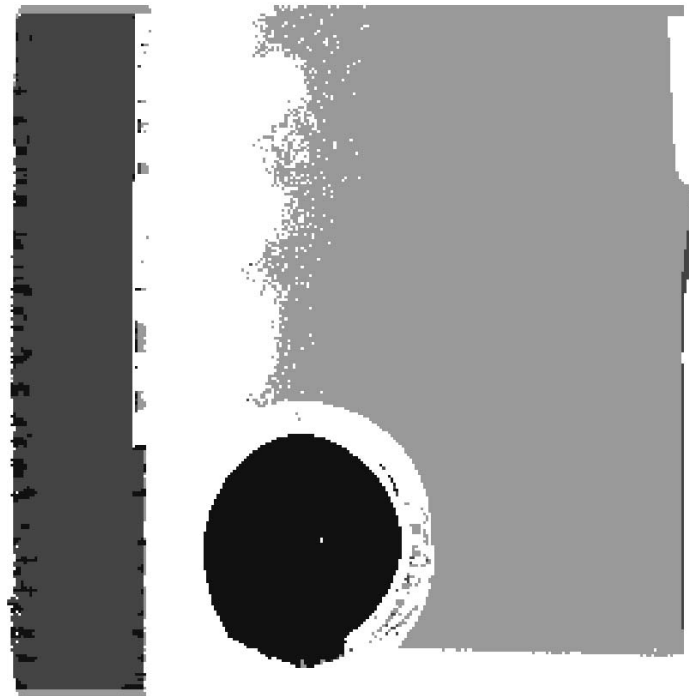
### 4.6.2 Beispiele aus der SAMPL Datenbank

Weiterführende Untersuchungen wurden an ausgewählten Datensätzen aus der SAMPL Datenbank Campbell und Flynn (1998) durchgeführt. Diese Bilddatenbank enthält einige Distanzbilder von standardisierten Szenen. Sie ist frei zugänglich im Internet verfügbar und war der Ausgangspunkt mehrerer Untersuchungen und Tests zum Thema Segmentierung und Objekterkennung und eignet sich daher, um Vergleiche zwischen den Verfahren zu ziehen.

Aus der Datenbank wurden Distanzbilder von Objekten mit gekrümmten Flächen ausgesucht. Zudem wurden Datensätze gewählt, zu denen ebenfalls CAD-Daten verfügbar sind. Die CAD-Daten wurden in Pro/Engineer eingelesen und entsprechend der in Abschnitt 4.3 beschriebenen Methode behandelt. Leider sind die meisten Modelle sehr ähnlich und im Wesentlichen aus Zylindern mit unterschiedlichen Radien aufgebaut. Daher beschränken sich die Tests auf zwei Datensätze mit unterschiedlicher Komplexität.

Der erste Datensatz mit der Bezeichnung "Adapter" besteht aus zwei unterschiedlich gekrümmten Zylindern und drei ebenen Flächen. Der Datensatz ist in Abbildung 4.19 dargestellt. Abbildung 4.19 (a) zeigt das CAD-Modell in einer schattierten Darstellung. Die übrigen Bilder zeigen jeweils das Distanzbild des Objektes in variierender Lage. Dem Distanzbild sind die Ränder der extrahierten Regionen überlagert. Mit einer Maskengröße von  $7 \times 7$  können die Krümmungen ausreichend berechnet werden, um die Zylinder zu unterscheiden. Das Verfahren detektiert alle relevanten Flächen und beschreibt die Regionen bis an ihre Ränder.

Für den zweiten Datensatz mit der Bezeichnung "Bigwe" sind die Ergebnisse in Abbildung 4.20 gegeben. Wiederum ist das schattierte CAD-Modell dargestellt und vier Ansichten des Objektes in Form von Tiefenbildern. Das Objekt enthält 5 zylindrische Flächen und die entsprechenden ebenen Flächen. Die Radien der Zylinder sind sehr viel ähnlicher und damit schwieriger zu unterscheiden. Eine Maskengröße vom  $15 \times 15$  ist notwendig, um die Krümmungen mit der erforderlichen Genauigkeit zu berechnen. Die wichtigen Flächen werden fehlerfrei detektiert. Aufgrund der Maskengröße werden jedoch die Ränder der Regionen wesentlich schlechter erfasst. Sehr kleine Flächen können nicht detektiert werden.



(a)



(b)

Abbildung 4.18: Resultat der modellbasierten Segmentierung. (a) Ergebnis nach reiner Klassifikation. (b) Klassifikation mit anschließendem Regionewachstum.

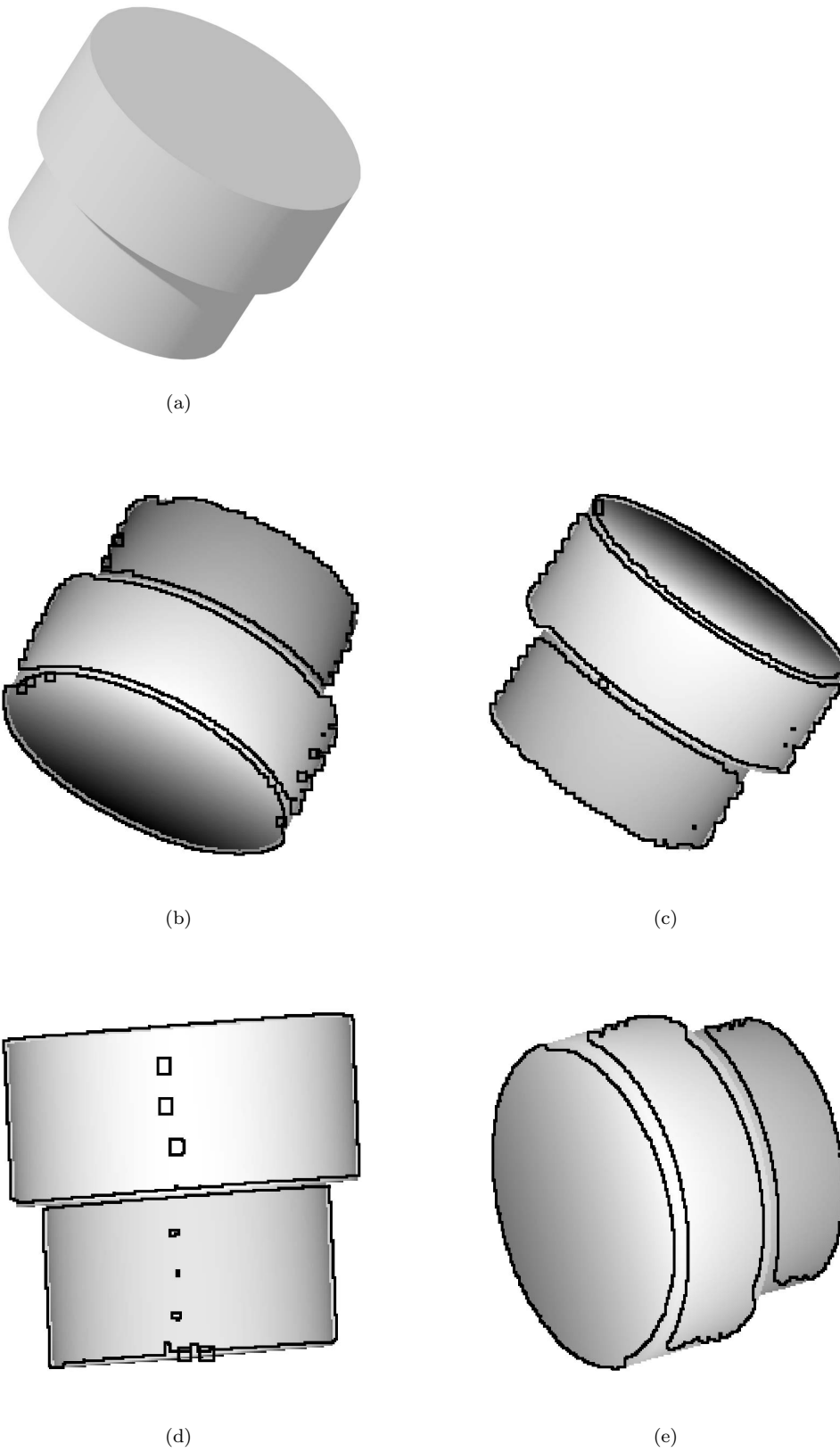
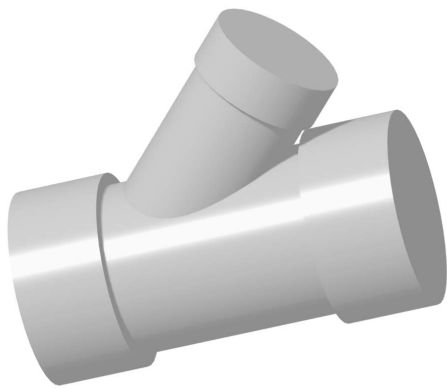
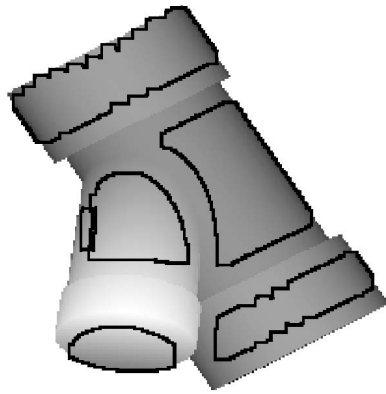


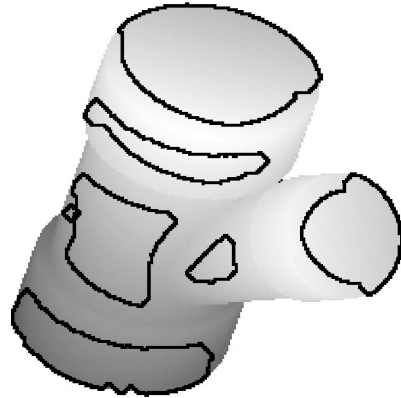
Abbildung 4.19: (a) Das CAD-Modell des Datensatzes "Adapter" in einer schattierten Darstellung. (b)-(e) Segmentierungsergebnisse aus vier Distanzbildern. Das Resultat der Klassifikation ist unabhängig von der Lage des Objektes.



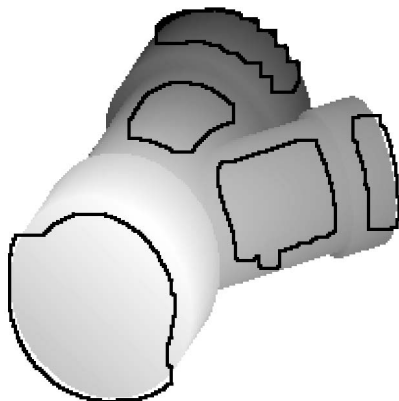
(a)



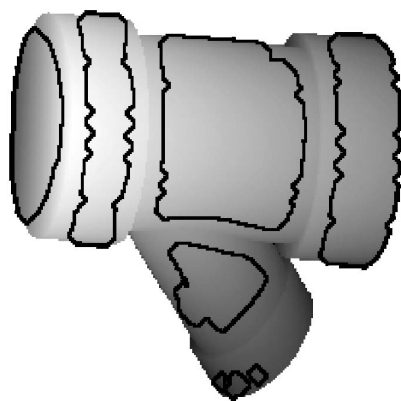
(b)



(c)



(d)



(e)

Abbildung 4.20: (a) Das CAD-Modell des Datensatzes “Bigwye” in einer schattierten Darstellung. (b)-(e) Segmentierungsergebnisse aus vier Distanzbildern. Das Resultat der Klassifikation ist unabhängig von der Lage des Objektes.

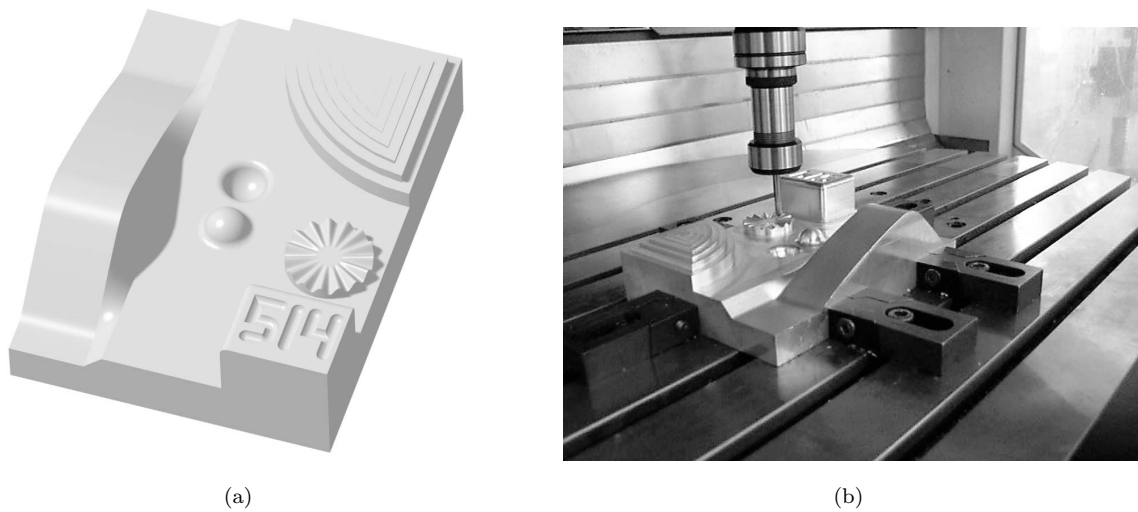


Abbildung 4.21: Selbst entwickelter Prüfkörper. (a) CAD-Modell des Körpers. (b) Fertigung des Körpers.

### 4.6.3 Prüfkörper des Sonderforschungsbereichs “Aktive Exploration”

Im Rahmen des Sonderforschungsbereiches “Aktive Exploration mittels Sensor/Aktor-Kopplung zur adaptiven Mess- und Prüftechnik” wurde ein eigener Prüfkörper entwickelt, der die Grundlage zu den Untersuchungen verschiedenster Verfahren und Sensortechnologien bildete. Da das CAD-Modell zu diesem Prüfkörper im Zusammenhang mit den hier dargelegten Arbeiten selbst entworfen wurde, konnten speziell einige interessante Merkmale eingearbeitet werden. Dies betrifft insbesondere aneinander grenzende Freiformflächen mit stetigem Übergang. Da der Körper jedoch verschiedenen Tests genügen muss, wurden verschiedene Merkmale eingearbeitet, die zum einen aus Betrachtungen praktischer Fehlerklassen in der industriellen Produktion entstammen, zum anderen aus den Erfahrungen in der Entwicklung eigener Sensorsysteme.

Zur Untersuchung der Tiefenauflösung räumlicher Sensoren wurde eine Treppenstufe integriert die Auflösungsschritte von einem Zentimeter bis zu einem hundertstel Millimeter aufweist. Zusätzlich wurde eine Rosette eingebracht, die einem zweidimensionalen Siemensstern ähnelt, und die laterale Auflösung in Verbindung mit der Tiefenauflösung eines Sensorsystems aufzeigen soll. Darüber hinaus wurden verschiedene Nuten aufgebracht, die eine Faltenbildung bei der Blechumformung nachbilden sollen. Ebenso sind kugelförmige Elemente vorhanden die eine Beulen- oder Dellenbildung nachahmen. Das schattierte dargestellte CAD-Modell ist in Abbildung 4.21 (a) gegeben.

Das konstruierte CAD-Modell dient als Vorlage zur Produktion des Prüfkörpers auf einer computergesteuerten Fräsmaschine mit hoher Fertigungsgenauigkeit. Die Fertigung des Prüfkörpers zeigt Abbildung 4.21 (b). Für die anschließenden Untersuchungen dient das CAD-Modell als Referenz.

Der gefertigte Prüfkörper wurde mit einem Streifenprojektor aufgenommen und die räumlichen Koordinaten mit der Software der Firma ABW berechnet, die auf der Basis einer polynomiellen Kalibriermethode arbeitet. Das Tiefenbild ist in Abbildung 4.22 (a) gegeben. Die daraus mit den gegebenen CAD-Daten abgeleitete Segmentierung ist in 4.22 (b) dargestellt. Ein Ausschnitt von drei aneinander anschließenden Freiformflächen mit stetigem Übergang ist in Abbildung 4.23 dargestellt. Eine derartige Segmentierung, die der ebenfalls dargestellten Aufteilung der Flächen im CAD-Modell 1 : 1 entspricht ist mit anderen Segmentierungsverfahren nicht zu erreichen.



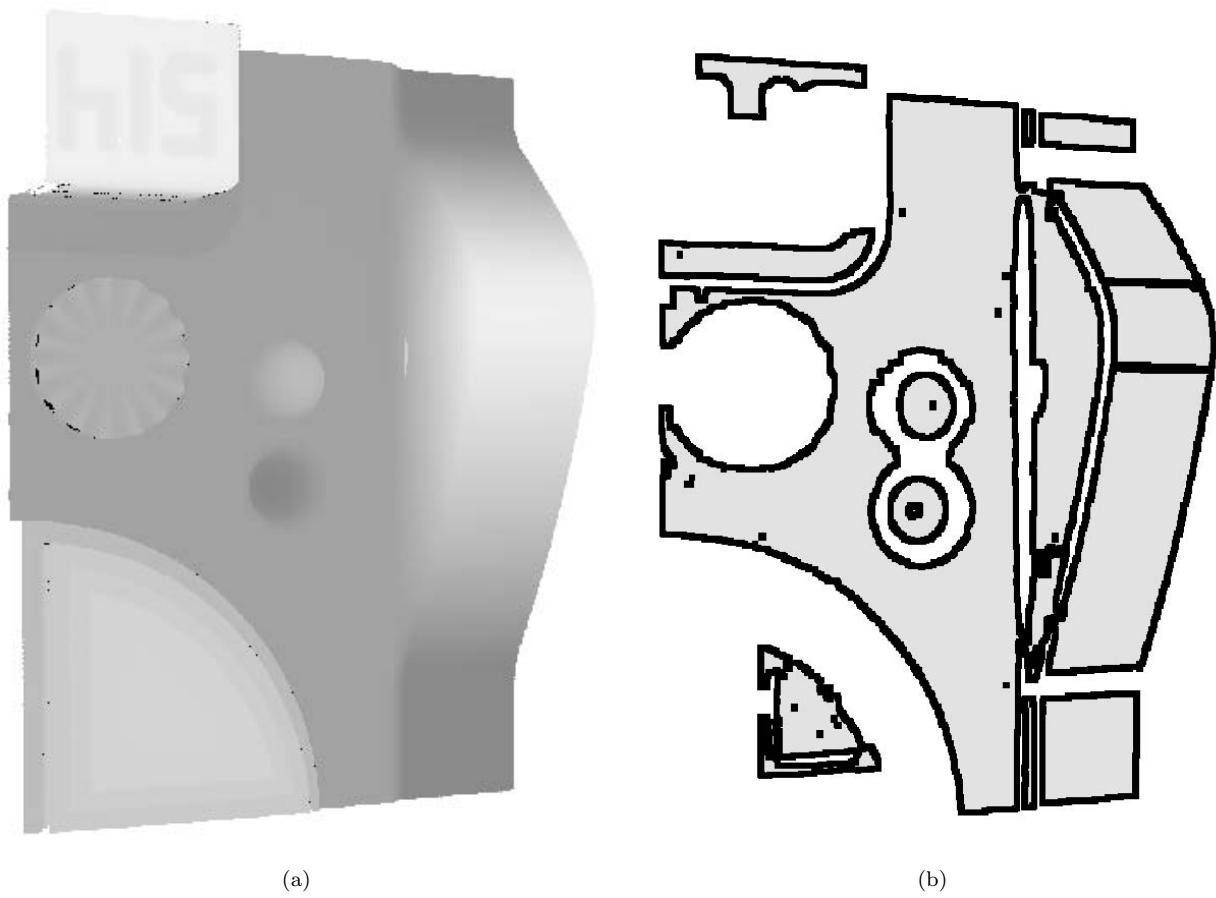


Abbildung 4.22: Segmentierung des Distanzbildes des Prüfkörpers. (a) Tiefenbild des Körpers. (b) Extrahierte Regionen.

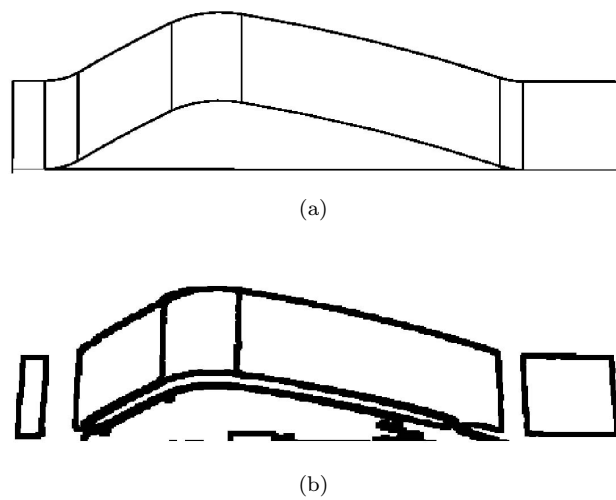


Abbildung 4.23: Ausschnitt aus der Segmentierung.

## 5 Zuordnungsanalyse und Transformationsbestimmung

### 5.1 Einleitung

Die im Rahmen dieser Arbeit dargestellten und untersuchten Verfahren verfolgen gemeinsam einen modellbasierten Ansatz. Entsprechend wird auch für die Objekterkennung eine modellbasierte Vorgehensweise gewählt. Die modellbasierte Objekterkennung beruht im Allgemeinen auf dem Vergleich eines geometrischen Modells mit Merkmalen, die aus Sensordaten extrahiert wurden. Tatsächlich wird eine Teilmenge der Merkmale, aus welchen sich das Modell zusammensetzt, mit wiederum einer Teilmenge der Merkmale der Szene derart zugeordnet, dass die so gebildeten Korrespondenzen einer gemeinsamen gültigen Transformation genügen. Die sich ergebende Transformation stellt im Falle eines 3D-Objekterkennungssystems eine räumliche Transformation dar, welche der Lage des Objektes im Koordinatensystem des Sensors entspricht.

Die folgenden Abschnitte stellen zunächst das grundsätzliche Problem der Korrespondenzanalyse dar. Anschließend wird die generelle Formulierung konkretisiert und auf die im vorangegangenen Kapitel dargestellte Merkmalsextraktion bezogen. Die Berechnung einer räumlichen Ähnlichkeitstransformation aus den gefundenen Korrespondenzen stellt den letzten Schritt des hier dargestellten Verfahrens zur Objekterkennung dar. Die Beispiele und Ergebnisse der Segmentierung werden wieder aufgegriffen und dazu verwendet, die entwickelten Verfahren zu demonstrieren.

### 5.2 Modellbasierte Objekterkennung

Es sei eine Menge  $S = \{f_1, f_2, \dots, f_n\}$  aus Merkmalen der Szene gegeben, sowie eine zweite Menge  $M = \{F_1, F_2, \dots, F_m\}$ , die das Modell beschreibt. In unserem Kontext entspricht jede Fläche einem Merkmal und umgekehrt. Jedes Merkmal trägt die berechneten Attribute. Das Ziel der Objekterkennung besteht darin, eine globale Zuordnung der Merkmale zu finden, das heißt eine Menge von Paaren  $\{(f_1, F_{m_1}), (f_2, F_{m_2}), \dots, (f_n, F_{m_n})\}$ , sodass jedes Merkmal aus dem Modell mit dem korrespondierenden Merkmal aus der Szene verbunden wird.

Theoretisch gibt es  $O(n^m)$  verschiedene Möglichkeiten solche Paarungen zu bilden. Natürlich stellt nur ein Bruchteil davon eine gültige Lösung dar. Es wurden verschiedene Verfahren entwickelt, um diese kombinatorische Vielfalt zu beherrschen. Dazu zählen etwa die Hough-Transformation, maximale Cliques, Graph-Matching, diskrete Relaxation und die Baumsuche mit Bedingungen.

Um die Korrektheit einer globalen Zuordnung zu überprüfen, kann man aus allen zugeordneten Merkmalen eine räumliche Transformation bestimmen und den geometrischen Abstand der korrespondierenden Merkmale berechnen. Auch um die Korrektheit einer lokalen Zuordnung zu überprüfen, kann man geometrische Kriterien heranziehen. Dabei kann man zwischen unären und binären Bedingungen unterscheiden.

Unäre Bedingungen müssen erfüllt sein zwischen einem Merkmal aus der Szene und dem korrespondierenden Merkmal aus dem Modell. Beispiele für unäre Bedingungen sind die Länge einer Linie, der Radius eines Kreises, Krümmungen, Umfang, die Größe und das Seitenverhältnis des minimal umschließenden Rechtecks, Flächeninhalt und Volumen.

Binäre Bedingungen überprüfen, ob eine Paarung aus Szene- und Modellmerkmal konsistent ist mit allen anderen Paarungen, die aufgestellt wurden. So müssen zum Beispiel der Winkel und die Distanz zwischen zwei Flächen in der Szene mit den entsprechenden Größen im Modell übereinstimmen.

Andere Untersuchungen (siehe zum Beispiel Messmer und Bunke, 1998) verwenden ein Graph-Matching Verfahren zur Objekterkennung, ein Verfahren, das auch im Zusammenhang zum Regionen Adjazenz Graphen (Fan, 1989) steht. Das Ergebnis einer regionenorientierten Segmentierung kann in einem Graphen repräsentiert werden, der für jede Region genau einen Knoten besitzt. Die Kanten zwischen den Knoten drücken die Nachbarschaft der Regionen aus. Sowohl die Szene als auch das Modell kann in dieser Form repräsentiert werden. Im Zuordnungsverfahren sucht man nach einer Projektion zwischen den beiden Graphen, welche die Topologie der Graphen erhält. Das heißt, zwei Knoten, die in der Szene (nicht) durch eine Kante verbunden sind, müssen auch im Modell (nicht) verbunden sein. Es wird also nach einem Isomorphismus gesucht.

Da in der Szene nicht unbedingt alle Merkmale des Objektes abgebildet sind, wird tatsächlich ein Isomorphismus zwischen Subgraphen gesucht, ein NP-vollständiges Problem. Die Kanten und Knoten der Graphen können mit Attributen versehen werden, um geometrische Eigenschaften abzulegen. Da nur benachbarte Regionen über Kanten verbunden sind, können nur die geometrischen Beziehungen dieser Nachbarschaften abgelegt werden. Beziehungen zwischen nicht benachbarten Regionen können in diesem Modell nicht ausgedrückt werden. Erfahrungen zeigen, dass beim Graph-Matching mit den geometrischen Eigenschaften kaum eine effektive Einschränkung des Suchraumes möglich ist. Zudem konnten in eigenen Experimenten nur extrem kleine Graphen zugeordnet werden, da der Rechenaufwand sehr schnell zu groß wurde. Außerdem ist festzustellen, dass die Nachbarschaft bei der räumlichen Segmentierung nicht immer ein stabiles Merkmal ist.

### 5.2.1 Zuordnungsanalyse durch Baumsuche

Aufgrund dieser Beobachtungen wird in dieser Arbeit das Verfahren der Baumsuche Grimson (1990) angewandt. Fortführend mit der oben eingeführten Notation, besitzt jedes Merkmal  $f_i$  der Szene eine Menge  $\Omega(f_i) = \{F_{i_1}, F_{i_2}, \dots, F_{i_j}\}$  möglicher Zuordnungen aus dem Modell. Elemente dieser Menge sind alle Merkmale des Modells, welche die unären Bedingungen zu  $f_i$  erfüllen. Die Richtung der Zuordnung, das heißt, ob einem Modellmerkmal ein Szenemerkmal zugeordnet wird oder umgekehrt, kann beliebig gewählt werden. Wir wählen hier in diesem Zusammenhang die Szene als Ausgangspunkt.

Man kann die Merkmale der Szene nach bestimmten Kriterien ordnen, zum Beispiel absteigend nach der Größe der Fläche. Man beginnt mit dem ersten Merkmal und wählt eine Zuordnung aus der Zuordnungsmenge. Dann fährt man fort mit dem zweiten Merkmal und erweitert so den Pfad der bisherigen Zuordnungen zu  $(f_1, F_{m_1}), (f_2, F_{m_2})$ . Diese Suche setzt sich in rekursiver Weise fort. Das Verfahren ist allgemein bekannt als Tiefensuche.

Da die Zuordnungsmengen bereits durch die unären Bedingungen eingeschränkt sind, ist der Suchraum deutlich kleiner als  $O(n^m)$ . Er kann weiter dadurch verkleinert werden, dass nach jedem Schritt die binären Bedingungen der aktuellen Zuordnung mit allen bisher gefundenen Zuordnungen überprüft werden.

Ist eine Bedingung nicht erfüllt kann der ganze folgende Teilbaum abgeschnitten werden. Da das reale Objekt vom idealen Modell abweichen kann und die Segmentierung Fehler enthalten kann, muss die Möglichkeit in Betracht gezogen werden, dass ein Merkmal keine Entsprechung besitzt. Daher wird das Symbol  $\lambda$  für die leere Zuweisung eingeführt. Da in der Szene durch Fehler in der Segmentierung typischerweise eine Vielzahl kleiner Splitterregionen (*sliver regions*) entsteht, kann man eine Reihe von Regionen von der Zuordnung ausschließen. Es hat sich bewährt die Merkmale nach der Größe zu ordnen, und die Anzahl der Merkmale aus der Szene zu begrenzen. Dies entspricht einer Begrenzung der Tiefe des Suchbaums. Zur abschließenden Bestimmung der Transformation sind mindestens drei Korrespondenzen erforderlich.

## 5.3 Zuordnungsmengen aus der Segmentierung

Im Rahmen des hier dargestellten Verfahrens wurde bereits im Segmentierungsschritt die Modellinformation genutzt, den Verfahrensablauf zu stützen. Konkret wurden die Krümmungseigenschaften einer bestimmten Teilfläche der Randbeschreibung des Modells genutzt, um eine Klassifikation der Sensordaten durchzuführen. Also wurde bereits während der Segmentierung eine Zuordnung von Teilflächen des Modells zu Regionen des Distanzbildes erzeugt. Auf eben diese Zuordnung kann nun im Rahmen der Zuordnungsanalyse zurückgegriffen werden.

Jede Teilfläche des Modells, welche eine Region "erzeugt" hat, wird in die Zuordnungsmenge dieser Region aufgenommen. Waren mehrere Flächen des Modells bezüglich ihrer Krümmungseigenschaften identisch oder hinreichen ähnlich, werden alle diese Flächen in die Zuordnungsmenge aufgenommen. Durch das Aufstellen der Zuordnungsmengen sind bereits die unären Bedingungen in den Prozess integriert. Weist das Modell eine ausreichende Vielfalt an gekrümmten Flächen auf, so sind die Zuordnungsmengen entsprechen klein. Somit ist auch der Suchraum deutlich eingeschränkt und die Lösung kann effizient berechnet werden.

Dieser Zusammenhang wird nun an einigen Beispielen verdeutlicht. Grundlage bilden die im vorangegangenen Kapitel erzielten Ergebnisse des Segmentierungsverfahrens. Ein einfaches Beispiel ergibt sich für den Datensatz "Adapter" aus der SAMPL Datenbank. Das Modell besteht aus zwei Zylindern mit unterschiedlichen Radien  $F_1$  und  $F_2$  und drei ebenen Flächen  $F_3$ ,  $F_4$  und  $F_5$ . Anders als in der Implementierung des Verfahrens werden hier zur besseren Anschaulichkeit nicht die internen Identifikationsnummern der Teilflächen aus dem CAD-System

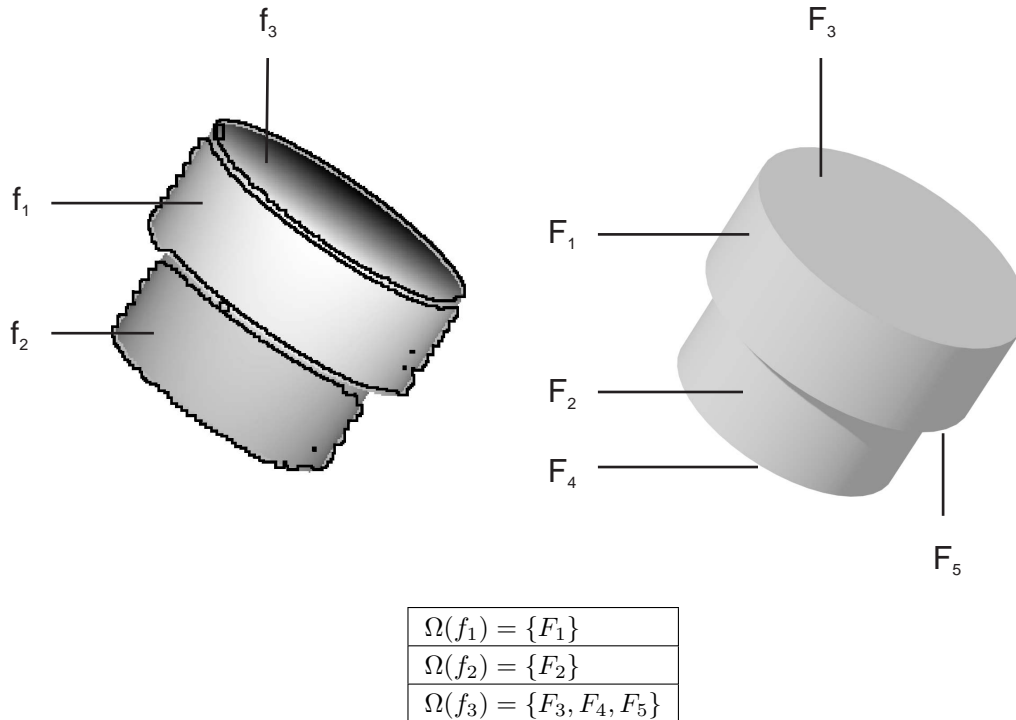


Abbildung 5.1: Ein einfaches Beispiel für die Ableitung der Zuordnungsmengen aus der Segmentierung. Die Segmentierung des Distanzdatensatzes “Adapter” erzeugt drei Regionen  $f_1$  bis  $f_3$ . Das CAD-Modell besteht aus fünf Flächen  $F_1$  bis  $F_5$ . Die Korrespondenzen der Merkmale in der Szene und der Merkmale im Modell ergeben sich aus den Krümmungseigenschaften der Flächen. Die Zuordnungsmengen jeder der drei Regionen sind in der Tabelle aufgeführt.

verwendet um die Flächen eindeutig zu kennzeichnen, sondern Symbole der Form  $F_i$  beziehungsweise  $f_i$ . Die Segmentierung der zweiten Ansicht erzeugte drei Regionen  $f_1$  bis  $f_3$ . Dabei wurde die erste Region  $f_1$  erzeugt durch die Fläche  $F_1$ . Dementsprechend beinhaltet die Zuordnungsmenge von  $f_1$  diese Fläche. Die einzelnen Flächen beziehungsweise Regionen und die resultierenden Zuordnungsmengen sind in Abbildung 5.1 dargestellt. Da zwei der Zuordnungsmengen nur ein Element enthalten, ergibt sich ein entsprechend einfacher Suchbaum. Für die Region  $f_1$  kann es nur eine mögliche Zuordnung geben, die Fläche  $F_1$ , für  $f_2$  existiert nur eine Zuordnung zu  $F_2$ . Erst ab der dritten Stufe verzweigt der Baum in drei Teilbäume. Es gibt also für diesen Suchbaum drei Pfade von der Wurzel zu einem Blatt des Baumes. Allein auf Basis der bisher eingeflossenen Information kann noch nicht entschieden werden, welcher der drei Pfade einer (global) korrekten Zuordnung entspricht. Diese Frage wird durch die globale Transformationsbestimmung beantwortet. Der Suchbaum ist in Abbildung 5.2 gegeben. Zur besseren Übersichtlichkeit wurden die Leerzuweisungen zum Symbol  $\lambda$  nicht aufgeführt. Wie oben erwähnt, muss zu jeder Zuordnungsmenge noch das Symbol  $\lambda$  hinzugefügt werden. In der Implementierung des Verfahrens geschieht dies implizit. Ebenso wurde die Unterteilung der Zylinder in zwei Teilflächen (siehe 4.3) zur besseren Übersichtlichkeit nicht aufgeführt.

Entsprechende Darstellungen lassen sich auch für den Datensatz “Bigwye” herleiten. Den fünf Regionen der Segmentierung der vierten Ansicht stehen acht Flächen des CAD-Modells gegenüber. Aus der Segmentierung ergeben sich eindeutige Zuordnungen für die drei Regionen  $f_1$ ,  $f_4$  und  $f_5$ , die alle unterschiedliche Zylinder beschreiben. Zwei der Regionen,  $f_2$  und  $f_3$ , beschreiben identische Zylinder, die im Modell den Flächen  $F_2$  und  $F_3$  entsprechen. Die ebene Region  $f_6$  hat drei mögliche Zuordnungspartner  $F_6$ ,  $F_7$  und  $F_8$ . Eine Übersicht der Flächen und Regionen sowie die Zuordnungsmengen sind in Abbildung 5.3 gegeben. Der Suchbaum ist in Abbildung 5.4 dargestellt. Es ergeben sich sechs verschiedene Pfade von der Wurzel zu den Blättern des Baums.

Dass das Verfahren der Baumsuche auch für ungünstigere Konstellationen mit zahlreichen gleichartigen Flächen hinreichend funktioniert, soll an dem folgenden Beispiel gezeigt werden, das aus einer früheren gemeinsamen Arbeit stammt (Böhm et al., 1999). Aus einem zusammengesetzten Distanzdatensatz wurden mittels eines Regionenwachstumsverfahrens für ebene Flächen 598 Regionen extrahiert. Das zu dem Datensatz korrespondierende CAD-Modell mit der Bezeichnung “Halter” enthält 42 Teilflächen. Eine Darstellung des Objektes ist in Abbildung 5.5(a) gegeben, der segmentierte Distanzdatensatz und das segmentierte Modell jeweils in Abbildung 5.5(b) beziehungsweise (c). Da die Regionen ausschließlich ebene Flächen repräsentieren, können keine Einschränkungen für die Zuordnungsmengen aus den Krümmungseigenschaften abgeleitet werden. Daher ist

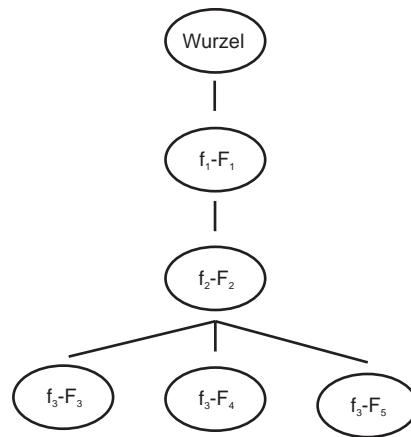
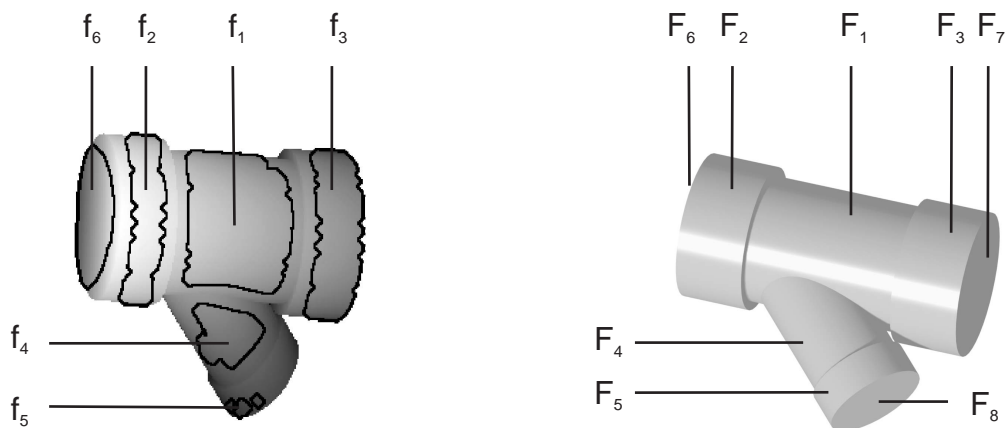


Abbildung 5.2: Der Suchbaum, der sich aus den in Abbildung 5.1 aufgeführten Zuordnungsmengen ergibt.



$\Omega(f_1) = \{F_1\}$
$\Omega(f_2) = \{F_2, F_3\}$
$\Omega(f_3) = \{F_2, F_3\}$
$\Omega(f_4) = \{F_4\}$
$\Omega(f_5) = \{F_5\}$
$\Omega(f_6) = \{F_6, F_7, F_8\}$

Abbildung 5.3: Ein weiteres Beispiel für die Ableitung der Zuordnungsmengen aus der Segmentierung. Die Segmentierung des Distanzdatensatzes "Bigwye" erzeugt sechs Regionen  $f_1$  bis  $f_6$ . Das CAD-Modell besteht aus acht Flächen  $F_1$  bis  $F_8$ . Die Zuordnungsmengen jeder der sechs Regionen sind in der Tabelle aufgeführt.

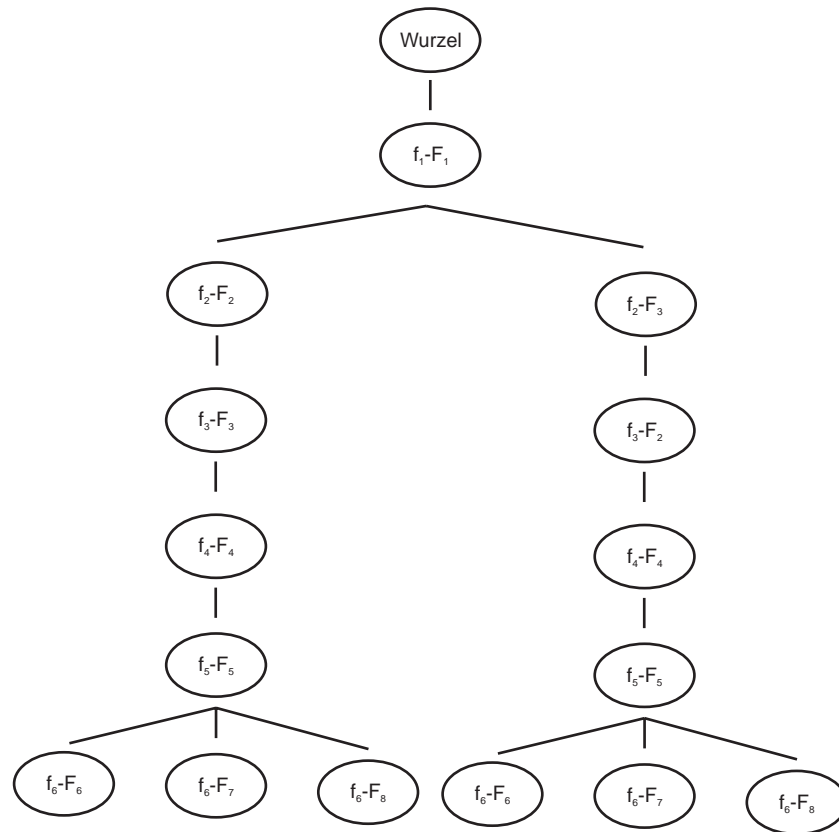


Abbildung 5.4: Der Suchbaum, der sich aus den in Abbildung 5.3 aufgeführten Zuordnungsmengen für eine Ansicht des Datensatzes "Bigwye" ergibt.

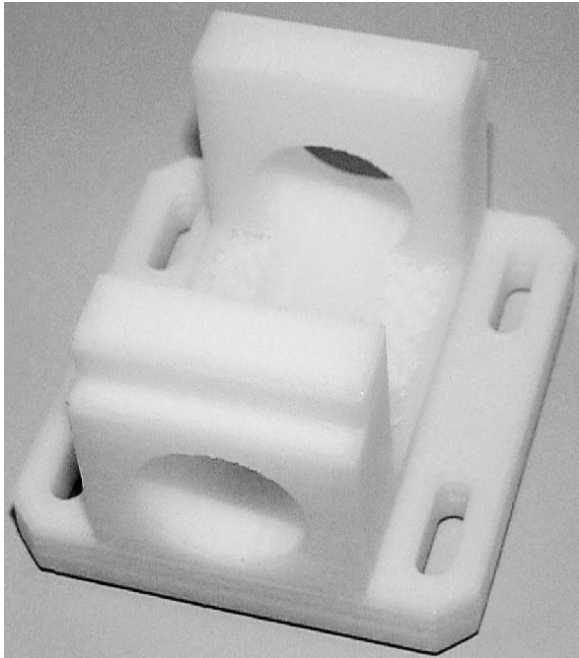
zunächst jede Fläche des CAD-Modells ein potenzieller Kandidat für die Zuordnung. Aus der Kombinatorik ergäbe sich so ein Suchraum von enormer Größe, der nicht mehr vollständig durchsucht werden kann.

Um den Suchraum einzuschränken, wurden zusätzliche Merkmale für jede extrahierte Region berechnet und ebenso für jede Modellfläche. Die zusätzlichen Merkmale beruhen auf einfachen geometrischen Eigenschaften der Randkurven der Flächen. Bei einem Polyeder sind diese Randkurven einfach zu beschreiben, da es sich bei den einzelnen Flächen um Ebenen, und bei den Randkurven somit ausschließlich um ebene Kurven handelt. Daher sind auch einfache geometrische Parameter ausreichend, um die Kurven hinlänglich unterscheiden zu können. Konkret wurden der maximale Abstand innerhalb des Randpolygons, der Umfang des Polygons und die Breite und Höhe des umschließenden kanonischen Rechtecks gewählt.

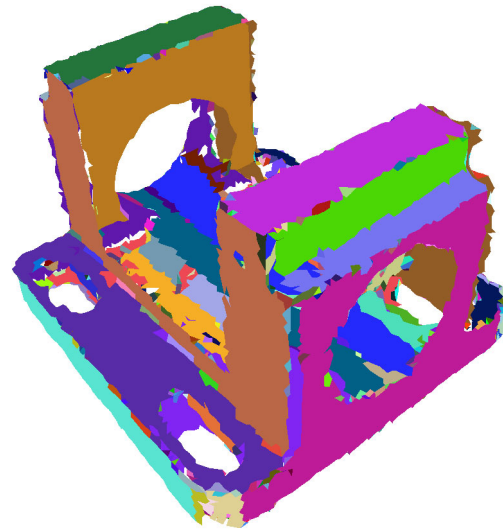
Die gewählten Merkmale ergaben in Verbindung mit der Begrenzung der Tiefe eine hinreichende Reduktion des Suchbaums. Die Tiefe des Suchbaums kann begrenzt werden, da zur Berechnung der Transformation lediglich mindestens drei korrespondierende Flächen bekannt sein müssen (siehe Abschnitt 5.4). Ein Ausschnitt des reduzierten Suchbaums ist in Abbildung 5.6 dargestellt. Die aus den Korrespondenzen hergeleitete Transformation ermöglicht die Überlagerung des Distanzdatensatzes mit dem CAD-Modell, wie in Abbildung 5.5(d) dargestellt. Es zeigt sich, dass die Objekterkennung korrekt durchgeführt wurde.

Bei der Wahl der zusätzlichen Merkmale muss natürlich beachtet werden, dass diese rotations- und translationsinvariant sind. Die Möglichkeit die soeben beschriebenen Merkmale als unäre Bedingungen hinzuzuziehen, hängt von der Güte der Segmentierung bezüglich des Flächenrands ab. Insbesondere bei einer Ebenensegmentierung eines Polyeders bilden diese Merkmale eine wichtige Stütze, die einzelnen Teilflächen zu unterscheiden, da hier ja keine Krümmungsunterschiede vorliegen. Bei einem allgemeinen Segmentierungsverfahren jedoch, das auch in der Lage sein soll Objekte mit gekrümmten Flächen und stetigem Flächenübergang zu behandeln, sind die Randkurven im Allgemeinen nur unzuverlässig extrahierbar. Hinzukommt, dass es sich um räumliche Kurven handelt, und somit die Definition geeigneter Merkmale schwieriger werden kann.

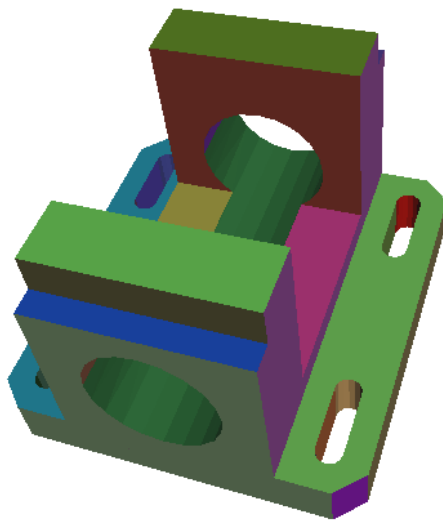
Im nächsten Abschnitt soll nun dargestellt werden, wie die Korrektheit der im Suchbaum möglichen Pfade mittels einer globalen geometrischen Transformation entschieden werden kann.



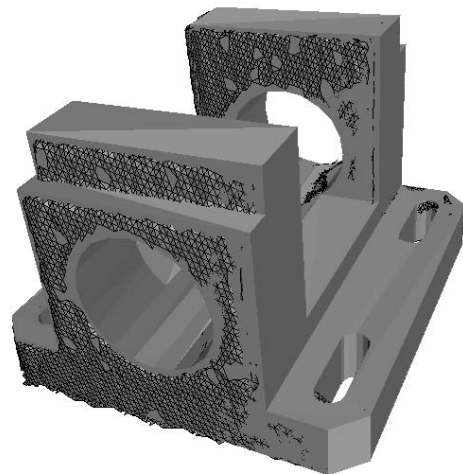
(a)



(b)



(c)



(d)

Abbildung 5.5: Untersuchung zur Zuordnungsanalyse ebener Flächensegmente. (a) Ein Intensitätsbild des Objektes "Halter". (b) Die Segmentierung eines integrierten Distanzdatensatzes in ebene Flächenstücke. (c) Die Segmentierung des entsprechenden CAD-Modells. (d) Die Überlagerung des CAD-Modells mit den triangulierten Distanzdaten.

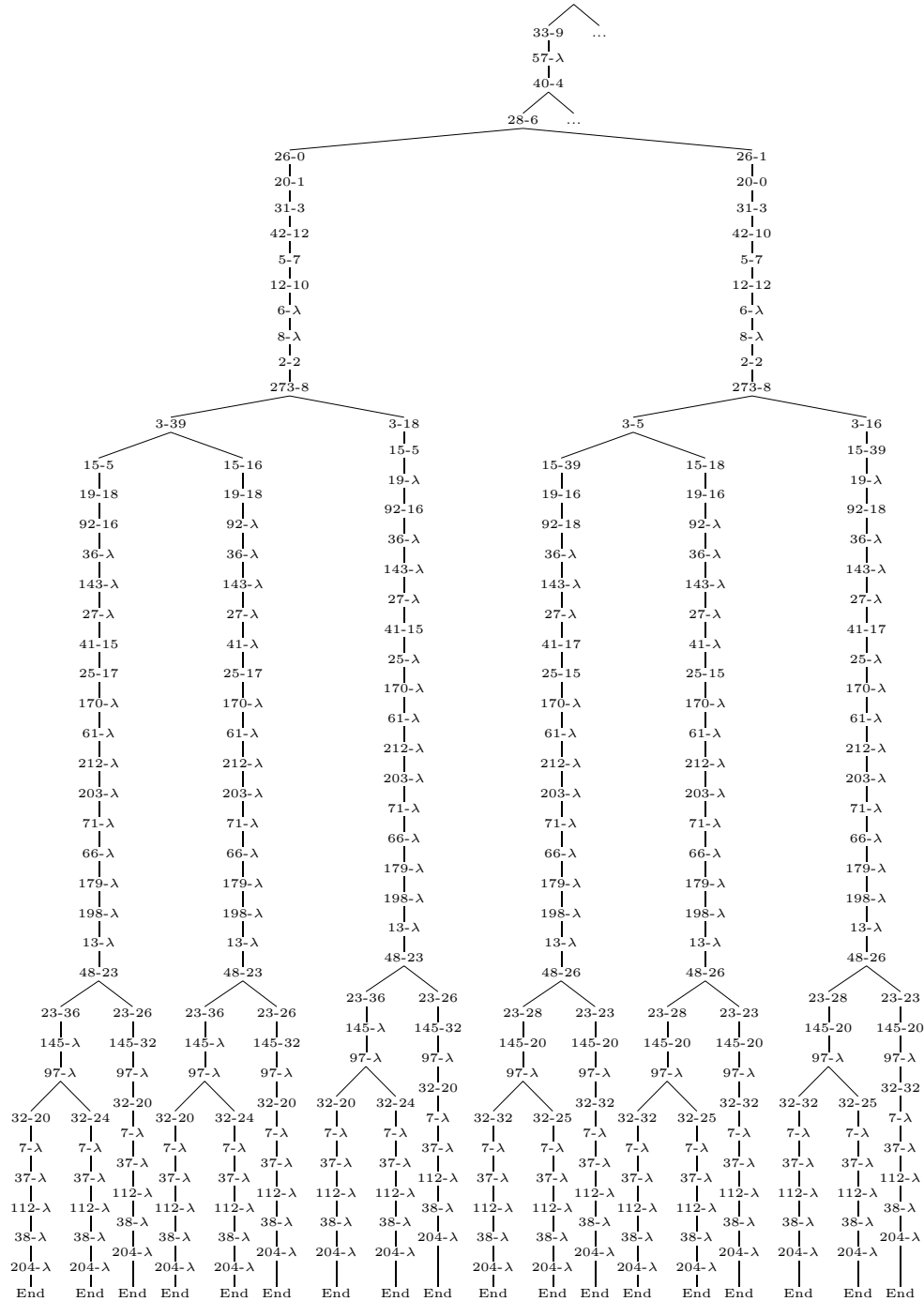


Abbildung 5.6: Ausschnitt aus einem Suchbaum, der durch zusätzliche Merkmale und durch eine Begrenzung der Tiefe reduziert wurde.



## 5.4 Transformationsbestimmung

Um die Lage des Objektes im Distanzdatensatz zu bestimmen, muss die räumliche Transformation des Modellkoordinatensystems in das Sensorkoordinatensystem bestimmt werden. Diese Transformation kann mithilfe der aus der Zuordnungsanalyse bekannten Korrespondenzen der einzelnen Flächen jeweils aus dem CAD-Modell und dem segmentierten Distanzdatensatz berechnet werden.

Die zunächst vorliegenden Korrespondenzen beziehen sich dabei zunächst ausschließlich auf Flächenpaare. Es ist wichtig zu beachten, dass die verschiedenen Flächentypen, ohne Betrachtung ihrer Randkurven, einen unterschiedlichen Beitrag zur Transformationsbestimmung liefern. So kann zum Beispiel aus der Korrespondenz einer ebenen Fläche lediglich auf zwei Rotationsparameter und einen Translationsparameter geschlossen werden, aus der Korrespondenz einer Kugelfläche können drei Translationsparameter abgeleitet werden. Die vollständige Transformation lässt sich nur aus Kombinationen von Korrespondenzen ableiten. Die große Vielfalt an möglichen Flächentypen erschwert hierbei die Vorhersagbarkeit der Berechnungen. Zudem ist bei komplexen Flächentypen die direkte Ableitung der Transformation nicht möglich.

Ein wesentlich einfacheres Berechnungsschema ergibt sich aus der Korrespondenz von Punkten. Schon aus der Korrespondenz von drei nicht kollinearen Punktpaaren lässt sich die notwendige räumliche Transformation berechnen. Diese Transformation sowie das Schema ihrer Berechnung werden im Folgenden dargestellt.

### 5.4.1 Transformation aus Punktkorrespondenzen

Im allgemeinen Fall wird die räumliche Ähnlichkeitstransformation zweier korrespondierender Punktmengen  $\{\mathbf{p}_i\}$  und  $\{\mathbf{P}_i\}$  mit  $i = 1, \dots, N$  gesucht, die dargestellt wird durch  $\mathbf{P}_i = \lambda \mathbf{R} \mathbf{p}_i + \mathbf{t}$ . Dabei repräsentiert  $\mathbf{R}$  eine  $3 \times 3$  Rotationsmatrix und  $\mathbf{t}$  einen räumlichen Translationsvektor.

In der praktischen Anwendung besteht keine exakte Übereinstimmung zwischen den beiden Punktmengen und somit wird die oben-stehende Gleichung um den Vektor  $\mathbf{v}_i$  ergänzt, der das Fehlerrauschen repräsentiert. Somit erhält man  $\mathbf{P} = \lambda \mathbf{R} \mathbf{p}_i + \mathbf{t} + \mathbf{v}_i$ .

Im Kontext dieser Arbeit wird jedoch ein korrekt kalibrierter Sensor vorausgesetzt, der metrische Daten liefert. Somit ist die Bestimmung eines Skalierungsparameters nicht erforderlich. Es ergibt sich somit die einfachere Starrkörpertransformation (*rigid body transform*) in der Form

$$\mathbf{P}_i = \mathbf{R} \mathbf{p}_i + \mathbf{t} + \mathbf{v}_i \quad (5.1)$$

Zur Bestimmung der optimalen Transformation  $[\hat{\mathbf{R}}, \hat{\mathbf{t}}]$ , welche das Fehlerkriterium

$$\Sigma^2 = \sum_{i=1}^N \|\mathbf{P}_i - \mathbf{R} \mathbf{p}_i - \mathbf{t}\|^2 \quad (5.2)$$

minimiert, sind in der Literatur zahlreiche Ansätze bekannt. Man unterscheidet grob in geschlossene und iterierende Verfahren, wobei hier die geschlossenen Verfahren von besonderem Interesse sind. Die geschlossenen Verfahren zeichnen sich dadurch aus, dass zu ihrer Lösung keinerlei Näherungswerte benötigt werden. Zudem lassen sie sich meist deutlich schneller berechnen, was zu Vorteilen in der Laufzeit führt.

Die einzelnen Ansätze unterscheiden sich meist in der Repräsentation der Rotationsmatrix. Einer der ersten Vorschläge stammt von Sanso (1973). Das besonders für die Realisierungen des *Iterative Closest Point* Algorithmus und verwandter Algorithmen (Besl und McKay, 1992; Feldmar et al., 1997; Dorai et al., 1994) am meisten zitierte Verfahren stammt von Horn (1987). Weitere Ansätze wurden von Arun et al. (1987) und Walker et al. (1991) vorgeschlagen. Ein Vergleich von vier populären Verfahren zur geschlossenen Bestimmung der optimalen Transformation von Lorusso et al. (1995) zeigte dabei auf, dass die Unterschiede bezüglich der Genauigkeit der berechneten Transformationsparameter zu vernachlässigen sind. Das Verfahren nach Horn erwies sich in dieser Untersuchung als stabil und effizient bezüglich der Laufzeit und soll nun kurz dargestellt werden.

Die Rotationsmatrix  $\mathbf{R}$  wird nach den vier Elementen des normierten Quaternions  $\mathbf{q} = (q_0, q_1, q_2, q_3)$  mit  $\|\mathbf{q}\| = 1$  parametrisiert<sup>1</sup>:

<sup>1</sup>Alternativ können die Elemente auch mit  $(q_0, q_x, q_y, q_z)$  bezeichnet werden, um der Tatsache Ausdruck zu verleihen, dass es sich um eine Rotation mit dem Winkel  $q_0$  um die Achse  $(q_x, q_y, q_z)$  handelt.

$$\mathbf{R} = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{bmatrix} \quad (5.3)$$

In einem ersten Schritt werden nun die Schwerpunkte der beiden Punktmengen  $\{\mathbf{p}_i\}$  und  $\{\mathbf{P}_i\}$  aufgestellt

$$\bar{\mathbf{p}} = \frac{1}{N} \sum_{i=1}^N \mathbf{p}_i \quad \bar{\mathbf{P}} = \frac{1}{N} \sum_{i=1}^N \mathbf{P}_i \quad (5.4)$$

und anschließend die reduzierten Koordinaten

$$\mathbf{p}'_i = \mathbf{p}_i - \bar{\mathbf{p}} \quad \mathbf{P}'_i = \mathbf{P}_i - \bar{\mathbf{P}} \quad (5.5)$$

berechnet. Mithilfe dieser Koordinaten wird die Matrix

$$\mathbf{M} = \sum_{i=1}^N \mathbf{p}'_i \mathbf{P}'_i{}^T = \begin{bmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{bmatrix} \quad (5.6)$$

aufgestellt, wobei  $S_{xx} = \sum_{i=1}^N p'_{ix} p'_{ix}$ ,  $S_{xy} = \sum_{i=1}^N p'_{ix} p'_{iy}$  und so weiter.

Mit den Elementen  $S_{xx}, \dots, S_{zz}$  kann nun die symmetrische  $4 \times 4$  Matrix

$$\mathbf{N} = \begin{bmatrix} (S_{xx} + S_{yy} + S_{zz}) & S_{yz} - S_{zy} & S_{zx} - S_{xz} & S_{xy} - S_{yx} \\ S_{yz} - S_{zy} & (S_{xx} - S_{yy} - S_{zz}) & S_{xy} + S_{yx} & S_{zx} + S_{xz} \\ S_{zx} - S_{xz} & S_{xy} + S_{yx} & (-S_{xx} + S_{yy} - S_{zz}) & S_{yz} + S_{zy} \\ S_{xy} - S_{yx} & S_{zx} + S_{xz} & S_{yz} + S_{zy} & (-S_{xx} - S_{yy} + S_{zz}) \end{bmatrix}$$

aufgestellt werden.

Der zum größten Eigenwert korrespondierende Eigenvektor  $\mathbf{e}$  mit  $\|\mathbf{e}\| = 1$  dieser Matrix  $\mathbf{N}$  ist äquivalent zum gesuchten Quaternion  $\mathbf{q}$ , welches die gesuchte Rotation beschreibt.

Nachdem die Rotation  $\mathbf{R}$  bekannt ist, kann der Translationsvektor  $\mathbf{t}$  einfach nach der Vorschrift

$$\mathbf{t} = \bar{\mathbf{p}} - \mathbf{R}\bar{\mathbf{P}} \quad (5.7)$$

berechnet werden.

Nachdem Rotation und Translation berechnet wurden, lässt sich die Punktmenge  $\{\mathbf{P}_i\}$  nach der Rechenvorschrift 5.1 umbilden.

Aus dem Fehlerkriterium 5.2 lässt sich der Fehler der ermittelten Transformation aus den Abweichungen der korrespondierenden Punkte nach ihrer Transformation als Wurzel der mittleren Quadratsumme (*root mean square*) bestimmen mit

$$rms^2 = \frac{1}{N} \sum_{i=1}^N \|\mathbf{P}_i - \mathbf{R}\mathbf{p}_i - \mathbf{t}\|^2 \quad (5.8)$$

Diese Fehlergröße erlaubt die Beurteilung der Güte einzelner Transformationen relativ zueinander.

Weitere Formulierungen zur Berechnung der Transformation erlauben die Gewichtung einzelner Punktkorrespondenzen, durch die Einführung von Gewichtungsfaktoren bei der Berechnung der Schwerpunkte und bei der Erstellung der Matrix  $\mathbf{M}$  (vergleiche Horn, 1987). Drixler (1993) erweitert die Formulierung um die Stochastizität der Punktkoordinaten durch die Einführung der Matrix der differentiellen Ähnlichkeitstransformation. Durch den Übergang in ein Gauß-Helmert-Modell ist jedoch eine direkte Lösung des Problems nicht mehr möglich. Stattdessen wird ausgehend von bekannten Startwerten eine iterative Lösung berechnet. Aus dem Ergebnis lassen sich neben dem geschätzten Varianzfaktor, die Kofaktorenmatrix sowie die Genauigkeiten der geschätzten Transformationsparameter ableiten.

### 5.4.2 Ableitung der Punktkorrespondenzen

Wie in der oben erwähnt, erhält man als Ergebnis der Korrespondenzanalyse zunächst korrespondierende Flächenpaare. Um die Transformation nach dem oben erarbeiteten Ansatz zu berechnen, muss nun zu jedem Flächenpaar ein entsprechendes Punktepaar abgeleitet werden.

Während bei einigen einfachen Flächentypen stets ein eindeutiger Punkt identifiziert werden kann, zum Beispiel der Mittelpunkt eines Kreises oder die Spitze eines Kegels, ist bei komplexen Flächentypen unter Umständen kein eindeutiger Punkt feststellbar. Auch die Randkurve einer Fläche stellt nur bedingt geeignete Punkte zur Verfügung. Zum einen kann die Randkurve selbst gekrümmt sein, was zu einem ähnlichen Problem im zweidimensionalen Fall führt. Zum anderen sind die Ränder einer Fläche, bedingt durch die Unzulänglichkeiten in den Segmentierungsverfahren, nur sehr unzuverlässig aus den Distanzdaten ableitbar.

Als zuverlässiger ableitbar haben sich die Zentroide einer Fläche erwiesen. Bei ebenen und durch Polygone begrenzten Flächen ist eine einfache und eindeutige Rechenvorschrift zur Berechnung der Flächenschwerpunkte gegeben. Bei gekrümmten Flächen ist die Definition eines eindeutigen Zentrums schwieriger. Zur einfachen und effizienten Berechnung ergeben sich zwei offensichtliche Möglichkeiten. Die Berechnung des arithmetischen Mittels aller Punkte einer Fläche, sowie die Berechnung eines zweidimensionalen Zentrums und dessen Transformation in den Raum.

Zur Berechnung des arithmetischen Mittels wird die Summe über alle zur Fläche  $S$  gehörenden Punkte gebildet

$$\bar{\mathbf{p}} = \frac{1}{N} \sum_{i=1}^N \mathbf{p}_i$$

sofern die Fläche als Menge diskreter Punkte gegeben ist. Andernfalls kann eine solche Menge aus jeder parametrisierten Beschreibung abgeleitet werden. Es ist anzumerken, dass der Punkt  $\bar{\mathbf{p}}$  im Allgemeinen nicht auf der Fläche  $S$  liegt. Dies ist eine wenig anschauliche Definition eines eindeutigen Punktes, insbesondere da wir ja von Distanzdaten ausgehen, welche eben die Oberfläche eines Objektes erfassen.

Eine alternative Definition eines eindeutigen Punktes kann über die zweidimensionale Gestalt der Fläche  $S$  erfolgen. Das Distanzbild enthält die zweidimensionale Projektion jeder abgebildeten Fläche. Aus dem Distanzbild kann durch einfache zweidimensionale Bildverarbeitung der zweidimensionale Schwerpunkt jeder segmentierten Fläche berechnet werden. Durch das Abgreifen der  $X$ ,  $Y$  und  $Z$ -Koordinaten an eben diesem Schwerpunkt erhält man einen dreidimensionalen Punkt.

Da die Projektion einer Fläche aus dem Modell in ihre zweidimensionale Gestalt innerhalb eines Distanzbildes nicht bekannt ist, diese Projektion basiert ja auf der Transformation die zu bestimmen ist, muss für die Berechnungen im CAD-Modell eine andere Methode gewählt werden.

Eine Methode die hinreichend ähnlich ist basiert auf der Parametrisierung einer Fläche  $S$  im zweidimensionalen Parameterraum  $(u, v)$  als  $S(u, v)$ . Hierbei ist die Fläche durch ihre Randkurve begrenzt und somit ist auch ihre Ausdehnung im Parameterraum beschränkt mit

$$u_{min} < u < u_{max} \quad v_{min} < v < v_{max}$$

Ein eindeutiger Mittelpunkt der Fläche  $S$  lässt sich als Punkt im Zentrum des Parameterraums bestimmen mit

$$\mathbf{p} = S\left(u_{min} + \frac{u_{max} - u_{min}}{2}, v_{min} + \frac{v_{max} - v_{min}}{2}\right)$$

Der Punkt  $\mathbf{p}$  liegt dann auf der Fläche  $S$ , wenn  $u_{min} + \frac{u_{max} - u_{min}}{2}$  beziehungsweise  $v_{min} + \frac{v_{max} - v_{min}}{2}$  zulässige Werte besitzen. Dies kann und muss im CAD-System bei jeder Berechnung überprüft werden. Der Punkt kann zum Beispiel dann nicht gültig sein, wenn er auf eine Aussparung in der Fläche trifft oder wenn die Randkurve konkav ist und den Punkt ausschließt. In einem solchen Fall wird das Punktepaar von der Transformationsberechnung ausgeschlossen. Die Abbildung 5.7 zeigt die Integration dieser Funktionalität in das CAD-System Pro/Engineer.

### 5.4.3 Transformationsergebnisse

Im Folgenden werden die Resultate der Transformationsberechnungen beschrieben, wie sie erreicht werden, wenn die Zuordnungen aus der Baumsuche verwendet werden. Für jede Teilfläche des CAD-Modells und jede

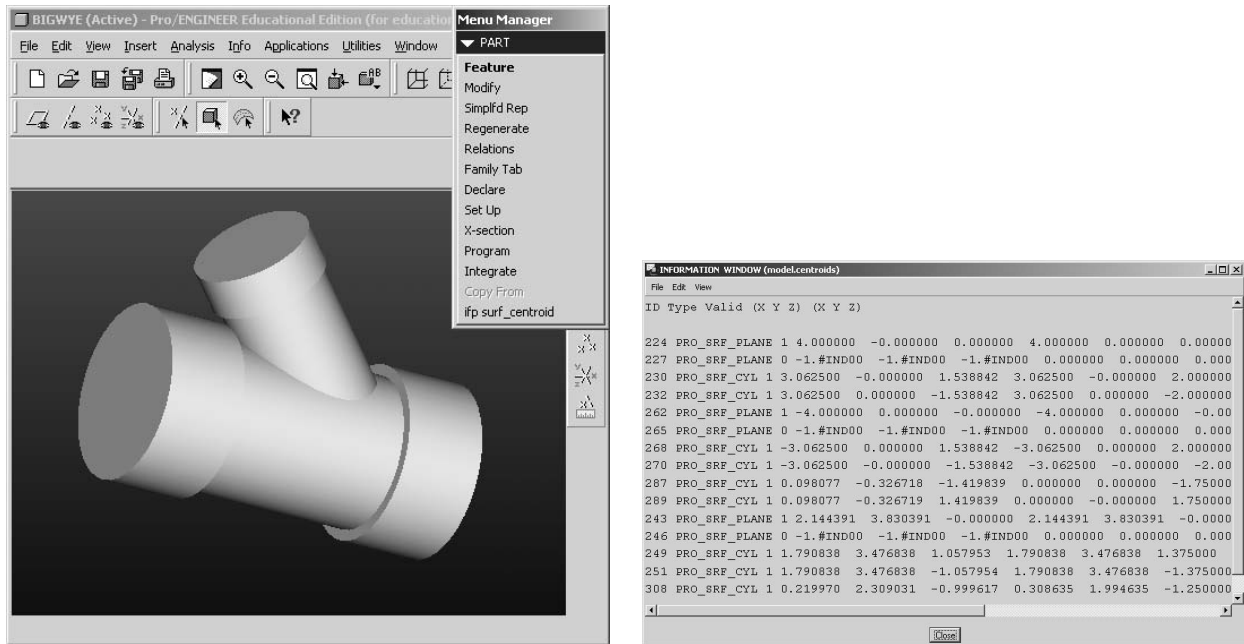


Abbildung 5.7: Integration der Zentroid-Bestimmung in Pro/Engineer. Links die grafische Benutzeroberfläche, rechts das Ergebnisfenster.

Region, die aus der Segmentierung resultiert werden die räumlichen Zentroide berechnet. Werden in einem Lösungspfad im Rahmen der Baumsuche zwei Flächen zur Korrespondenz gebracht, werden die entsprechenden räumlichen Punkte in die Transformationsschätzung mit einbezogen. Die dargestellten Ergebnisse folgen aus den Ergebnissen des Segmentierungsverfahrens, wie sie in Abschnitt 4.6.2 und 4.6.3 beschrieben wurden, und auf den Ergebnissen der Zuordnungsanalyse aus Abschnitt 5.3.

In Abbildung 5.8 werden die Ergebnisse für den Datensatz "Adapter" in der Übersicht gezeigt. Den vier verschiedenen Distanzbildern des Objektes werden jeweils die CAD-Modelle gegenübergestellt, in den verschiedenen Lagen, also Transformationen, die aus dem Suchbaum folgen. Angegeben werden jeweils die beiden besten Transformationen, also die Transformationen mit dem geringsten Fehler nach der Schätzung. Es zeigt sich, dass die korrekte Transformation auch zuverlässig den geringsten Fehler ergibt und die nächstbeste Transformation mit deutlich größerem Fehler folgt. Somit kann der korrekte Pfad im Suchbaum eindeutig identifiziert werden.

Abbildung 5.9 zeigt die Übersicht über die berechneten Transformationen für den Datensatz "Bigwe". Für jeweils vier verschiedene Distanzbilder, die das Objekt in vier verschiedenen Lagen zeigt, werden die drei besten Transformationen aufgezeigt. Wiederum kann die korrekte Transformation und somit die korrekte Lösung aus dem Suchraum anhand des Fehlerkriteriums zuverlässig identifiziert werden.

Die Abbildungen 5.10 und 5.11 zeigen nochmals im Detail die korrekten Transformationen. Es sind die zur Schätzung herangezogenen Zentroide sowohl im Distanzbild als auch im CAD-Modell eingezeichnet. Die berechnete Transformation wird anhand der Überlagerung des CAD-Modells mit den triangulierten Distanzdaten dargestellt. Dies ermöglicht es, die Güte der gewonnenen Transformation besser abzuschätzen.

Die Abbildungen 5.12 bis 5.14 zeigen die entsprechenden Resultate für das Testobjekt "Bigwe". Die Abbildung 5.15 schließlich zeigt das Ergebnis für das Prüfobjekt des Sonderforschungsbereichs 514. Hierbei wurde der Suchraum auf die fünf anhand ihrer Krümmungen eindeutig identifizierbaren gekrümmten Flächen beschränkt. Es zeigt sich somit, dass die gewählten Verfahrensschritte geeignet sind auch komplex strukturierte Objekte zu verarbeiten.

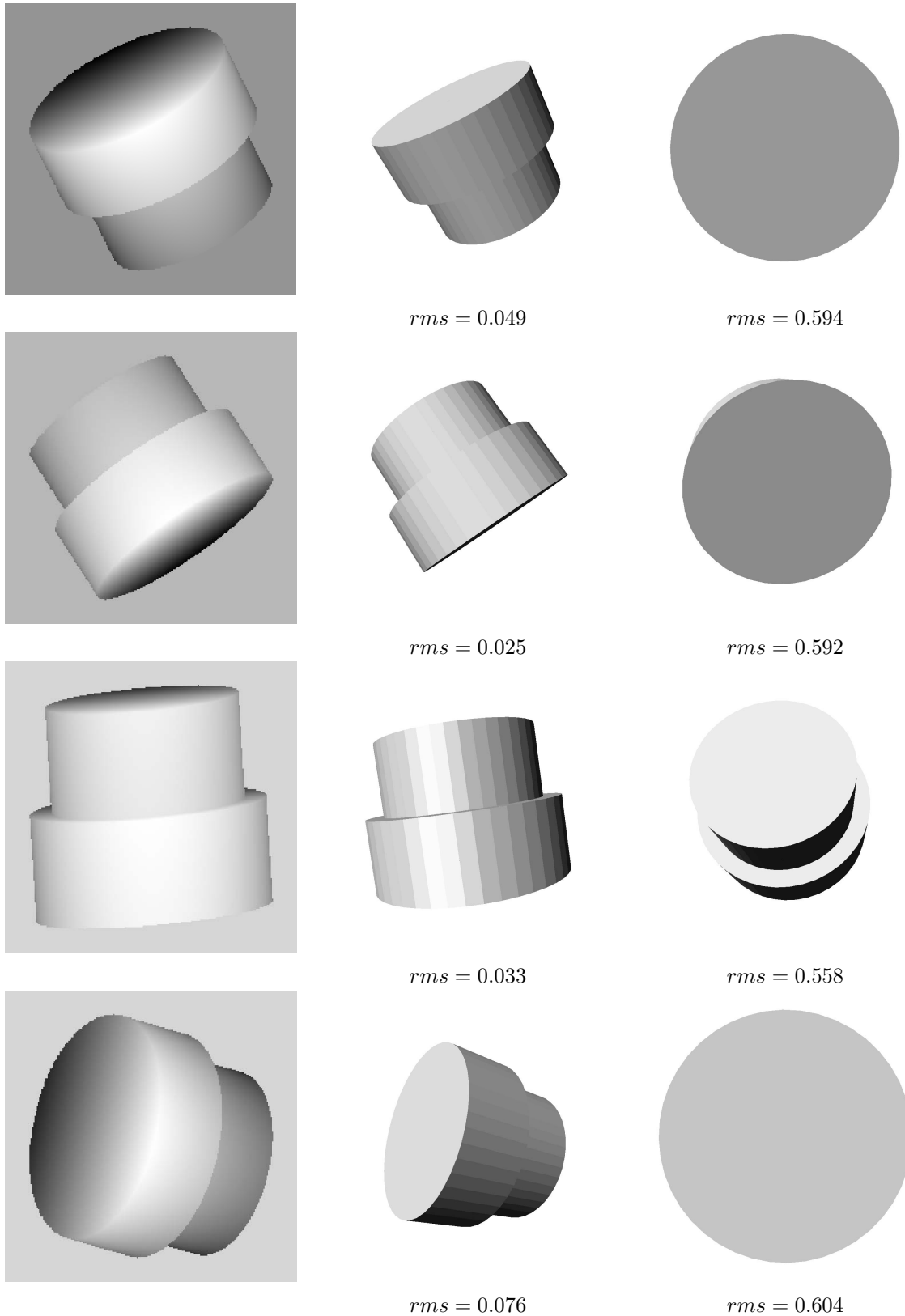


Abbildung 5.8: Ergebnis mit fünf verschiedenen Distanzbildern des Testdatensatzes “Adapter”. Jede Zeile entspricht einer Ansicht. In der ersten Spalte sind die Distanzbilder dargestellt. Die zweite und dritte Spalte zeigen das CAD-Modell in der jeweils berechneten Transformation. Die Transformationen sind dabei nach ihrem Fehler sortiert, das heißt, die beste Transformation steht in der zweiten Spalte. Unter den Abbildungen ist jeweils der Fehler der Transformationsschätzung angegeben.

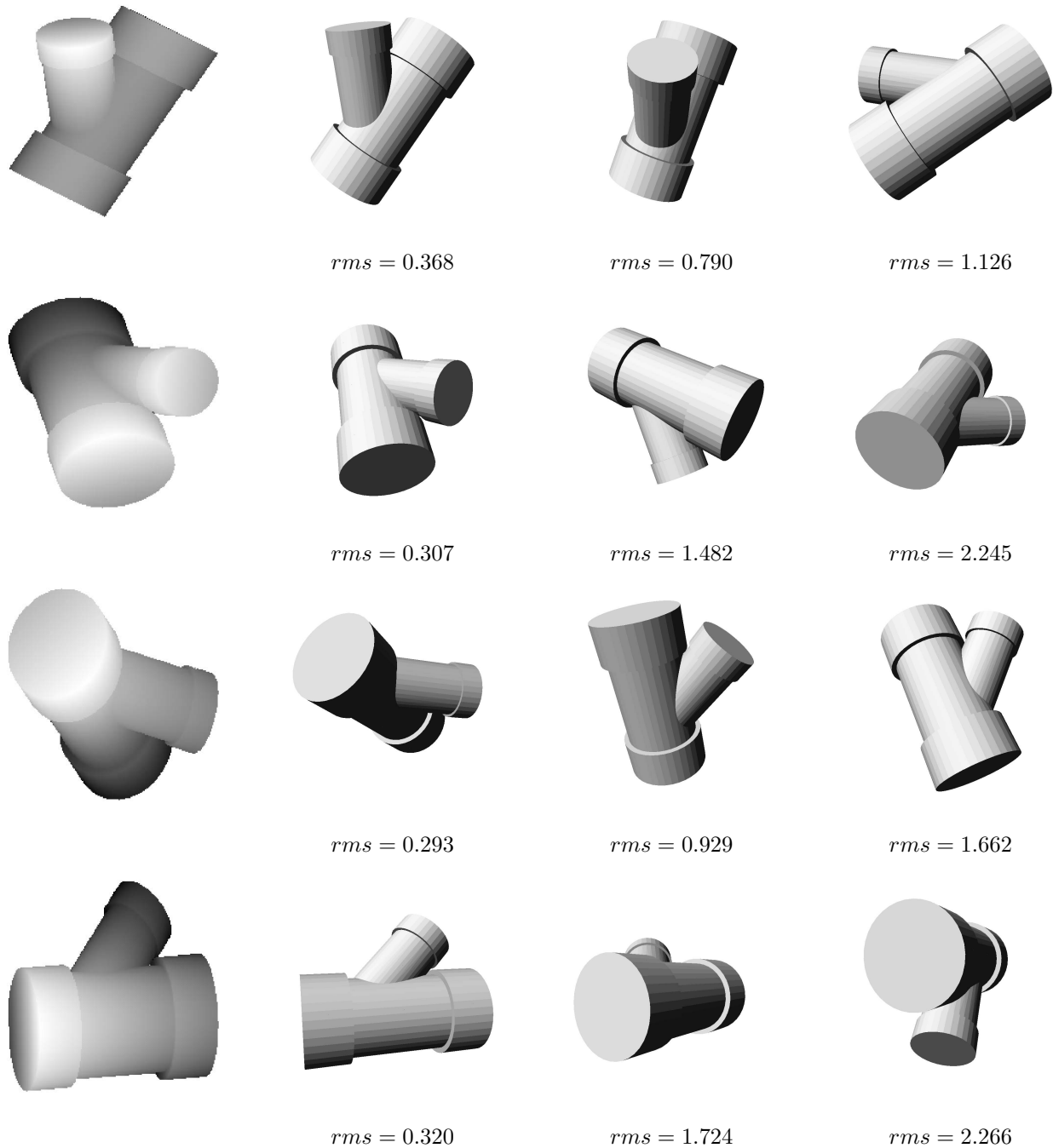


Abbildung 5.9: Ergebnis mit vier verschiedenen Distanzbildern des Testdatensatzes “Bigwye”. Jede Zeile entspricht einer Ansicht. In der ersten Spalte sind die Distanzbilder dargestellt. Die zweite, dritte und vierte Spalte zeigen das CAD-Modell in der jeweils berechneten Transformation. Die Transformationen sind dabei nach ihrem Fehler sortiert, das heißt, die beste Transformation steht in der zweiten Spalte. Unter den Abbildungen ist jeweils der Fehler der Transformationsschätzung angegeben.

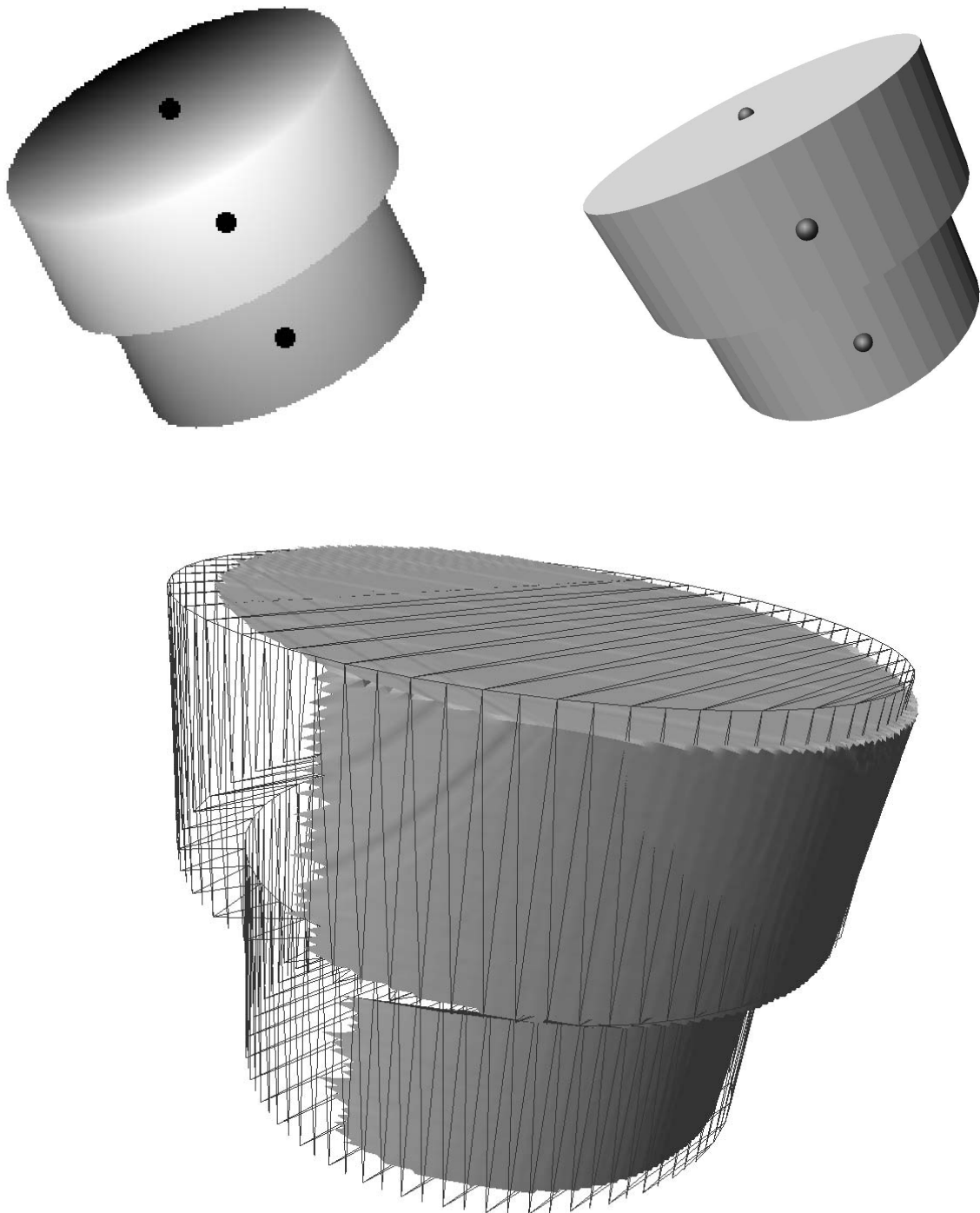


Abbildung 5.10: Die Abbildung zeigt die Transformation des CAD-Modells zur ersten Ansicht des Distanzdatensatzes "Adapter". In den oberen beiden Abbildungen ist links der Distanzdatensatz und rechts das CAD-Modell schattiert dargestellt. Zudem sind die Zentroide der korrespondierenden Flächen eingetragen. Die untere Abbildung zeigt eine Überlagerung beider Datensätze. Das CAD-Modell ist als Drahtgittermodell dargestellt und die Distanzdaten als schattierte Fläche.

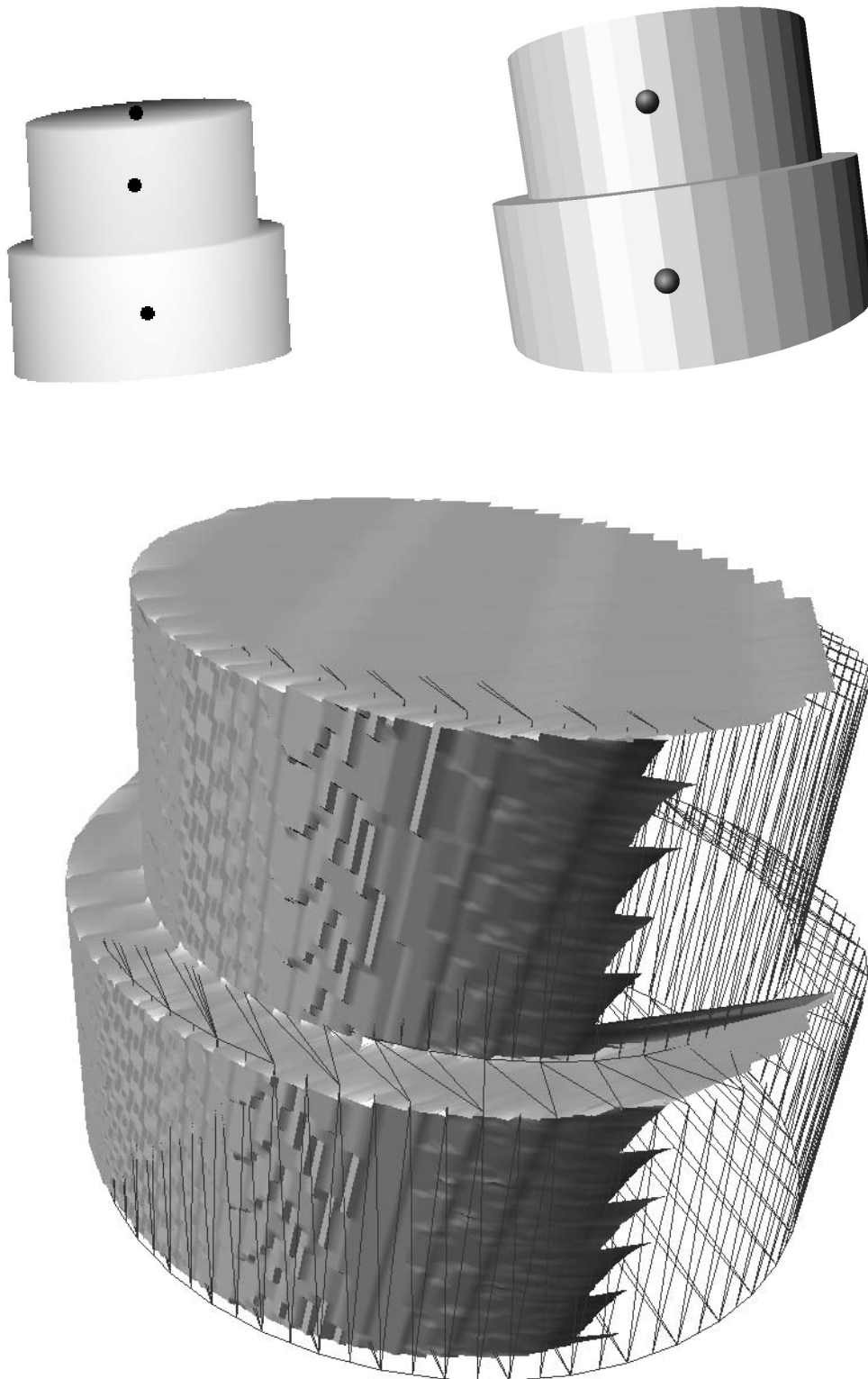


Abbildung 5.11: Die Abbildung zeigt die Transformation des CAD-Modells zur vierten Ansicht des Distanzdatensatzes "Adapter". In den oberen beiden Abbildungen ist links das Distanzbild und rechts das transformierte CAD-Modell dargestellt. Zudem sind die Zentroide der korrespondierenden Flächen eingetragen. Die untere Abbildung zeigt eine Überlagerung beider Datensätze. Das CAD-Modell ist als Drahtgittermodell dargestellt und die Distanzdaten als schattierte Fläche.



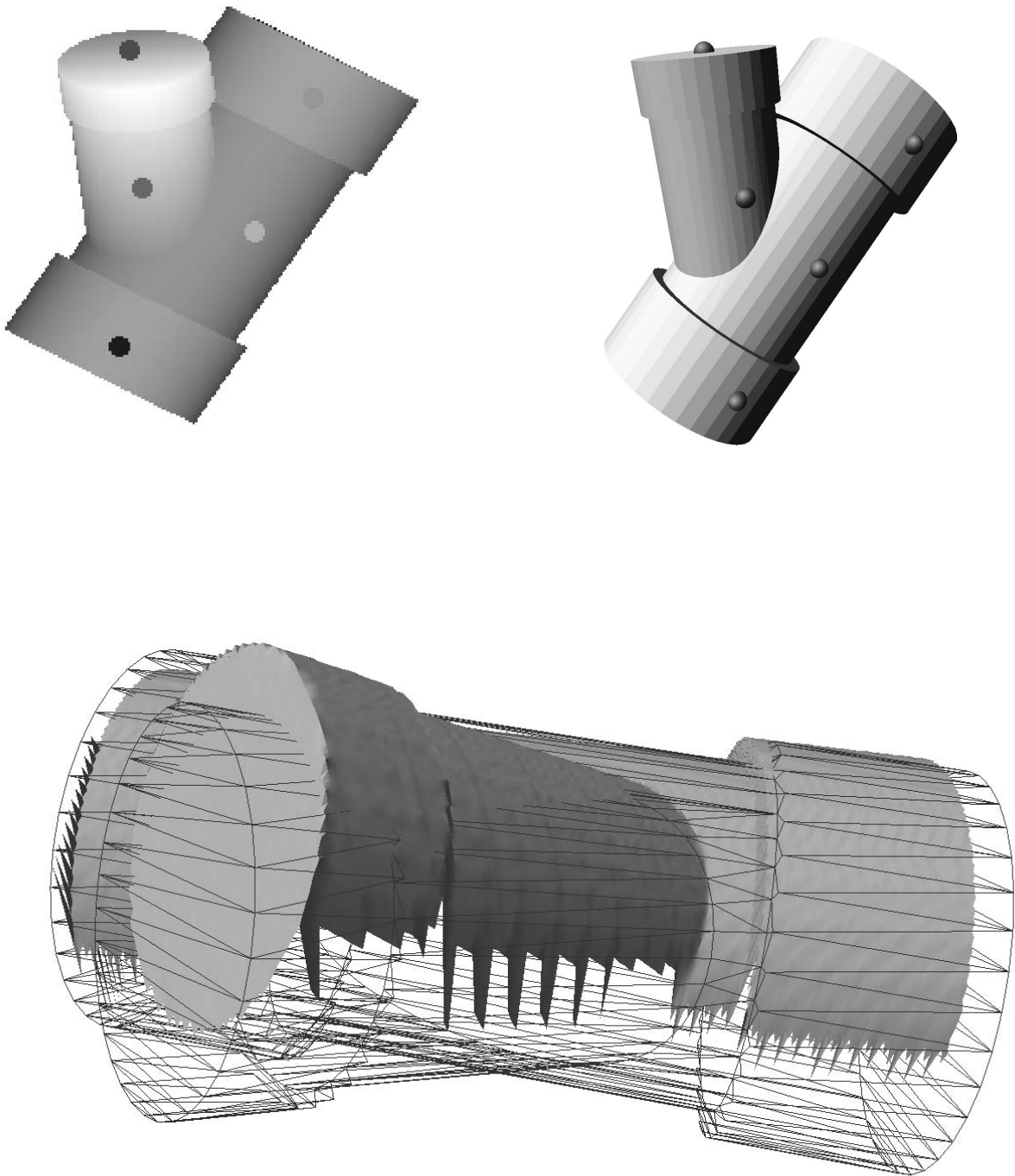


Abbildung 5.12: Die Abbildung zeigt die Transformation des CAD-Modells zur ersten Ansicht des Distanzdatensatzes "Big-*wye*". In den oberen beiden Abbildungen ist links der Distanzdatensatz und rechts das CAD-Modell schattiert dargestellt. Zudem sind die Zentroide der korrespondierenden Flächen eingetragen. Die untere Abbildung zeigt eine Überlagerung beider Datensätze. Das CAD-Modell ist als Drahtgittermodell dargestellt und die Distanzdaten als schattierte Fläche.

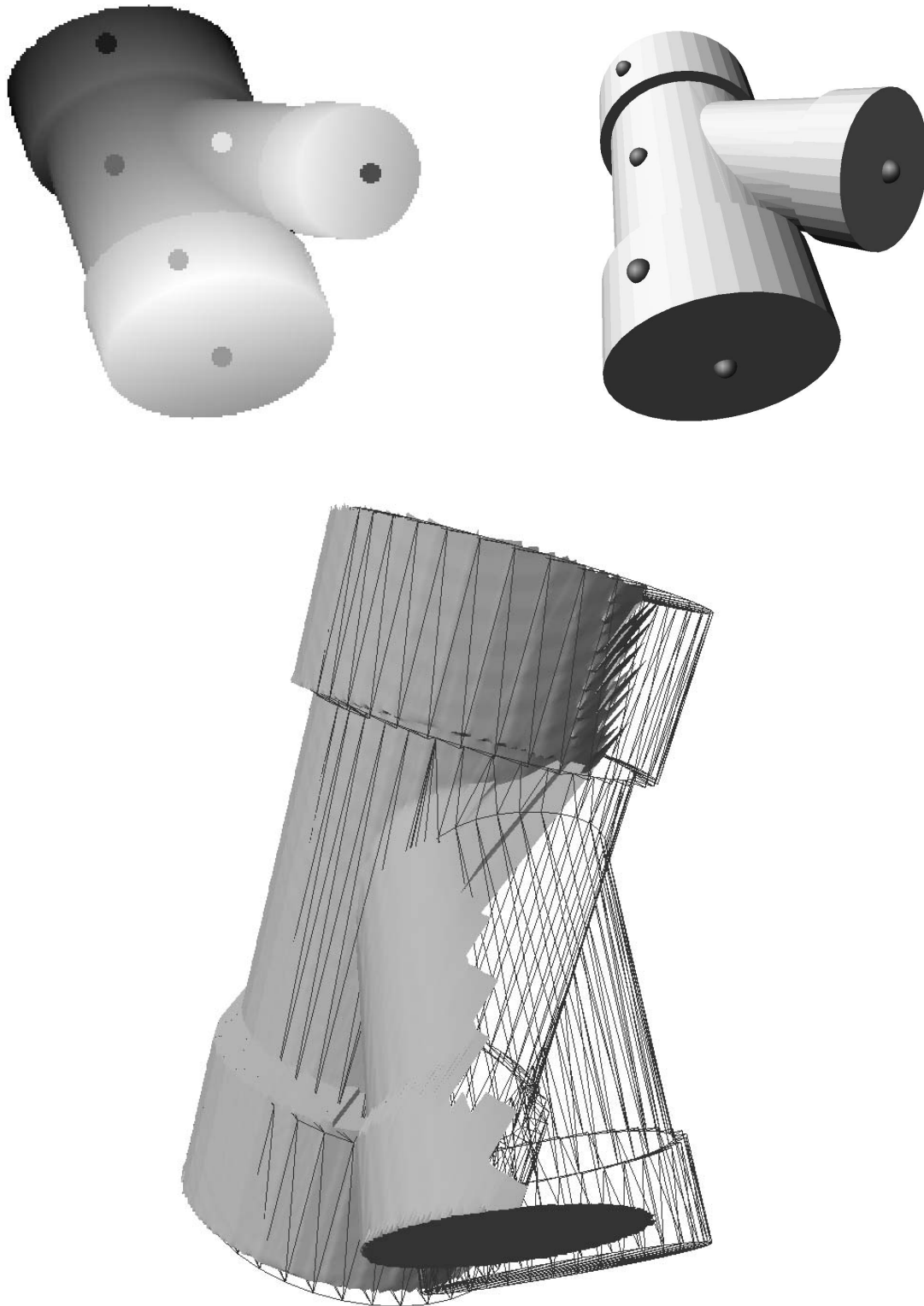


Abbildung 5.13: Die Abbildung zeigt die Transformation des CAD-Modells zur zweiten Ansicht des Distanzdatensatzes "Bigwye". In den oberen beiden Abbildungen ist links der Distanzdatensatz und rechts das CAD-Modell schattiert dargestellt. Zudem sind die Zentroide der korrespondierenden Flächen eingetragen. Die untere Abbildung zeigt eine Überlagerung beider Datensätze. Das CAD-Modell ist als Drahtgittermodell dargestellt und die Distanzdaten als schattierte Fläche.

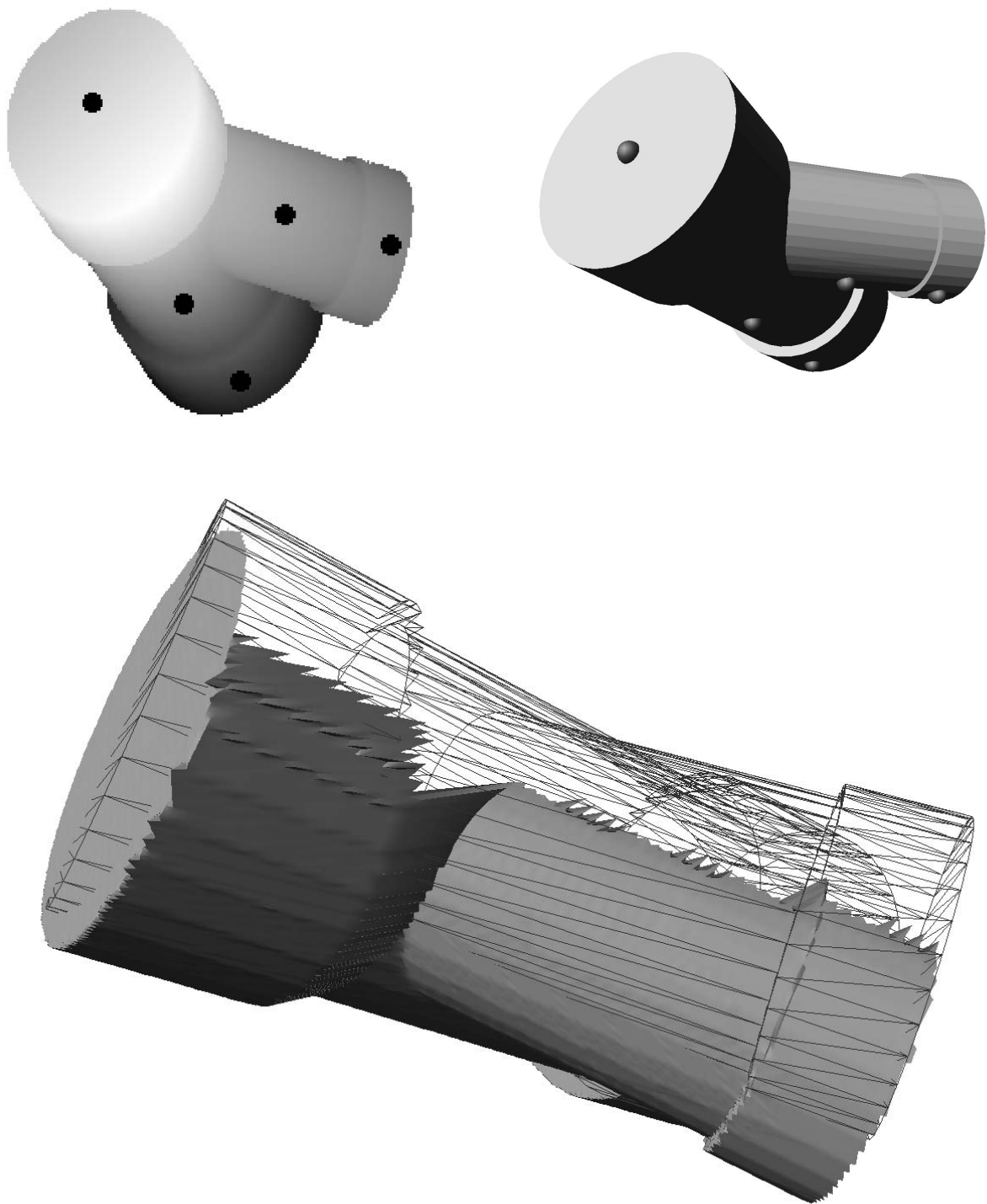


Abbildung 5.14: Die Abbildung zeigt die Transformation des CAD-Modells zur dritten Ansicht des Distanzdatensatzes "Bigwye". In den oberen beiden Abbildungen ist links der Distanzdatensatz und rechts das CAD-Modell schattiert dargestellt. Zudem sind die Zentroide der korrespondierenden Flächen eingetragen. Die untere Abbildung zeigt eine Überlagerung beider Datensätze. Das CAD-Modell ist als Drahtgittermodell dargestellt und die Distanzdaten als schattierte Fläche.

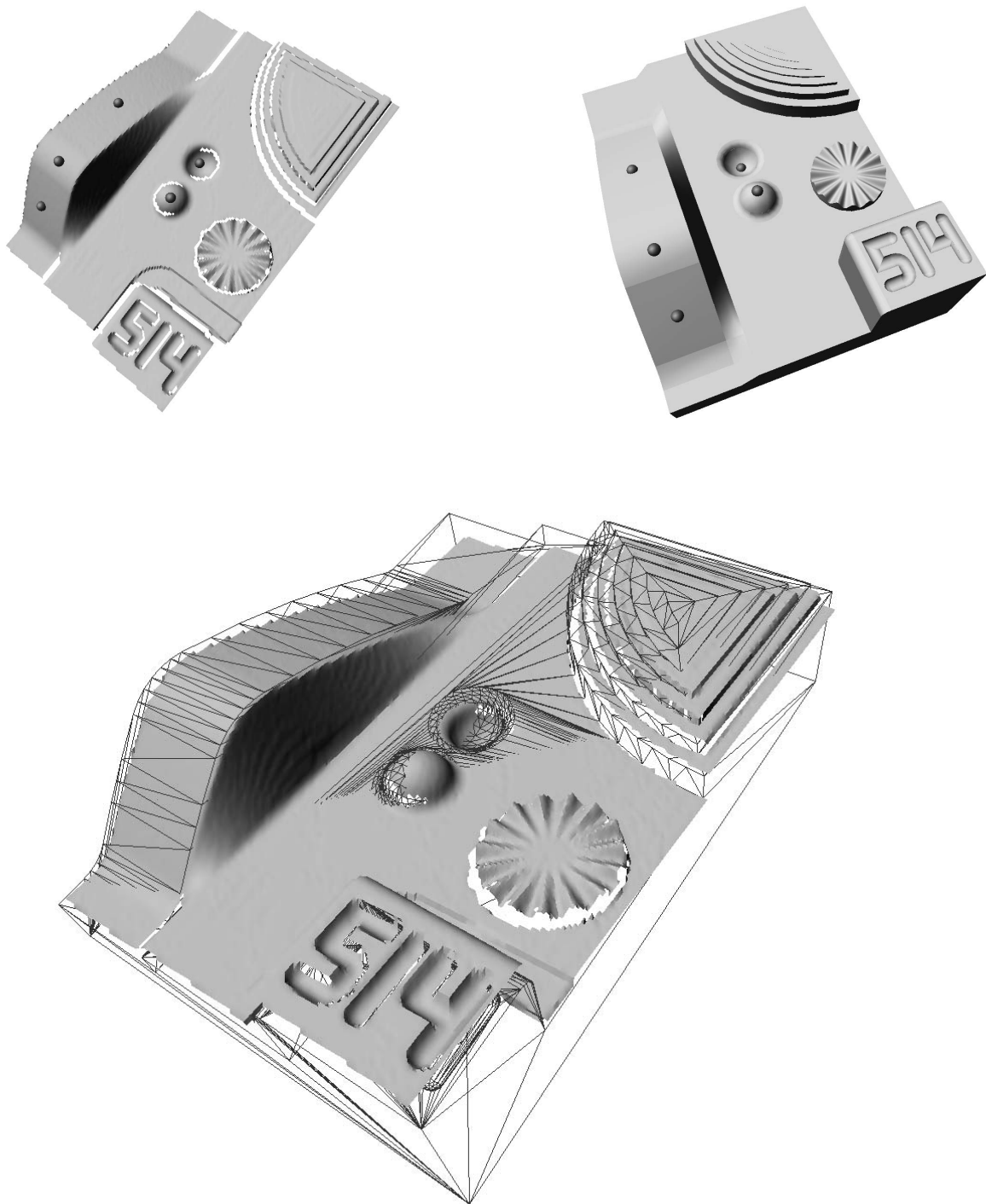


Abbildung 5.15: Die Abbildung zeigt die Transformation des CAD-Modells des Prüfobjektes des SFB 514 zum Distanzdatensatz. In den oberen beiden Abbildungen ist links der Distanzdatensatz und rechts das CAD-Modell jeweils schattiert dargestellt. Zudem sind die Zentroide der korrespondierenden Flächen eingetragen. Die untere Abbildung zeigt eine Überlagerung beider Datensätze. Das CAD-Modell ist als Drahtgittermodell dargestellt und die Distanzdaten als schattierte Fläche.

## 6 Zusammenfassung

Die Motivation der hier vorgelegten Arbeit stammt aus den Arbeiten zum Sonderforschungsbereich “Aktive Exploration mittels Sensor/Aktor-Kopplung zur adaptiven Mess- und Prüftechnik” an der Universität Stuttgart. Der Sonderforschungsbereich bewegte sich im Rahmen der optischen Mess- und Prüftechnik für industrielle Fertigungsprozesse. Das Ziel war durch Kombination verschiedener Sensoren und Auswerteverfahren ein flexibles System zu schaffen, das sich selbstständig an wechselnde Mess- und Prüfaufgaben anpassen kann. Hierzu ist es unter anderem notwendig, in das Messvolumen eingebrachte Prüflinge automatisch zu erkennen und deren Lage und Orientierung im Messsystem zu bestimmen.

Die vorgelegte Arbeit beschreibt ein modellbasiertes System zur Segmentierung und Objekterkennung von Objekten mit gekrümmten Flächen basierend auf Distanzbildern. Dabei wird ein weiteres Thema des Sonderforschungsbereiches in dieser Arbeit aufgegriffen, nämlich die Verwendung eines expliziten geometrischen Modells zur Speicherung aller für den Prüfprozess relevanten Daten. Die Repräsentation des Modells beruht auf einem CAD-Modell, das eine mathematisch exakte und zuverlässige Beschreibung beliebiger Körper erlaubt.

Das neu entwickelte Verfahren zur modellbasierten Segmentierung stellt einen Kern der Arbeit dar. Basierend auf der mittleren und der Gaußschen Krümmung, die als invariante Merkmale gewählt wurden, werden aus dem CAD-Modell Krümmungsklassen abgeleitet, die in einem überwachten Klassifikationsverfahren zur Segmentierung des Distanzbildes genutzt werden. Es konnte gezeigt werden, dass dieser Ansatz in der Lage ist gekrümmte Flächen aus dem Distanzbild zu extrahieren und dies auch in Fällen in denen benachbarte Flächen einen stetigen Übergang aufweisen. Besonders zu bemerken ist, dass nach diesem Ansatz eine Segmentierung des Distanzbildes in Regionen erfolgt, die der Aufteilung des CAD-Modells in seine Teilflächen entspricht.

Das Verfahren setzt dabei voraus, dass sich die Teilflächen des Körpers im Merkmalsraum, der durch die mittlere und die Gaußsche Krümmung gebildet wird, unterscheiden lassen. Dies ist jedoch nicht bei allen Objekttypen gegeben. Als Extremfall kann der Polyeder genannt werden, dessen Teilflächen sich allein durch die Krümmungen nicht unterscheiden lassen. Für solche speziellen Fälle existieren gesonderte Lösungsansätze, die im Einzelfall dem hier vorgestellten allgemeineren Ansatz vorzuziehen sind.

Besondere Betrachtung wurde der Krümmungsschätzung aus Distanzdaten gewidmet, die für das Verfahren der Segmentierung die invarianten Merkmale liefert. Wie zu erwarten war, ist die Krümmungsschätzung extrem empfindlich gegenüber dem Rauschen in den Sensordaten. Um dies zu kompensieren, müssen gesonderte Vorverarbeitungsschritte auf den Sensordaten durchgeführt werden. Diese Vorverarbeitungsschritte unterscheiden sich jedoch erheblich zwischen den verschiedenen Sensortypen. Die in dieser Arbeit verwendeten Daten eines Streifenprojektionssystems besitzen ein systematisches Wellenmuster, welches durch die Anwendung eines einfachen Glättungsfilters in einer Iteration kompensiert wurde. Es ist unmöglich allgemein-gültige Verfahren für alle Sensortypen aufzustellen. Mögliche Lösungsansätze könnten die von Boulanger et al. (2002) entwickelten speziellen Filterverfahren liefern, die auf einem physikalischen Sensormodell beruhen.

Die Korrespondenzanalyse, also die Zuordnung der Merkmale aus Szene und Modell, erfolgte durch ein Baum-suchverfahren. Aufgrund der frühzeitigen Einbindung der Modelldaten bei der Segmentierung, konnte die Anzahl der zur Zuordnung in Betracht kommenden Merkmale gering gehalten und somit die Größe des Suchraums reduziert werden.

Aus den korrespondierenden Flächen wurde mittels einer Starrkörpertransformation die Transformation des CAD-Modells auf das Objekt in der Szene berechnet. Das vorgestellte Verfahren der Objekterkennung arbeitet dabei völlig frei von Näherungswerten und in den Experimenten konnte gezeigt werden, dass die Objekte aus beliebiger Lage erkannt und transformiert werden können. Die Qualität der errechneten Transformation genügt voll den angestrebten Anwendungen und kann bei Bedarf ohne weitere Interaktion durch iterative Verfahren nach dem Ansatz des *Iterative Closest Point* verfeinert werden.

## A Eigenwert Verfahren

Ein einfaches lineares Gleichungssystem lässt sich wie folgt lösen :

$$\mathbf{Ax} = \mathbf{1} \quad (\text{A.1})$$

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{1} \quad (\text{A.2})$$

Wird das Gleichungssystem jedoch beschrieben durch implizite Gleichungen der Form  $F(a_1, \dots, a_n) = 0$ , so wird bezogen auf die obige Darstellung  $\mathbf{1}$  zum Nullvektor. In die Gleichung A.2 eingesetzt führt dies zu  $\mathbf{x} = 0$ :

$$\mathbf{Ax} = 0 \quad (\text{A.3})$$

$$\mathbf{x} = \mathbf{A}^{-1}0 = 0 \quad (\text{A.4})$$

Das Problem  $\mathbf{Ax} = 0$  lässt sich jedoch einfach lösen, wenn man die Definition von Eigenwert und Eigenvektor bedenkt:

Ein  $\lambda \in \mathfrak{R}$  heißt Eigenwert von  $\mathbf{L}$ , wenn es ein  $\mathbf{b} \in V(\mathfrak{R})$  gibt mit

$$\mathbf{Lb} = \lambda\mathbf{b}, \quad b \neq 0$$

Jedes solche  $\mathbf{b}$  heißt dann Eigenvektor von  $\mathbf{L}$  zum Eigenwert  $\lambda$ . Die Gleichung lässt sich umformen zu:

$$\mathbf{Lb} = \lambda\mathbf{b} \Leftrightarrow \mathbf{Lb} - \lambda\mathbf{b} = 0 \Leftrightarrow (\mathbf{L} - \lambda\mathbf{I})\mathbf{b} = 0$$

Existiert ein Eigenwert  $\lambda_0 = 0$  mit dem korrespondierenden Eigenvektor  $\mathbf{b}_{\lambda_0}$ , so gilt:

$$(\mathbf{L} - \lambda_0\mathbf{I})\mathbf{b}_{\lambda_0} = 0 \Leftrightarrow \mathbf{Lb}_{\lambda_0} = 0$$

Für das Gleichungssystem A.3 bedeutet dies, dass der Lösungsvektor  $\mathbf{x}$  berechnet werden kann als der Eigenvektor von  $\mathbf{A}$ , der zum Eigenwert 0 korrespondiert.

### Ausgleichung

Gibt es mehr Beobachtungen als Unbekannte, ist die Matrix  $\mathbf{A}$  nicht quadratisch und es existiert daher keine Inverse, beziehungsweise kein Eigenwert. Bei dieser Art der Problemstellung ist der Lösungsvektor  $x$  gesucht, der das Gleichungssystem mit geringst-möglichem Fehler löst:

$$\mathbf{Ax} = \mathbf{1} \quad (\text{A.5})$$

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{1} \quad (\text{A.6})$$

Ist  $\mathbf{1} = 0$  gilt das oben Gesagte analog. Die Lösung wird deutlich, wenn wir die Gleichung A.6 folgendermaßen umformen

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T0 \Leftrightarrow (\mathbf{A}^T\mathbf{A})\mathbf{x} = \mathbf{A}^T0 \Leftrightarrow (\mathbf{A}^T\mathbf{A})\mathbf{x} = 0$$

Der Lösungsvektor  $\mathbf{x}$  kann berechnet werden als derjenige Eigenvektor von  $(\mathbf{A}^T\mathbf{A})$ , der zum kleinsten Eigenwert korrespondiert. Zu beachten ist, dass jedes Vielfache des Lösungsvektors  $\rho\hat{\mathbf{x}}$  mit  $\rho \in \mathfrak{R}$  ebenfalls eine Lösung ist.

## B Alternative Herleitung der Krümmungen

### B.1 Erste Ableitung

Die erste Ableitung an der Stelle  $(u, v)$  wird beschrieben durch den Gradienten:

$$\nabla X(u, v) = \begin{bmatrix} X_u(u, v) \\ X_v(u, v) \end{bmatrix} \quad (\text{B.1})$$

oder in kurzer Schreibweise ohne Angabe der Koordinaten:

$$\nabla X = \begin{bmatrix} X_u \\ X_v \end{bmatrix} \quad (\text{B.2})$$

Der Gradient einer Fläche an einem Punkt trägt folgende Eigenschaften:

- Die Gradientenrichtung ist die Richtung des steilsten Anstiegs an diesem Punkt.
- Die Gradientenstärke ist die Steigung dieses Anstiegs.
- Die Gradientenrichtung ist senkrecht zur ebenen Kurve in diesem Punkt.
- Der Gradient definiert die Tangentialebene in diesem Punkt.
- Der Gradient kann verwendet werden, um die erste Ableitung in beliebiger Richtung zu berechnen:

$$X_r = \vec{r} \cdot \nabla X \quad (\text{B.3})$$

#### B.1.1 Normalkoordinaten erster Ordnung

Um die lokalen Eigenschaften einer Fläche unabhängig vom gewählten Koordinatensystem untersuchen zu können, ist es hilfreich ein lokales Koordinatensystem zu etablieren, das sich nur aus der lokalen Umgebung der Fläche selbst herleiten lässt.

Die Gradientenrichtung und ein dazu senkrecht stehender Vektor in der Tangentialebene bilden ein solches lokales Koordinatensystem, das unabhängig vom übergeordneten räumlichen Koordinatensystem ist. Selbstverständlich ist dieses Koordinatensystem für jeden Punkt der Fläche anders. Auch ändern sich die Werte der beteiligten Größen, wenn das übergeordnete Koordinatensystem gewechselt wird, aber die lokale Richtung bleibt dieselbe.

### B.2 Zweite Ableitung

Für die zweite Ableitung ist die *Hessesche* das Äquivalent zum Gradienten:

$$H = \begin{bmatrix} X_{uu} & X_{uv} \\ X_{vu} & X_{vv} \end{bmatrix} \quad (\text{B.4})$$

Da

$$X_{uv} = X_{vu}$$

ist die Matrix symmetrisch. Wiederum können die Ableitungen in beliebiger Richtung berechnet werden:

$$X_{rs} = \vec{r}^T H \vec{s} \quad (\text{B.5})$$

oder mit nur einer Richtung

$$X_{rr} = \vec{r}^T H \vec{r} \quad (\text{B.6})$$

Die Hessesche hat folgende Eigenschaften:

- Ihre Determinante ist gleich dem Produkt ihrer Eigenwerte und ist unabhängig von der Wahl von  $u$  und  $v$ .
- Ihre Spur ist ebenfalls unabhängig von der Wahl von  $u$  und  $v$ .
- Die geeignete Wahl von  $u$  und  $v$  erzeugt eine Diagonalmatrix - dies sind die Eigenvektoren von  $H$  und die Diagonalelemente sind die Eigenwerte.

### B.2.1 Hauptkrümmungen und Hauptkrümmungsrichtungen

Die Eigenwerte und Eigenvektoren der Hesseschen haben eine geometrische Bedeutung für den Punkt der Fläche:

- Der Eigenvektor, der zum größten Eigenwert korrespondiert, zeigt in die Richtung der größten Krümmung (zweite Ableitung).
- Der Eigenvektor, der zum kleinsten Eigenwert korrespondiert, zeigt in die Richtung der kleinsten Krümmung.
- Die korrespondierenden Eigenwerte sind die Krümmungen.

Die beiden Eigenvektoren heißen Hauptkrümmungsrichtungen. In ihrer Richtung existiert keine gemischte partielle Ableitung. Sie stehen immer senkrecht zueinander.

Die beiden Eigenwerte heißen Hauptkrümmungen. Sie sind translations- und rotations-invariant. Sie werden mit  $k_1$  und  $k_2$  abgekürzt und haben immer reale Werte.

### B.2.2 Normalkoordinaten zweiter Ordnung

Genau wie für die Normalkoordinaten erster Ordnung, können wir die Hauptkrümmungsrichtungen verwenden, um die Normal-Koordinaten zweiter Ordnung herzuleiten. Nimmt man noch den Normalenvektor hinzu, erhält man ein lokales Dreibein, welches unabhängig vom übergeordneten Koordinatensystem ist.



# Abbildungsverzeichnis

1.1	Ein allgemeines Schema des modellbasierten maschinellen Sehens in Anlehnung an Arman und Aggarwal (1993). . . . .	8
2.1	Die Stereo-Anordnung von zwei Kameras zum Objekt. . . . .	13
2.2	Sensorsystem mit Musterprojektion und drei Kameras der Firma CogniTens. . . . .	14
2.3	Mechanische Scanbewegungen für punkt- oder linienhafte Sensoren. . . . .	14
2.4	Kommerzielle Streifenprojektionssysteme. (a) COMET-400 Vario Zoom der Firma STEINBICHLER. (b) ATOS IIe der Firma GOM. . . . .	15
2.5	Das Prinzip des Lichtschnitts mit einer Kamera und einem Linienprojektor. . . . .	16
2.6	Die Projektionssequenzen für den Gray-Code und die beiden Alternativen zur Subpixel Bestimmung, Phasen-Shift und Line-Shift. . . . .	17
2.7	Aufbau und Anwendung von Streifenprojektionssystemen. . . . .	18
2.8	Ein Datensatz, wie er mit der Multi-Sensor-Messmaschine aus Abbildung 2.7 erzeugt wurde. . . . .	18
2.9	Leica Laser Radar System LR 200 für mittlere Volumina in industriellen Anwendungen mit 100 $\mu\text{m}$ Messgenauigkeit. . . . .	19
2.10	Typische terrestrische Laser Scanner für geodätische Anwendungen. (a) Leica HDS 3000. (b) Riegl LMS-Z420. (c) Zoller + Fröhlich Imager 5003. . . . .	19
2.11	Eindeutige Zerlegung eines Würfels in seine sechs Teilflächen. . . . .	23
2.12	Mehrere gültige Unterteilungen der Randfläche für denselben Körper. . . . .	23
2.13	Ordnung der Repräsentationsformen nach ihrer Dimension. . . . .	24
3.1	Darstellung einer nach Gleichung 3.3 parametrisierten Kugel mit den Kreiskurven entlang der Parameter. . . . .	28
3.2	Beispiel einer NURBS-Freiformfläche. . . . .	33
4.1	Beispiel für eine Segmentierung durch Gruppierung der Abtastzeilen. . . . .	40
4.2	Segmentierung durch Bereichswachstum. . . . .	41
4.3	Beispiel für eine vollständige Segmentierung durch Bereichswachstum in gekrümmte Flächen. . . . .	41
4.4	Veränderter Ablauf eines allgemeinen Schemas des modellbasierten künstlichen Sehens in Anlehnung an Arman und Aggarwal (1993). . . . .	44
4.5	Definition und Krümmungseigenschaften einer Ebene. . . . .	44
4.6	Definition und Krümmungseigenschaften einer Kugel. . . . .	45
4.7	Definition und Krümmungseigenschaften eines Zylinders. . . . .	45
4.8	Definition und Krümmungseigenschaften eines Kegels. . . . .	46
4.9	Definition und Krümmungseigenschaften eines Torus. . . . .	46
4.10	Eine Regelfläche erzeugt aus zwei identischen Kurven $c_1$ und $c_2$ , zwischen denen linear interpoliert wird. . . . .	48
4.11	Eine Regelfläche entstanden aus einem Zugkörper mit der erzeugenden Kurve $c$ . . . . .	48
4.12	Eine kubische NURBS-Fläche mit einem 4x4 Gitter von Kontrollpunkten. . . . .	49
4.13	Einbindung der Krümmungsextraktion in das CAD-System Pro/Engineer. . . . .	50
4.14	Die Normalenvektoren $\mathbf{n}_1$ und $\mathbf{n}_2$ zweier benachbarter Punkte $\mathbf{p}_1$ und $\mathbf{p}_2$ . . . . .	51
4.15	Artefakte einzelner Fehlpixel bei der Filtermethode. . . . .	54

4.16	Untersuchungen zur Genauigkeit eines Steifenprojektionssystems. . . . .	56
4.17	Untersuchungen an einer Testszene aus verschiedenen Flächentypen. . . . .	57
4.18	Resultat der modellbasierten Segmentierung. (a) Ergebnis nach reiner Klassifikation. (b) Klassifikation mit anschließendem Regionenwachstum. . . . .	59
4.19	(a) Das CAD-Modell des Datensatzes "Adapter" in einer schattierten Darstellung. (b)-(e) Segmentierungsergebnisse aus vier Distanzbildern. Das Resultat der Klassifikation ist unabhängig von der Lage des Objektes. . . . .	60
4.20	(a) Das CAD-Modell des Datensatzes "Bigwye" in einer schattierten Darstellung. (b)-(e) Segmentierungsergebnisse aus vier Distanzbildern. Das Resultat der Klassifikation ist unabhängig von der Lage des Objektes. . . . .	61
4.21	Selbst entwickelter Prüfkörper. (a) CAD-Modell des Körpers. (b) Fertigung des Körpers. . . . .	62
4.22	Segmentierung des Distanzbildes des Prüfkörpers. (a) Tiefenbild des Körpers. (b) Extrahierte Regionen. . . . .	63
4.23	Ausschnitt aus der Segmentierung. . . . .	63
5.1	Ableitung der Zuordnungsmengen aus der Segmentierung des Distanzdatensatzes "Adapter". . . . .	66
5.2	Der Suchbaum, der sich aus den in Abbildung 5.1 aufgeführten Zuordnungsmengen ergibt. . . . .	67
5.3	Ableitung der Zuordnungsmengen aus der Segmentierung des Distanzdatensatzes "Bigwye". . . . .	67
5.4	Der Suchbaum, der sich aus den in Abbildung 5.3 aufgeführten Zuordnungsmengen für eine Ansicht des Datensatzes "Bigwye" ergibt. . . . .	68
5.5	Untersuchung zur Zuordnungsanalyse ebener Flächensegmente. . . . .	69
5.6	Ausschnitt aus einem Suchbaum, der durch zusätzliche Merkmale und durch eine Begrenzung der Tiefe reduziert wurde. . . . .	70
5.7	Integration der Zentroid-Bestimmung in Pro/Engineer. Links die grafische Benutzeroberfläche, rechts das Ergebnisfenster. . . . .	74
5.8	Ergebnis mit fünf verschiedenen Distanzbildern des Testdatensatzes "Adapter". . . . .	75
5.9	Ergebnis mit vier verschiedenen Distanzbildern des Testdatensatzes "Bigwye". . . . .	76
5.10	Transformation des CAD-Modells zur ersten Ansicht des Distanzdatensatzes "Adapter". . . . .	77
5.11	Transformation des CAD-Modells zur vierten Ansicht des Distanzdatensatzes "Adapter". . . . .	78
5.12	Transformation des CAD-Modells zur ersten Ansicht des Distanzdatensatzes "Bigwye". . . . .	79
5.13	Transformation des CAD-Modells zur zweiten Ansicht des Distanzdatensatzes "Bigwye". . . . .	80
5.14	Transformation des CAD-Modells zur dritten Ansicht des Distanzdatensatzes "Bigwye". . . . .	81
5.15	Transformation des CAD-Modells des Prüfobjektes des SFB 514 zum Distanzdatensatz. . . . .	82

## Literaturverzeichnis

- Abdelmalek, N. N. (1990). Algebraic Error Analysis for Surface Curvatures and Segmentation of 3-D Range Images. *Pattern recognition*, 23, S. 807–817.
- Abramowski, S. und Müller, H. (1992). *Geometrisches Modellieren*. BI Wissenschaftsverlag.
- Ahn, S. J., Rauh, W., Cho, H. S. und Warnecke, H.-J. (2002). Orthogonal Distance Fitting of Implicit Curves and Surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), S. 620–638.
- Altschuler, M., Altschuler, B. R. und Taboada, J. (1979). Measuring surfaces spacecoded by a laser-projected dot matrix. In: *Imaging Applications for Automated Industrial Inspection and Assembly*, Bd. 182. SPIE, S. 187–191.
- Alvertos, Nicolas und D’Cunha, Ivan (1991). Curvature Effect of Median Filtering of Range Images. In: *Proceedings of Southeastcon*. IEEE, Williamsburg, VA, USA, S. 910–914.
- Arman, F. und Aggarwal, J. K. (1993). Model-Based Object Recognition in Dense Range Images. *ACM Computing Surveys*, 25(1), S. 5–43.
- Arun, K. S., Huang, T. S. und Blostein, S. D. (1987). Least-Squares fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5), S. 698–700.
- Ballard, D. H. und Brown, C. M. (1982). *Computer Vision*. Prentice-Hall, Englewood-Cliffs, New Jersey.
- Barr, A. H. (1981). Superquadrics and angle-preserving transformations. *IEEE Computer Graphics Applications*, 1(1), S. 11–23.
- Besl, P. J. (1988a). Active, Optical Range Imaging Sensors. *Machine Vision and Applications*, 1(2), S. 127–152.
- Besl, P. J. (1988b). *Surfaces in Range Image Understanding*. Springer series in perception engineering, Springer-Verlag. Revised ed. of Ph.D. thesis, University of Michigan.
- Besl, P. J. (1990). The free-form surface matching problem. In: Freeman, H. (Hg.), *Machine Vision for Three-Dimensional Scenes*. Academic Press, San Diego, S. 25–71.
- Besl, P. J. und Jain, R. C. (1985). Three-Dimensional Object Recognition. *ACM Computing Surveys*, 17(1), S. 75–145.
- Besl, P. J. und Jain, R. C. (1986). Invariant Surface Characteristics for 3D Object Recognition in Range Images. *Computer Vision, Graphics, Image Processing*, 33(1), S. 33–80.
- Besl, P. J. und Jain, R. C. (1988). Segmentation through variable-order surface fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(2), S. 167–192.
- Besl, P. J. und McKay, N. D. (1992). A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), S. 239–256.
- Böhm, J., Brenner, C., Gühring, J. und Fritsch, D. (1999). CAD-basierte Objekterkennung für ein multisensorielles Meßsystem. In: *DAGM-Symposium*. Informatik Aktuell, Springer, S. 62–69.
- Böhm, J., Gühring, J. und Brenner, C. (2001). Implementation and calibration of a multi sensor measuring system. In: El-Hakim, Sabry und Grün, Armin (Hg.), *Videometrics and Optical Methods for 3D Shape Measurement*, Bd. 4309 von *Proceedings of SPIE*. SPIE, S. 154–161.
- Binford, T. O. (1971). Visual Perception by Computer. In: *IEEE Conference on Systems Science and Cybernetics*. IEEE, Miami, FL, USA.
- Blais, François (2003). A Review of 20 Years of Range Sensor Development. In: *Videometrics VII, Proceedings of SPIE-IS&T Electronic Imaging*, Bd. 5013. SPIE, S. 62–76.
- Bolles, R. C. und Horaud, P. (1986). 3DPO: A Three-Dimensional Part Orientation System. *International Journal of Robotics Research*, 5(3), S. 3–26.
- Bookstein, F. L. (1979). Fitting Conic Sections to Scattered Data. *Computer Graphics and Image Processing*, 9(1), S. 56–71.

- Boulanger, P., Jokinen, O. und Beraldin, A. (2002). Intrinsic Filtering of Range Images Using a Physically Based Noise Model. In: *VI02*. S. 320–326.
- Brenner, C., Böhm, J. und Gühring, J. (1999). Photogrammetric calibration and accuracy evaluation of a cross-pattern stripe projector. *Proc. SPIE Videometrics VI*, 3641, S. 164–172.
- Brenner, C. und Fritsch, D. (1997). Object recognition for a multi-sensor measuring system. In: *Videometrics V*. SPIE, San Diego, S. 197–206.
- Brenner, Claus (2000). *Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen*. Dissertation, Universität Stuttgart.
- Breuckmann, Bernd (1993). *Bildverarbeitung und optische Meßtechnik*. Franzis.
- Bricault, I. und Monga, O. (1997). From Volume Medical Images to Quadratic Surface Patches. *Computer Vision and Image Understanding*, 67(1), S. 24–38.
- Campbell, R. J. und Flynn, P. J. (1998). A WWW-Accessible Database for 3D Vision Research. In: *Proc. IEEE Workshop on Empirical Evaluation Methods in Computer Vision*. S. 148–152.
- Campbell, R. J. und Flynn, P. J. (1999). Eigenshapes for 3D Object Recognition in Range Data. In: *IEEE Conf. on Computer Vision and Pattern Recognition*. S. II:505–510.
- Campbell, R. J. und Flynn, P. J. (2002). Recognition of free-form objects in dense range data using local features. In: *Proceedings of the 16th International Conference on Pattern Recognition*. S. 607–610.
- Cantzler, C. und Fisher, R. B. (2001). Comparison of HK and SC curvature description methods. In: *Third international conference on 3-D Digital Imaging and Modeling*. IEEE, Quebec City, Canada, S. 285–291.
- do Carmo, M. (1976). *Differential Geometry of Curves and Surfaces*. Prentice-Hall.
- Chen, Frank, Brown, Gordon M. und Song, Mumin (2000). Overview of three-dimensional shape measurement using optical methods. *Optical Engineering*, 39(1), S. 10–22.
- Dorai, C., Weng, J. und Jain, A. K. (1994). Optimal Registration of Multiple Range Views. In: *International Conference on Pattern Recognition*. S. A:569–571.
- Drixler, Erwin (1993). *Analyse der Form und Lage von Objekten im Raum*. Dissertation, Universität Fridericiana zu Karlsruhe, München.
- Duda, R. O. und Hart, P. E. (1973). *Pattern Classification and Scene Analysis*. Wiley.
- El-Hakim, S. F., Beraldin, J.-A. und Blais, F. (1995). A Comparative Evaluation of the Performance of Passive and Active 3-D Vision Systems. In: *St. Petersburg Conference on Digital Photogrammetry*. S. 14–25.
- Faber, Petko und Fisher, R. B. (2001). Euclidean Fitting Revisited. *Lecture Notes in Computer Science*, 2059, S. 165–173.
- Fan, T., Medioni, G. und Neviata, R. (1986). Description of Surfaces from Range Data using Curvature Properties. In: *Conference on Computer vision and pattern recognition*. S. 86–91.
- Fan, T. J. (1989). *Describing and Recognizing 3-D Objects Using Surface Properties*. Springer.
- Feddema, John T. und Little, Charles Q. (1997). Rapid World Modeling: Fitting Range data to Geometric Primitives. In: *IEEE International Conference on Robotics and Automation*, Bd. 4. IEEE, Albuquerque, New Mexico, USA, S. 2807–2812.
- Feldmar, J., Declerck, J., Malandain, G. und Ayache, N. (1997). Extension of the ICP Algorithm to Nonrigid Intensity-Based Registration of 3D Volumes. *CVIU*, 66(2), S. 193–206.
- Fitzgibbon, A. W., Pilu, M. und Fisher, R. B. (1999). Direct Least Square Fitting of Ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5), S. 476–480.
- Flynn, P. J. und Jain, A. K. (1989). On Reliable Curvature Estimation. In: *IEEE Conference on Pattern Recognition*. S. 110–116.
- Flynn, P. J. und Jain, A. K. (1991). CAD-Based Computer Vision: From CAD Models to Relational Graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2), S. 114–132.

- Frigui, H. und Krishnapuram, R. (1999). A Robust Competitive Clustering Algorithm With Applications in Computer Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5), S. 450–465.
- Gühring, J., Brenner, C., Böhm, J. und Fritsch, D. (2000). Data processing and calibration of a cross-pattern stripe projector. In: *Int. Archives of Photogrammetry and Remote Sensing*, Bd. 33. ISPRS, Amsterdam, Netherlands, S. 327–338.
- Gonzalez, R. C. und Woods, R.E. (1992). *Digital Image Processing*. Addison-Wesley, Reading, Massachusetts.
- Greenspan, M. A. und Boulanger, P. (1999). Efficient and Reliable Template Set Matching for 3D Object Recognition. In: *3DIM99 : Proceedings of the 2nd International Conference on 3-D Digital Imaging and Modeling, Ottawa, Canada, Oct. 4-8*. S. 230–239.
- Grimson, W. E. L. (1990). *Object Recognition by Computer*. MIT Press.
- Gruen, A., Zhang, L. und Visnovcova, J. (2001). Automatic Reconstruction and Visualization of a Complex Buddha Tower of Bayon. In: *Fifth Conference on Optical 3-D Measurement Techniques*. S. 166–178.
- Guy, G. und Medioni, G. (1997). Inference of Surfaces, 3D Curves, and Junctions from Sparse, Noisy, 3D Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11), S. 1265–1277.
- Hamann, Bernd (1992). Curvature Approximation for Triangulated Surfaces. In: Farin, Gerald E., Hagen, Hans, Noltemeier, Hartmut und Knödel, Walter (Hg.), *Geometric Modelling*, Bd. 8 von *Computing Supplement*. Springer, S. 139–153.
- Haralick, R. M. und Shapiro, L. G. (1992). *Computer and Robot Vision*. Addison-Wesley, Reading, Massachusetts.
- Hetzl, Guenter, Leibe, Bastian, Levi, Paul und Schiele, Bernt (2001). 3D Object Recognition from Range Images using Local Feature Histograms. In: *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'01)*, Bd. 2. S. 394–399.
- Hilton, A.D.M., Illingworth, J. und Windeatt, T. (1995). Statistics of Surface Curvature Estimates. *Pattern Recognition*, 28(8), S. 1201–1221.
- Hoffman, R. und Jain, A. K. (1986). Segmentation and Classification of Range Images. In: *IEEE Conf. on Computer Vision and Pattern Recognition*. S. 446–451.
- Hoover, A., Jean-Baptiste, G., Goldgof, D., Boywer, K.W., Flynn, P.J., Jiang, X. und Bunke, H. (1995). Range Image Segmentation: The User's Dilemma. In: *Proc. of International Symposium on Computer Vision*. S. 323–328.
- Hoover, A., Jean-Baptiste, G., Jiang, X.Y., Flynn, P.J., Bunke, H., Goldgof, D.B., Bowyer, K., Eggert, D.W., Fitzgibbon, A.W. und Fisher, R.B. (1996). An Experimental Comparison of Range Image Segmentation Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(7), S. 673–689.
- Horn, B. K. P. (1984). Extended Gaussian Images. In: *Proceedings of the IEEE*, Bd. 72. S. 1671–1686.
- Horn, Berthold K. P. (1987). Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4), S. 629–642.
- Horowitz, S. L. und Pavlidis, T. (1974). Picture Segmentation by directed split and merge procedure. In: *Proceedings of 2nd International Conference on Pattern Recognition*. S. 424–433.
- Häusler, Gerd (1999). Three-Dimensional Sensors - Potentials and Limitations. In: *Handbook of Computer Vision and Applications*, Bd. 1, Kap. 19. Academic Press, S. 485–506.
- Jain, A. K., Murty, M. N. und Flynn, P. J. (1999). Data Clustering: A Review. *ACM Computing Surveys*, 31(3), S. 264–323.
- Jarvis, R. A. (1983). A Perspective on Range Finding Techniques for Computer Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2), S. 122–139.
- Jiang, X., Bowyer, K. W., Morioka, Y., Hiura, S., Sato, K., Inokuchi, S., Bock, M., Guerra, C., Loke, R. E. und du Buf, J. M. H. (2000). Some Further Results of Experimental Comparison of RangeImage Segmentation Algorithms. In: *ICPR00*. S. Vol IV: 877–881.

- Jiang, X. und Bunke, H. (1998). Range image segmentation: Adaptive grouping of edges into regions. In: R. Chin, T.-C. Pong (Hg.), *Proc. of the 3rd Asian Conference on Computer Vision*, Bd. 1352 von *Lecture Notes in Computer Science*. Springer, Hongkong, S. 299 – 306.
- Jiang, Xiaoyi und Bunke, Horst (1994). Fast Segmentation of Range Images into Planar Regions by Scan Line Grouping. *Machine Vision and Applications*, 7(2), S. 115–122.
- Jiang, Xiaoyi und Bunke, Horst (1997). *Dreidimensionales Computersehen*. Springer, Berlin.
- Karbacher, S., Labourex, X., Schoen, N. und Haeusler, G. (2001). Processing Range Data for Reverse Engineering and Virtual Reality. In: *Third international conference on 3-D Digital Imaging and Modeling*. IEEE, IEEE Computer Society, Los Alamitos, CA, USA, S. 314–321.
- Koenderink, J. J. und van Doorn, A. (1992). Surface shape and curvature scales. *Image and Vision Computing*, 10(8), S. 557–565.
- Koenderink, J. J. und van Doorn, A. (1994). Two-plus-one-dimensional differential geometry. *Pattern Recognition Letters*, 5, S. 439–443.
- Koster, K. und Spann, M. (2000). MIR: An Approach to Robust Clustering-Application to Range Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(5), S. 430–444.
- Krsek, P., Lukacs, G. und Martin, R. R. (1998). Algorithms for Computing Curvatures from Range Data. In: *The Mathematics of Surfaces VIII*. S. 1–16.
- Lorusso, A., Eggert, D. W. und Fisher, R. B. (1995). A Comparison of Four Algorithms for Estimating 3-D Rigid Transformation. In: *British Machine Vision Conference*. S. 237–246.
- Lukacs, G., Martin, R. R. und D., Marshall A. (1998). Faithful least-squares fitting of spheres, cylinders, cones and tori for reliable segmentation. In: Burkardt, H. und Neumann, B. (Hg.), *Proc. of European Conference on Computer Vision*, Bd. 1. S. 671–686.
- Messmer, B. T. und Bunke, H. (1998). A New Algorithm for Error Tolerant Subgraph Isomorphism Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(5), S. 493–504.
- Mortenson, Michael E. (1997). *Geometric Modeling*. John Wiley & Sons.
- Page, D. L., Koschan, A., Sun, Y., Paik, J. und Abidi, M.A. (2001). Robust Crease Detection and Curvature Estimation of Piecewise Smooth Surfaces from Triangle Mesh Approximations Using Normal Voting. In: *IEEE Conf. on Computer Vision and Pattern Recognition*. S. 162–167.
- Parvin, B. und Medioni, G.G. (1986). Segmentation of Range Images into Planar Surfaces by Split and Merge. In: *IEEE Conf. on Computer Vision and Pattern Recognition*. S. 415–417.
- Petitjean, Sylvain (2002). A Survey of Methods for Recovering Quadrics in triangle Meshes. *ACM Computing Surveys*, 34(2), S. 211–262.
- Powell, M. W., Bowyer, K. W., Jiang, X. und Bunke, H. (1998). Comparing Curved-Surface Range Image Segmenters. In: *International Conference on Computer Vision 1998*. S. 286–291.
- Redding, Nicholas J. (2000). Implicit Polynomials, Orthogonal Distance Regression, and the Closest Point on a Curve. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(2), S. 191–199.
- Requicha, A. A. G. (1980). Representations for rigid solids: theory, methods and systems. *ACM Computing Surveys*, 12, S. 437–464.
- Rioux, M. (1984). Laser range finder based on synchronized scanners. *Applied Optics*, 23, S. 3837–3844.
- Sacchi, R., Poliakoff, J., Häfele, K. und Thomas, P. (1999). Curvature Estimation for Segmentation of Triangulated Surfaces. In: *3DIM99 : Proceedings of the 2nd International Conference on 3-D Digital Imaging and Modeling, Ottawa, Canada, Oct. 4-8*. IEEE Computer Society, Los Alamitos, CA, USA, S. 536–544.
- Sanso, F. (1973). An exact solution of the roto-translation problem. *Photogrammetria*, 29, S. 203–216.
- Spapidis, N. S. und Besl, J. P. (1995). Direct Construction of Polynomial Surfaces from Dense Range Images through Region Growing. *TG*, 14(2), S. 171–200.

- Stokely, E. M. und Wu, S. Y. (1992). Surface Parameterization and Curvature Measurement of Arbitrary 3-D Objects: Five Practical Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8), S. 833–840.
- Tang, C. K. und Medioni, G. (1999). Robust Estimation of Curvature Information from Noisy 3D Data for Shape Description. In: *Proc. of International Conference on Computer Vision*. S. 426–433.
- Taubin, G. (1991). Estimation of Planar Curves, Surfaces, and Nonplanar Space Curves Defined by Implicit Equations with Applications to Edge and Range Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(11), S. 1115–1138.
- Taubin, Gabriel (1995). Estimating the Tensor of Curvature of a Surface from a polyhedral Approximation. In: *Proc. of International Conference on Computer Vision*. S. 902–907.
- Thirion, Jean-Philippe (1993a). The Extremal Mesh and the Understanding of 3D Surfaces. Techn. Ber. 2149, INRIA.
- Thirion, Jean-Philippe (1993b). New Feature Points based on Geometric Invariants for 3D Image Registration. Techn. Ber. 1901, INRIA.
- Tiziani, H. J. (1997). Optical Metrology of Engineering Surfaces - Scope and Trends. In: Rastogi, P. K. (Hg.), *Optical Measurement Techniques and Applications*. Artech House, Boston Aufl., S. 15–50.
- Trucco, E. und Fisher, R. B. (1995). Experiments in Curvature-Based Segmentation of Range Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2), S. 177–182.
- VDI (2002). *Optische 3D-Messtechnik - Blatt 2*. VDI/VDE-Richtlinie 2634.
- Walker, M. W., Shao, L. und Volz, R. A. (1991). Estimating 3-D Location Parameters Using Dual Number Quaternions. *CVGIP: Image Understanding*, 54(3), S. 358–367.
- Werghi, N., Fisher, R. B., Ashbrook, A. und Robertson, C. (1999). Faithful Recovering of Quadric Surfaces from 3D Range Data. In: *3DIM99 : Proceedings of the 2nd International Conference on 3-D Digital Imaging and Modeling, Ottawa, Canada, Oct. 4-8*. IEEE Computer Society, Los Alamitos, CA, USA, S. 280–289.
- Yuen, Peter, Khalili, Nasser und Mokhtarian, Farzin (1999). Curvature Estimation on Smoothed 3-D Meshes. In: *British Machine Vision Conference*. S. 133–142.

# Lebenslauf

## Persönliche Angaben

Jan Böhm

Geboren 9. August 1970 in Stuttgart

## Schulbildung

1977 - 1981 Eichendorffschule, Stuttgart

1981 - 1990 Elly-Heuss-Knapp-Gymnasium, Stuttgart

17. Mai 1990 Abitur

## Studium

1991 Diplomstudium Informatik an der Universität Stuttgart

1995 - 1996 Masterstudium an der University of Massachusetts at Dartmouth

1. September 1996 Abschluss Master of Science, Computer Science

29. September 1997 Abschluss als Diplom Informatiker an der Universität Stuttgart

## Beruf

Seit November 1997 wissenschaftlicher Mitarbeiter am Institut für Photogrammetrie der Universität Stuttgart

1998 - 2001 Mitarbeiter im Sonderforschungsbereich 514