



HNI-VERLAGSSCHRIFTENREIHE
Prof. Dr. rer. nat Georg Hartmann (Hrsg.)

Ralph Trapp

***Stereoskopische
Korrespondenzbestimmung
mit impliziter
Detektion von Okklusionen***

Band

43

HEINZ NIXDORF INSTITUT
Universität-GH Paderborn

Der Einsatz binokularer Sehsysteme eröffnet sowohl in der Natur als auch in der Technik die Möglichkeit zum räumlichen Sehen. Das Grundprinzip bildet hierbei eine passive Triangulation, deren Ausgangspunkte die korrespondierenden Positionen darstellen, auf die ein Raumpunkt in die Stereobilder projiziert wird. Das zentrale Problem besteht bei dieser Technik darin, die korrespondierenden Bildpunkte eindeutig einander zuzuordnen. Dieses sogenannte Korrespondenzproblem ist einerseits aufgrund mehrerer ähnlicher Strukturen in der betrachteten Szene oft stark mehrdeutig und besitzt andererseits nicht immer eine Lösung, da Bereiche in der Szene auftreten können, die nur aus einer der beiden Perspektiven zu sehen sind. Weiterhin wird eine eindeutige Zuordnung korrespondierender Bildbereiche durch interokuläre Differenzen wie perspektivische Verzerrungen, Beleuchtungsunterschiede und Rauschprozesse zusätzlich erschwert.

In der vorliegenden Arbeit werden die einzelnen Komponenten eines Gesamtsystems vorgestellt, die zur stereoskopischen Rekonstruktion der räumlichen Struktur einer Szene erforderlich sind. Den Schwerpunkt der Arbeit bildet ein Selbstorganisationsprozeß, der in Verbindung mit weiteren Verfahrensschritten eine eindeutige Zuordnung korrespondierender Bildpunkte erlaubt. Darüber hinaus werden hierbei einseitig sichtbare Bildbereiche, die eine wesentliche Fehlerursache in der Stereoskopie darstellen, detektiert und vom Zuordnungsprozeß ausgeschlossen.

Ralph Trapp

geboren 1968, studierte Elektrotechnik mit der Vertiefungsrichtung Automatisierungstechnik an der Universität-GH Paderborn. Von 1995 bis 1998 war er als Stipendiat des Graduiertenkollegs am Heinz Nixdorf Institut der Universität-GH Paderborn im Fachgebiet Grundlagen der Elektrotechnik (Leitung Prof. Dr. rer. nat. Georg Hartmann) wissenschaftlich tätig. Bis zu seiner Promotion im September 1998 beschäftigte sich Herr Trapp mit der Entwicklung eines Gesamtsystems zur Stereobildanalyse.

ISBN 3 - 931466 - 42 - 6

Trapp, Ralph

Stereoskopische Korrespondenzbestimmung mit impliziter Detektion von Okklusionen / Trapp, Ralph. - 1. Aufl.- Paderborn: HNI-Verlagsschriftenreihe, 1998

(HNI-Verlagsschriftenreihe, Bd. 43; Grundlagen der Elektrotechnik; Herausgeber: Hartmann, Georg) Zugl.: Paderborn, Univ.-GH, Diss., 1998

ISBN 3 - 931466 - 42 - 6

© Heinz Nixdorf Institut, Universität-GH Paderborn - Paderborn - 1998.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Herausgebers und des Verfassers unzulässig und strafbar. Das gilt insbesondere für Vervielfältigung, Übersetzungen, Mikroverfilmungen und Einspeicherung und Verarbeitung in elektronischen Systemen.

Satz und Gestaltung: Ralph Trapp

Druck und Einband: Druckerei Reike, Paderborn

Umschlaggestaltung: A&O Marketing, Paderborn

Printed in Germany

Geleitwort des Herausgebers

Die Erforschung und Nachbildung kognitiver Leistungen hat in einer Zeit fast unbegrenzt wachsender Rechnerleistung große Fortschritte gemacht. Dies gilt auch für die automatische visuelle Erkennung und Vermessung von Objekten. Allerdings zeigt sich, daß derzeit die Anwendungen vorwiegend auf die Sichtkontrolle und Fertigungsüberwachung konzentriert sind. Die naheliegende Anwendung im Bereich der Robot-Vision, also der sichtgestützten Manipulation, hat noch nicht den oft vorhergesagten Durchbruch erreicht. Bei immer gleichbleibenden Fertigungsvorgängen wird Robot-Vision oft als überflüssig betrachtet, bei stark variierenden Manipulationsaufgaben wie etwa der Demontage erscheint die Automatisierung oft zu aufwendig oder nicht realisierbar.

Ein wichtiges Ziel unserer Arbeitsgruppe war es daher, anhand der Demontage von Altautos exemplarisch zu zeigen, daß mit einem lernfähigen Objekterkennungssystem eine sichtgeführte, wissensbasierte Manipulation mit vertretbarem Aufwand machbar ist. Für die Anwendbarkeit war von hervorragender Bedeutung, daß es gelang, die Vermessung der Demontageobjekte in sechs Freiheitsgraden mit hinreichender Genauigkeit auf der Basis stereoskopischer Information durchzuführen. Dies war so wichtig, weil ein über einen Meter entferntes Kamerapaar - im Gegensatz zu Sensoren an der Roboterhand - den Manipulationsprozeß nicht stört und keine überflüssigen Roboterbewegungen zur Sensorpositionierung erfordert. Dies war auch deshalb wichtig, weil einunddieselbe Sensorik Objekterkennung und Vermessung übernimmt und Sensorfusionsprobleme vermieden werden.

Obwohl nun Stereosysteme seit vielen Jahren ausführlich untersucht wurden, waren sie in der bekannten Form für die genannte Aufgabenstellung nicht geeignet. So ändern sich in dem von uns untersuchten Beispiel der Demontage von Autorädern mit jedem Manipulationsvorgang räumliche Lage, Größe, Form, Bauart und Zustand der Objekte, damit aber auch Konturstrukturen und räumliche Verdeckungen. Herkömmliche Stereosysteme zeigen unter diesen Bedingungen nicht die gewünschte Robustheit und führen bei räumlichen Verdeckungen oft zu unvorhersagbaren Meßfehlern.

Mit seiner Dissertation leistete Herr Trapp einen maßgeblichen Beitrag zur qualitativen und quantitativen Verbesserung von Stereosystemen im Allgemeinen und für die genannte Anwendung. Er setzt auf einem Verfahren von Haken und Reimann auf, bei dem die geschätzten Disparitäten unter den Randbedingungen Eindeutigkeit und Kontinuität iterativ optimiert werden. Es gelingt ihm, die Detektion von Okklusionen in diesen Iterationsprozeß einzubeziehen, falsche Distanzzuweisungen an verdeckten Bereichen auszuschließen und die Ausbreitung verdeckungsbedingter Fehler in sichtbare Gebiete hinein zu unterbinden. Zusätzlich konnte die Möglichkeit zu einer subpixelgenauen Disparitätsschätzung geschaffen werden, durch die die Genauigkeit des Verfahrens erheblich erhöht wurde. Herr Trapp hat damit nicht nur neue theoretische Ergebnisse vorgelegt, sondern diese auch in einem sehr robusten System der Anwendung zugänglich gemacht. Schließlich hat die technische Realisierung eines aus Standard-Kameras aufgebauten kostengünstigen Stereosystems die praktische Bedeutung der Arbeit weiter gesteigert.

Stereoskopische Korrespondenzbestimmung mit impliziter Detektion von Okklusionen

Zur Erlangung des akademischen Grades eines

DOKTOR - INGENIEUR

vom Fachbereich Elektrotechnik der

Universität Paderborn

- Gesamthochschule -

genehmigte Dissertation

von

Dipl.-Ing. Ralph Trapp

aus Winterberg

Referent: Prof. Dr. rer. nat. Georg Hartmann

Korreferent: Prof. Dr.-Ing. Ulrich Rückert

Tag der mündlichen Prüfung: 15. September 1998

Paderborn 1998

D 14 - 140

Vorwort

Die vorliegende Arbeit entstand im Rahmen meiner wissenschaftlichen Tätigkeit im Fachgebiet Grundlagen der Elektrotechnik an der Universität-GH Paderborn. An dieser Stelle möchte ich all denen herzlich danken, die mich in vielfältiger Weise bei der Anfertigung der Arbeit unterstützt haben.

Mein ganz besonderer Dank gilt Herrn Prof. Dr. rer. nat. Georg Hartmann, der diese Arbeit mit großem Interesse angeregt, betreut und gefördert hat. Sein Engagement, aber auch der mir eingeräumte Handlungsfreiraum trugen entscheidend zum Gelingen der Arbeit bei.

Herrn Prof. Dr.-Ing. Ulrich Rückert danke ich für die bereitwillige Übernahme des Korreferats und das damit gezeigte Interesse an der Arbeit.

Besonderer Dank gilt auch Herrn Dr.-Ing. Siegbert Drüe, der mir während meiner gesamten Tätigkeit seine uneingeschränkte Unterstützung in jeder Hinsicht zuteil werden ließ.

Meinen Kollegen des Fachgebietes Grundlagen der Elektrotechnik möchte ich für die freundliche Zusammenarbeit und ihre ständige Hilfsbereitschaft danken. Hervorheben möchte ich hierbei Ralf Stemmer, Jörg Thiem und Carsten Wolff, die mich durch das Korrekturlesen des Manuskriptes und viele wertvolle Anregungen beim Anfertigen der Arbeit unterstützt haben.

Ferner danke ich den studentischen Hilfskräften sowie allen Diplomanden, die durch ihr engagiertes Mitwirken wichtige Beiträge zur Arbeit geleistet haben.

Paderborn, im September 1998

Ralph Trapp

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlegende Aspekte und Strategien in der Stereoskopie	5
2.1	Das Prinzip der passiven Triangulation.....	5
2.2	Das Korrespondenzproblem	7
2.3	Das Blendenproblem.....	8
2.4	Einschränkungen in der Stereoskopie	9
2.4.1	Die Epipolareinschränkung	10
2.4.2	Die Eindeutigkeitseinschränkung	11
2.4.3	Die Kontinuitätseinschränkung	11
2.4.4	Die Ordnungseinschränkung	12
2.4.5	Die Disparitätsgradientengrenze	13
2.5	Strategien zur Korrespondenzbestimmung	14
2.5.1	Gebietsbasierte Verfahren	14
2.5.2	Merkmalsbasierte Verfahren	15
2.5.3	Phasenbasierte Verfahren.....	16
2.5.4	Regularisierungsmethoden.....	21
2.5.5	Dynamische Programmierung	22
2.5.6	Relaxationsverfahren	23
3	Der Ansatz von Reimann und Haken	25
3.1	Die Haken'sche Mustererkennungsgleichung.....	25
3.2	Anwendung auf die Stereoskopie	26
3.2.1	Ähnlichkeit von Bildregionen	26
3.2.2	Kontinuität durch Kopplung von Mustererkennungsgleichungen.....	27
4	Vom Kamerabild zur Tiefenkarte	31
4.1	Das aktive binokulare Sehsystem	31
4.1.1	Systemüberblick.....	32
4.1.2	Besonderheiten der Konstruktion.....	34
4.2	Die Kalibrierung des Kamerasystems	36
4.2.1	Das mathematische Modell	37
4.2.2	Bestimmung der Modellparameter.....	40
4.2.3	Einfluß der optischen Freiheitsgrade auf die Modellparameter.....	44

4.3	Entzerrung und Rektifikation der Bilddaten	45
4.3.1	Kompensation von Verzeichnungsfehlern	45
4.3.2	Rektifikation der Stereobilddaten	46
4.4	Bilddatendekomposition durch Gaborfilter	51
4.4.1	Vorüberlegungen zum Filterentwurf	52
4.4.2	Ortsfrequenzeinteilung	54
4.4.3	Orientierungseinteilung	57
4.4.4	Die Auflösungsrampe	61
4.5	Das Ähnlichkeitsmaß	63
4.5.1	Lokale Korrelation	63
4.5.2	Eigenschaften des Ähnlichkeitsmaßes	65
4.6	Korrespondenz durch Selbstorganisation	68
4.6.1	Initialisierung des Selbstorganisationsprozesses	69
4.6.2	Unterdrückung von Okklusionskorrespondenzen	71
4.6.3	Unterdrückung von Okklusionspaaren	74
4.6.4	Stabilitätsbetrachtungen	76
4.6.5	Das Prinzip des Wettbewerbs	82
4.6.6	Numerische Realisierungsaspekte	83
4.6.7	Parameterwahl	86
4.7	Subpixelgenaue Disparitätsschätzung	88
4.8	Rekonstruktion der Weltkoordinaten	89
4.8.1	Räumliche Rekonstruktion unter idealen Voraussetzungen	90
4.8.2	Kompensation der Projektionszentrumsdrift	91
4.9	Abschließende Betrachtungen	93
4.9.1	Verfahrensgrenzen	94
4.9.2	Einordnung des Verfahrens	95
5	Experimentelle Ergebnisse	97
5.1	Okklusionsdetektion am Beispiel eines Zufallspunktstereogramms	97
5.2	Dynamik der Variablen	99
5.3	Einfluß der Umfeldkopplung	100
5.4	Auswirkung der subpixelgenauen Disparitätsschätzung	101
5.5	Natürliche Stereogramme	103
5.6	Meßgenauigkeit bei der Positionsbestimmung	104

6	Demontage von Altautos - ein Anwendungsbeispiel	107
6.1	Lageerfassung des Rades	108
6.2	Detektion des Schraubenkranzes	109
6.3	Vermessung einzelner Schrauben	111
6.4	Laufzeiten und spezielle Lösungen	112
7	Zusammenfassung	113
8	Symbolverzeichnis	117
9	Literaturverzeichnis	121

1 Einleitung

Die fortschreitende Entwicklung der Computertechnologie und anderer technischer Disziplinen bietet zunehmend die Möglichkeit, komplizierte Aufgaben auf Automaten zu übertragen, um den Menschen auf diese Weise zu entlasten. Nahezu vollkommen automatisiert sind bei dem derzeitigen Stand der Technik Prozesse, die im wesentlichen aus periodisch wiederkehrenden Handlungsabläufen bestehen. Als Beispiel hierfür seien die Fertigungsstraßen in der Automobilindustrie genannt. Die immer gleichbleibende Situation in der Serienproduktion erlaubt es hier, die Montageaufgaben durch vordefinierte Bewegungen der Montageroboter zu bewältigen. Problematischer sind dahingegen Prozesse, die nicht vorhersehbaren Änderungen unterliegen. In solchen Situationen müssen diese Änderungen sensorisch erfaßt werden, damit der Automat entsprechend reagieren kann. Speziell in der Robotik sind die zu erfassenden Änderungen oft von räumlicher Natur. Dies ist zum Beispiel der Fall, wenn ein Robotersystem die Aufgabe hat, Objekte zu greifen oder zu manipulieren, deren relative Position in Bezug auf das Robotersystem nicht a priori bekannt ist. Um eine Interaktion des Roboters mit seiner Umgebung zu ermöglichen, ist es in diesem Zusammenhang notwendig, die räumliche Struktur der betreffenden Umgebung sensorisch zu erfassen. Eine ähnliche Situation herrscht beispielsweise bei der Navigation autonomer Fahrzeuge oder bei der sogenannten „Hand-Auge-Koordination“ in der Robotik. Weitere Anwendungsgebiete, in denen die Vermessung räumlicher Strukturen ebenfalls von zentraler Bedeutung ist, sind die Photogrammetrie oder die Oberflächenanalyse.

In Verbindung mit den obengenannten Aufgaben haben sich in der Praxis seit langem sogenannte aktive Verfahren bewährt. Als aktiv werden in diesem Zusammenhang Techniken bezeichnet, bei denen zur räumlichen Vermessung spezielle Strahlungsquellen eingesetzt werden. Die Reflexion der entsendeten Strahlung läßt bei diesen Verfahren Rückschlüsse auf die Position der bestrahlten Objekte zu. In diese Kategorie fallen sowohl Laufzeitmessungen mit Ultraschall oder Laserlicht als auch Triangulationsverfahren wie die Lasertriangulation oder Ansätze mit strukturiertem Licht [EY85, PA82]. Besteht die Aufgabe jedoch darin, selektiv Objekte in einer variablen Umgebung zu manipulieren, so ergibt sich die Notwendigkeit diese Objekte zunächst zu identifizieren. Abgesehen von einigen Triangulationsansätzen, bei denen die Empfangseinheit eine Kamera darstellt, lassen die Sensoreinheiten aktiver Verfahren im allgemeinen eine eindeutige Identifikation nicht zu. Des weiteren ist der Einsatz geeigneter Strahlungsquellen, wie z. B. Laserlicht, aus sicherheitstechnischen Gründen in der Praxis oft kritisch.

Sogenannte passive Verfahren, die allein die diffuse Umgebungsbeleuchtung ausnutzen, sind seit etwa drei Jahrzehnten Gegenstand intensiver Forschung. Das wachsende Interesse an diesen Verfahren ist nicht zuletzt darin begründet, daß auch die Natur im Zuge der Evolution passive Strategien hervorgebracht hat, die effizient zur Identifikation und Lokalisation eingesetzt werden. Als Beispiel sei an dieser Stelle das Sehsystem des Menschen genannt. Zahlreiche psychophysikalische Experimente bewiesen die Existenz mehrerer auf unterschiedlichen physikalischen Effekten beruhender Mechanismen, die der Mensch und auch andere höhere Lebe-

wesen zur Erfassung räumlicher Strukturen nutzen. Im Zuge der funktionellen technischen Nachbildung dieser Prozesse entstanden die sogenannten „Form durch X“-Techniken, bei denen die räumliche Struktur durch den Hinweis X rekonstruiert wird. Zu den bekanntesten Techniken dieser Art zählen: Form durch Stereoskopie, Bewegung, Fokussierung, Schattierung, Textur und Kontur.

Die Eingangsinformationen passiver Verfahren bilden im allgemeinen die Bilder oder Bildfolgen eines monokularen oder polyokularen Kamerasystems. Da die Tiefeninformation der betrachteten Szene durch die Projektion des dreidimensionalen Raumes auf die zweidimensionale Bildebene während der bildgebenden Prozesse einer Kamera verloren geht, ist die Rekonstruktion dieser Information die zentrale Aufgabe, die es bei diesen Verfahren zu lösen gilt. Die enormen Datenmengen, die bei diesem Problem verarbeitet werden müssen und die inhärente Unsicherheit, die mit der Rekonstruktion der Tiefeninformation verbunden ist, sind Gründe dafür, daß passive Verfahren zum gegenwärtigen Stand der Technik bei der Erfassung räumlicher Strukturen in der Praxis eine eher untergeordnete Rolle spielen.

Während alle anderen der obengenannten passiven Verfahren monokulare Techniken sind, ist beim Stereosehen ein binokularer oder polyokularer Sehapparat notwendig. Durch die aus unterschiedlichen Perspektiven aufgenommenen Bilder ist bei dieser Methode die Möglichkeit zu einer passiven Triangulation gegeben. Abgesehen von der sehr ungenauen Rekonstruktion, die mit Hilfe der Fokussierung durchgeführt werden kann, erlaubt nur diese passive Technik eine direkte quantitative Aussage über die räumliche Position der in der betrachteten Szene vorhandenen Strukturen. Auch in der Natur nimmt die Stereoskopie eine gewisse Sonderstellung ein. Obwohl das Zusammenwirken mehrerer der obengenannten passiven Mechanismen bei der räumlichen Wahrnehmung des Menschen eine wissenschaftlich gesicherte Tatsache ist, gelang es Julesz [Jul60] durch psychophysikalische Experimente auf eindrucksvolle Art nachzuweisen, daß der Mensch allein aufgrund stereoskopischer Informationen zur räumlichen Wahrnehmung befähigt ist.

In der Stereoskopie, die auch ein zentraler Bestandteil der vorliegenden Arbeit ist, wird die zu vermessende Szene mit Hilfe von Kameras aus zwei oder mehreren verschiedenen Perspektiven aufgenommen. Aufgrund der unterschiedlichen Positionen, auf die ein Raumpunkt in die Bildebenen projiziert wird, kann die Position des Punktes im Raum rekonstruiert werden. Im Gegensatz zu aktiven Triangulationsmethoden ist das zur Rekonstruktion notwendige Dreieck in der Stereoskopie nicht a priori bekannt, sondern es muß vielmehr aus der Bildinformation ermittelt werden. Zu diesem Zweck muß eine eindeutige Korrespondenz zwischen den Punkten der einzelnen Bilder hergestellt werden. Da die Zuordnung der Bildpunkte zum einen stark mehrdeutig sein kann und zum anderen nicht immer möglich ist, wird dieses sogenannte Korrespondenzproblem als ein nach der Hadamard'schen Definition „schlecht gestelltes“ Problem bezeichnet [BPT88, PTK85]. Die Lösung des Korrespondenzproblems wird darüber hinaus durch perspektivische Verzerrungen, Rauschprozesse, Beleuchtungs- und Kontrastunterschiede zwischen den Bildern zusätzlich erschwert. Diese für passive Rekonstruktionstechniken charakteristischen Probleme können im allgemeinen nur unter Ausnutzung geeigneten Vorwissens bewältigt werden.

Beginnend mit dem Ansatz von Marr und Poggio [MP76] entwickelte sich in den letzten zwei Jahrzehnten eine nahezu unübersehbare Vielzahl von unterschiedlichen stereoskopischen Ansätzen. Diese Verfahren können hinsichtlich der verwendeten Kamerageometrie, der Merkmale, die zur Korrespondenzbestimmung herangezogen werden, und der Methoden zur Auflösung der Mehrdeutigkeiten unterschieden werden. In der Literatur sind Konstruktionen mit den klassischen zwei bis zu vier Kamerasystemen zu finden, die entweder in einer statischen Konstellation betrieben werden oder aktiv ihre relative Lage verändern können. Weiterhin wird zwischen gebietsbasierten, merkmalsbasierten und phasenbasierten Stereoverfahren unterschieden. Um Korrespondenzen zwischen den Stereobildern zu bestimmen, werden bei gebietsbasierten Ansätzen lokale Gebiete der Grauwertbilder miteinander korreliert, wohingegen merkmalsbasierte Verfahren zunächst Merkmale aus den Grauwertbildern extrahieren, die anschließend einander zugeordnet werden. Wird die relative Verschiebung korrespondierender Bildbereiche, die sogenannte Disparität, mit Hilfe der Phaseninformation in den Antworten von Quadraturfilterpaaren ermittelt, so wird von phasenbasierten Verfahren gesprochen. Bei der Ausnutzung von Vorwissen zur Auflösung der Mehrdeutigkeiten kommen im wesentlichen Relaxationstechniken, Regularisierungsmethoden, stochastische Optimierung und Methoden der dynamischen Programmierung zum Einsatz.

Eine der bedeutendsten Fehlerursachen in der stereoskopischen Korrespondenzbestimmung stellen Bereiche in einer betrachteten Szene dar, die nur aus einer Kameraperspektive zu sehen sind. Zu den Bildbereichen, in die diese Regionen der Szene abgebildet werden, existieren keine Korrespondenzpartner im anderen Stereobild. Diese Bildbereiche werden als Okklusionen bezeichnet. Werden Okklusionen bei der Korrespondenzbestimmung nicht geeignet berücksichtigt, so kommt es je nach Ansatz zu mehr oder weniger starken Fehlkorrespondenzen und damit zu einer fehlerbehafteten Tiefenrekonstruktion. Okklusionen stellen daher ein schwerwiegendes Problem in der Stereoskopie dar, das erstaunlicherweise nur in sehr wenigen Ansätzen eine explizite Behandlung erfährt. Die gemeinsame Basis der Verfahren, die Okklusionen berücksichtigen, bildet ein dualer oder bidirektionaler Zuordnungsprozeß [GLY92, HA89, JM92, LB95, WAH88]. Auf eine berechnungsintensive Weise werden in diesen Ansätzen die Bilddaten des linken Stereokanals denen des rechten und zusätzlich in umgekehrter Richtung zugeordnet. Dies geschieht in zwei identischen, aber weitestgehend unabhängigen Prozessen.

Gegenstand der vorliegenden Arbeit ist ein Gesamtsystem, das die mechanisch-geometrischen Voraussetzungen, die Bildverarbeitungstechniken und die Verfahren zur stereoskopischen Korrespondenzbestimmung umfaßt, die notwendig sind, um die räumlichen Struktur einer betrachteten Szene anhand von Stereobildern zu rekonstruieren. Auf der mechanischen Seite behandelt die Arbeit den Aufbau und die Kalibrierung eines binokularen Sehsystems, das es erlaubt, zwei Kameras in vier rotatorischen Freiheitsgraden zu bewegen. Mit Hilfe dieser Freiheitsgrade können die Kameras auf nahezu beliebig positionierte Raumpunkte ausgerichtet werden. Das aktive Sehsystem, das nicht zuletzt wegen seines modularen und kostengünstigen Aufbaus bereits an der Universität Hamburg mit Erfolg nachgebaut wurde, ermöglicht einen effizienten Einsatz des stereoskopischen Verfahrens. Weiterhin behandelt die vorliegende Arbeit die Vorverarbeitung der Bilddaten in Form einer Filterbank, die einerseits die Grund-

lage zur Merkmalsextraktion im Rahmen eines Mustererkennungssystems bildet und andererseits die Eingangsinformation für die Korrespondenzbestimmung liefert. Den zentralen Bestandteil der Arbeit bildet das Verfahren zur Korrespondenzbestimmung, das im wesentlichen aus zwei Verarbeitungsschritten besteht. Zunächst erfolgt die Berechnung eines komplexwertigen Ähnlichkeitsmaßes, das die Übereinstimmung des linken und rechten gefilterten Bildsignals widerspiegelt. Dieses Ähnlichkeitsmaß dient anschließend zur Initialisierung eines auf der Arbeit von Reimann und Haken [RH94] basierenden Selbstorganisationsprozesses, der unter Ausnutzung geeigneten Vorwissens die Korrespondenzen zwischen den Stereobildern errechnet. Eine besondere Eigenschaft dieses Verfahren stellt die implizite Detektion von Okklusionen dar. Spezielle Kopplungsstrukturen innerhalb des Selbstorganisationprozesses verhindern hierbei Korrespondenzen mit okkludierten Bildbereichen, ohne daß ein aufwendiger bidirektionaler Zuordnungsprozeß notwendig ist.

Nachdem die mit der Stereoskopie verbundenen Probleme kurz dargestellt worden sind, erfolgt in Kapitel zwei eine Beschreibung der Einschränkungen, die als a priori Wissen zur Lösung des Korrespondenzproblems im allgemeinen herangezogen werden und denen daher in der Stereoskopie eine zentrale Bedeutung zukommt. Die unterschiedlichen Verfahren zur stereoskopischen Korrespondenzbestimmung und deren Taxonomie sind ebenfalls Gegenstand dieses Kapitels. Besonderen Stellenwert haben hierbei phasenbasierte Verfahren, da das in dieser Arbeit behandelte Verfahren teilweise eine gewisse Verwandtschaft zu diesen Techniken aufweist. Ein Relaxationsverfahren, das die Grundlage des in dieser Arbeit behandelten Selbstorganisationsprozesses liefert, ist der im dritten Kapitel ausführlich beschriebene Ansatz von Reimann und Haken. Das auf der Haken'schen Mustererkennungsgleichung [Hak79, Hak87, Hak91] beruhende Verfahren ermöglicht eine effiziente, leicht parallelisierbare Lösung des Korrespondenzproblems. Das im Rahmen dieser Arbeit entwickelte aktive Sehsystem, seine Kalibrierung und die einzelnen Verfahrensschritte der Korrespondenzbestimmung sind Gegenstand des vierten Kapitels. Im Vordergrund steht hierbei eine Erweiterung des Relaxationsansatzes von Reimann und Haken um eine effiziente Strategie zur Detektion der in der Szene vorhandenen Okklusionen. Die Leistungsfähigkeit des Verfahrens wird im fünften Kapitel sowohl an synthetischen als auch an realen Stereogrammen demonstriert. Weiterhin wird in diesem Kapitel die Genauigkeit des Gesamtsystems bei der räumlichen Rekonstruktion diskutiert. Eine konkrete Anwendung in Form eines Demontageprozesses, bei dem die Bewegungen eines Roboterarmes koordiniert werden müssen, wird in Kapitel sechs gegeben, ehe im siebten Kapitel die wesentlichen Aspekte der Arbeit zusammengefaßt werden.

2 Grundlegende Aspekte und Strategien in der Stereoskopie

Die verschiedenen Faktoren, die in der Stereoskopie zum Tragen kommen, angefangen bei der Kamerageometrie über die Bildverarbeitungstechniken bis hin zu modellbasierten Optimierungsverfahren, bieten Raum für die unterschiedlichsten Lösungsansätze. Diese Tatsache, aber auch die rasante Entwicklung in der Video- und der Computertechnologie, sind Gründe für die Vielfalt der bislang entwickelten Verfahren. Da selbst eine detaillierte Beschreibung der wesentlichen Verfahren den Rahmen der Arbeit sprengen würde, wird in diesem Kapitel nur eine Übersicht und eine Bewertung hinsichtlich der prinzipiellen Eigenschaften einzelner Verfahrensklassen gegeben, nachdem kurz das Prinzip der passiven Triangulation und die für diese Technik charakteristische Problematik erörtert wurden. Für eine ausführliche, wenn auch nicht mehr ganz aktuelle Übersicht über stereoskopische Verfahren sei der Leser an [DA89] verwiesen.

2.1 Das Prinzip der passiven Triangulation

Zur stereoskopischen Positionsbestimmung werden im allgemeinen Kamerasysteme eingesetzt, welche die elektromagnetische Strahlungsintensität im Bereich des sichtbaren Spektrums erfassen. Durch die physikalischen Prozesse der Strahlenreflexion wird eine dreidimensionale Szene, die sich aus mehreren komplexen Objekten zusammensetzen kann, mit Hilfe eines solchen Kamerasystems in eine zweidimensionale Strahlungsintensitätsverteilung umgesetzt. Zur digitalen Weiterverarbeitung wird diese Verteilung durch einen Sensor in elektrische Signale umgewandelt und sowohl im Ort als auch in der Amplitude quantisiert. Die sich auf diese Weise ergebende werte- und ortsdiskrete Intensitätsverteilung wird im allgemeinen als Grauwertbild bezeichnet.

Die bildgebenden Prozesse in einer Kamera bewirken, daß die Tiefeninformation der betrachteten Szene verlorengeht. Die Rekonstruktion dieser Information ist die zentrale Aufgabe in der Stereoskopie und erfolgt durch eine sogenannte passive Triangulation. Anstelle des bei aktiven Verfahren eingesetzten Senders tritt in der Stereoskopie eine zweite Kamera. Mit Hilfe der beiden räumlich versetzt angeordneten Kameras wird die zu analysierende Szene aus unterschiedlichen Perspektiven aufgenommen. Die Entfernung zwischen einem Raumpunkt und dem Kamerasystem kann so aus der Differenz der korrespondierenden Positionen, auf die der Raumpunkt in die Bildebenen projiziert wird, rekonstruiert werden. Die Differenz korrespondierender Positionen im linken und rechten Stereobild wird als Disparität oder auch Parallaxe bezeichnet. In Bild 2-1 wird das Prinzip der passiven Triangulation am Beispiel einer Stereogeometrie, bei der die optischen Achsen der beiden Kamerasysteme parallel ausgerichtet sind, verdeutlicht. Der Abbildungsprozeß wird hier durch die Zentralprojektion des klassischen Lochkameramodells beschrieben [DH73, GW93, Jäh93]. Zur Veranschaulichung wird die Triangulation im folgenden am zweidimensionalen Beispiel erläutert.

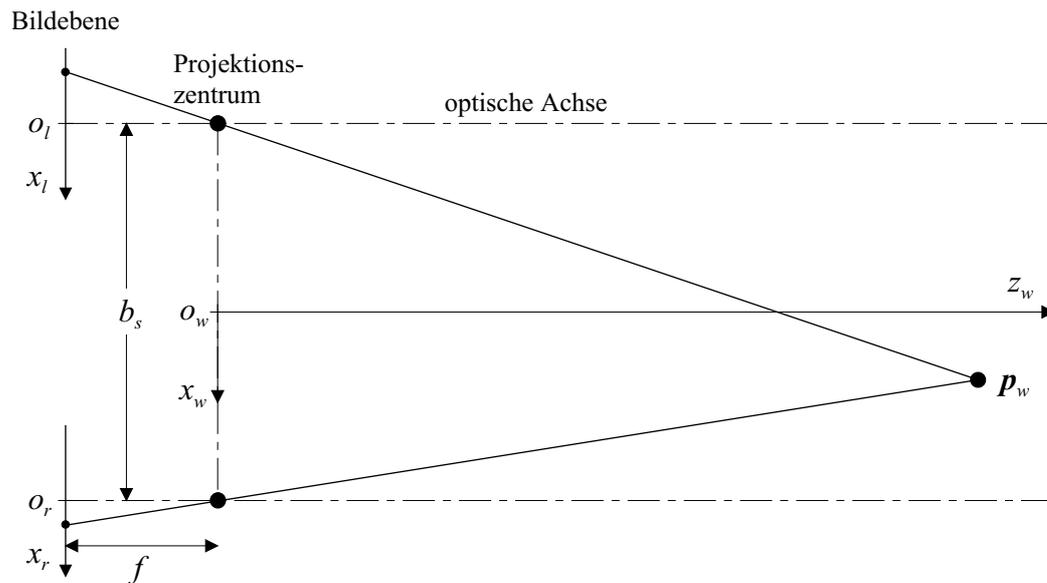


Bild 2-1: Stereogeometrie mit parallelen optischen Achsen am Beispiel des Lochkameramodells

Die Koordinaten (x_w, z_w) des Punktes p_w beziehen sich hierbei auf ein Weltkoordinatensystem, dessen Ursprung sich in der Mitte der beiden Projektionszentren befindet. Der Abstand der beiden Projektionszentren wird als Stereobasis b_s bezeichnet. Die Bildebenen sind in einem Abstand f , der sogenannten effektiven Brennweite, von den Projektionszentren parallel zur Stereobasis angeordnet. Vereinfachend soll hier angenommen werden, daß die effektiven Brennweiten der beiden Kameras identisch sind. Die in diesem Fall eindimensionalen Bildkoordinaten x , auf die der betrachtete Raumpunkt p_w im linken und rechten Bild projiziert wird, ergeben sich nach der Abbildungsgleichung der Zentralprojektion zu:

$$x_l = -\frac{f\left(x_w + \frac{b_s}{2}\right)}{z_w}; \quad x_r = -\frac{f\left(x_w - \frac{b_s}{2}\right)}{z_w} \quad (2-1)$$

Durch einfache Umformungen können die Weltkoordinaten des Punktes p_w aus diesen Beziehungen rekonstruiert werden:

$$x_w = \frac{b_s(x_l + x_r)}{2d_x}; \quad z_w = -\frac{f b_s}{d_x} \quad \text{mit} \quad d_x = x_l - x_r, \quad (2-2)$$

wobei d_x die Disparität der Bildkoordinaten ist, auf die der Punkt p_w projiziert wird.

Obwohl das Prinzip der passiven Triangulation bei Kenntnis der korrespondierenden Bildpunkte von einfacher Natur ist, stellt die praktische Umsetzung dieser Technik ein nicht unerhebliches Problem dar. Ein Grund hierfür ist, daß die Kamerageometrie in den seltensten Fällen vollständig bekannt ist. Um die Kameras in eine definierte Position und Orientierung bringen zu können, müssen daher zunächst die intrinsischen Kameraparameter, welche die Abbildungsvorgänge beschreiben, und die extrinsischen Kameraparameter, welche die Position des Projektionszentrums und die Orientierung der optischen Achse definieren, ermittelt werden. Dies geschieht im allgemeinen durch eine Kalibrierung. Bei dieser Technik werden spezielle Meßpunkte, deren Positionen im Raum und deren zugehörige Bildpositionen, auf die sie projiziert werden, genau bekannt sind, analysiert, um die gesuchten Parameter zu bestimmen. Die Parameter müssen hierbei aus einem nichtlinearen Gleichungssystem, das durch die sogenannten Kollinearitätsgleichungen aufgestellt wird, ermittelt werden [AK89, Kra84]. Bei einfachen Kameramodellen kann dieses Gleichungssystem zwar geschlossen gelöst werden (siehe z.B. [FT86, Len87, Tsa87]), jedoch ist insbesondere bei komplexeren Modellen eine genaue Ermittlung der Parameter nur durch eine aufwendige nichtlineare Optimierung möglich [Gen79, Li94, WCH92]. Aufgrund des nichtlinearen Modells müssen darüber hinaus diese Optimierungsalgorithmen geeignet initialisiert werden, da sonst die Gefahr besteht, daß die Verfahren zu lokalen Minima im jeweils verwendeten Kostenfunktional konvergieren. Liegen keine Herstellerangaben bezüglich der Kameraparameter vor, so können gewisse Initialisierungswerte mit Hilfe von Zoomingverfahren oder Fluchtpunktberechnungen ermittelt werden [CT90, Ech90, LT87, Li94].

2.2 Das Korrespondenzproblem

Zur Rekonstruktion der Tiefeninformation ist es in der Stereoskopie erforderlich, das zur Triangulation notwendige Dreieck aus der Bildinformation zu ermitteln. Zu diesem Zweck müssen die Bildpunkte oder Bildregionen der beiden Stereokanäle in eindeutiger Weise einander zugeordnet werden. Die unvollständige Information in den beiden Bildern über den abgebildeten Raum führt dazu, daß die Zuordnung korrespondierender Bildpunkte im allgemeinen ein im Hadamard'schen Sinn „schlecht gestelltes“ Problem ist [BPT88]. Dies bedeutet, daß die Lösung des Problems entweder nicht eindeutig ist, überhaupt nicht existiert oder nicht kontinuierlich von den Meßdaten abhängt. Da in beiden Bildern durchaus mehrere ähnliche oder sogar periodische Strukturen auftreten können, ist die Lösung dieses sogenannten Korrespondenzproblems im allgemeinen stark mehrdeutig.

Eine oft angewandte Strategie, um die Mehrdeutigkeit des Korrespondenzproblems zu reduzieren, besteht in dem Vergleich von komplexeren Bildstrukturen anstelle der einzelnen Bildpunkte. Zwar kann dieses Vorgehen, das in Kapitel 2.5.2 noch näher erläutert wird, das Korrespondenzproblem je nach Komplexität der verwendeten Bildstrukturen mehr oder weniger vereinfachen, aber eine eindeutige Zuordnung kann auch durch diese Strategie nicht immer gewährleistet werden (vgl. Bild 2-2).

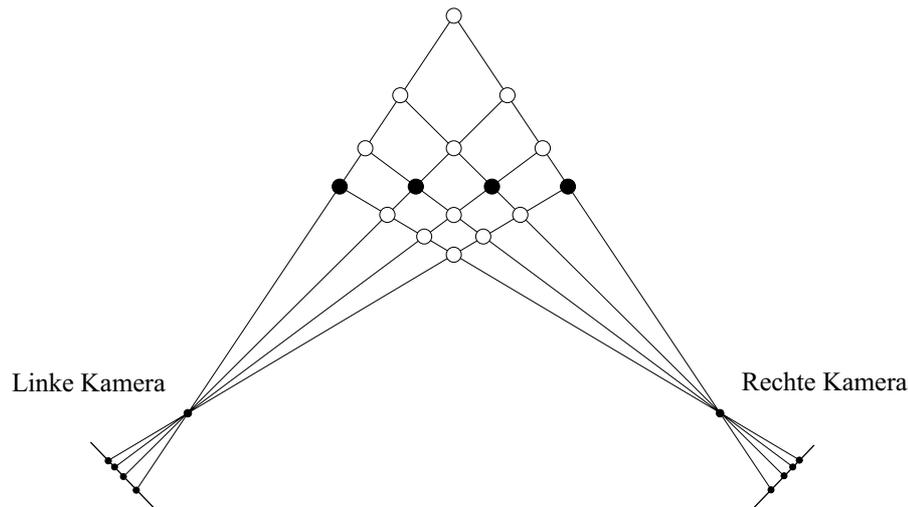


Bild 2-2: Mehrdeutigkeit bei der Zuordnung von korrespondierenden Bildpunkten. Vier identische Objekte (ausgefüllte Kreise) werden in die beiden Bildebenen projiziert. Jeder der vier Projektionen des einen Kamerabildes kann jede der Projektionen im anderen Kamerabild zugeordnet werden. Die entsprechenden Positionsrekonstruktionen ergeben vier korrekte (ausgefüllte Kreise) und zwölf falsche (leere Kreise) Lösungen (in Anlehnung an [Jul71]).

Ein weiteres spezifisches Problem stellen Bereiche in der zu analysierenden Szene dar, die nur aus einer der beiden Perspektiven gesehen werden können. Diese sogenannten Okklusionen führen bei den meisten existierenden Stereoverfahren zu erheblichen Fehlern in der Korrespondenzbestimmung und damit in der Objektlokalisierung, da in diesen Bereichen das Korrespondenzproblem keine Lösung besitzt. Die Zuordnung korrespondierender Bildpunkte wird darüber hinaus durch interokuläre Differenzen wie perspektivische Verzerrungen, Beleuchtungs- und Kontrastunterschiede zwischen den Bildern und Rauschprozesse zusätzlich erschwert.

Um dennoch eine eindeutige Lösung zu erhalten, wird in der Stereoskopie wie auch bei anderen schlecht gestellten Problemen im allgemeinen Vorwissen über die physikalische Natur des Problems zur Einschränkung des Lösungsraumes genutzt [BPT88, MMP87, PTK85] (siehe Kapitel 2.4).

2.3 Das Blendenproblem

Bei einer Stereogeometrie mit parallelen optischen Achsen, wie sie in Bild 2-1 zu sehen ist, ist die Verschiebung korrespondierender Bildpunkte in einem Stereobildpaar immer parallel zur Stereobasis gerichtet. Bei einer genauen Kenntnis der Kamerageometrie kann aufgrund dieser Tatsache die Richtung der Disparität im Vorfeld ermittelt und somit die Suche nach korrespondierenden Bildpunkten erheblich erleichtert werden. Problematisch sind jedoch Bildregionen,

in denen keine Strukturen oder Intensitätsänderungen in Richtung der Stereobasis auftreten. In diesem Fall kann eine Verschiebung korrespondierender Bildpunkte nicht festgestellt werden (vgl. Bild 2-3). Da die Detektion der Verschiebung in der Stereoskopie meist durch einen lokalen Operator erfolgt, wobei der übrige Teil der Szene ausgeblendet wird, wird dieses Problem auch als ein Spezialfall des sogenannten Blendenproblems betrachtet, dem in der Bewegungsanalyse eine besondere Bedeutung zukommt (siehe z.B. [Jäh93]).

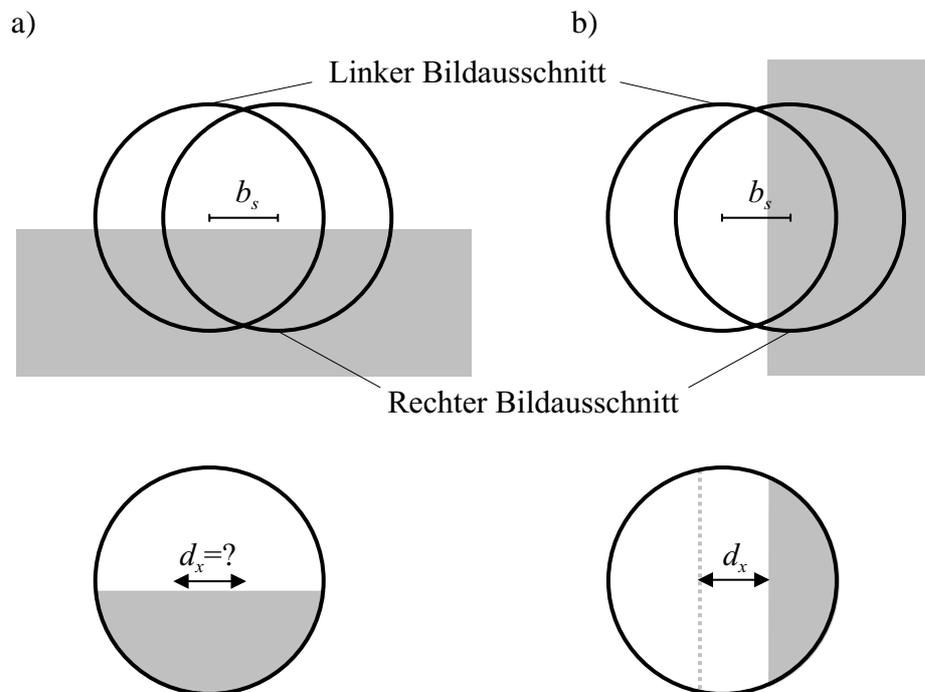


Bild 2-3: Schematische Darstellung des Blendenproblems in der Stereoskopie: a) Keine eindeutige Aussage bei einer parallel zur Stereobasis gerichteten Grauwertkante. b) Eindeutige Verschiebung in den Stereobildern bei einer Grauwertkante, die orthogonal zur Stereobasis gerichtet ist

Eine Lösung dieses Problems kann durch den Einsatz eines oder mehrerer zusätzlicher Kamerasysteme erzielt werden. Sind die Basisstrecken eines solchen Systems nicht parallel zueinander angeordnet, so können alle Grauwertänderungen zur Tiefenrekonstruktion herangezogen werden [AL87, HAL90, II86, PH86, SP96].

2.4 Einschränkungen in der Stereoskopie

Bedingt durch seine spezielle Natur kann das Korrespondenzproblem, ebenso wie viele andere schlecht gestellte Probleme, nur unter Ausnutzung geeigneten Vorwissens eindeutig gelöst werden. Mit Hilfe dieses Vorwissens wird der Lösungsraum entsprechend eingeschränkt und das Problem in ein gut gestelltes Problem umgewandelt. In den folgenden Unterkapiteln werden einige dieser Einschränkungen, die eine besondere Relevanz für die vorliegende Arbeit zeigen, näher erläutert.

2.4.1 Die Epipolareinschränkung

Die Epipolareinschränkung ist eine Einschränkung des Suchraumes, in dem der Korrespondenzpartner eines bestimmten Bildpunktes im jeweils anderen Bild zu finden ist. Ist die Geometrie der Kamerasysteme hinreichend genau bekannt, so muß der Korrespondenzpartner nicht in der gesamten Bildebene des anderen Bildes gesucht werden, sondern nur entlang einer sogenannten Epipolarlinie. Im Gegensatz zu einer Stereogeometrie mit parallelen optischen Achsen verläuft diese Linie im allgemeinen Fall nicht parallel zur Stereobasis. Die Epipolarlinie ist vielmehr die Schnittlinie der sogenannten Epipolarebene, die durch die zwei optischen Zentren der Kameras und den betrachteten Raumpunkt definiert wird, mit der betreffenden Bildebene.

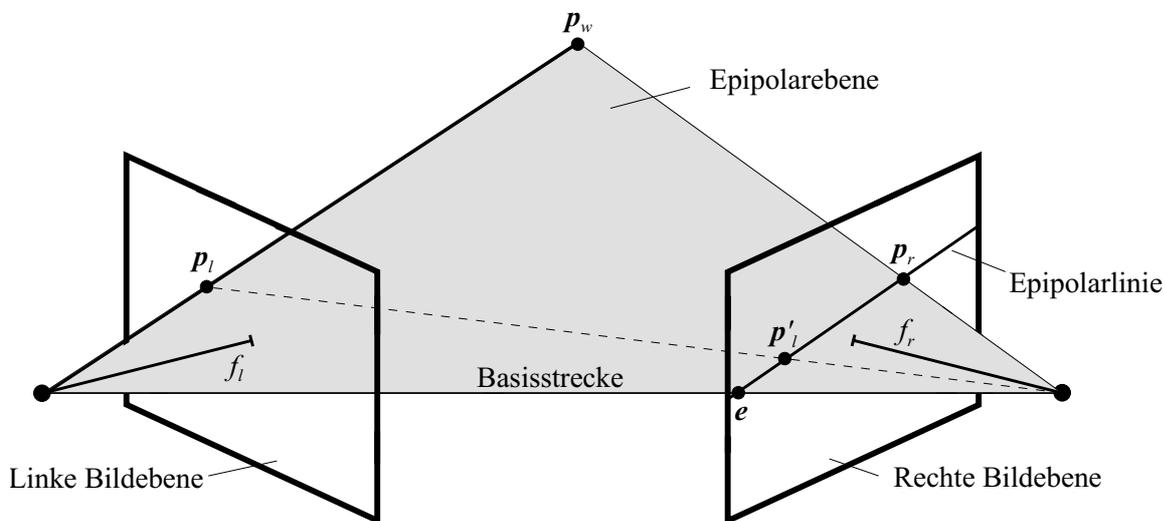


Bild 2-4: Graphische Konstruktion von Epipolarlinien bei nichtparallelen optischen Achsen. Die zum linken Bildpunkt p_l korrespondierende Epipolarlinie im rechten Bild wird durch die Projektion p'_l des betreffenden Bildpunktes in die rechte Bildebene und den Epipol e definiert. Zur Veranschaulichung sind die Bildebenen in diesem Beispiel vor den Projektionszentren angeordnet.

Sind die optischen Achsen der beiden Kameras exakt parallel zueinander und die Bildzeilen parallel zur Stereobasis ausgerichtet, ein Zustand, der in der Praxis schwer realisierbar ist aber dennoch bei vielen theoretischen Ansätzen vorausgesetzt wird (siehe z.B. [HSY96, KA87, LB95]), so sind die Epipolarlinien in den beiden Bildebenen kollinear und identisch mit den Bildzeilen.

Bei Kamerasystemen, bei denen die optischen Achsen zur Vergrößerung des Überlappungsbereiches der beiden Stereobilder aktiv ausgerichtet werden können, stellt dieser Zustand jedoch nur einen Ausnahmefall dar. Die nichtparallelen optischen Achsen führen bei diesen Architekturen zu vertikalen Disparitätskomponenten, die ungleich Null sind und von der Position in der

Bildebene abhängen. Im Zuge der Korrespondenzbestimmung muß in diesem Fall somit entweder eine berechnungsintensive zweidimensionale Suche nach korrespondierenden Bildpunkten durchgeführt werden, oder die Epipolarlinien müssen für jeden Bildpunkt neu geschätzt werden, um so das Problem wieder auf eine eindimensionale Suche zu reduzieren [BB85, Ols92, YZ88].

Eine aufwandsgünstigere Möglichkeit, die Suche nach Korrespondenzpartnern zu vereinfachen, besteht darin, die Bildinformation der beiden Bilder in eine gemeinsame Ebene, die sogenannte Rektifikationsebene, zu projizieren. Durch diese Transformation kann eine nicht-parallele Stereokonfiguration in eine parallele Anordnung überführt werden [AH88, GG93, PD96].

2.4.2 Die Eindeutigkeitseinschränkung

Eine der beiden Einschränkungen, die zuerst von Marr und Poggio [MP76] postuliert wurden, ist die Eindeutigkeitseinschränkung. Sie basiert auf der physikalischen Eigenschaft, daß ein bestimmter Punkt auf einer Oberfläche, die ein Objekt begrenzt, eine eindeutige Position in Raum und Zeit aufweist. Diese physikalische Einschränkung formulierten Marr und Poggio für die Korrespondenzbestimmung in Stereobilddaten in Form der folgenden Regel:

Jedem Merkmal in den beiden Bildern kann maximal eine Disparität zugeordnet werden.

In der Praxis kann diese Regel nicht uneingeschränkt auf die Grauwerte der Stereobilder angewendet werden. Konflikte können insbesondere dann auftreten, wenn in der Szene transparente Oberflächen vorhanden sind. Um dieses Problem zu umgehen, werden in vielen Verfahren zunächst eindeutig lokalisierbare Merkmale aus den Stereobildpaaren extrahiert, auf die dann, wie auch schon von Marr und Poggio vorgeschlagen [MP79], die Eindeutigkeitseinschränkung angewendet wird. Ein Kantenmerkmal kann beispielsweise nicht auf einer transparenten Oberfläche liegen.

2.4.3 Die Kontinuitätseinschränkung

Die Kontinuitätseinschränkung ist die zweite von Marr und Poggio postulierte Einschränkung [MP76]. Sie beruht auf der Tatsache, daß Materie kohärent ist. Sie kann in einzelne Objekte, die durch ihre Oberflächen begrenzt sind, unterteilt werden. Weiterhin sind diese Objektoberflächen im allgemeinen glatt im Vergleich zu ihrer Entfernung zum Beobachter. Die Regel, die Marr und Poggio aus diesen physikalischen Eigenschaften ableiteten, lautet:

Der Disparitätsverlauf ist fast überall stetig. Nur in einem kleinen Bruchteil der Bildfläche treten Objektbegrenzungen und damit Diskontinuitäten in der Tiefe auf, die in Unstetigkeitsstellen im Disparitätsverlauf resultieren.

2.4.4 Die Ordnungseinschränkung

Der sogenannten Ordnungs- oder Monotonieeinschränkung liegt die Annahme zugrunde, daß die Abfolge, in der Objekte in die Bildebenen abgebildet werden, im linken und rechten Bild identisch ist [BB81, Fau93, OK85]. Wie die nachfolgende Abbildung zeigt, hat diese Einschränkung jedoch nur begrenzte Gültigkeit.

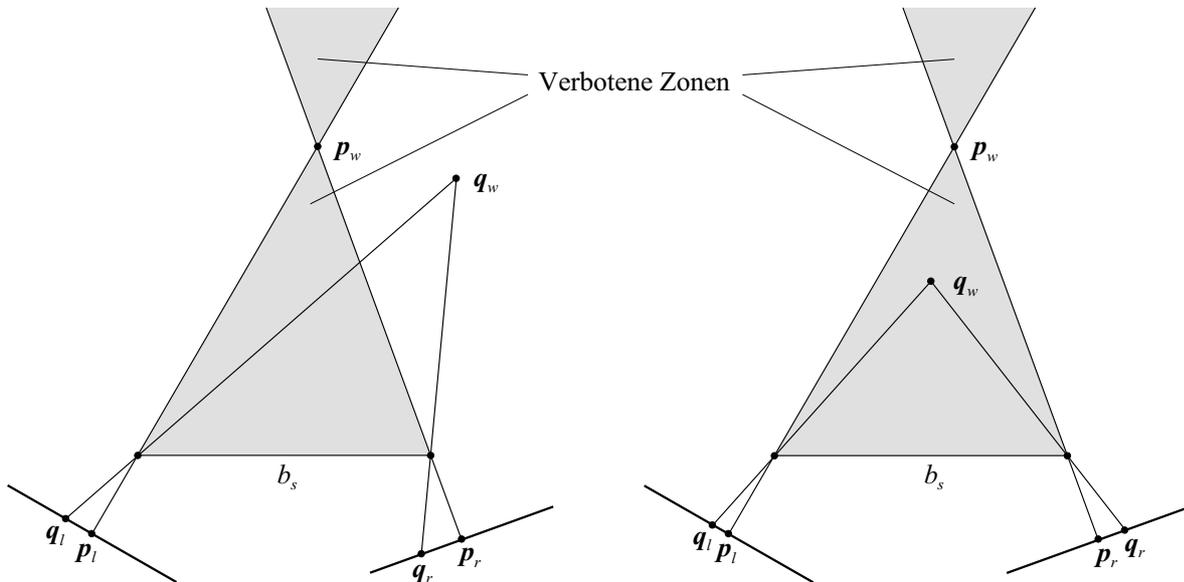


Bild 2-5: Verbotene Zonen bei der Ordnungseinschränkung (Erläuterung siehe Text)

Liegt der Punkt q_w , wie in Bild 2-5 im linken Beispiel zu sehen ist, bei einem bereits zugeordneten Korrespondenzpaar (p_l, p_r) außerhalb der grau unterlegten Bereiche, so ist die Reihenfolge der Projektionen von links nach rechts in beiden Bildern identisch und die Ordnungseinschränkung ist erfüllt. Eine Verletzung der Ordnungseinschränkung liegt im rechten Beispiel in Bild 2-5 vor. Der Punkt q_w befindet sich hier innerhalb eines grau unterlegten Bereiches und die Reihenfolge der Projektionen im linken und im rechten Bild ist umgekehrt. Die Korrespondenzpaare (p_l, p_r) und (q_l, q_r) schließen sich in diesem Fall gegenseitig aus. Die grau unterlegten Bereiche werden daher auch als „verbotene Zonen“ bezeichnet.

Obwohl es durchaus reale Szenen gibt, in denen die Ordnungseinschränkung verletzt wird, existieren zwei Argumente, die für diese Einschränkung sprechen [Fau93]:

1. Der Abstand der betrachteten Punkte zu den Kameras ist im allgemeinen groß im Vergleich zur Basisstrecke b_s , wodurch die verbotenen Zonen sehr klein werden.
2. Die Ordnungseinschränkung kann genutzt werden, um Okklusionen zu detektieren, da Punkte, die nur aus einer Kameraperspektive zu sehen sind, generell in den verbotenen Zonen korrekter Korrespondenzpaare liegen.

2.4.5 Die Disparitätsgradientengrenze

Die Einschränkung des sogenannten Disparitätsgradienten wurde aus psychophysischen Untersuchungen von Burt und Julesz [BJ80] abgeleitet. Im Rahmen dieser Untersuchungen konnte festgestellt werden, daß das menschliche Sehsystem nur dann zu einer binokularen Fusion fähig ist, wenn der Disparitätsgradient den Wert 1,0 nicht überschreitet. Zur approximativen Berechnung des Disparitätsgradienten wird im allgemeinen eine virtuelle Kamera eingeführt, die in der Mitte der beiden Kameras angeordnet ist. Diese Kamera wird auch als zyklische Kamera bezeichnet [Jul71].

Seien \mathbf{p}_w und \mathbf{q}_w zwei Raumpunkte und $\mathbf{p}_{l/r} = [x_p \ y_p]_{l/r}^T$ bzw. $\mathbf{q}_{l/r} = [x_q \ y_q]_{l/r}^T$ die zugehörigen Bildpunkte im linken bzw. rechten Stereokanal eines binokularen Kamerasystems mit parallelen optischen Achsen, so ergeben sich die korrespondierenden zyklischen Bildkoordinaten zu:

$$\mathbf{p}_c = \frac{\mathbf{p}_l + \mathbf{p}_r}{2} \quad \text{bzw.} \quad \mathbf{q}_c = \frac{\mathbf{q}_l + \mathbf{q}_r}{2} \quad (2-3)$$

Sind weiterhin

$$d_{xp} = x_{pl} - x_{pr} \quad \text{und} \quad d_{xq} = x_{ql} - x_{qr} \quad (2-4)$$

jeweils die horizontalen Disparitäten der Bildpunkte, so kann der Disparitätsgradient durch die dimensionslose Größe

$$\nabla d_{pq} = \frac{|d_{xp} - d_{xq}|}{|\mathbf{p}_c - \mathbf{q}_c|} \quad (2-5)$$

approximiert werden. Wird nun für alle Bildpunktombinationen (i, j) mit $i \neq j$ eines Bildes ein Disparitätsgradient mit $\nabla d_{ij} < \nabla d_g$ gefordert, so ergeben sich ähnlich wie bei der Ordnungseinschränkung verbotene Zonen um jeden rekonstruierten Raumpunkt, in denen sich kein weiterer Raumpunkt befinden darf (siehe z.B. [Fau93, LH96, PMF85, PP+85]).

Diese Einschränkung der möglichen Korrespondenzen ist allgemeiner als die Ordnungseinschränkung aus Kapitel 2.4.4 und kann je nach Disparitätsgradientengrenze die Eindeutigkeits-einschränkung einschließen. Die Disparitätsgradientengrenze geht für $\nabla d_g = 2$ in die Ordnungseinschränkung über [Fau93].

2.5 Strategien zur Korrespondenzbestimmung

Die Disparitätsmessung und die damit verbundene Korrespondenzbestimmung in der Stereoskopie kann, wie die Vielzahl der verschiedenen Ansätze zeigt, auf unterschiedliche Art und Weise erfolgen. Obwohl auf eine detaillierte Beschreibung der einzelnen Lösungsansätze hier aus Übersichtsgründen verzichtet wird, sollen dennoch die Prinzipien und die grundlegenden Eigenschaften der einzelnen Lösungsstrategien in diesem Abschnitt dargestellt werden, um das in dieser Arbeit vorgestellte Verfahren zur Korrespondenzbestimmung taxonomisch einordnen und von den existierenden stereoskopischen Verfahren abgrenzen zu können. Hierbei handelt es sich keineswegs um eine vollständige Auflistung aller existierenden Strategien, sondern es werden nur solche betrachtet, die eine gewisse Relevanz in der Stereoskopie oder eine Verwandtschaft zu dem in dieser Arbeit vorgestellten Verfahren aufweisen.

2.5.1 Gebietsbasierte Verfahren

Bei sogenannten gebietsbasierten Verfahren erfolgt die Zuordnung einzelner Bildbereiche der Stereobilder anhand der Grauwerte. Zur Bestimmung eines Ähnlichkeitsmaßes werden die Grauwerte der Stereobilder jeweils in einem begrenzten Gebiet miteinander verrechnet [BT80, CM92, Mor77, YC78]. Die Disparität ergibt sich im einfachsten Fall durch die Verschiebung der Bildregionen im linken und rechten Bild, die sich durch das größte Maß an Übereinstimmung auszeichnen. Einige dieser Verfahren selektieren durch einen sogenannten Interest-Operator vor der Korrespondenzbestimmung zunächst aus jedem Bild Gebiete mit besonderen Eigenschaften, die anschließend miteinander verrechnet werden [BT80, Mor77]. Eine Untergruppe dieser Verfahren bilden die intensitätsbasierten Methoden [Gen88, Hor86, LB95]. In diesen Ansätzen werden lediglich punktweise die Intensitäts- bzw. Grauwertdifferenzen der Stereobilder ohne die Verrechnung eines größeren lokalen Gebietes zur Korrespondenzbestimmung verwendet. Die Fehleranfälligkeit der intensitätsbasierten Verfahren ist jedoch sehr hoch, so daß sie eine relativ geringe Rolle in der Stereoskopie spielen.

Je nach Interest-Operator berechnen gebietsbasierte Verfahren für sehr viele Bildpunkte eine Korrespondenz, wodurch sich sehr dicht besetzte Disparitätskarten, in welche die Verschiebungen der korrespondierenden Bildpunkte eingetragen werden, ergeben. Eine Interpolation der Disparitätskarten gestaltet sich daher als sehr einfach oder kann ganz entfallen. Im allgemeinen wird bei den gebietsbasierten Stereoverfahren davon ausgegangen, daß das eine Bild lediglich eine verschobene Version des anderen ist. In der Praxis gestaltet sich das Problem jedoch als wesentlich komplizierter. So können Unterschiede bezüglich der Beleuchtung und des Kontrastes zwischen den beiden Bildern auftreten. Weiterhin können relativ zur Stereobasis geneigte Oberflächen bewirken, daß sich die beiden Grauwertverläufe durch eine Dehnung bzw. Stauchung unterscheiden.

Durch die direkte Verrechnung der Intensitäten sind gebietsbasierte Stereoverfahren äußerst empfindlich bezüglich der obengenannten interokulären Differenzen. Je nach Größe des berücksichtigten Grauwertgebietes können die Korrespondenzen bei den gebietsbasierten Ver-

fahren stark mehrdeutig sein, so daß im allgemeinen ein zusätzlicher Verarbeitungsschritt notwendig ist, der die Mehrdeutigkeiten auflöst.

2.5.2 Merkmalsbasierte Verfahren

Die meisten der existierenden stereoskopischen Ansätze können in die Kategorie der merkmalsbasierten Verfahren eingeordnet werden. Bei dieser Technik werden zunächst Merkmale aus den Bilddaten extrahiert, die das Bild auf einer abstrakteren Ebene beschreiben. Anschließend wird in einem weiteren Schritt eine Korrespondenzbestimmung auf der Ebene der Merkmale durchgeführt. Häufig verwendete Merkmale sind hierbei Kanten-, Linien- oder Eckenpunkte oder auch Kanten- bzw. Liniensegmente. Große Bedeutung kommt bei diesen Verfahren den differenzierenden Filtern zu, die Grauwertänderungen wie Kanten oder Linien aus den Bildsignalen herausfiltern. Eines der bedeutendsten Filter ist hierbei der Marr-Hildreth-Operator [Mar82, MH80]. Da rein differenzierende Filter Rauschprozesse erheblich verstärken können, werden diese bei dem Marr-Hildreth-Operator zunächst durch ein gaußförmiges Tiefpaßfilter unterdrückt. Das Ergebnis wird anschließend zweifach räumlich differenziert. In der Praxis wird diese Operation durch den Einsatz eines linearen Bandpaßfilters, dessen Impulsantwort $\Delta g(\mathbf{x})$ sich aus der zweifachen räumlichen Ableitung der Gaußfunktion ergibt, in einem Schritt durchgeführt.

$$\Delta g(\mathbf{x}) = \frac{-1}{\pi\sigma^4} \left(1 - \frac{x^2 + y^2}{2\sigma^2} \right) e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (2-6)$$

Das Übertragungsverhalten dieses Filters wird hierbei durch den Parameter σ bestimmt. Die Merkmalspunkte, die in diesem Ansatz Kantenpunkte repräsentieren, werden durch die Nulldurchgänge im Ergebnis $r(\mathbf{x})$ einer zweidimensionalen Faltung von Grauwertbild $i(\mathbf{x})$ und Operator $\Delta g(\mathbf{x})$ definiert.

$$r(\mathbf{x}) = \Delta g(\mathbf{x}) * i(\mathbf{x}) \quad (2-7)$$

Die Richtung der Kante kann durch den Gradienten des Faltungsproduktes bestimmt werden. Der isotrope Operator $\Delta g(\mathbf{x})$ und vereinfachte Versionen hiervon bilden die Grundlage für viele merkmalsbasierte Stereoverfahren [Gri81a, Gri81b, Gri85, GST89, HA89, HSY96, KA87, MF81].

Andere merkmalsbasierte Verfahren nutzen mehrere anisotropische Filter, die gerichtete Grauwertänderungen aus dem Bildsignal extrahieren [JM92, MP79, PH86]. Auf diese Weise ergeben sich höher spezifizierte Merkmale, welche die Anzahl der möglichen Korrespondenzpaare

drastisch reduzieren können. Bei einer weiteren Kategorie von Verfahren werden die extrahierten Punktmerkmale vor der Korrespondenzbestimmung zu höheren Merkmalen weiterverarbeitet, indem sie zu sogenannten linearen Kantensegmenten verknüpft werden [AF87, AL87, HS89, LB87, MN85, MMN89, PH86, RBS87, SP96].

Aufgrund der selektiven Extraktion von Bildmerkmalen reduzieren merkmalsbasierte Verfahren im allgemeinen die Anzahl der möglichen Korrespondenzpaare. Hierdurch gestaltet sich die Lösung des Korrespondenzproblems meist weniger aufwendig als bei den gebietsbasierten Verfahren. Weiterhin sind die extrahierten Merkmale, je nach Verfahren und Merkmalsart, unempfindlicher bezüglich interokulären Differenzen, wie perspektivischen Verzerrungen, Kontrast- und Beleuchtungsunterschieden usw. Darüber hinaus kann die Korrespondenzbestimmung durch das einfache Vergleichen von Merkmalsattributen im allgemeinen schneller erfolgen als bei gebietsbasierten Methoden. Da die Merkmale jedoch unter Umständen relativ selten im Bild vorhanden sind, ergeben merkmalsbasierte Verfahren oft nur sehr spärlich besetzte Disparitätskarten, so daß zusätzlich eine Interpolation durchgeführt werden muß. Genau wie bei den gebietsbasierten Verfahren ist bei diesen Ansätzen im allgemeinen ein weiterer Verarbeitungsschritt notwendig, um die Mehrdeutigkeiten aufzulösen.

2.5.3 Phasenbasierte Verfahren

Die Grundlage für die sogenannten phasenbasierten Methoden zur Disparitätsmessung liefert der Verschiebungssatz der Fouriertransformation. Obwohl die zu verarbeitenden Grauwertbilder im allgemeinen in ortsdiskreter Form mit einem durch die Bildabmessungen begrenzten Definitionsintervall vorliegen, sollen die weiteren theoretischen Betrachtungen zur Vereinfachung der Darstellung für den ortskontinuierlichen Fall mit unendlichen Definitionsintervallen erfolgen. Das im weiteren verwendete zweidimensionale ortskontinuierliche Fouriertransformationspaar sei durch die Hintransformation

$$F(\mathbf{k}) = \mathcal{F}\{f(\mathbf{x})\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathbf{x}) e^{-j\mathbf{k}^T \mathbf{x}} d^2 \mathbf{x} \quad (2-8)$$

und durch die Rücktransformation

$$f(\mathbf{x}) = \mathcal{F}^{-1}\{F(\mathbf{k})\} = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\mathbf{k}) e^{j\mathbf{k}^T \mathbf{x}} d^2 \mathbf{k} \quad (2-9)$$

definiert. Hierbei kennzeichnen $\mathbf{x} = [x \ y]^T$ die Orts-, $\mathbf{k} = [k_x \ k_y]^T$ die Ortsfrequenzvariablen und $j = \sqrt{-1}$ die imaginäre Einheit.

Unterscheiden sich die zwei Eingangsbilder $i_l(\mathbf{x})$ und $i_r(\mathbf{x})$ nur durch eine Verschiebung $\mathbf{d} = [d_x \ d_y]^T$,

$$i_l(\mathbf{x}) = i_r(\mathbf{x} + \mathbf{d}) \quad (2-10)$$

so gilt nach dem Verschiebungssatz der Fouriertransformation (siehe z.B. [Jäh93])

$$\mathcal{F}\{f(\mathbf{x} + \mathbf{d})\} = F(\mathbf{k})e^{j\mathbf{k}^T \mathbf{d}} \quad (2-11)$$

für die Fouriertransformierten der Bilder:

$$I_l(\mathbf{k}) = I_r(\mathbf{k})e^{j\mathbf{k}^T \mathbf{d}} \quad (2-12)$$

Die Spektren $I(\mathbf{k})$ der Bildsignale können durch ihren Betrag $|I(\mathbf{k})|$ und ihre Phase $\phi(\mathbf{k})$ mit

$$I(\mathbf{k}) = |I(\mathbf{k})|e^{j\phi(\mathbf{k})} \quad (2-13)$$

beschrieben werden. Werden nun die Spektren der beiden Bilder auf die folgende Weise miteinander verrechnet

$$\frac{I_l(\mathbf{k})I_r^*(\mathbf{k})}{|I_l(\mathbf{k})||I_r(\mathbf{k})|} = \frac{|I_l(\mathbf{k})||I_r(\mathbf{k})|e^{j(\phi_l(\mathbf{k}) - \phi_r(\mathbf{k}))}}{|I_l(\mathbf{k})||I_r(\mathbf{k})|} = e^{j\mathbf{k}^T \mathbf{d}}, \quad (2-14)$$

so reduziert sich die Messung der Disparität auf eine Suche nach Dirac-Impulsen in der inversen Fouriertransformation von Beziehung (2-14), da

$$\mathcal{F}^{-1}\left\{e^{j\mathbf{k}^T \mathbf{d}}\right\} = \delta(\mathbf{x} + \mathbf{d}). \quad (2-15)$$

Diese einfache Art der Disparitätsmessung wird auch als Phasenkorrelation bezeichnet [KH75, OP89]. In der Stereobildverarbeitung ist jedoch im allgemeinen nicht mit einer ausschließlich globalen Verschiebung zwischen den Bildern zu rechnen, da unterschiedlich weit vom Kame-

rasystem entfernte Objekte unterschiedliche Disparitätswerte in den Stereobildern aufweisen (vgl. Beziehung (2-2)). Folglich müssen die Verschiebungen korrespondierender Bildbereiche in dem Stereobildpaar durch lokale Operatoren ermittelt werden, so daß eine Phasenkorrelation generell nur in Verbindung mit einer gefensterten Fouriertransformation sinnvoll ist. Auf einem ähnlichen Prinzip basieren die Verfahren, bei denen die Disparität durch die sogenannten Cepstral-Funktion bestimmt wird [CB91, CB93, OP89, YS89].

Die größte Bedeutung innerhalb der phasenbasierten Verfahren kommt den sogenannten Phasendifferenzmethoden zu. Die Phaseninformation ergibt sich bei diesen Techniken aus der Antwort von komplexen Filterpaaren, mit denen die Eingangsbilder gefiltert werden. Eine wesentliche Voraussetzung für dieses Verfahren ist, daß die Phase der Filterantworten näherungsweise eine lineare Funktion des Ortes ist. Diese Eigenschaft kann dann erreicht werden, wenn die Filterübertragungsfunktion keinen Gleichanteil aufweist und für negative Frequenzen verschwindet. Diese Eigenschaft wird als Quadraturverhalten bezeichnet (siehe [WK+95]). Die überwiegende Anzahl der Verfahren verwendet zur Extraktion der Phaseninformation sogenannte Gaborfilter [Gab46], die jedoch nur unter bestimmten Voraussetzungen eine Quadratureigenschaft besitzen (siehe [FJ93, FJJ91, JJ89, JJ94, MBE95, San88, TM94]). Eine Übersicht über Verfahren mit exakter Quadratureigenschaft ist in [WK+95] gegeben.

Als Gaborfilter wird im allgemeinen ein Filter bezeichnet, dessen Impulsantwort durch das Produkt einer komplexen Schwingung mit einer Gaußfunktion beschrieben werden kann. Für den eindimensionalen Fall, der hier zur Erläuterung des Prinzips betrachtet wird, sei sie wie folgt definiert:

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} e^{jk_0x} \quad (2-16)$$

Hierbei sei k_0 die Modulationsortsfrequenz und σ der Streuungsparameter der Gaußfunktion. Die Übertragungsfunktion $G(k)$ des Filters ergibt sich somit zu:

$$G(k) = e^{-\frac{\sigma^2(k-k_0)^2}{2}} \quad (2-17)$$

Das wachsende Interesse an Gaborfiltern im Bereich der Bildverarbeitung begründet sich auf zwei Eigenschaften: In neueren neurophysiologischen Untersuchungen konnte gezeigt werden, daß das Übertragungsverhalten des visuellen Systems höherer Lebewesen in einer bestimmten Verarbeitungsstufe, im Bereich der sogenannten einfachen Zellen, erstaunlich genau durch Gaborfunktionen modelliert werden kann [Dau85, JP87a, JP87b, JSP87, KMB82, Mar80, PR81, Sab96]. Die zwei bislang existierenden Theorien über das visuelle System, auf der einen

Seite das Modell der lokalen rezeptiven Felder von Hubel und Wiesel [HW62, BCH73] und auf der anderen Seite das visuelle System als Fourieranalysator [CR68, DDY79, KB81, MF77, PLT71], konnten durch dieses Modell erweitert und in gewisser Weise vereint werden.

Die zweite wesentliche Eigenschaft der Gaborfilter ergibt sich durch die spezielle Form der Filterfunktion. Gabor [Gab46] konnte zeigen, daß die nach ihm benannten Elementarsignale die Unschärferelation der Informationstheorie durch ihr minimales „Zeit-Bandbreite-Produkt“ an der unteren Schranke erfüllen. Für die Bildverarbeitung ergibt sich hieraus die Möglichkeit, Bilder frequenzselektiv bei optimaler Auflösung im Orts- und Ortsfrequenzbereich zu filtern. Sanger [San88] zeigte erstmals, daß es mit Hilfe von Gaborfiltern auf einfache Art möglich ist, Disparitäten in Stereobildern zu messen.

Seien $r_l(x)$, und $r_r(x)$ die Antworten der Gaborfilter im linken und rechten Stereokanal, die sich aus der Faltung der Filterfunktion $g(x)$ mit dem Eingangssignal $i(x)$ ergeben:

$$r(x) = g(x) * i(x) = \int_{-\infty}^{\infty} g(\tilde{x})i(x - \tilde{x})d\tilde{x} \quad (2-18)$$

Die Filterantworten können durch ihren Betrag $|r(x)|$ und den Phasenwinkel $\phi(x)$ beschrieben werden.

$$r(x) = |r(x)|e^{j\phi(x)} \quad (2-19)$$

Sind die Eingangssignale der beiden Stereokanäle und damit auch die Filterantworten relativ zueinander um die Disparität d_x verschoben mit $r_l(x) = r_r(x + d_x)$, so kann diese Verschiebung durch die Phasendifferenz der Filterantworten approximiert werden:

$$d_x \approx \frac{\phi_l(x) - \phi_r(x)}{k_0} \quad (2-20)$$

Die Disparität kann nur dann exakt durch Beziehung (2-20) errechnet werden, wenn die verwendeten Filter eine infinitesimale Bandbreite aufweisen [San88]. Bei größeren Bandbreiten ergibt Beziehung (2-20) nur dann eine korrekte Disparitätsmessung, wenn die Eingangssignale sinusförmig sind und die Filtermodulationsortsfrequenz k_0 im Nenner durch die Frequenz der Sinusschwingung ersetzt wird (siehe [Fle94]). Einerseits ist jedoch der Grauwertverlauf in einem Bild meist nicht sinusförmig und andererseits kann eine gute Auflösung im Ortsbereich nur durch relativ breitbandige Filter erreicht werden.

Je nach gewählter Bandbreite der Filter kann die Näherung in Beziehung (2-20) daher zu erheblichen Fehlern bei der Disparitätsbestimmung führen. Um dies zu vermeiden, verwenden viele Verfahren die mittlere Phasenableitung der linken und rechten Filterantwort als Näherung der lokalen Ortsfrequenz anstelle der Filtermodulationsortsfrequenz [FJ93, FJJ91, JJ89, JJ94, JJT91, TM94].

$$d_x \approx \frac{\phi_l(x) - \phi_r(x)}{\bar{k}} \quad \text{mit} \quad \bar{k} = \frac{1}{2}(k_l + k_r) \approx \frac{1}{2} \left(\frac{\partial}{\partial x} \phi_l(x) + \frac{\partial}{\partial x} \phi_r(x) \right) \quad (2-21)$$

Die Phasenableitung wird auch als Momentanfrequenz bezeichnet [Boa92, Pap65]. Der Mittelwert der linken und rechten lokalen Ortsfrequenzen wird hier eingesetzt, um die Auswirkungen von relativen Stauchungen bzw. Dehnungen der Grauwertverläufe, die durch relativ zur Stereobasis geneigte Oberflächen entstehen können, zu vermindern.

Sowohl gebiets- als auch die meisten merkmalsbasierten Verfahren erlauben eine Disparitätsmessung nur in Form von ganzzahligen Vielfachen des Sensorraasterabstandes. Die Genauigkeit der Tiefenrekonstruktion, die bei diesen Verfahren mit Pixelgenauigkeit bezeichnet wird, wird durch diese Diskretisierung erheblich eingeschränkt. Phasendifferenzmethoden haben dahingegen den Vorteil, daß Disparitäten genauer als das zugrundeliegende Abtastraster bestimmt werden können. Diese Verfahren sind somit subpixelgenau. Da die Phaseninformation invariant bezüglich der Amplitude der Filterantworten ist, sind phasenbasierte Verfahren darüber hinaus relativ robust bezüglich interokulären Beleuchtungs- und Kontrastdifferenzen.

Durch die Mehrdeutigkeit bei der Phasenberechnung können jedoch nur Disparitätswerte bis zur halben Modulationswellenlänge des verwendeten Filters gemessen werden. Um große Disparitätswerte messen zu können, müssen daher Filter mit sehr niedriger Modulationsortsfrequenz verwendet werden. Da die Bandbreite dieser Filter relativ gering sein muß, damit die Quadratureigenschaft erhalten bleibt, ist die Ortsauflösung entsprechend gering. Wenn diese Verfahren dennoch eine gute Auflösung im Ortsbereich aufweisen sollen, sind sogenannte Grob-Fein-Strategien notwendig [FJJ91, San88, TM94]. Bei diesen Strategien werden die Bilder mit mehreren Filtern unterschiedlicher Modulationsortsfrequenz gefiltert, wobei die in einer bestimmten Filterebene ermittelten Disparitäten dazu dienen, die Filterergebnisse der nächsten, feiner auflösenden Filterebene so zu verschieben, daß die jeweils auftretenden Disparitäten in den aktuellen Meßbereich fallen und somit genauer erfaßt werden können usw.

Im Gegensatz zu merkmals- und gebietsbasierten Verfahren benötigen phasenbasierte Verfahren keine zusätzlichen Strategien zur Auflösung von Mehrdeutigkeiten, da das Korrespondenzproblem durch die Grob-Fein-Strategie implizit gelöst wird. Wird jedoch in einer Auflösungsebene eine falsche Verschiebung ermittelt, ist es möglich, daß diese Strategie zu falschen Disparitätsmessungen führt. Weiterhin wird die Phase im allgemeinen instabil, wenn die Ausgangssignale der Filter in der Nähe des Ursprungs der komplexen Ebene liegen, so daß an diesen Stellen die Disparitätsmessung äußerst ungenau wird. Einige Verfahren nutzen des-

halb sogenannte Vertrauenskarten, die aus den Amplituden der Filterantworten errechnet werden, um die Güte der Disparitätsmessung zu bewerten [TM94, VT95, WK+95]. Wie viele andere Ansätze reagieren auch phasenbasierte Verfahren sehr empfindlich auf Okklusionen.

2.5.4 Regularisierungsmethoden

Die im wesentlichen auf Tikhonov [TA77] beruhende Theorie der Regularisierung wird sowohl bei gebiets- als auch bei merkmalsbasierten Verfahren dazu genutzt, um die Mehrdeutigkeiten des Korrespondenzproblems mit Hilfe von a priori Wissen aufzulösen [Mor93]. Das Korrespondenzproblem wird bei der sogenannten Standardregularisierung in Form der Minimierung einer Energie- oder Kostenfunktion e_d formuliert.

$$e_d = \iint (h_l(x, y) - h_r(x + d_x(\mathbf{x}), y))^2 + \lambda (f(d_x(\mathbf{x})))^2 d^2 \mathbf{x} \quad (2-22)$$

Je nach Ansatz können hierbei die Bilddaten $h(\mathbf{x})$ des linken und rechten Stereokanals reine Grauwerte [Gen88, Hor86] oder bereits bandpaßgefilterte Intensitäten sein [Bar89, DP86, HSY96, MMP87]. Der erste quadratische Term in Beziehung (2-22) repräsentiert die Übereinstimmung der beiden Stereokanäle bei der entsprechenden Verschiebung d_x an der Bildkoordinate $\mathbf{x} = [x \ y]^T$. Durch den zweiten Term, der als Stabilisierungsfunktional bezeichnet wird, fließt die Kontinuitätseinschränkung (siehe Kapitel 2.4.3) mit in die Kostenfunktion e_d ein. Um eine möglichst glatte Disparitätskarte $d_x(\mathbf{x})$ zu erhalten, wird die Funktion f meist als Betrag eines Gradienten [BPT88, Gen88] oder auch als Betrag eines Laplaceoperators [Bar89, Hor86, MMP87, PTK85] realisiert. Der sogenannte Regularisierungsparameter λ ist eine positive Konstante, die den Einfluß der Kontinuitätseinschränkung auf die Lösung bestimmt. Vertikale Disparitätskomponenten werden bei diesen Ansätzen selten modelliert.

Eine alternative Möglichkeit zur Ausnutzung der Kontinuitätseinschränkung besteht in der Modellierung der Disparitätskarten durch sogenannte Markov'sche Zufallsfelder [Bar89, CC92, DP86, HSY96, LB95, YGB91]. Dieses Feld ist eine auf einem diskreten Gitter definierte Zufallsverteilung, bei der die bedingte Wahrscheinlichkeit einer Korrespondenz nur von einer kleinen Untermenge in ihrer lokalen Nachbarschaft abhängt. Mit Hilfe einer Bayes'schen Schätzung wird bei diesen Ansätzen die a posteriori Wahrscheinlichkeit einer bestimmten Lösung errechnet. Das Ziel bei diesen Verfahren ist es, diese a posteriori Wahrscheinlichkeit zu maximieren oder eine entsprechende Fehlerfunktion zu minimieren. Die Standardregularisierung kann als ein Spezialfall dieser Bayes'schen Methode angesehen werden (siehe z.B. [PTK85]).

Die Minimierung der Kosten- oder Fehlerfunktionen kann auf verschiedene Weise erfolgen. Ist die Kostenfunktion konvex, so können über das Variationsprinzip die Euler-Lagrange-Differentialgleichungen ermittelt und anschließend iterativ gelöst werden [Gen88, Hor86, Ter86, TWK87]. Eine Voraussetzung hierfür ist allerdings, daß $d_x(\mathbf{x})$ in Beziehung (2-22) stetig und

differenzierbar ist [BPT88]. Der Disparitätsverlauf $d_x(x)$ hat in der Stereoskopie jedoch die Eigenschaft, daß er im allgemeinen nicht kontinuierlich ist. Vielmehr treten in natürlichen Szenen an Objektkonturen grundsätzlich Diskontinuitäten auf, wenn sich mehrere Objekte in unterschiedlichen Entfernungen im Blickfeld befinden. Darüber hinaus sind gerade in den zuvor geschilderten Szenarien meist Okklusionen vorhanden, in denen $d_x(x)$ nicht definiert ist. Weiterhin kann nicht gewährleistet werden, daß das Funktional in Beziehung (2-22) konvex ist. Die direkte Minimierung über die Euler-Lagrange-Gleichungen führt daher nicht immer zu einer optimalen Lösung.

Ansätze, die im Fall eines nichtkonvexen Funktionals die Konvergenz zu lokalen Minima vermeiden, können in die stochastischen Relaxationsverfahren untergliedert werden. Diese stochastischen Ansätze ermöglichen unter gewissen Umständen die Bestimmung des globalen Optimums durch Strategien wie „Simulated Annealing“ [Bar89, GG84, KGV83, KD88]. Allerdings sind diese Verfahren mit einem erheblichen Berechnungsaufwand verbunden. Wesentlich schneller hingegen, jedoch nur suboptimal, sind deterministische Verfahren wie die „Mean-Field-Theory“ [CC92, YGB91] und die „Graduated-Non-Convexity-Methode“ [LB95].

Obwohl diese Verfahren teilweise in der Lage sind, das globale Optimum der jeweiligen Kostenfunktionale zu ermitteln, können die im allgemeinen willkürlich gewählten Modelle der Disparitätskarten, die als a priori Wissen einfließen, dazu führen, daß die optimale Lösung mit der Realität nicht viel gemein hat. Um diesen Effekt zu vermindern, werden bei einigen Verfahren zusätzlich sogenannte Linienprozesse eingeführt, die in Form von binären Variablen Diskontinuitäten im Disparitätsverlauf modellieren [GG84, LB95, MMP87, YGB91]. Ein nicht unwesentlicher Nachteil der Regularisierung ist der erhebliche Berechnungsaufwand, der erforderlich ist, um die Mehrdeutigkeiten des Korrespondenzproblems auf diese Weise aufzulösen.

2.5.5 Dynamische Programmierung

Die dynamische Programmierung ist ein effektiver Weg, um Kostenfunktionale mit einer großen Anzahl von diskreten Variablen zu minimieren bzw. zu maximieren. Bei Verfahren, die diese Technik zur stereoskopischen Korrespondenzbestimmung nutzen, wird zunächst jedes Merkmal oder jeder Bildpunkt entlang einer Epipolarlinie von einem Bildrand zum anderen mit einem aufsteigenden Index versehen. Dies geschieht sowohl im linken Bild mit einem Index $m \in [0..M]$ als auch im rechten Bild mit einem Index $n \in [0..N]$ entlang der jeweils korrespondierenden Epipolarlinien. Anschließend wird eine Matrix konstruiert, deren Einträge an der Stelle (m, n) die Kosten repräsentieren, die eine Zuordnung der Merkmale m und n hervorrufen würde. Die Kosten jedes Matrixeintrages werden hierbei im allgemeinen durch die Ähnlichkeit der entsprechenden Merkmale und gegebenenfalls durch die Kontinuität der resultierenden Disparitätskarte bestimmt.

Ziel dieser Verfahrens ist es, ausgehend vom linken oberen Matrixeintrag $(0, 0)$ einen Pfad bis zum rechten unteren Matrixeintrag (M, N) zu suchen, der in der Summe die geringsten Kosten

hervorrufen. Die Lösung des Korrespondenzproblems kann auf diese Weise als eine Suche nach dem optimalen Pfad in Bezug auf bestimmte Kostenfunktionale formuliert werden [BB81, GLY92, OK85, RBS87]. Ein häufig gewählter Ansatz zur Errechnung des optimalen Pfades ist der sogenannte Viterbi-Algorithmus (siehe z.B. [RBS87]).

Eine wesentliche Voraussetzung für diese Verfahren ist, daß die korrespondierenden Merkmale oder Bildpunkte in der gleichen Reihenfolge im linken und rechten Stereokanal auftreten. Diese Voraussetzung, die gleichbedeutend mit der in Kapitel 2.4.4 beschriebenen Ordnungseinschränkung ist, ist jedoch nicht immer erfüllt (vgl. Bild 2-5). Darüber hinaus reagieren diese Verfahren besonders anfällig auf periodische Strukturen und Okklusionen (siehe [DA89, Shi87]).

2.5.6 Relaxationsverfahren

Eine weitere Verfahrenskategorie zur Lösung des Korrespondenzproblems bilden die sogenannten Relaxationsverfahren. Der diesen Verfahren gemeinsame Ansatz hat seinen Ursprung im Bereich der Bildsegmentierung und ist im wesentlichen auf die Arbeiten von Rosenfeld, Hummel und Zucker [RHZ76] zurückzuführen. Bei den meisten Ansätzen, die dieses Verfahren in der Stereoskopie anwenden, werden zunächst Merkmale oder Bildregionen mit besonderen Eigenschaften aus den Grauwertbildern extrahiert. Den Bildkoordinaten, an denen die erwähnten Elemente auftreten, werden anschließend sogenannte Knoten zugeordnet. Jeder dieser Knoten wird weiterhin mit einem Satz von Variablen versehen, die jeweils eine Korrespondenz des betreffenden Knotens mit unterschiedlichen Elementen im anderen Bild repräsentieren. Diese Variablen werden je nach Ansatz als Wahrscheinlichkeit [BT80, KA87] oder als Neuronenaktivität [LH96, MPP78, MP76, MP79] bezeichnet.

Zu Beginn des eigentlichen Relaxationsprozesses werden die Variablen in Abhängigkeit der Ähnlichkeit der entsprechenden Merkmale oder Grauwerte initialisiert. Anschließend werden die Variablenwerte in einem dynamischen Prozeß iterativ aktualisiert. Der Zuwachs der Variablenwerte hängt dabei meist von den Variablenwerten anderer Knoten in einem lokal begrenzten Umfeld in der Bildebene ab, bei denen die korrespondierenden Merkmale die gleiche Disparität aufweisen. Diese Variablen wirken verstärkend auf den betreffenden Variablenwert, wodurch Variablenwerte, die lokal häufiger auftretende Disparitäten repräsentieren, schneller wachsen. Auf diese Weise wird bei Relaxationsansätzen die Kontinuitätseinschränkung (siehe Kapitel 2.4.3) zur Auflösung der Mehrdeutigkeiten ausgenutzt. Der Iterationsprozeß wird im allgemeinen beendet, wenn eine stationäre Lösung oder eine vorgegebene Anzahl von Iterationen erreicht wurde. Da nach der Eindeutigkeitseinschränkung (siehe Kapitel 2.4.2) nur maximal eine Korrespondenz pro Knoten existieren darf, bestimmt schließlich die Variable mit dem höchsten Wert die Disparität des betreffenden Knotens. Neben den verstärkend wirkenden Kopplungen sind bei einigen Ansätzen die Variablen zusätzlich mit allen anderen Variablen des betreffenden Knotens dämpfend gekoppelt. Auf diese Weise kann die Eindeutigkeitseinschränkung explizit in den Iterationsprozeß integriert werden [RH94, MPP78, MP76, MP79]. Die Variablen sind je nach Ansatz entweder wertdiskret [MPP78, MP76, MP79] oder wertkontinuierlich [BT80, KA87, RH94, MN85, Pra85].

Relaxationsverfahren sind aufgrund der verstärkend wirkenden lokalen Kopplung relativ robust bezüglich interokulärer Differenzen und weisen im allgemeinen eine wesentlich höhere Konvergenzgeschwindigkeit als Regularisierungsmethoden auf. Da die Kopplungen bei den Relaxationsverfahren in der Stereoskopie immer lokal begrenzt sind, haben diese Verfahren darüber hinaus den Vorteil, daß sie sehr einfach zu parallelisieren sind. Als Vorbild für diese Art von Ansatz können die lokalen Nachbarschaftsbeziehungen von Neuronen, die in biologischen Sehsystemen nachgewiesen werden konnten, angesehen werden (siehe z.B. [Jul71, MP76]).

3 Der Ansatz von Reimann und Haken

Ein Stereoverfahren, das in diesem Kapitel gesondert betrachtet wird, ist der in die Kategorie der Relaxationsverfahren einzuordnende Ansatz von Reimann und Haken [RH94]. Der Grund für die detailliertere Betrachtung an dieser Stelle ist, daß ein wesentlicher Teilaspekt der vorliegenden Arbeit auf diesem Ansatz basiert.

Grundlage für diesen Relaxationsansatz, der im weiteren als Selbstorganisationsprozeß bezeichnet wird, ist die im ersten Teil dieses Kapitels beschriebene Haken'sche Mustererkennungsgleichung [Hak87, Hak91]. Im weiteren Verlauf des Kapitels werden die einzelnen Erweiterungen erläutert, die notwendig sind, um diese allgemeine Theorie auf die Stereoskopie anzuwenden.

3.1 Die Haken'sche Mustererkennungsgleichung

Der Forschungsschwerpunkt des von Haken begründeten interdisziplinären Wissenschaftszweigs der Synergetik ist die Untersuchung von Vielteilchensystemen fern vom thermischen Gleichgewicht [Hak79, Hak81, Hak82]. Ein interessantes Phänomen, das unter gewissen Voraussetzungen bei solchen Vielteilchensystemen auftritt, ist die Struktur- oder Musterbildung durch spontane Selbstorganisation. Die einzelnen Bestandteile, die bei diesen Systemen in Wechselwirkung miteinander stehen, können dabei z. B. Atome, Flüssigkeitspartikel oder Neurone sein. Durch das Konzept der sogenannten Ordnungsparameter gelang es, einen gemeinsamen mathematischen Rahmen für die spontane Entstehung von Ordnung für Systeme aus sehr unterschiedlichen wissenschaftlichen Disziplinen zu schaffen. Diese Ordnungsparameter bestimmen die Struktur oder das Muster, das sich innerhalb der dynamischen Prozesse der Vielteilchensysteme herausbildet.

Mit Hilfe der sogenannten Mustererkennungsgleichung, die aus der Ordnungsparametergleichung des Benard-Problems der Hydrodynamik abgeleitet werden kann, konnte Haken eine formale Analogie zwischen Musterbildung und Mustererkennung aufzeigen [Hak87, Hak91].

$$\dot{\xi}_k = \lambda_k \xi_k - (b + c) \sum_{k' \neq k} \xi_{k'}^2 \xi_k - c \xi_k^3 \quad (3-1)$$

Die Parameter b und c sind hierbei positive Konstanten. Für $\lambda_k > 0$ findet ein Wettkampf zwischen den verschiedenen Ordnungsparametern ξ_k statt. Es kann gezeigt werden, daß nur der Ordnungsparameter mit dem betragsmäßig größten Startwert die stabile stationäre Lösung

$$\xi_{k_s} = \pm \sqrt{\frac{\lambda_k}{c}} \quad \text{mit} \quad \dot{\xi}_{k_s} = 0 \quad (3-2)$$

annehmen kann und auf diese Weise als Gewinner aus dem Wettbewerb hervorgeht, während alle anderen Ordnungsparameter die stationäre Lösung $\xi_{k_s} = 0$ annehmen (siehe z. B. [Hak91]). Für negative Startwerte wird dabei die negative stationäre Lösung angenommen und für positive die positive stationäre Lösung. Für $\lambda_k > 0$ repräsentiert der erste Term auf der rechten Seite von Beziehung (3-1) exponentielles Wachstum, wobei der dritte Term dieses Wachstum begrenzt. Der zweite Term in Beziehung (3-1) bewirkt schließlich eine dämpfende Wirkung aller übrigen in den Prozeß involvierten Ordnungsparameter auf den betrachteten Wert ξ_k und ist somit die eigentliche Ursache für den Wettbewerb.

Die Ordnungsparametergleichung konnte von Haken auf Probleme der Mustererkennung erfolgreich angewendet werden. Haken zeigte, daß unter gewissen Voraussetzungen unvollständige oder gestörte Muster rekonstruiert werden können, wenn die Startwerte der Ordnungsparameter in Beziehung (3-1) die Übereinstimmung eines Testmustervektors \mathbf{q} mit einem als Prototyp bezeichneten Vergleichsmuster \mathbf{v}_k repräsentieren. Der Wettbewerb muß in diesem Fall nur zwischen einer Anzahl von Ordnungsparametern, die der Anzahl der Prototypen entspricht, ausgetragen werden. Mit Hilfe von Beziehung (3-1) kann das Testmuster \mathbf{q} über Zwischenzustände in das Vergleichsmuster \mathbf{v}_k , das die größte Ähnlichkeit mit dem Testmuster aufweist, überführt werden. In diesem Zusammenhang werden die Ordnungsparameter auch als Neurone bezeichnet [Hak91].

3.2 Anwendung auf die Stereoskopie

Zur Anwendung der Haken'schen Mustererkennungsgleichung auf das stereoskopische Korrespondenzproblem sind einige zusätzliche Mechanismen notwendig. Da in der Stereoskopie generell keine Prototypen vorliegen, muß die Initialisierung des Selbstorganisationsprozesses auf einem anderen Wege geschehen als im vorangegangenen Abschnitt. Weiterhin ist die Lösung des Korrespondenzproblems an gewisse Einschränkungen gebunden, die zusätzlich in den Selbstorganisationsprozeß integriert werden müssen.

3.2.1 Ähnlichkeit von Bildregionen

Im Ansatz von Reimann [Rei95] geht dem Selbstorganisationsprozeß zunächst ein gebietsbasiertes Verfahren voraus, mit dem die Ähnlichkeit $s(\mathbf{x}, \mathbf{d})$ zwischen Bildregionen in den beiden Stereokanälen ermittelt wird. Da keine besondere Kamerageometrie in diesem Ansatz vorausgesetzt wird, ist eine zweidimensionale Suche nach Korrespondenzpartnern erforderlich. Um eine lokale Messung der Ähnlichkeit zu erhalten, wird hier eine zweidimensionale Fensterfunktion f verwendet, die bewirkt, daß nur Werte in einem kleinen Umfeld um den betreffenden Bildpunkt zur Messung der Ähnlichkeit beitragen.

$$s(\mathbf{x}, \mathbf{d}) = s_0 - \frac{\iint |f(\mathbf{x} - \tilde{\mathbf{x}})(i_l(\tilde{\mathbf{x}}) - i_r(\tilde{\mathbf{x}} + \mathbf{d}))| d^2 \tilde{\mathbf{x}}}{\iint f(\mathbf{x}) d^2 \mathbf{x}} \quad (3-3)$$

Die Konstante s_0 wird hierbei so gewählt, daß das Ähnlichkeitsmaß keine negativen Werte annimmt und maximal bei größter Übereinstimmung der Bildregionen wird. Um eine gewisse Invarianz bezüglich lokal auftretender Beleuchtungsunterschiede zwischen linkem und rechtem Bild zu erreichen, werden die Grauwerte $i(\mathbf{x})$ vor der Berechnung des Ähnlichkeitsmaßes von ihrem lokalen Mittelwert befreit.

$$\dot{i}(\mathbf{x}) = i(\mathbf{x}) - \frac{\iint f(\mathbf{x} - \tilde{\mathbf{x}}) i(\tilde{\mathbf{x}}) d^2 \tilde{\mathbf{x}}}{\iint f(\mathbf{x}) d^2 \mathbf{x}} \quad (3-4)$$

Wie bei den meisten Relaxationsverfahren, dient das Ähnlichkeitsmaß dazu, den Relaxationsprozeß zu initialisieren, der in einem weiteren Schritt die Mehrdeutigkeiten des Korrespondenzproblems unter Ausnutzung der Einschränkungen auflöst.

3.2.2 Kontinuität durch Kopplung von Mustererkennungsgleichungen

Der in Kapitel 3.1 beschriebene Formalismus wird in dem Ansatz von Reimann [Rei95] durch eine geeignete Erweiterung auf das Korrespondenzproblem in der Stereoskopie angewendet. In diesem Ansatz wird jedem Bildpunkt in einem der Stereobilder ein Satz von Mustererkennungsgleichungen zugeordnet. Jeder der pro Bildpunkt vorgesehenen Ordnungsparameter $\xi(\mathbf{x}, \mathbf{d}, t)$, die von Reimann [Rei95] als binokulare Neurone bezeichnet werden, repräsentiert hier eine Korrespondenz des betreffenden Bildpunktes mit einem Bildpunkt im anderen Stereokanal. Über Beziehung (3-1) wirken nun alle binokularen Neurone eines Bildpunktes dämpfend aufeinander, so daß ein Wettbewerb zwischen den Neuronen mit den gleichen Bildkoordinaten entsteht. Im weiteren Verlauf der Arbeit werden die Ordnungsparameter als Variablen bezeichnet.

Werden die Variablen mit Hilfe des in Kapitel 3.2.1 beschriebenen Ähnlichkeitsmaßes initialisiert, so würde nach dem in Kapitel 3.1 beschriebenen Mechanismus die Variable, welche die größte Ähnlichkeit repräsentiert, den Wettbewerb gewinnen und somit die Korrespondenz des betreffenden Bildpunktes bestimmen. Die Eindeutigkeitsbeschränkung der Stereoskopie (siehe Kapitel 2.4.2) würde auf diese Weise implizit erfüllt. Da jedoch allein die Ähnlichkeit von Bildregionen aufgrund der speziellen Natur des Korrespondenzproblems (siehe Kapitel 2.2) in den seltensten Fällen zu einer eindeutigen Zuordnung führt, wird im Reimann'schen Ansatz zusätzlich jede Variable mit den Variablen, welche die gleiche Disparität repräsentieren, in einem kleinen räumlichen Umfeld U um die betreffende Bildkoordinate gekoppelt.

Durch dieses verstärkend wirkende Umfeld können auch Variablen den Wettbewerb gewinnen, die nicht mit dem höchsten Ähnlichkeitsmaß initialisiert wurden, wenn im Umfeld die gleiche Disparität entsprechend stark durch hohe Variablenwerte vertreten ist. Diese Kopplung bewirkt, daß Variablen mit konformem Umfeld im Wettbewerb bevorzugt werden und kontinuierliche Disparitätskarten, wie sie in der Kontinuitätseinschränkung gefordert werden (siehe Kapitel 2.4.3), entstehen.

Für jede Bildkoordinate $\mathbf{x} = [x \ y]^T$ und jede mögliche Disparität $\mathbf{d} = [d_x \ d_y]^T$ ergeben sich auf diese Weise die sogenannten gekoppelten Mustererkennungsgleichungen:

$$\begin{aligned} \dot{\xi}(\mathbf{x}, \mathbf{d}, t) = & \lambda(\mathbf{x}, \mathbf{d}) \xi(\mathbf{x}, \mathbf{d}, t) - (b + c) \sum_{d' \neq d} \xi^2(\mathbf{x}, \mathbf{d}', t) \xi(\mathbf{x}, \mathbf{d}, t) - c \xi^3(\mathbf{x}, \mathbf{d}, t) \\ & + \sum_{\mathbf{x}' \in U} e(\mathbf{x}') \xi(\mathbf{x}', \mathbf{d}, t) \xi(\mathbf{x}, \mathbf{d}, t) \end{aligned} \quad (3-5)$$

Über den vierten Term auf der rechten Seite in Gleichung (3-5) wirken Variablen, welche die gleiche Disparität repräsentieren wie die betreffende Variable $\xi(\mathbf{x}, \mathbf{d}, t)$, innerhalb einer lokalen Umgebung U um die Bildkoordinate \mathbf{x} verstärkend auf $\xi(\mathbf{x}, \mathbf{d}, t)$. Der erste und vierte Term verkörpern somit einen zeitabhängigen Zuwachsfaktor, der einerseits aufgrund des begrenzten Umfeldes einen gewissen Gradienten im Disparitätsverlauf zuläßt und andererseits verhindert, daß grundsätzlich die Variable mit dem höchsten Initialisierungswert den Wettbewerb gewinnt. Die stationären Lösungen von Beziehung (3-5) ungleich Null ergeben sich zu:

$$\xi_s(\mathbf{x}, \mathbf{d}) = \pm \sqrt{\frac{1}{c} \left(\lambda(\mathbf{x}, \mathbf{d}) - (b + c) \sum_{d' \neq d} \xi_s^2(\mathbf{x}, \mathbf{d}') + \sum_{\mathbf{x}' \in U} e(\mathbf{x}') \xi_s(\mathbf{x}', \mathbf{d}) \right)} \quad (3-6)$$

Jede Variable wird zu Beginn des Wettbewerbs über das Ähnlichkeitsmaß aus Beziehung (3-3) mit einem positiven Anfangswert, der kleiner als die positive stationäre Lösung ist, initialisiert. Reimann konnte sicherstellen, daß die grundlegende Eigenschaft der Mustererkennungsgleichung, daß grundsätzlich nur eine Variable den Wettbewerb gewinnen kann, bei geeigneter Wahl der Parameter auch für Beziehung (3-5) erhalten bleibt [Rei95]. Die negativen Lösungen in (3-6) können aufgrund der positiven Initialisierung in diesem Ansatz nicht erreicht werden. Weiterhin ist der Grad des verstärkend wirkenden Terms eine Ordnung niedriger als der Grad der dämpfend wirkenden Terme, so daß ein exponentielles Wachstum der Variablen ausgeschlossen ist.

Ein wesentlicher Nachteil des Ansatzes von Reimann und Haken ist, daß im stationären Zustand an jeder Bildkoordinate eine Variable einen Wert ungleich Null annimmt. Somit exi-

stiert für jede Bildkoordinate ein Gewinner des Wettbewerbs, der eine Korrespondenz repräsentiert. Dies gilt auch für Okklusionsbereiche, in denen keine Korrespondenz existiert.

Zur Beurteilung der Leistungsfähigkeit stereoskopischer Korrespondenzbestimmungen werden im weiteren sogenannte Zufallspunktstereogramme verwendet. Diese synthetischen Bilder wurden im wesentlichen von Julesz [Jul60] eingeführt und haben sich seit langem als Testmuster in der Stereoskopie etabliert. Bei einem Stereogramm dieser Art werden mit Hilfe eines Zufallsprozesses zwei identische Bilder erzeugt, die aus zufällig verteilten Nullen und Einsen bestehen. Anschließend wird ein zweites kleineres Bild beliebiger Form auf dem gleichen Wege erzeugt und mit einem relativen Versatz in die beiden zuvor generierten Bilder eingefügt. Diese sich nur durch den versetzten Bereich unterscheidenden Bilder repräsentieren das linke und rechte Bild des Stereogramms.

Daß der Mensch allein aufgrund stereoskopischer Informationen zu räumlichem Sehen befähigt ist, läßt sich anhand eines Zufallspunktstereogramms leicht feststellen. Wird das linke Bild in Abbildung 3-1 mit dem linken Auge betrachtet und das rechte Bild mit dem rechten Auge, eine Aufgabe, die etwas Übung erfordert und durch ein Blatt Papier, das aufrecht zwischen die beiden Bilder gestellt wird, erleichtert werden kann, so entsteht nach kurzer Zeit der Eindruck, daß ein Quadrat aus dem Hintergrund hervortritt.

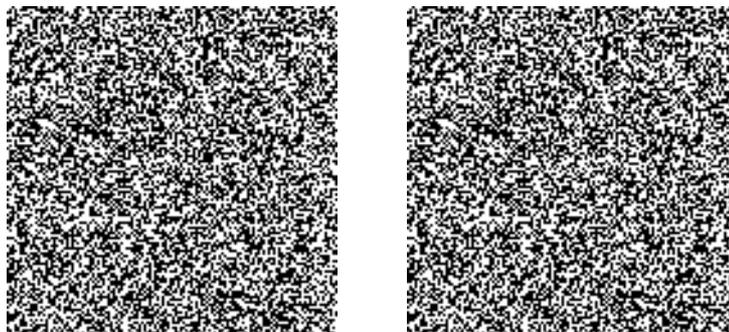


Bild 3-1: Zufallspunktstereogramm nach Julesz. Bei stereoskopischer Betrachtung entsteht der Eindruck eines aus dem Hintergrund hervortretenden Quadrats.

Die Probleme, die durch Okklusionen bei fast allen Verfahren zur stereoskopischen Korrespondenzbestimmung hervorgerufen werden, sollen anhand des in Bild 3-1 gezeigten Stereogramms verdeutlicht werden. Zu diesem Zweck wird der Ansatz von Reimann und Haken auf dieses Beispiel angewendet. Als Resultat sind die Disparitätskarten in der nachfolgenden Grafik zu sehen. Die Disparitätswerte sind hier als Grauwerte kodiert dargestellt. Hierbei bedeuten hellere Grautöne eine größere Disparität als dunklere. Die Disparitätskarten beziehen sich auf das linke Grauwertbild in Bild 3-1.

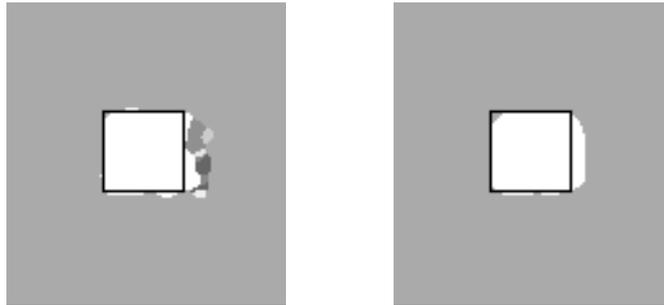


Bild 3-2: Nach dem Ansatz von Reimann und Haken berechnete Disparitätskarten zum Stereogramm in Bild 3-1. a) Disparitätskarte bei kleinem verstärkend wirkendem Umfeld. b) Disparitätskarte bei großem Umfeld

Der Umriß des verschobenen Quadrates ist in Bild 3-2 zur Verdeutlichung nachträglich als schwarzer Rahmen in die Disparitätskarten eingefügt worden. Da hier das Quadrat nur in horizontaler Richtung verschoben wurde, ist auch nur die Disparitätskarte der d_x -Komponente dargestellt. Auffallend an den Disparitätskarten ist der Bereich rechts neben dem Quadrat. Hier treten Disparitätswerte auf, die im Stereogramm nicht vorhanden sind. Es handelt sich bei diesem Bildbereich um ein Okklusionsgebiet, das im linken Bild zu sehen ist und im rechten Bild durch das eingefügte Quadrat verdeckt wird. In diesem Bereich ist die Disparität nicht definiert. Der Ansatz von Reimann und Haken ergibt hier Disparitätswerte, deren zugehörige Variablenwerte bei der Initialisierung im Mittel etwas größer als die übrigen Variablenwerte waren. Die Größe der Umgebung U in Beziehung (3-5) bewirkt, daß sich bei kleinerer Umgebung lokale Gruppen von falschen Disparitätswerten herausbilden oder bei größerer Umgebung, die Disparitätswerte von Vorder- und Hintergrund auf ebenso falsche Weise zusammenwachsen (siehe Bild 3-2a bzw. b). Um Korrespondenzfehler zu verhindern, wird in dem Ansatz von Reimann und Haken der Parameter $\lambda(\mathbf{x}, \mathbf{d})$ individuell für jede Variable gleich dem Wert des Ähnlichkeitsmaßes $s(\mathbf{x}, \mathbf{d})$ gewählt. Auf diese Weise erreichen Variablenwerte, deren zugehörige Bildbereiche einander sehr ähnlich sind, schneller die stationäre Lösung als andere Variablen und gewinnen auf diese Weise den Wettbewerb leichter. Die Werte, welche die Variablen im stationären Zustand annehmen, sind in diesem Modell um so größer je größer das Ähnlichkeitsmaß $s(\mathbf{x}, \mathbf{d})$ ist (siehe Beziehung (3-6)). Eine nachträglich auf die Variablenwerte angewandte Schwellwertoperation extrahiert anschließend nur die Disparitäten, die eine entsprechend hohe stationäre Lösung angenommen haben.

Die Annahme, daß Zuordnungen von Okklusionen mit Hilfe eines Ähnlichkeitsmaßes detektiert werden können, ist jedoch nicht allgemeingültig. Darüber hinaus sind aufgrund der geringeren verstärkenden Wirkung des Umfeldes die Variablenwerte im stationären Zustand entlang von Diskontinuitäten im Disparitätsverlauf generell niedriger als bei Variablen mit konformem Umfeld, so daß auch diese Werte von der Schwellwertoperation unterdrückt werden (siehe Beziehung (3-6)).

4 Vom Kamerabild zur Tiefenkarte

Zur Ermittlung der räumlichen Struktur einer Szene anhand von Stereobildern müssen generell mehrere Verfahrensschritte und Systemkomponenten miteinander verknüpft und aufeinander abgestimmt werden. So ist ein Stereoverfahren beispielsweise grundsätzlich an die Geometrie der eingesetzten Kameras gebunden. Ist diese nicht bekannt, so steigt einerseits der Aufwand bei der stereoskopischen Korrespondenzbestimmung beachtlich und andererseits ist eine quantitative Vermessung der betrachteten Szene nicht direkt möglich. Weiterhin sind die Kamerabilder im allgemeinen durch Linsenfehler verzerrt und mit Rauschprozessen überlagert, so daß die reinen Grauwerte zu einer direkten Korrespondenzbestimmung wenig geeignet sind. Um Störeinflüsse dieser Art zu unterdrücken, ist es daher sinnvoll, die Bilddaten vor der Korrespondenzbestimmung durch spezielle Kompensationstechniken und Filterverfahren geeignet aufzubereiten.

Obwohl der zentrale Gegenstand der vorliegenden Arbeit die stereoskopische Korrespondenzbestimmung und die Behandlung der damit verbundenen Probleme ist, entfällt ein nicht unwesentlicher Teil der Arbeit auf die obengenannten zusätzlichen Aspekte, die bei der Realisierung eines Systems zur räumlichen Vermessung zu berücksichtigen sind. In diesem Kapitel werden nun die einzelnen Systemkomponenten betrachtet, die im Rahmen eines Gesamtsystems erforderlich sind, um aus zwei Kamerabildern die räumliche Struktur einer Szene zu rekonstruieren. Die Reihenfolge, in der die einzelnen Komponenten und Verfahrensschritte aufgeführt werden, richtet sich dabei im wesentlichen nach dem Informationsfluß, angefangen bei der Bildaufnahme bis hin zur Tiefenkarte.

4.1 Das aktive binokulare Sehsystem

Seitdem gezeigt werden konnte, daß gewisse Aufgaben im Bereich der Bildverarbeitung durch Verwendung von Kamerasystemen, deren Abbildungseigenschaften und deren Lage im Raum aktiv einflußbar sind, wesentlich vereinfacht werden können [AWB87, Baj85], wurden an mehreren Universitäten unterschiedliche Prototypen von aktiven Sehsystemen entwickelt [AA92, Chr92, Fer92, JMT92, PE92]. Viele dieser Konstruktionen lehnen sich an biologische Systeme an und nutzen diese als Vorbild.

Da das Blickfeld einer Kamera aufgrund der endlichen Ausdehnung der Sensorfläche begrenzt ist, haben aktive Sehsysteme gegenüber statischen Systemen den Vorteil, daß sie ein wesentlich größeres Erfassungsgebiet explorieren können. Darüber hinaus können mit einem solchen System Aufgaben aus dem Bereich der Bewegungsanalyse, wie z. B. die Bewegungsverfolgung, effizient gelöst werden. In der Stereoskopie hat ein Sehsystem mit aktiv ausrichtbaren optischen Achsen den Vorteil, daß der Überlappungsbereich der Stereobilder maximiert und gleichzeitig der Wertebereich der auftretenden Disparitäten und damit der Suchaufwand minimiert werden kann. Um mit einem solchen System genaue Positionsbestimmungen durchführen zu können, muß jedoch die Lage der Kameras mit entsprechender Genauigkeit

meßtechnisch erfaßt werden können. Sind weiterhin die optischen Eigenschaften des Kamerasystems, wie z.B. Fokus, Zoom oder Blende aktiv beeinflussbar, so kann das System sehr flexibel an seine Umgebung angepaßt werden.

Die aktiven Sehsysteme, die zum gegenwärtigen Zeitpunkt kommerziell verfügbar sind, besitzen bei weitem nicht die Positionierungsgenauigkeiten, die für Anwendungen wie die passive Triangulation erforderlich sind. Die bislang existierenden Prototypen sind dahingegen entweder extrem aufwendig [JMT92, PE92] und dementsprechend nur unter erheblichem finanziellen Aufwand zu fertigen, oder sie stellen nur reine Experimentierumgebungen mit ebenfalls fehlender Genauigkeit dar [AA92, Chr92, Fer92]. Aus diesen Gründen wurde im Verlauf der Arbeit speziell unter den Gesichtspunkten der Positionierungsgenauigkeit und des Kostenaufwandes ein aktives Sehsystem entwickelt. In diesem Kapitel erfolgt nun eine kurze Beschreibung der Systemarchitektur und ihrer Besonderheiten. Eine ausführliche Dokumentation der technischen Details des Systems ist in [Tra96b, TD96] zu finden.

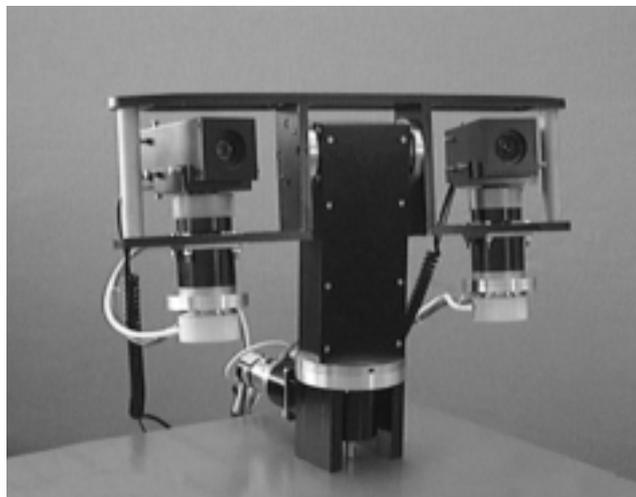


Bild 4-1: Aktives binokulares Kamerasystem zur Erfassung räumlicher Strukturen anhand von Stereobilddaten

4.1.1 Systemüberblick

Das in Bild 4-1 dargestellte Sehsystem besitzt vier mechanische Freiheitsgrade, die es ermöglichen, die optischen Achsen der beiden Kameras flexibel und sehr präzise auszurichten. Die Motoren, die zu diesem Zweck die Kameras antreiben, werden nach einem sogenannten Microstepping-Verfahren angesteuert und über optische Meßwertgeber in einem geschlossenen Regelkreis betrieben. Beide Kameras können unabhängig voneinander um eine vertikale Vergenzachse bewegt werden, wodurch eine Objektfixation mit unsymmetrischen Vergenzwinkeln möglich ist. Weiterhin können die Kameras um eine gemeinsame Neigeachse bewegt werden, die zur Ausrichtung der optischen Achsen in vertikaler Richtung dient. Eine zusätzliche gemeinsame Schwenkachse, die in der Mitte zwischen den beiden Vergenzachsen angeordnet

ist, ermöglicht eine Objektfixation mit stets symmetrischen Vergenzwinkeln und kann perspektivische Verzerrungen bei extremen Konstellationen verringern (vgl. Bild 4-2).

Während die Motoren für die Vergenzbewegungen die Kameras direkt antreiben, werden die Schwenk- und Neigebewegungen indirekt über losearme Winkelgetriebe ausgeführt. Diese Realisierungsmöglichkeit zur Erzeugung des für die obengenannten Bewegungen erforderlichen Drehmomentes wurde trotz der Ungenauigkeit durch die verbleibende Getriebelose zugunsten eines einfachen modularen Aufbaus gewählt.

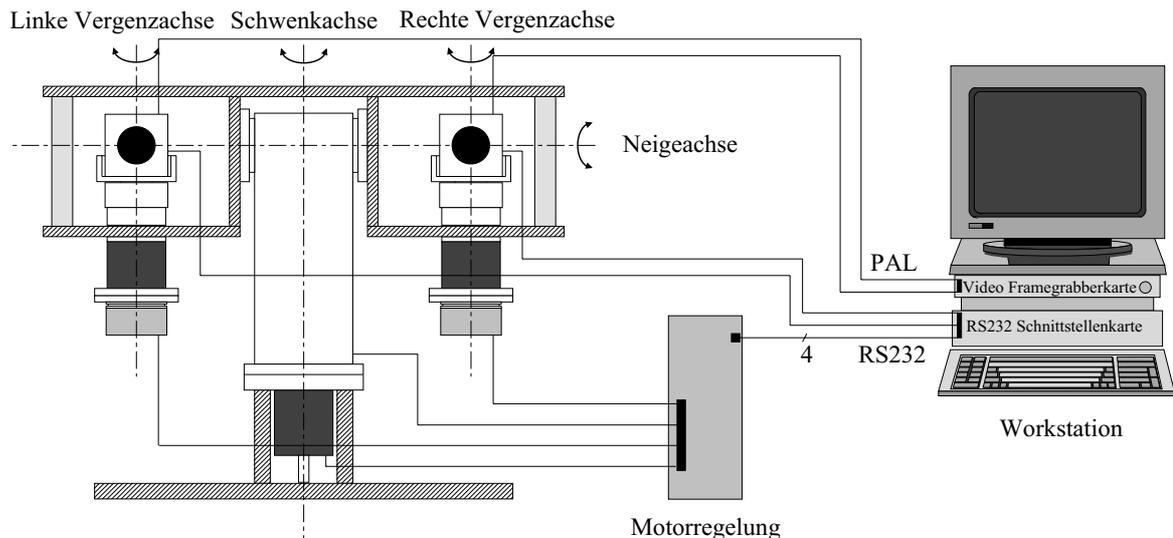


Bild 4-2: Konfiguration der Systemkomponenten

Auf einen weiteren Freiheitsgrad in Form einer Zyklotorsionsbewegung, welche die Möglichkeit bietet, beide Kameras um die optischen Achsen zu drehen, wurde bei dieser Konstruktion aufgrund des erforderlichen Mehraufwandes verzichtet. Eine Zyklotorsionsbewegung dient zur aktiven Ausrichtung der Epipolarlinien (siehe Kapitel 2.4.1) und wird sowohl in technischen als auch in biologischen Sehsystemen zur Aufwandsreduzierung bei stereoskopischen Anwendungen genutzt [JMT92, JT94]. Weitere Freiheitsgrade zur automatischen Kompensation der Projektionszentrumsdrift, die durch Änderung der optischen Freiheitsgrade hervorgerufen wird, sind in dieser Konstruktion ebenfalls nicht vorgesehen. Im Gegensatz zum sehr aufwendigen KTH-System [PE92], bei dem diese Drift mechanisch kompensiert wird, erfolgt bei der hier vorgestellten Konstruktion zunächst eine Kalibrierung, mit der die Drift im Vorfeld ermittelt wird. Während des Betriebs wird die Projektionszentrumsdrift abhängig vom Kamerazustand aus Tabellen ermittelt und z. B. bei der Tiefenrekonstruktion berücksichtigt.

Das Gesamtsystem ist für einen statischen „Look-and-Move“-Betrieb konzipiert, bei dem die Sollwerte für die Positionierung und die Bewegungsparameter über ein RS232-Protokoll von einer Sun-Workstation vorgegeben werden. Diese kaskadische Kontrollstruktur ist so ausgelegt, daß jeder Zyklus von Bildaufnahme und Bewegung sein eigenes von vorhergehenden Zyklen unabhängiges dynamisches Verhalten aufweist.

Die Kameramodule besitzen flexible Mehrlinsensysteme mit jeweils drei optischen Freiheitsgraden. Über das RS232-Protokoll können Fokus, Zoom und Blende kontrolliert werden. Neben der Regelung der Linsen- und Blendenpositionen verfügt ein auf dem Kameraboard integrierter Mikrocontroller über Verfahren zur automatischen Fokus- und Blendeneinstellung, die bei Bedarf zugeschaltet werden können. Die Kameras liefern ein PAL-Signal, das mit Hilfe eines SBus Frame-Grabbers digitalisiert wird.

Die Positionierungsgenauigkeit der einzelnen Antriebe ist in der nachfolgenden Tabelle dargestellt. Die Winkelauflösung der Achsen mit Getriebe entspricht hierbei lediglich der minimalen Schrittweite. Da die Drehgeber aus Platzgründen nicht auf die Ausgangswelle der Getriebe montiert werden konnten, wurden sie an der Motorwelle befestigt. Winkeldifferenzen, die durch die Getriebelose entstehen, können daher nicht durch die Motorregelung kompensiert werden. Die reale Positionierungsgenauigkeit der Schwenk- und Neigebewegung ist deshalb im allgemeinen geringer als die in der Tabelle angegebene Winkelauflösung.

Achse	Winkelauflösung	Arbeitsbereich	Getriebeübersetzung
Vergenzachsen	0,02197°	±60°	Direktantrieb
Neigeachse	0,009°	±90°	10:1
Schwenkachse	0,009°	±90°	10:1

Tabelle 4-1: Positionierungsgenauigkeiten und Arbeitsbereiche der einzelnen Achsen

Ein wesentlicher Vorteil dieser Konstruktion gegenüber den meisten existierenden Prototypen und den kommerziell verfügbaren Systemen besteht in dem Direktantrieb der Vergenzachsen. Um eine möglichst hohe Winkelauflösung zu erzielen, werden bei den meisten aktiven Sehsystemen die Kameras über entsprechende Getriebe um die Vergenzachsen gedreht. Eine Rückkopplung der Winkelposition auf der Getriebeseite ist dabei aus Platzgründen im allgemeinen nicht möglich, wodurch die Getriebelose nicht erfaßt werden kann (siehe z.B. [AA92, Chr92, Fer92, JMT92]). Da die Genauigkeit, mit der die Vergenzwinkel gemessen werden können, erheblich die Genauigkeit der Tiefenrekonstruktion beeinflußt, wurde bei diesem System ein Direktantrieb verwendet. Durch den Einsatz von Hybrid-Schrittmotoren, die über ein Microstepping-Verfahren angesteuert werden, und hochauflösenden Drehgebern konnte trotz Direktantrieb eine angemessene Winkelauflösung erzielt werden.

4.1.2 Besonderheiten der Konstruktion

Um die optische Achse einer Kamera abhängig von der Bildinformation auf einen Raumpunkt ausrichten zu können, ist es von Vorteil, wenn die Kamera um das Projektionszentrum des Lochkameramodells gedreht wird. Nur in dieser Konstellation wird eine rein rotatorische Bewegung des Kameramodells durchgeführt und die Ausrichtung der optischen Achse kann unabhängig von der Entfernung des betreffenden Raumpunktes erfolgen.

Da einerseits die Position des Projektionszentrums im allgemeinen nicht in den Herstellerangaben zu finden ist, insbesondere dann nicht, wenn es sich um kostengünstige Kameramodule handelt, und andererseits die Lage des Projektionszentrums von den optischen Freiheitsgraden des Kamerasystems abhängt, ist eine flexible Konstruktion notwendig, wenn das System in der oben beschriebenen Konstellation betrieben werden soll. Zu diesem Zweck wurde die Konstruktion mit einer besonders flexiblen Kameraaufhängung versehen. So ermöglichen drei translatorische Justier Vorrichtungen eine präzise Positionierung des Projektionszentrums der Kamera im Schnittpunkt der Vergenz- und der Neigeachse.

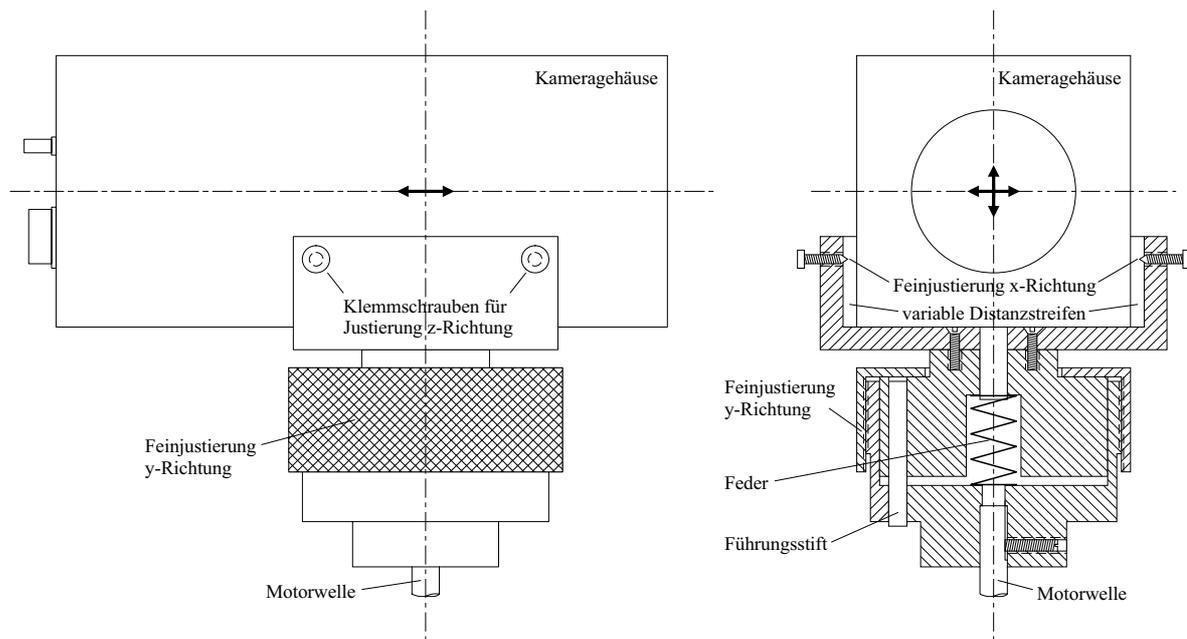


Bild 4-3: Translatorische Positionierung des Kameramoduls

Da die absolute Position des Projektionszentrums von Zoom und Fokus abhängt, kann eine Justierung im allgemeinen nur für einen bestimmten Kamerazustand erfolgen. Wie die Kalibrierung des Systems jedoch zeigte, verschiebt sich das Projektionszentrum bei Veränderung des Fokus nur minimal, so daß bei konstantem Zoom seine Lage als näherungsweise konstant angesehen werden kann.

Über einen speziellen Mechanismus kann das Projektionszentrum der Kameras in vertikaler Richtung positioniert werden. In einer Führung, die gegen Verwindung gesichert ist, kann hierzu das Kameragehäuse über ein Gewinde entlang der Vergenzachse verschoben werden. In Querrichtung läßt sich das Kameragehäuse durch unterschiedliche Distanzstreifen und in Längsrichtung innerhalb einer Klemmvorrichtung positionieren (vgl. Bild 4-3).

Drei weitere rotatorische Justierungseinrichtungen erlauben es, die optischen Achsen der Kameras so auszurichten, daß sie näherungsweise in einer Ebene liegen. Innerhalb der Gehäuse sind die Kameras zu diesem Zweck flexibel befestigt, so daß die Schwenk-, Neige- und Kippwinkel mit Hilfe der Justierungsschrauben eingestellt werden können (vgl. Bild 4-4).

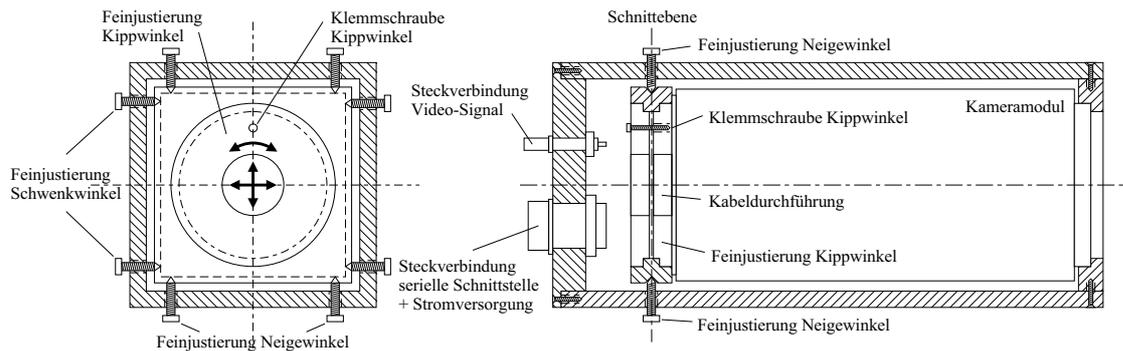


Bild 4-4: Rotatorische Justierungsmöglichkeiten

Diese flexiblen Justierungsmöglichkeiten erlaubten bei der hier vorgestellten Konstruktion den Einsatz von weniger hochwertigen und daher kostengünstigen Kameramodulen, obwohl deren Abbildungsgeometrie von Exemplar zu Exemplar teilweise stark variierte.

4.2 Die Kalibrierung des Kamerasystems

Die hier eingesetzten Kamerasysteme haben den Vorteil, daß sie sehr flexibel an die Umwelt angepaßt werden können. So können z. B. Objekte automatisch fokussiert werden oder durch eine verstellbare Blende können schwankende Beleuchtungsverhältnisse kompensiert werden. Derartig komplexe Mehrlinsensysteme, die diese Möglichkeiten aufweisen, lassen sich jedoch nicht mehr mathematisch geschlossen beschreiben. Zur Modellierung der bildgebenden Prozesse solcher Linsensysteme wird im allgemeinen die Zentralprojektion des Lochkameramodells herangezogen. Dieses weit verbreitete Modell ist zwar mathematisch sehr einfach handhabbar, jedoch gelten die jeweiligen Modellparameter immer nur für einen bestimmten Kamerazustand. Da sich darüber hinaus die Kamerasysteme nur sehr grob durch die einfache Zentralprojektion beschreiben lassen, sind in der Fachliteratur zahlreiche Erweiterungen dieses Modells zu finden, die zusätzlich diverse Linsenverzeichnungen berücksichtigen [CR93, Len87, LT87, Li94, MC+93, SA94, Tsa87, WCH92].

Da das in Kapitel 4.1 beschriebene aktive Kamerasystem unter anderem die Grundlage für die räumliche Vermessung bilden soll, ist eine genaue Modellierung der bildgebenden Prozesse dieses Systems unumgänglich. Im folgenden werden daher die Modellierung der Kameras und die Ermittlung der Modellparameter im Rahmen einer aufwendigen Kalibrierungsprozedur näher erläutert. Die Kalibrierung der Modellparameter muß auch bei den hier eingesetzten Kamerasystemen für die verschiedenen Systemzustände erfolgen, da das verwendete Modell von den optischen Freiheitsgraden der Kameras abhängt. Eine Kalibrierung der Modellparameter während der Laufzeit würde sehr hohe Anforderungen an das Kalibrierungsverfahren stellen und müßte in Echtzeit erfolgen. Andererseits werden die Bewegungen des Kamerasystems im Rahmen der Regelung meßtechnisch erfaßt, so daß die veränderbaren extrinsischen Parameter relativ zu einer Referenzposition gemessen werden können. Die Kalibrierung muß daher nur für eine Kameraposition durchgeführt werden.

Die optischen Freiheitsgrade des eingesetzten Mehrlinsensystems sind zwar ebenfalls meßbar, sie entsprechen jedoch nicht den intrinsischen oder den extrinsischen Parametern des Lochkameramodells. Diese Parameter müssen deshalb für verschiedene Systemzustände bestimmt und tabelliert werden.

4.2.1 Das mathematische Modell

Die Modellierung der bildgebenden Prozesse der Kamera erfolgt durch eine erweiterte Version des klassischen Lochkameramodells. Das Modell beschreibt den Prozeß, durch den ein Punkt \mathbf{q}_w , dessen Koordinaten sich auf ein dreidimensionales Weltkoordinatensystem beziehen, in die Bildebene der Kamera projiziert wird. Dieser Abbildungsprozeß erfolgt über mehrere Zwischenschritte, die im folgenden näher erläutert werden. Die im weiteren verwendeten Koordinatensysteme sind, soweit nicht anders vermerkt, kartesische rechtshändige Koordinatensysteme.

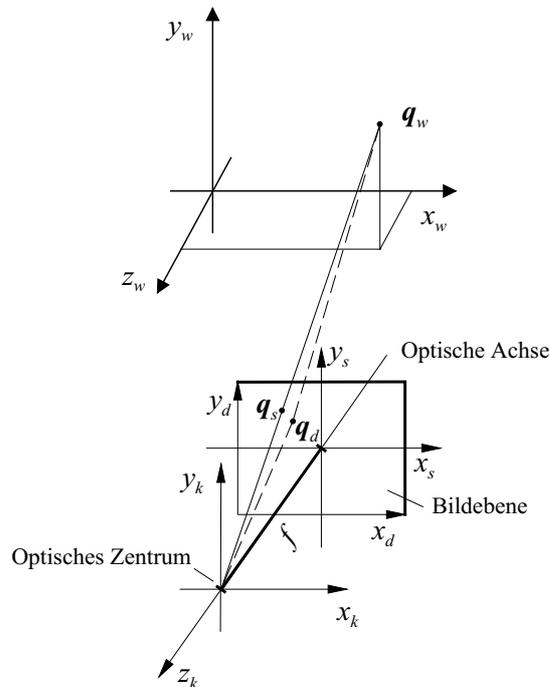


Bild 4-5: Erweitertes Lochkameramodell mit den einzelnen Koordinatensystemen

Um die Position in der Bildebene zu bestimmen, auf die ein beliebiger Punkt in der betrachteten Szene abgebildet wird, werden zunächst die Weltkoordinaten $\mathbf{x}_w = [x_w \ y_w \ z_w]^T$ mittels einer affinen Abbildung in die sogenannten Kamerakoordinaten $\mathbf{x}_k = [x_k \ y_k \ z_k]^T$ überführt.

$$\mathbf{x}_k = \mathbf{R}\mathbf{x}_w + \mathbf{t} \quad (4-1)$$

Der Ursprung dieses Koordinatensystems liegt im optischen Zentrum der betreffenden Kamera, wobei die z_k -Achse identisch mit der optischen Achse des Kameramodells ist (vgl. Bild 4-5). Die Matrix \mathbf{R} ist hierbei eine orthonormale Rotationsmatrix mit

$$\mathbf{R} = \begin{bmatrix} \cos \gamma \cos \beta & \cos \gamma \sin \beta \sin \alpha + \sin \gamma \cos \alpha & \sin \gamma \sin \alpha - \cos \gamma \sin \beta \cos \alpha \\ -\sin \gamma \cos \beta & \cos \gamma \cos \alpha - \sin \gamma \sin \beta \sin \alpha & \sin \gamma \sin \beta \cos \alpha + \cos \gamma \sin \alpha \\ \sin \beta & -\cos \beta \sin \alpha & \cos \beta \cos \alpha \end{bmatrix} \quad (4-2)$$

und $\mathbf{t} = [t_x \ t_y \ t_z]^T$ ist ein geeigneter Translationsvektor. Durch die Abbildungsgleichung des Lochkameramodells (vgl. Beziehung (2-1)) werden die Kamerakoordinaten \mathbf{x}_k in einem weiteren Schritt in die zweidimensionalen Sensorebenenkoordinaten $\mathbf{x}_s = [x_s \ y_s]^T$ abgebildet.

$$x_s = -f \frac{x_k}{z_k}; \quad y_s = -f \frac{y_k}{z_k} \quad (4-3)$$

Im Fall des idealen Linsensystems werden die Sensorebenenkoordinaten \mathbf{x}_s anschließend durch den Bildsensor auf die unverzerrten diskreten Bildkoordinaten $\mathbf{x}_u = [x_u \ y_u]^T$ überführt.

$$x_u = x_0 + a_x^{-1} x_s; \quad y_u = y_0 + a_y^{-1} y_s \quad (4-4)$$

Die Parameter x_0 und y_0 sind hierbei die unverzerrten Bildkoordinaten des Punktes, an dem die optische Achse des Kameramodells die Sensorebene durchstößt. Dieser Punkt wird als Hauptpunkt bezeichnet. Die Parameter a_x und a_y stellen den horizontalen bzw. vertikalen Abstand zweier benachbarter Sensorelemente dar. Durch die Normierung der Sensorparameter in (4-4) auf diese Größen, haben die Bild- und Hauptpunktkoordinaten die normierte Einheit Pixel (Picture elements). Der Ursprung des Bildkoordinatensystems ist die linke untere Bildecke. Zur Vereinfachung werden die effektive Brennweite f und die Sensorelementabstände a_x , a_y zu den normierten effektiven Brennweiten f_x und f_y zusammengefaßt.

$$f_x = a_x^{-1} f; \quad f_y = a_y^{-1} f \quad (4-5)$$

Da die eingesetzten Kameras die Möglichkeit bieten, auch relativ kleine Brennweiten einzustellen, treten insbesondere in diesem Zustand beachtliche Verzeichnungsfehler auf. Neben der kleinen Brennweite ist die niedrige Qualität der verwendeten Objektive, die hier aus Kostengründen eingesetzt werden, eine weitere Ursache für diese Fehler. Um die Verzeichnungsfehler

in einem späteren Verarbeitungsschritt kompensieren zu können, wird deshalb zusätzlich die sogenannte radiale Linsendistorsion modelliert. Weitere, hier aus Aufwandsgründen nicht berücksichtigte Verzeichnungsarten sind die tangentielle Linsendistorsion und die dünne Prismadistorsion (siehe [CR93, Li94, MC+93, SA94, WCH92]).

Die durch die Linsendistorsion verzerrten Bildkoordinaten \mathbf{x}_d ergeben sich aus den unverzerrten Bildkoordinaten zu:

$$\mathbf{x}_d = \mathbf{x}_u + \boldsymbol{\tau} \quad (4-6)$$

Die radiale Linsendistorsion wird in der Literatur im allgemeinen durch die folgende implizite Darstellung modelliert (siehe z.B. [Len87, LT87, Li94]):

$$\boldsymbol{\tau} = \begin{bmatrix} (x_d - x_0)(\kappa_1 r_t^2 + \kappa_2 r_t^4 + \dots) \\ (y_d - y_0)(\kappa_1 r_t^2 + \kappa_2 r_t^4 + \dots) \end{bmatrix} \quad \text{mit:} \quad r_t = \sqrt{(x_d - x_0)^2 + (y_d - y_0)^2} \quad (4-7)$$

Da jedoch die Linsendistorsion durch eine Entzerrung der Bilddaten im Rahmen der Bildaufbereitung kompensiert werden soll, wie in Kapitel 4.3.1 noch näher erläutert wird, ist eine explizite Formulierung der verzerrten Bildkoordinaten mit

$$\mathbf{x}_d = f(\mathbf{x}_u) \quad (4-8)$$

erforderlich. Um diesen Zusammenhang aus den Beziehungen (4-6) und (4-7) zu erhalten, müßte ein gekoppeltes Gleichungssystem dritten Grades für jeden Bildpunkt und nach jeder Bildaufnahme gelöst werden. Dieser zusätzliche Berechnungsaufwand kann vermieden werden, wenn die radiale Linsendistorsion durch die folgende explizite Darstellung modelliert wird.

$$\boldsymbol{\tau} = \kappa \begin{bmatrix} (x_u - x_t)((x_u - x_t)^2 + (y_u - y_t)^2) \\ (y_u - y_t)((x_u - x_t)^2 + (y_u - y_t)^2) \end{bmatrix} \quad (4-9)$$

Die Terme höherer Ordnung (siehe Beziehung (4-7)) werden in diesem Ansatz vernachlässigt, da sich einerseits durch diesen zusätzlichen Mehraufwand nur geringe Verbesserungen in der

Modellierung erzielen lassen und andererseits die teilweise sehr hohen Beträge dieser Komponenten leicht zur numerischen Instabilität bei der Ermittlung der Parameter führen können (siehe [Tsa87]). Eine weitere Besonderheit dieses Ansatzes besteht in der Modellierung des Torsionszentrums $\mathbf{x}_t = [x_t \ y_t]^T$. Während das Torsionszentrum \mathbf{x}_t im allgemeinen gleich dem Hauptpunkt \mathbf{x}_0 gewählt wird [CR93, LZB95, Len87, LT87, Li94, MC+93, NM+92, SA94, Tsa87, WCH92], werden hier zwei weitere Parameter vorgesehen, die den Ort des Torsionszentrums separat beschreiben. Durch diese zusätzlichen Parameter können die Ungenauigkeiten, die bei der Montage und der Fertigung der Linsen auftreten, berücksichtigt und später kompensiert werden (siehe Kapitel 4.3.1).

Aus der Kombination der Beziehungen (4-1) - (4-6) ergeben sich die sogenannten Kollinearitätsgleichungen, die den gesamten Abbildungsprozeß des Kameramodells beschreiben.

$$x_d = x_0 + \tau_x - f_x \frac{r_{11}x_w + r_{12}y_w + r_{13}z_w + t_x}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} \quad (4-10)$$

$$y_d = y_0 + \tau_y - f_y \frac{r_{21}x_w + r_{22}y_w + r_{23}z_w + t_y}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} \quad (4-11)$$

Die Parameter r_{ij} sind hierbei die Komponenten der Rotationsmatrix \mathbf{R} aus Beziehung (4-2). Als extrinsische Parameter sind somit die drei Komponenten des Translationsvektors \mathbf{t} und die drei Eulerwinkel α , β und γ zu kalibrieren. Die zu bestimmenden intrinsischen Parameter sind die Hauptpunktkoordinaten \mathbf{x}_0 , die normierten effektiven Brennweiten f_x , f_y und die radiale Linsendistorsion τ mit ihren Zentrumskoordinaten \mathbf{x}_t . Die Parameter werden einem $(M \times 1)$ Vektor \mathbf{p} wie folgt zugeordnet:

$$\mathbf{p} = [x_0 \ y_0 \ f_x \ f_y \ t_x \ t_y \ t_z \ \alpha \ \beta \ \gamma \ \kappa \ x_t \ y_t]^T \quad (4-12)$$

4.2.2 Bestimmung der Modellparameter

Die Parameter des Kameramodells können aus den Kollinearitätsgleichungen (4-10) und (4-11) ermittelt werden, wenn eine geeignete Menge von Meßpunkten vorliegt. Hierzu müssen die Weltkoordinaten der Meßpunkte und die Bildkoordinaten der zugehörigen Abbildungen bekannt sein. Die Meßpunktmenge muß darüber hinaus so gewählt werden, daß ein nicht unterbestimmtes Gleichungssystem aufgestellt werden kann. Die zur Kalibrierung der hier eingesetzten Kameras erforderlichen Meßpunkte liefert eine eigens zu diesem Zweck konstruierte Meßvorrichtung.

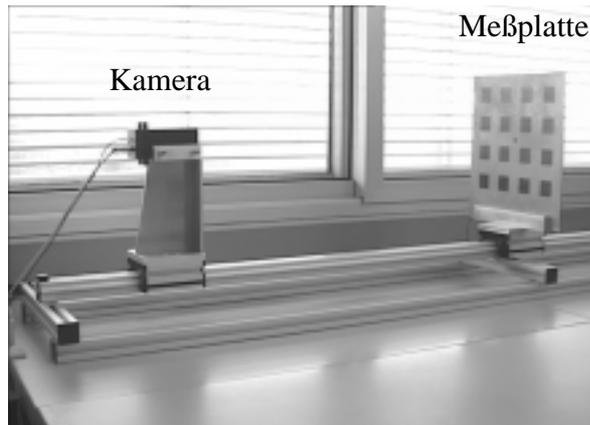


Bild 4-6: Meßvorrichtung zur Kalibrierung von Kamerasystemen. Die Ecken der schwarzen Quadrate bilden die Meßpunkte für die Kalibrierung.

Die Meßpunkte bilden bei dieser Vorrichtung die Ecken schwarzer Quadrate, die phototechnisch auf eine Meßplatte aufgebracht wurden (vgl. Bild 4-6). Während die Weltkoordinaten der Meßpunkte relativ genau vermessen werden können, müssen die zugehörigen Bildkoordinaten, auf welche die Meßpunkte abgebildet werden, in den Grauwertbildern zunächst detektiert werden. Um bei dieser Prozedur eine möglichst hohe Genauigkeit zu erreichen, sind aufwendige Verfahren notwendig, die eine subpixelgenaue Detektion der Meßpunkte erlauben (siehe [Tra96b, TD96]). Die unbekanntes Modellparameter werden unter Berücksichtigung des Meßrauschens durch ein im folgenden beschriebenes Optimierungsverfahren ermittelt.

Sei $\mathbf{x}_m = [x_1 \ y_1 \ x_2 \ y_2 \ \dots \ x_N \ y_N]_d^T$ der störungsfreie $(2N \times 1)$ Koordinatenvektor von N Meßpunkten in verzerrten Bildkoordinaten und sei $\mathbf{x}'_m = \mathbf{x}_m + \mathbf{v}$ der korrespondierende mit dem $(2N \times 1)$ Fehlervektor \mathbf{v} behaftete Meßvektor, so gilt:

$$\mathbf{x}'_m = \mathbf{f}(\mathbf{p}) + \mathbf{v} \quad (4-13)$$

Wobei $\mathbf{f}(\mathbf{p})$ einen Vektor darstellt, dessen Komponenten sich aus den Kollinearitätsgleichungen durch Einsetzen der jeweiligen Meßpunktweltkoordinaten und des Parametervektors \mathbf{p} errechnen. Das Ziel ist es nun den $(M \times 1)$ Parametervektor \mathbf{p} so zu bestimmen, daß der Fehler \mathbf{v} minimiert wird.

Dieses Problem der nichtlinearen Optimierung kann durch Minimierung der quadratischen Form

$$e_k = \mathbf{v}^T \mathbf{W} \mathbf{v} \quad (4-14)$$

beschrieben werden. Die Matrix \mathbf{W} ist hierbei eine positiv definite, symmetrische $(2N \times 2N)$ Gewichtsmatrix. Im allgemeinen sollte \mathbf{W} die Inverse der Kovarianzmatrix $\mathbf{\Sigma}_v$ des Meßrauschens \mathbf{v} sein. Wenn das Meßrauschen \mathbf{v} mittelwertfrei ist und die einzelnen Meßwerte unkorreliert sind, wie vernünftigerweise angenommen werden kann, so ist die Kovarianzmatrix $\mathbf{\Sigma}_v$ eine Diagonalmatrix. Werden alle Messungen mit der gleichen Genauigkeit durchgeführt, so entspricht die Matrix \mathbf{W} einer Multiplikation der Einheitsmatrix \mathbf{I} mit einer skalaren Größe. Die Minimierung von e_k wird auch als ein verallgemeinertes Verfahren der kleinsten Fehlerquadrate bezeichnet.

Eine Möglichkeit den Parametervektor \mathbf{p} zu bestimmen, stellt das Gauß-Newton-Verfahren für nichtlineare Ausgleichsprobleme [FH84] dar. Dieses Verfahren wird in der Photogrammetrie auch als Bündelausgleichung bezeichnet [AK89, Kra84]. Bei diesem Ansatz wird die Funktion f zunächst durch eine Taylor-Reihe um einen Parameterwert \mathbf{p}_0 linearisiert, indem Terme höherer Ordnung vernachlässigt werden.

$$f(\mathbf{p}) \approx f(\mathbf{p}_0) + \left. \frac{\partial}{\partial \mathbf{p}} f(\mathbf{p}) \right|_{\mathbf{p}=\mathbf{p}_0} (\mathbf{p} - \mathbf{p}_0) \quad (4-15)$$

Für den Meßvektor aus (4-13) folgt somit:

$$\mathbf{x}'_m = f(\mathbf{p}_0) + \mathbf{J}(\mathbf{p}_0)(\mathbf{p} - \mathbf{p}_0) + \mathbf{v} \quad (4-16)$$

Hierbei ist $\mathbf{J}(\mathbf{p}_0)$ die Jacobi-Matrix der Funktion f an der Stelle \mathbf{p}_0 . Aus der notwendigen Bedingung für ein Minimum von e_k

$$\frac{\partial e_k}{\partial \mathbf{p}} = 2\mathbf{v}^T \mathbf{W} \frac{\partial \mathbf{v}}{\partial \mathbf{p}} = 0 \quad (4-17)$$

und durch Auflösen von Beziehung (4-16) nach dem Fehlervektor \mathbf{v} mit anschließender Ableitung nach dem Parametervektor \mathbf{p}

$$\frac{\partial \mathbf{v}}{\partial \mathbf{p}} = -\mathbf{J}(\mathbf{p}_0) \quad (4-18)$$

kann die explizite Darstellung für den Parametervektor \mathbf{p}

$$(\mathbf{p} - \mathbf{p}_0) = \Delta \mathbf{p} = (\mathbf{J}^T(\mathbf{p}_0) \mathbf{W} \mathbf{J}(\mathbf{p}_0))^{-1} \mathbf{J}^T(\mathbf{p}_0) \mathbf{W}(\mathbf{x}'_m - \mathbf{f}(\mathbf{p}_0)) \quad (4-19)$$

gewonnen werden. Eine Invertierung der Matrix $\mathbf{J}^T(\mathbf{p}_0) \mathbf{W} \mathbf{J}(\mathbf{p}_0)$ ist jedoch nur dann möglich, wenn die Meßpunkte nicht in einer gemeinsamen Ebene liegen. Aus diesem Grund ist es notwendig, bei jeder Parameterbestimmung die in Bild 4-6 gezeigte Meßplatte zur Generierung der Meßpunktmenge mehrfach verschoben aufzunehmen.

Aus Beziehung (4-19) ergibt sich ein iterativer Optimierungsalgorithmus, wenn der Korrekturvektor $\Delta \mathbf{p}$ nach jedem Iterationsschritt i zum aktuellen Parametervektor \mathbf{p}_i addiert wird.

$$(\mathbf{p}_{i+1} - \mathbf{p}_i) = \Delta \mathbf{p}_i = (\mathbf{J}^T(\mathbf{p}_i) \mathbf{W} \mathbf{J}(\mathbf{p}_i))^{-1} \mathbf{J}^T(\mathbf{p}_i) \mathbf{W}(\mathbf{x}'_m - \mathbf{f}(\mathbf{p}_i)) \quad (4-20)$$

Die gesuchten Modellparameter \mathbf{p} ergeben sich auf diese Weise nach Erfüllung eines geeigneten Konvergenzkriteriums

$$|\mathbf{p}_{i+1} - \mathbf{p}_i| < \varepsilon; \quad \text{mit} \quad \varepsilon \geq 0. \quad (4-21)$$

Aufgrund der Nichtlinearität der Kollinearitätsgleichungen kann im allgemeinen nicht garantiert werden, daß das Verfahren zum globalen Optimum konvergiert. Ob die optimale Lösung ermittelt werden kann, hängt davon ab, ob der Initialisierungsvektor \mathbf{p}_0 im Konvergenzgebiet des Verfahrens liegt. Zur Ermittlung der meisten Initialisierungsparameter wurde bei der Kalibrierung der Kameras das Tsai-Verfahren [Tsa87] eingesetzt, das eine direkte suboptimale Lösung der Kollinearitätsgleichungen erlaubt. Die Hauptpunktkoordinaten, die nicht durch das Tsai-Verfahren errechnet werden können, wurden vorab durch ein Zoomingverfahren bestimmt [LT87]. Die Kalibrierung zeigte, daß die auf diese Weise ermittelten Initialisierungsparameter sehr nahe bei der optimalen Lösung liegen, so daß die Konvergenz des Gauß-Newton-Verfahrens kein Problem darstellte [Luk98, Tra96b]. Die Kombination dieser einzelnen Verfahrensschritte erlaubte es, die Kameras ohne Vorkenntnisse und unabhängig von Herstellerangaben zu kalibrieren.

Das Gauß-Newton-Verfahren ist eine Approximation des Newton-Verfahrens und besitzt lineare Konvergenz. Das Newton-Verfahren besitzt zwar quadratische Konvergenz [ERS77, Hor79], jedoch wurde dieses Verfahren hier nicht eingesetzt, da es durch die zusätzliche Berechnung der Hesse-Matrix in jedem Iterationsschritt wesentlich aufwendiger ist und im Fall der Konvergenz zur gleichen Lösung konvergiert wie das Gauß-Newton-Verfahren [Gen79]. Das Konvergenzgebiet des Gauß-Newton-Verfahrens kann durch den Einsatz der Levenberg-Marquardt-Methode vergrößert werden [FH84]. Der Einsatz dieser Methode war jedoch hier, aufgrund der relativ genauen Initialisierungsparameter, nicht erforderlich.

4.2.3 Einfluß der optischen Freiheitsgrade auf die Modellparameter

Im Rahmen der Kalibrierung der Kamerasysteme wurde in zahlreichen Messungen der Einfluß der unterschiedlichen optischen Freiheitsgrade auf die Modellparameter ermittelt. Während sich die Blendeneinstellung der verwendeten Kameramodule nur in einem vernachlässigbaren Maß auf die Modellparameter auswirkt, zeigte sich ein teilweise erheblicher Einfluß der Fokus- und insbesondere der Zoomeinstellungen auf das Kameramodell. Besonders groß ist dieser Einfluß auf die effektiven Brennweiten, die radiale Linsendistorsion und die Position des Projektionszentrums in Richtung der optischen Achse.

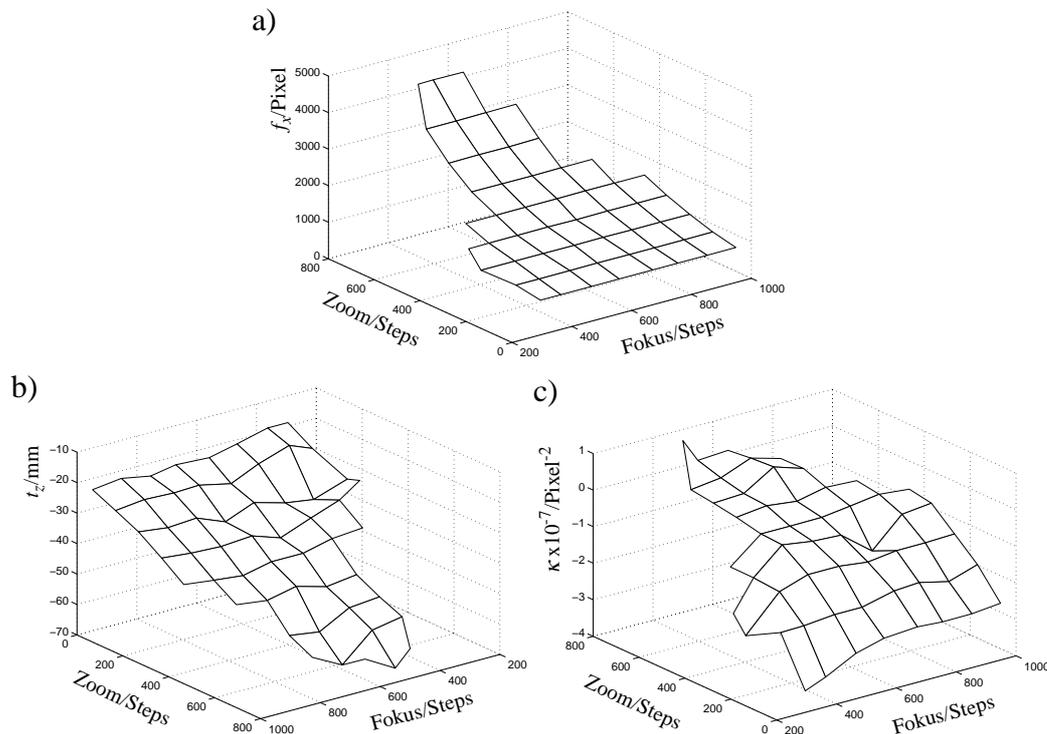


Bild 4-7: Einfluß von Zoom und Fokus auf die effektive Brennweite (a), die Position des Projektionszentrums in Richtung der optischen Achse (b) und die radiale Linsendistorsion (c) einer Kamera. Die Fokus- und Zoomwerte sind in Encoderschritten angegeben.

Zur Bestimmung der Modellparameter für die unterschiedlichen Kamerazustände wurden durchschnittlich 450 Meßpunkte pro Kamerazustand ausgewertet. Der normalisierte Kalibrierungsfehler nach Weng lag bei fast allen Messungen um den Wert 1,0. Dies entspricht einer Kalibrierung, deren verbleibender Fehler im Vergleich zu den Quantisierungsfehlern, die durch das diskrete Abtastraster der Sensoren hervorgerufen werden, vernachlässigbar gering ist [WCH92]. Die Modellierung des Distorsionszentrums durch zusätzliche Parameter (vgl. Kapitel 4.2.1) wirkte sich in einer Verbesserung des normalisierten Kalibrierungsfehlers bis zu 10% gegenüber der herkömmlichen Modellierung aus [Luk98].

Wie in Abbildung 4-7 zu sehen ist, konnten nicht alle möglichen Kombinationen von Zoom und Fokus kalibriert werden. Der Grund hierfür ist, daß die Kameras eine begrenzte Tiefenschärfe besitzen, die mit zunehmendem Zoom geringer wird. Die Meßpunkte der hier verwendeten Meßvorrichtung (vgl. Bild 4-6) konnten daher nicht bei allen Kamerazuständen scharf abgebildet werden.

Um die Auswirkungen der Parameterschwankungen auf die Tiefenrekonstruktion kompensieren zu können, wurden die Modellparameter für die einzelnen Kamerazustände tabelliert. Die optischen Freiheitsgrade der Kamerasysteme werden während des Betriebs meßtechnisch erfaßt und die jeweiligen Modellparameter werden in Abhängigkeit dieser Meßwerte mit Hilfe der „Look-up“-Tabellen ermittelt und in nachfolgenden Verarbeitungsschritten berücksichtigt.

4.3 Entzerrung und Rektifikation der Bilddaten

Die genaue Modellierung der Abbildungsprozesse der Kameras in Kapitel 4.2 liefert die Voraussetzung für die Aufbereitung der Stereobilddaten in einem ersten Verarbeitungsschritt. Die in diesem Zusammenhang eingesetzten Strategien dienen dazu, die Genauigkeit der Tiefenrekonstruktion zu erhöhen und die weiteren Verarbeitungsschritte des Stereoverfahrens erheblich zu vereinfachen.

4.3.1 Kompensation von Verzeichnungsfehlern

In welcher Form Linsenverzeichnungen in weiteren Verarbeitungsschritten Berücksichtigung erfahren, ist in der Literatur nur selten erläutert. In der Stereoskopie ist es denkbar, die nach der herkömmlichen Methode modellierte Linsendistorsion bei der Tiefenrekonstruktion oder der Schätzung von Epipolarlinien zu berücksichtigen. Da die Epipolarlinien jedoch hier aufgrund des hohen Berechnungsaufwandes nicht explizit geschätzt, sondern die Stereobilddaten über eine Rektifikationstransformation in Stereobilder mit kollinearen Epipolarlinien überführt werden sollen, werden die Bilddaten über eine Entzerrungsoperation zunächst von der Linsendistorsion befreit. Ein zusätzlicher vorteilhafter Nebeneffekt dieses Vorgehens ist, daß auch ein Mustererkennungsverfahren, das diese Bilddaten als Eingangsinformation verwendet, nicht durch die Verzeichnungsfehler beeinflusst wird.

Eine Möglichkeit zur Entzerrung der Bilddaten besteht darin, die Grauwerte des verzerrten Bildes an den Koordinaten x_d auf die unverzerrten Bildkoordinaten x_u abzubilden. Bei dieser Transformation ergeben sich jedoch im allgemeinen Zielkoordinaten, die zwischen den diskreten Bildkoordinaten liegen. Die Grauwerte müßten somit in einer geeigneten Weise auf die jeweils benachbarten Koordinaten verteilt werden. Wesentlich einfacher ist es dahingegen, diesen Abbildungsprozeß in umgekehrter Richtung zu durchlaufen, indem zunächst für jeden Bildpunkt im unverzerrten Bild die entsprechenden verzerrten Bildkoordinaten x_d errechnet werden. Die Voraussetzung hierfür liefert die in Beziehung (4-9) vorgestellte Modellierung der radialen Linsendistorsion. Aus den Beziehungen (4-6) und (4-9) ergibt sich folgender expliziter Zusammenhang zwischen den verzerrten und den unverzerrten Bildkoordinaten.

$$\mathbf{x}_d = \begin{bmatrix} x_u + \kappa(x_u - x_t)((x_u - x_t)^2 + (y_u - y_t)^2) \\ y_u + \kappa(y_u - y_t)((x_u - x_t)^2 + (y_u - y_t)^2) \end{bmatrix} \quad (4-22)$$

Die auf diese Weise ermittelten verzerrten Bildkoordinaten liegen zwar im allgemeinen ebenfalls zwischen den diskreten Bildpunkten, jedoch können die gesuchten Grauwerte $i(\mathbf{x}_u)$ in diesem Fall relativ einfach durch eine bilineare Interpolation der benachbarten vier Grauwerte um die Koordinate \mathbf{x}_d berechnet werden.

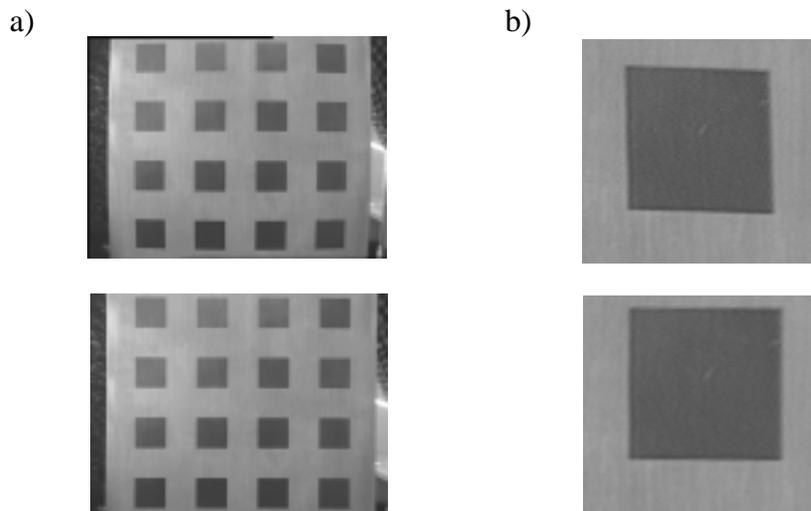


Bild 4-8: Kompensation der radialen Linsendistorsion am Beispiel einer Kalibrierungsmeßplatte. a) Verzerrtes und entzerrtes Bild. b) Bildausschnitt um das rechte obere Quadrat

4.3.2 Rektifikation der Stereobilddaten

Das in Kapitel 4.1 vorgestellte binokulare Sehsystem bietet die Möglichkeit, die optischen Achsen der Kameras aktiv auf eine Szene auszurichten und auf diese Weise Objekte im Blickfeld zu fixieren. Neben den schon erwähnten Vorteilen, die hierdurch erzielt werden können, sind jedoch auch gewisse Nachteile mit dieser Fähigkeit verbunden. Die Ausnutzung der Epipolareinschränkung (siehe Kapitel 2.4.1) zur Reduktion des Aufwandes bei der Suche nach Korrespondenzpartnern gestaltet sich bei einem System mit den obenerwähnten Eigenschaften etwas komplizierter. Der Grund hierfür ist, daß die Epipolarlinien nicht wie bei einer Kamera-konstellation mit parallelen optischen Achsen identisch mit den Bildzeilen sind, sondern im allgemeinen schräg zu diesen verlaufen. Bei bekannter Kamerageometrie können die Epipolarlinien zwar ermittelt werden (vgl. Bild 2-4), jedoch muß diese Berechnung für jeden einzelnen Bildpunkt durchgeführt werden und ist somit relativ aufwendig. Aufgrund des diskreten Abstrasters bedarf es bei dieser Technik weiterhin einer geeigneten Interpolation der Grauwerte.

Eine andere Möglichkeit besteht darin, das Bildsignal derart zu transformieren, daß die Epipolarlinien stets kollinear und identisch mit den Bildzeilen sind. Dies kann durch eine Projektion des rechten und des linken Bildes in eine gemeinsame Ebene, die sogenannte Rektifikationsebene, geschehen [AH88, GG93, PD96, Tra96c]. Die Rektifikationsebene sei hierzu in einem Abstand f_n von der Basislinie b_s angeordnet und parallel zu den y_k -Achsen der Kamerakoordinatensysteme ausgerichtet.

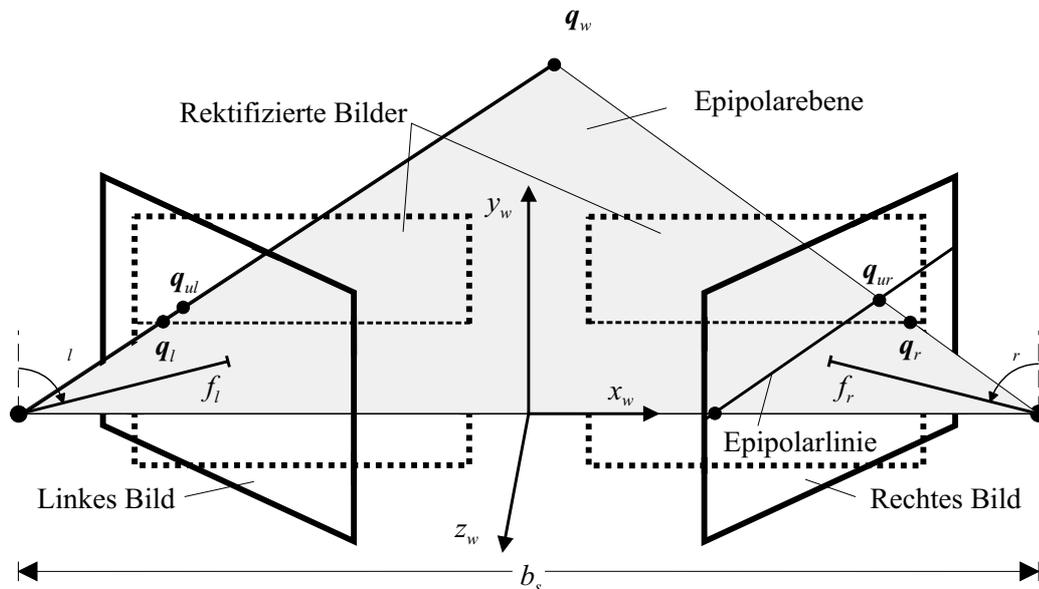


Bild 4-9: Rektifikation der Stereobilddaten durch die Projektion der Bilder in eine gemeinsame Rektifikationsebene

In einem Weltkoordinatensystem, dessen Ursprung sich in der Mitte der Basislinie befindet (vgl. Bild 4-9), kann die Rektifikationsebene durch die Normalenform

$$\mathbf{n}^T \mathbf{x}_w + f_n = 0 \quad (4-23)$$

beschrieben werden. Der Normalenvektor $\mathbf{n} = [0 \ 0 \ 1]^T$ entspricht hierbei dem Einheitsvektor des Weltkoordinatensystems in z_w -Richtung. Die neuen Koordinaten eines Punktes im rektifizierten Bild ergeben sich aus dem Durchstoßpunkt der Geraden, die durch den Kamerakoordinatenursprung und den betreffenden Bildpunkt in der Sensorebene definiert ist, und der Rektifikationsebene (vgl. Bild 4-9). Um diesen Punkt zu ermitteln, wird im folgenden die entsprechende Geradengleichung im Weltkoordinatensystem aufgestellt. Hierbei werden die Kamerakoordinaten \mathbf{x}_k zunächst auf das linke der beiden Kamerasysteme bezogen.

Die Weltkoordinaten \mathbf{x}_w sind über die folgende Transformation mit den Kamerakoordinaten \mathbf{x}_k verknüpft:

$$\mathbf{x}_w = \mathbf{R}\mathbf{x}_k + \mathbf{t} \quad (4-24)$$

Der Translationsvektor ergibt sich für die linke Kamera zu $\mathbf{t} = [-b_s/2 \ 0 \ 0]^T$. Die weiteren Berechnungen erfolgen für den Fall, daß die beiden Kameras so justiert sind, daß die optischen Achsen der Kameras in einer Ebene liegen und daß die Bildzeilen bei parallel ausgerichteten optischen Achsen kollinear zur Basislinie sind. Für diese Konfiguration, die durch die in Kapitel 4.1.2 beschriebenen Vorrichtungen leicht eingestellt werden kann, hat die Rotationsmatrix \mathbf{R} die vereinfachte Form:

$$\mathbf{R} = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \quad (4-25)$$

Der Winkel β entspricht unter idealen Voraussetzungen dabei dem meßtechnisch erfaßbaren Vergenzwinkel β_v der betrachteten Kamera. Ein beliebiger Punkt \mathbf{q} in der Sensorebene wird im entsprechenden Kamerakoordinatensystem durch $\mathbf{q} = [x_s \ y_s \ -f]^T$ definiert. Werden dieser Punkt und der Ursprung des Kamerakoordinatensystems über Beziehung (4-24) in die Weltkoordinaten überführt, so kann eine Geradengleichung für die oben beschriebene Gerade in Parameterform zu

$$\mathbf{x}_w = \begin{bmatrix} -\frac{b_s}{2} \\ 0 \\ 0 \end{bmatrix} + \lambda \mathbf{R} \begin{bmatrix} x_s \\ y_s \\ -f \end{bmatrix} \quad (4-26)$$

angegeben werden. Da der Durchstoßpunkt der Geraden und der Rektifikationsebene die gesuchten Koordinaten bestimmt, wird Beziehung (4-26) in Beziehung (4-23) eingesetzt. Somit folgt für den freien Parameter λ :

$$\lambda = \frac{f_n}{f \cos\beta + x_s \sin\beta} \quad (4-27)$$

Durch Substituieren von λ in Beziehung (4-26) ergeben sich die Weltkoordinaten der rektifizierten Bilddaten zu:

$$\mathbf{x}_w = \begin{bmatrix} f_n \frac{x_s \cos \beta - f \sin \beta}{f \cos \beta + x_s \sin \beta} - \frac{b_s}{2} \\ \frac{f_n y_s}{f \cos \beta + x_s \sin \beta} \\ -f_n \end{bmatrix} \quad (4-28)$$

Die rektifizierten Bildkoordinaten \mathbf{x} können nun direkt aus den ersten beiden Komponenten des Vektors \mathbf{x}_w in Beziehung (4-28) ermittelt werden. Hierzu wird die x_w -Komponente geeignet verschoben, so daß der Ursprung des Sensorebenenkoordinatensystems auf einen Punkt \mathbf{x}'_0 im rektifizierten Bild abgebildet wird.

$$\mathbf{x} - \mathbf{x}'_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{x}_w + \begin{bmatrix} \frac{b_s}{2} + f_n \tan \beta \\ 0 \end{bmatrix} = f_n \begin{bmatrix} \frac{x_s}{f(\cos \beta)^2 + x_s \sin \beta \cos \beta} \\ \frac{y_s}{f \cos \beta + x_s \sin \beta} \end{bmatrix} \quad (4-29)$$

Der Punkt \mathbf{x}'_0 wird in diesem Ansatz so gewählt, daß der Ursprung des Sensorebenenkoordinatensystems auf das neue Bildzentrum abgebildet wird. Der Ursprung des neuen Koordinatensystems ist die linke untere Bildecke.

Da die Sensorebenenkoordinaten \mathbf{x}_s nicht direkt zur Verfügung stehen, werden diese durch die unverzerrten Bildkoordinaten \mathbf{x}_u , die mit Hilfe der Entzerrungstransformation aus Kapitel 4.3.1 rekonstruiert werden, angenähert. Unter Verwendung der Beziehungen (4-4) und (4-5) ergibt sich somit:

$$\mathbf{x} - \mathbf{x}'_0 = f_n \begin{bmatrix} \frac{(x_u - x_0)}{f_x (\cos \beta)^2 + (x_u - x_0) \sin \beta \cos \beta} \\ \frac{f_x (y_u - y_0)}{f_x f_y \cos \beta + f_y (x_u - x_0) \sin \beta} \end{bmatrix} \quad (4-30)$$

Da diese Rektifikationstransformation ebenso wie die Entzerrung in Kapitel 4.3.1 auf diskrete Koordinaten angewendet werden muß, ist auch in diesem Zusammenhang eine bilineare Interpolation der Grauwerte notwendig. Die explizite Auflösung von Beziehung (4-30) nach den unverzerrten Bildkoordinaten, die zu diesem Zweck erforderlich ist, stellt jedoch kein Problem

dar. Zur Verminderung des Berechnungsaufwandes werden die Entzerrung und die Rektifikation verkettet, wodurch nur eine bilineare Interpolation pro Bildpunkt erforderlich ist.

Durch den noch zu bestimmenden Abstand f_n der Rektifikationsebene zur Basislinie wird im wesentlichen die lokale Dehnung bzw. Stauchung der Bilder beeinflusst, durch die die Bilddaten bei der Rektifikationstransformation verzerrt werden. Um diese Verzerrungen im Mittel möglichst gering zu halten, wird der Abstand f_n daher gleich dem arithmetischen Mittelwert der auf die Ebenennormale \mathbf{n} projizierten normierten effektiven Brennweiten f_{xl} und f_{xr} gewählt.

$$f_n = \frac{1}{2}(f_{xl} \cos \beta_l + f_{xr} \cos \beta_r) \quad (4-31)$$

Bei dieser Wahl des Parameters f_n verläuft die Rektifikationsebene bei symmetrischen Vergenzwinkeln und gleicher normierter effektiver Brennweite durch die Hauptpunkte des Kamerasystems, wodurch die Bildmitte der rektifizierten Bilder nicht verzerrt wird.

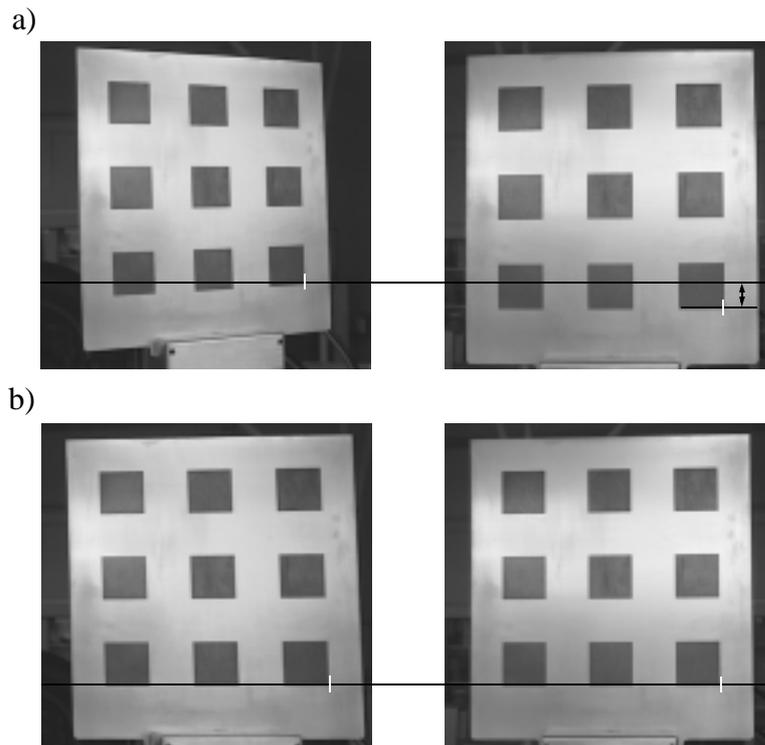


Bild 4-10: a) Originales und b) rektifiziertes Stereogramm einer Kalibrierungsmeßplatte. Im Originalstereogramm weist die Disparität der im linken Bild markierten Position eine vertikale Komponente auf. Durch die Rektifikation wird diese kompensiert.

4.4 Bilddatendekomposition durch Gaborfilter

Die enorm großen Datenmengen, die in der digitalen Bildverarbeitung manipuliert werden müssen, stellen trotz der Fortschritte auf dem Gebiet der Computertechnologie immer noch ein erhebliches Problem dar. Zur Reduktion des Berechnungsaufwandes kommen daher in der Bildverarbeitung oft Strategien zum Einsatz, die sich im wesentlichen am biologischen Vorbild orientieren. So werden allgemein in der Mustererkennung aber auch speziell in der Stereoskopie in vielen Ansätzen Merkmale aus dem Grauwertbild extrahiert, welche die Bildinformation auf einer abstrakten Ebene beschreiben und auf diese Weise die Datenmenge erheblich reduzieren können. Als Beispiel sei an dieser Stelle der Marr-Hildreth-Operator genannt, der sowohl in der Stereoskopie als auch in der Mustererkennung zur Merkmalsextraktion eingesetzt wird (vgl. Kapitel 2.5.2). Dieser Operator simuliert das Übertragungsverhalten der sogenannten Zentrum-Umfeld-Zellen, die in der Retina vieler höherer Lebewesen nachgewiesen werden konnten (siehe z.B. [HW62]). Auch der Einsatz sogenannter Auflösungspyramiden [Bur84, BA83] dient in vielen Ansätzen zur Aufwandsreduktion und ist durch das inhomogene Auflösungsvermögen des Auges und die nachgewiesenen unterschiedlichen Ortsfrequenzkanäle im Sehsystem höherer Lebewesen ebenfalls biologisch motiviert [AP79, DDY79, MF77, MTT78, PLT71].

Obwohl durch die vorangegangenen Verarbeitungsschritte bereits einige prinzipielle Voraussetzungen für eine effiziente Korrespondenzbestimmung gegeben sind, werden die Bilddaten einem weiteren Verarbeitungsschritt in Form einer Filteroperation unterzogen. Ziel dieser Filterstrategie ist es, analog zu den einfachen Zellen des primären visuellen Kortex höherer Lebewesen ein zweidimensionales monochromes Bildsignal mit Hilfe einer Filterbank orientierungs- und frequenzselektiv zu filtern. Der Berechnungsaufwand soll dabei durch eine Auflösungspyramide erheblich reduziert werden. Die Filterbank wird durch einen Satz von Gaborfiltern, die bereits prinzipiell in Kapitel 2.5.3 beschrieben wurden, gebildet. Das minimale „Zeit-Bandbreite-Produkt“ dieser Filter bewirkt dabei eine optimale Auflösung im Orts- und im Ortsfrequenzbereich.

Neben der Aufwandsreduktion, die durch die Filterung in Verbindung mit einer Auflösungs- pyramide erzielt werden kann, existieren zwei weitere Gründe für diesen Verarbeitungsschritt. Da parallel zur Korrespondenzbestimmung ein Mustererkennungsverfahren eine selektive Identifikation einzelner Objekte erlauben soll, deren räumliche Position anschließend durch die passive Triangulation bestimmt wird, dient die im folgenden näher beschriebene Filteroperation als Vorstufe für eine Extraktion von Konturmerkmalen. Anhand dieser orientierten Konturmerkmale werden schließlich einzelne Objekte im Rahmen einer Erkennungsstrategie identifiziert [SH97].

Der wesentliche Aspekt für die stereoskopische Korrespondenzbestimmung ist jedoch, daß mit Hilfe dieser Filterstrategie orientierungsselektiv Intensitätsänderungen, die eine Komponente in Richtung der Epipolarlinien aufweisen, aus den Grauwertbildern extrahiert werden können. Bedingt durch das stereoskopische Blendenproblem tragen allein diese Komponenten die Information für die passive Triangulation (siehe Kapitel 2.3). Weiterhin werden Rauschpro-

zesse, welche die Grauwertbilder im allgemeinen überlagern und die zu erheblichen Problemen in der Korrespondenzbestimmung führen können, durch die Filterung effizient unterdrückt. Darüber hinaus liefern die hier eingesetzten Gaborfilter die Voraussetzung für eine subpixelgenaue Disparitätsbestimmung in Kapitel 4.7.

Erkennung und Korrespondenzbestimmung geht somit der gleiche Vorverarbeitungsprozess voraus, wodurch kein zusätzlicher Aufwand durch unterschiedliche Vorverarbeitungsstrategien betrieben werden muß.

4.4.1 Vorüberlegungen zum Filterentwurf

Da die Gaborfilter auf zweidimensionale Grauwertbilder angewendet werden sollen, wird die Beziehung (2-16) zunächst um eine Dimension erweitert. Die komplexe Impulsantwort $g(\mathbf{x})$ des im weiteren verwendeten Gaborfilters ergibt sich in Bezug auf die zweidimensionalen Bildkoordinaten \mathbf{x} zu:

$$g(\mathbf{x}) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\mathbf{x}^T A \mathbf{x}} e^{j\mathbf{k}_0^T \mathbf{x}} \quad (4-32)$$

Die Matrix A setzt sich hierbei aus einer zweidimensionalen Rotationsmatrix \mathbf{R} und einer Parametermatrix \mathbf{S} zusammen.

$$\mathbf{A} = \mathbf{R}^T \mathbf{S} \mathbf{R} = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} \sigma_x^{-2} & 0 \\ 0 & \sigma_y^{-2} \end{bmatrix} \begin{bmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{bmatrix} \quad (4-33)$$

In der zweidimensionalen Formulierung der Gaborfunktionen treten zusätzlich der Parameter σ_y , der die Breite der gaußförmigen Einhüllenden in der zusätzlichen Dimension bestimmt, und der Drehwinkel φ , der die Lage der Einhüllenden im zweidimensionalen Raum beschreibt, auf. Die zusätzliche Komponente des Modulationsortsfrequenzvektors $\mathbf{k}_0 = [k_{01} \ k_{02}]^T$ bildet einen weiteren Entwurfsfreiheitsgrad. Es läßt sich leicht zeigen, daß das Gaborfilter auch für den zweidimensionalen Fall die Unschärferelation an der unteren Schranke erfüllt [Dau85].

Aus Symmetriegründen wird für den folgenden Filterentwurf gefordert, daß der Modulationsfrequenzvektor parallel zu einer der Hauptachsen der gaußförmigen Einhüllenden ist. Weiterhin soll nach dem biologischen Vorbild die Modulationsrichtung orthogonal zur längeren Hauptachse der Einhüllenden sein [JP87a, JP87b, JSP87]. Mit $\sigma_x < \sigma_y$ folgt somit:

$$\mathbf{k}_0 = k_0 \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} \quad (4-34)$$

Mit dem Ähnlichkeitssatz der mehrdimensionalen Fouriertransformation

$$\mathcal{F}\{f(\mathbf{A}\mathbf{x})\} = \frac{1}{|\det(\mathbf{A})|} F((\mathbf{A}^{-1})^T \mathbf{k}) \quad (4-35)$$

ergibt sich die Übertragungsfunktion des Gaborfilters aus Beziehung (4-32) zu:

$$G(\mathbf{k}) = e^{-\frac{1}{2}(\mathbf{k}-\mathbf{k}_0)^T (\mathbf{A}^{-1})^T (\mathbf{k}-\mathbf{k}_0)} \quad (4-36)$$

Mit Hilfe dieses orientierungsselektiven Bandpaßfilters soll eine Filterbank gebildet werden, die das Grauwertbild in N Ortsfrequenzbänder und M Orientierungen zerlegt. Die einzelnen Ortsfrequenzbänder sollen dabei den Einsatz einer AuflösungsPyramide ermöglichen. Diese Strategie soll es erlauben, sehr feine Details, die im allgemeinen mit hochfrequenten Signalanteilen verbunden sind, im Ortsbereich sehr gut aufzulösen, während niederfrequente Signalanteile zur Aufwandsreduktion entsprechend grob aufgelöst werden. Des weiteren sollen Konturmerkmale wie Kanten oder Linien sehr einfach aus den Filterantworten extrahiert werden können. Die Antwort von Bandpaßfiltern schwingt im allgemeinen in der Nähe von Grauwertsprüngen, die in der Bildverarbeitung gewöhnlich die Merkmale bilden, wodurch eine eindeutige Detektion und Lokalisierung der Merkmale erschwert wird. Werden die Gaborfilter jedoch so ausgelegt, daß sie eine gewisse Quadratureigenschaft aufweisen, kann dieses Problem erheblich vereinfacht werden.

Die Antwort $r(\mathbf{x})$ eines Gaborfilters ist durch die zweidimensionale Faltung eines Grauwertbildes $i(\mathbf{x})$ mit der Impulsantwort $g(\mathbf{x})$ des Filters definiert.

$$r(\mathbf{x}) = i(\mathbf{x}) * g(\mathbf{x}) \quad (4-37)$$

Diese Antwort kann als ein Produkt eines komplexen Tiefpaßsignals $r_t(\mathbf{x})$ mit einer komplexen Schwingung beschrieben werden [AKM95].

$$r(\mathbf{x}) = r_t(\mathbf{x}) e^{j\mathbf{k}_0^T \mathbf{x}} \quad (4-38)$$

Wird nun der Betrag der Filterantwort gebildet, so verschwindet aufgrund der Quadratureigenschaft des Filters die komplexe Schwingung und das Resultat ist die Amplitudeneinhüllende der Filterantwort.

$$|r(\mathbf{x})| = |r_t(\mathbf{x})| = \sqrt{(\operatorname{Re}\{r(\mathbf{x})\})^2 + (\operatorname{Im}\{r(\mathbf{x})\})^2} \quad (4-39)$$

In der Bildverarbeitung ergibt sich durch dieses Vorgehen der Vorteil einer schwingungsarmen Antwort auf eine orientierte Intensitätsänderung im Grauwertbild. Eine Extraktion von Kanten- oder Linienmerkmalen kann somit einfach durch eine Maximumsuche im Betragsergebnis erfolgen.

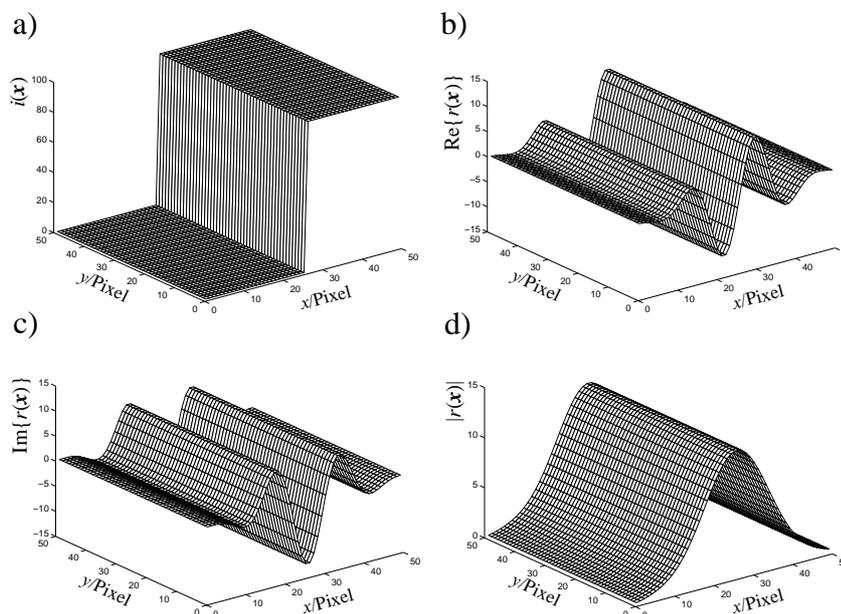


Bild 4-11: a) Sprungförmiger Grauwertverlauf. b) Realteil, c) Imaginärteil und d) Betrag der zugehörigen Antwort eines Gaborfilters

Um näherungsweise eine Quadraturbeziehung zwischen Real- und Imaginärteil der Filterantworten zu erreichen, ist es notwendig, daß Gleichanteile im Grauwertbild möglichst stark bedämpft werden, bzw. daß die Filterimpulsantwort näherungsweise mittelwertfrei ist.

4.4.2 Ortsfrequenzeinteilung

Bei der Einteilung des Eingangsspektrums auf einzelne Filterkanäle mit unterschiedlichen Beträgen der Modulationsortsfrequenz wird gefordert, daß sich die Übertragungsfunktionen benachbarter Kanäle in Richtung der Modulationsortsfrequenz bei einer bestimmten Ortsfrequenz k_r überschneiden. Bei dieser Ortsfrequenz sollen beide Filterübertragungsfunktionen

den gleichen Wert r aufweisen. Die Übertragungsfunktion aus Beziehung (4-36) wird hierzu bei einem Ortsfrequenzvektor $\mathbf{k}_r = k[\cos\varphi \sin\varphi]^T$ betrachtet, der mit der k_x -Achse der Ortsfrequenzebene den gleichen Winkel φ wie die Modulationsortsfrequenz \mathbf{k}_0 einschließt. Die Übertragungsfunktion aus Beziehung (4-36) reduziert sich damit zu:

$$G(k) = e^{-\frac{1}{2}(k-k_0)^2\sigma_x^2} \quad (4-40)$$

Um den Ortsfrequenzraum gleichmäßig auf die Filter aufzuteilen, besteht die Möglichkeit, die absolute Bandbreite Δk in Richtung der Modulationsortsfrequenz mit

$$\Delta k = \left[\frac{\int_{-\infty}^{\infty} G(k)G^*(k)(k-k_0)^2 dk}{\int_{-\infty}^{\infty} G(k)G^*(k) dk} \right]^{\frac{1}{2}} \quad (4-41)$$

für alle N Filter konstant zu wählen.

$$\Delta k_n = \frac{1}{\sigma_{xn}\sqrt{2}} = \text{konst} \quad \forall n \in [0, \dots, N-1] \quad (4-42)$$

Diese Einstellung führt dazu, daß alle Filter eine gleich gute bzw. gleich schlechte Auflösung im Ortsbereich aufweisen. Sinnvoller erscheint es jedoch im Bereich der Bildverarbeitung, hochfrequente Signalanteile, die meist Objektdetails im Grauwertbild repräsentieren, besser im Ortsbereich aufzulösen als niederfrequente Anteile, die allmähliche Grauwertänderungen repräsentieren. Um dieses Verhalten zu realisieren, wird die relative Bandbreite Δb konstant für alle N Filter gewählt.

$$\Delta b = \frac{\Delta k_n}{k_{0n}} = \frac{1}{\sigma_{xn}k_{0n}\sqrt{2}} = \text{konst} \quad \forall n \in [0, \dots, N-1] \quad (4-43)$$

Diese Einstellung führt zu einer logarithmischen Bandaufteilung des Ortsfrequenzspektrums. Im weiteren wird die relative Bandbreite in Oktaven angegeben. Die allgemein übliche Berechnungsweise lautet hierbei:

$$\Delta b_{okt} = \log_2 \left(\frac{k_{0n} + \Delta k_n / 2}{k_{0n} - \Delta k_n / 2} \right) \quad (4-44)$$

Die Bandbreite der Filter nimmt durch die Vorgabe in Beziehung (4-43) mit wachsender Modulationsortsfrequenz logarithmisch zu, wodurch die Ortsauflösung ebenfalls zunimmt.

Um eine Unterabtastung der Filterantworten auf den diskreten Bildkoordinaten mit Abtastfaktoren, die sich aus Potenzen von zwei ergeben, zu ermöglichen, werden die Filter in Oktavbänder eingeteilt. Die Beträge benachbarter Modulationsortsfrequenzen weisen in diesem Fall die folgende Beziehung auf:

$$k_{0n} = 2k_{0n+1} \quad (4-45)$$

Aus der Forderung, daß zwei benachbarte Filterübertragungsfunktionen bei einer bestimmten Ortsfrequenz k_r in Richtung des Modulationsvektors einen Wert r aufweisen sollen

$$G_n(k_r) = G_{n+1}(k_r) = r, \quad (4-46)$$

ergibt sich nach einigen Umformungen:

$$\sigma_{xn} k_{0n} = 3\sqrt{-2\ln(r)} \quad (4-47)$$

Der Wert, den die Filterübertragungsfunktionen für Gleichanteile annehmen, läßt sich aus den Beziehungen (4-40) und (4-47) zu

$$G_n(k=0) = e^{-\frac{1}{2}k_{0n}^2\sigma_{xn}^2} = e^{9\ln(r)} = \text{konst} \quad \forall n \in [0, \dots, N-1] \quad (4-48)$$

errechnen. Bei einer Filterbank, bei der beispielsweise $r = 0,5$ gewählt wird, beläuft sich dieser Wert bereits auf weniger als 0,2% des Maximalwertes der Übertragungsfunktion. Der Betrag der Modulationsortsfrequenz wird im Rahmen der Oktavbandeinteilung zu

$$k_{0n} = \frac{\pi}{2^{n+1}}; \quad \text{mit} \quad n \in [0 \dots N-1] \quad (4-49)$$

gewählt. Werden die Filterantworten im Rahmen einer Datenkomprimierung mit einem Faktor $s_n = 2^n$ abgetastet, so ist die kritische Ortsfrequenz, oberhalb der die Filterantwort nach dem Shannon'schen Abtasttheorem keine Signalanteile mehr aufweisen darf, die Modulationsortsfrequenz des benachbarten Filterkanals. Der Wert, den die Übertragungsfunktion an der kritischen Ortsfrequenz annimmt, ist somit gleich dem Wert, den das Filter bei Gleichanteilen aufweist. Die Filterantworten sind bei $r = 0,5$ somit näherungsweise mittelwertfrei und auch der Fehler, der durch die Unterabtastung entsteht, ist vernachlässigbar gering. Das nachfolgende Beispiel zeigt die Übertragungsfunktionen von Gaborfiltern für $N = 3$ Kanäle mit $r = 0,5$. Die relative Bandbreite der Filter entspricht hierbei etwa $\Delta b_{okt} \approx 0,3$ Oktaven.

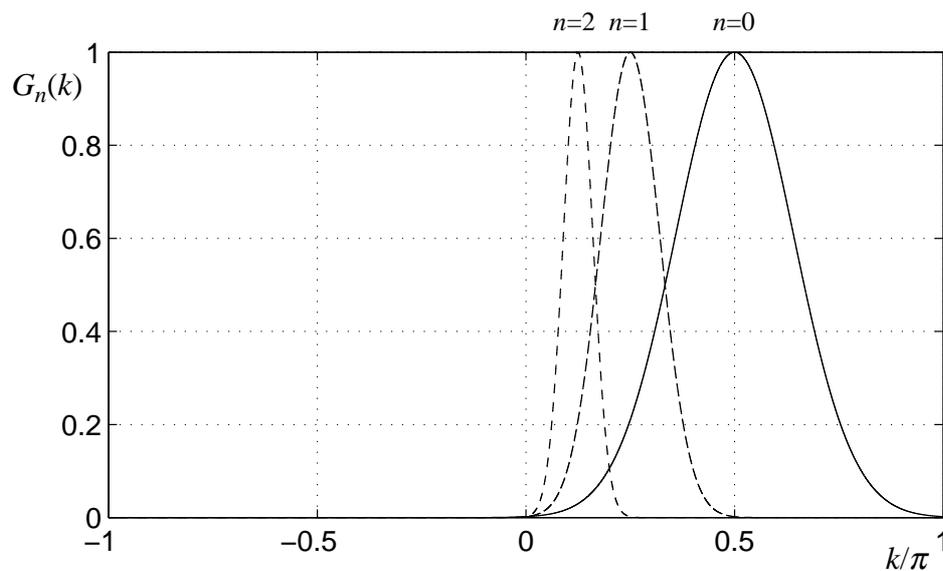


Bild 4-12: Schnitt durch die Übertragungsfunktionen der Filterbank in Richtung der Modulationsortsfrequenz für $N = 3$

4.4.3 Orientierungseinteilung

Als noch festzulegende Entwurfsparameter verbleiben an dieser Stelle der Parameter σ_y und der Winkel φ . Der Parameter σ_y beschreibt die Ausdehnung der einhüllenden Gaußfunktion orthogonal zur Modulationsrichtung \mathbf{k}_0 . Hierdurch bestimmt er direkt die Orientierungsselektivität des Filters. Durch den Winkel φ wird die Richtung der Modulationsortsfrequenz und damit die Orientierung des Filters festgelegt (vgl. Beziehung (4-33) und (4-34)). Die Winkelauflösung $\Delta\varphi$ ergibt sich aus der Anzahl M der Filter mit unterschiedlicher Orientierung, die für jeden Filterkanal n vorgesehen werden.

$$\Delta\varphi = \frac{2\pi}{M} \quad (4-50)$$

Die Anzahl M der Filter sei für alle N Filterkanäle identisch. Der Winkel φ_m der einzelnen Filter ergibt sich unter diesen Voraussetzungen zu:

$$\varphi_m = m\Delta\varphi; \quad \text{mit} \quad m \in [0 \dots M-1] \quad (4-51)$$

Analog zum Kapitel 4.4.2 wird der Parameter σ_{yn} so bestimmt, daß sich die Übertragungsfunktionen benachbarter Filter auf einem Kreis, dessen Radius gleich dem Betrag k_{0n} der Modulationsortsfrequenz k_{0n} ist, mit gleichem Wert r überschneiden. Da die Filter symmetrisch zur Modulationsrichtung sind, errechnet sich der Frequenzvektor \mathbf{k}_m für diesen Schnittpunkt zu:

$$\mathbf{k}_m = k_{0n} \left[\cos\left(\varphi_m + \frac{\Delta\varphi}{2}\right) \quad \sin\left(\varphi_m + \frac{\Delta\varphi}{2}\right) \right]^T \quad (4-52)$$

Durch Einsetzen der Ortsfrequenz \mathbf{k}_m in Beziehung (4-36) ergibt sich unter Ausnutzung verschiedener trigonometrischer Additionstheoreme:

$$G_{mn}(\mathbf{k}_m) = e^{-\frac{1}{2}k_{0n}^2 \left(\sigma_{xn}^2 \left(\cos\frac{\Delta\varphi}{2} - 1 \right)^2 + \sigma_{yn}^2 \left(\sin\frac{\Delta\varphi}{2} \right)^2 \right)} = r \quad (4-53)$$

Unter Ausnutzung von (4-47) und (4-49) errechnet sich der Parameter σ_{yn} aus Beziehung (4-53) zu:

$$\sigma_{yn} = \frac{\sigma_{xn} \sqrt{1 - 9 \left(\cos\frac{\Delta\varphi}{2} - 1 \right)^2}}{3 \sin\frac{\Delta\varphi}{2}} \quad (4-54)$$

Wie diese Beziehung zeigt, wird das Verhältnis der Hauptachsen der einhüllenden Gaußfunktion nur durch die Winkelauflösung $\Delta\varphi$ bestimmt. Das Verhältnis ist somit konstant für alle $N \times M$ Filterkanäle. Die vorzugebenden Parameter der Filterbank beschränken sich somit auf die Anzahl N und M der Filter mit unterschiedlicher Modulationsortsfrequenz und Orientierung und den Parameter r .

Da das zu filternde Grauwertbild $i(\mathbf{x})$ reell ist, können Symmetrieeigenschaften der Fouriertransformation ausgenutzt werden, so daß die Filterung nur mit der Hälfte der M Orientierungen durchgeführt werden muß. Da alle Übertragungsfunktionen mit gleichem Betrag der

Modulationsortsfrequenz abgesehen von ihrer Orientierung identisch sind, gilt folgende Symmetrie:

$$G_{mn}(\mathbf{k}) = G_{(m+M/2)n}(-\mathbf{k}) \quad (4-55)$$

Für die Fouriertransformierte $R_{mn}(\mathbf{k})$ der Filterantwort $r_{mn}(\mathbf{x})$ gilt somit:

$$R_{(m+M/2)n}(\mathbf{k}) = G_{(m+M/2)n}(\mathbf{k})I(\mathbf{k}) = G_{mn}(-\mathbf{k})I^*(-\mathbf{k}) = R_{mn}^*(-\mathbf{k}) \quad (4-56)$$

Unter Ausnutzung der Symmetrieeigenschaft

$$\mathcal{F}\{f^*(\mathbf{x})\} = F^*(-\mathbf{k}) \quad (4-57)$$

ergibt sich für die komplexen Filterantworten im Ortsbereich:

$$r_{(m+M/2)n}(\mathbf{x}) = r_{mn}^*(\mathbf{x}) \quad (4-58)$$

Die Antworten der jeweils um den Winkel π versetzten Filterorientierungen können auf diese Weise leicht aus den ursprünglichen Filterantworten durch eine komplexe Konjugation errechnet werden. Da somit nur die Filterantworten der Orientierung $m \in [0 \dots M/2 - 1]$ für jeden Filterkanal n direkt berechnet werden müssen, läßt sich der Aufwand an dieser Stelle entsprechend reduzieren.

Im Rahmen der Mustererkennung werden die Filterantworten aller Orientierungen ohne eine Gewichtung genutzt, um Konturmerkmale aus den Bildern zu extrahieren (siehe [SH97]). Aufgrund des stereoskopischen Blendenproblems (siehe Kapitel 2.3) sind jedoch für die Stereoskopie nur die Grauwertänderungen von Bedeutung, die eine Komponente in Richtung der Epipolarlinien aufweisen. Bedingt durch die Rektifikation in Kapitel 4.3.2 liegen die spektralen Anteile, durch die diese Grauwertänderungen repräsentiert werden, auf der k_x -Achse in der Ortsfrequenzebene. Zur Korrespondenzbestimmung werden daher auch nur die Filterkanäle mit der Orientierung $\varphi_m = 0$ und zusätzlich, um dichter besetzte Disparitätskarten zu erhalten, die zwei Nachbarorientierungen herangezogen.

In Abbildung 4-13 ist die Aufteilung der Ortsfrequenzebene in die einzelnen Kanäle der Filterbank zu sehen. Durch die Ellipsen sind jeweils die Ortsfrequenzen gekennzeichnet, an denen

die Übertragungsfunktionen den Wert $G_{mn}(\mathbf{k}) = 0,5$ annehmen. Die Winkelauflösung beträgt in diesem Beispiel $\Delta\varphi = 30^\circ$. Alle übrigen Parameter entsprechen denen aus Kapitel 4.4.2.

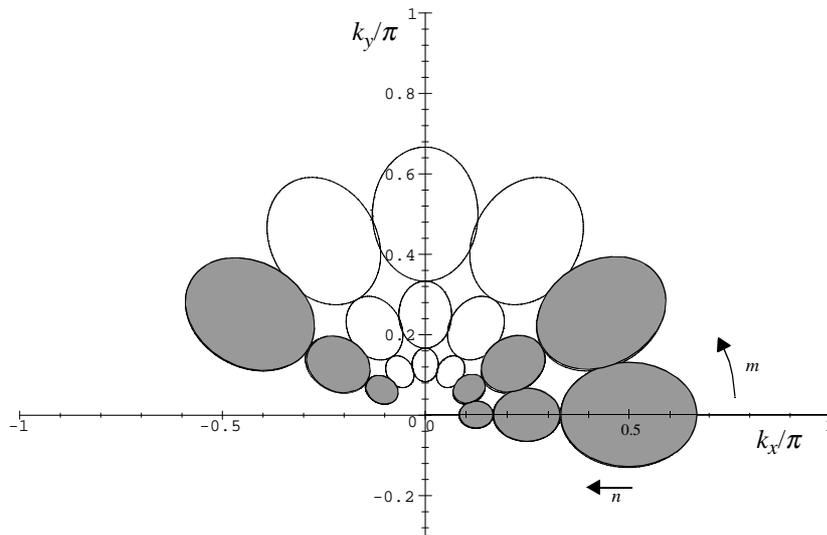


Bild 4-13: Halbwertskonturen der Filterbank in der Ortsfrequenzebene. Während zur Mustererkennung alle Kanäle ausgewertet werden, dienen nur die Filterantworten der grau unterlegten Filterkanäle zur stereoskopischen Korrespondenzbestimmung.

Die Impulsantwort der zweidimensionalen Gaborfilter weist bei der Orientierung $\varphi = 0$ einen geraden Realteil und einen ungeraden Imaginärteil in horizontaler Richtung auf. In Anlehnung an die klassische Bildverarbeitung kann der Realteil somit als Linienfilter und der Imaginärteil als Kantenfilter bezeichnet werden [HW62, Wah84]. Als ein interessanter Aspekt sei an dieser Stelle vermerkt, daß neuere neurophysiologische Untersuchungen eine erstaunliche Übereinstimmung zwischen der Form der rezeptiven Felder einfacher Zellen höher entwickelter Lebewesen und der Impulsantwort zweidimensionaler Gaborfilter zeigen [JP87a, JP87b, JSP87, PR81].

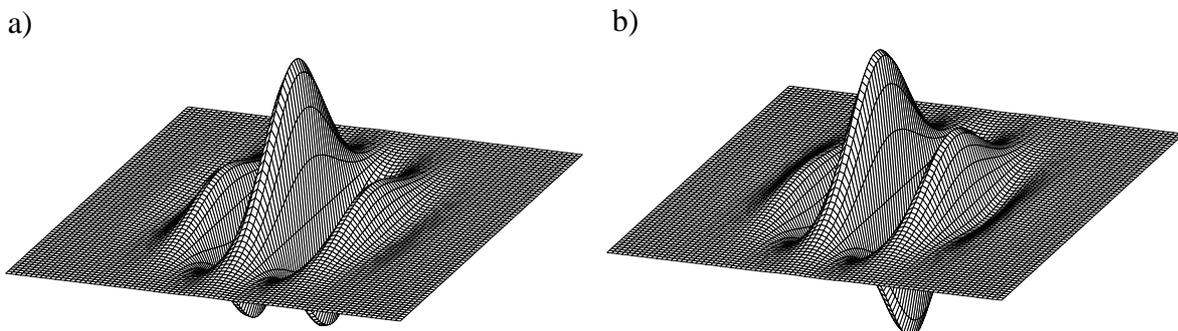


Bild 4-14: Impulsantwortpaar des Gaborfilters. a) Realteil und b) Imaginärteil der Filterimpulsantwort

Soweit nicht anders erwähnt, werden zur Berechnung aller weiteren Ergebnisse und Beispiele die in diesem Kapitel angegebenen Filterparameter verwendet.

4.4.4 Die Auflösungspyramide

Zur Reduktion des Berechnungsaufwandes werden bei vielen klassischen Bildanalyseverfahren mehrstufige Auflösungspyramiden wie Gauß- oder Laplacepyramiden eingesetzt [Bur84, BA83, Jäh93]. Bei dieser Technik wird das Grauwertbild unter Berücksichtigung des Abtasttheorems kaskadisch mit isotropen Tief- oder Bandpaßfiltern gefiltert und abgetastet. Das Ergebnis einer Filterung mit anschließender Unterabtastung dient dabei als Eingangsinformation für die jeweils nächste Auflösungsebene. Diese Art von Vorverarbeitung basiert in vielen Fällen auf separablen Faltungsmasken und wird im allgemeinen mit einer speziellen Faltungshardware durchgeführt [Jäh93]. Da Hardware dieser Art nicht zur Verfügung stand, und da die hier eingesetzten Filterimpulsantworten weder in polaren noch in kartesischen Koordinaten vollständig separabel sind, wurde der gesamte Filtervorgang als Multiplikation des Eingangsspektrums mit den Übertragungsfunktionen im Fourierbildbereich mit anschließender Rücktransformation implementiert.

Die Auslegung der Filter in Kapitel 4.4.2 erlaubt es, die Antworten $r_{mn}(\mathbf{x})$ einer Filterebene n unter Berücksichtigung des Shannon'schen Abtasttheorems (siehe z.B. [Jäh93]) mit dem Faktor $s_n = 2^n$ abzutasten, ohne daß ein wesentlicher Informationsverlust auftritt.

$$\tilde{r}_{mn}(\mathbf{x}) = r_{mn}(s_n \mathbf{x}) \quad (4-59)$$

Die auf diese Weise komprimierten Filterantworten $\tilde{r}_{mn}(\mathbf{x})$ werden im Gegensatz zu den klassischen Auflösungspyramiden hier nicht kaskadisch, sondern unabhängig voneinander berechnet. Diese Technik hat einerseits den Vorteil, daß die Filterbank leicht auf einer Mehrprozessormaschine parallelisiert werden kann und andererseits bietet sie die Möglichkeit, unterschiedliche Bildbereiche pro Auflösungskanal zu filtern. Im weiteren wird diese Eigenschaft genutzt, um in den einzelnen Auflösungsebenen Bildbereiche unterschiedlicher Größe zu verarbeiten. Die Bildbereiche sind dabei konzentrisch und überlappend angeordnet und in ihrer Größe so gewählt, daß nach dem Filtervorgang und der Abtastung in jedem Kanal die gleiche Anzahl von Abtastpunkten vorliegt. Auf diese Weise ergibt sich eine Auflösungspyramide mit einem kartesischen Abtastgitter, deren Auflösung mit steigendem Abtastfaktor s_n ab-, und deren Eingangsbildbereich gleichzeitig zunimmt. Damit der Basis-2-Algorithmus bei der Fouriertransformation angewendet werden kann, wird im weiteren für die Dimensionen der Eingangsbilder eine Potenz von zwei gewählt.

Ähnlich den biologischen Sehsystemen besitzt somit die Auflösungspyramide einen kleinen Bereich, in dem noch sehr feine Strukturen, die sich durch hochfrequente spektrale Anteile auszeichnen, aufgelöst werden können. In Richtung der Peripherie werden nur noch zuneh-

mend grobe und damit niederfrequente Strukturen erfaßt. Eine solche Architektur setzt zwangsläufig einen bewegungsfähigen Sehapparat voraus, der es durch Ausrichten der optischen Achsen erlaubt, beliebige Bereiche der betrachteten Szene in den hochauflösenden Bereich abzubilden. In Abbildung 4-15 sind die Antworten von drei Auflösungskanälen auf ein Grauwertbild zu sehen. Die Grauwerte der Bilder repräsentieren hierbei jeweils die Beträge der Filterantworten. Zur übersichtlicheren Darstellung sind die einzelnen orientierungsselektiven Kanäle jeweils in einem Bild überlagert gezeigt. Als orientierte Konturmerkmale werden bei der Mustererkennung jeweils lokale Maxima in den Beträgen der einzelnen Filterkanäle detektiert und zur Identifikation herangezogen [SH97].

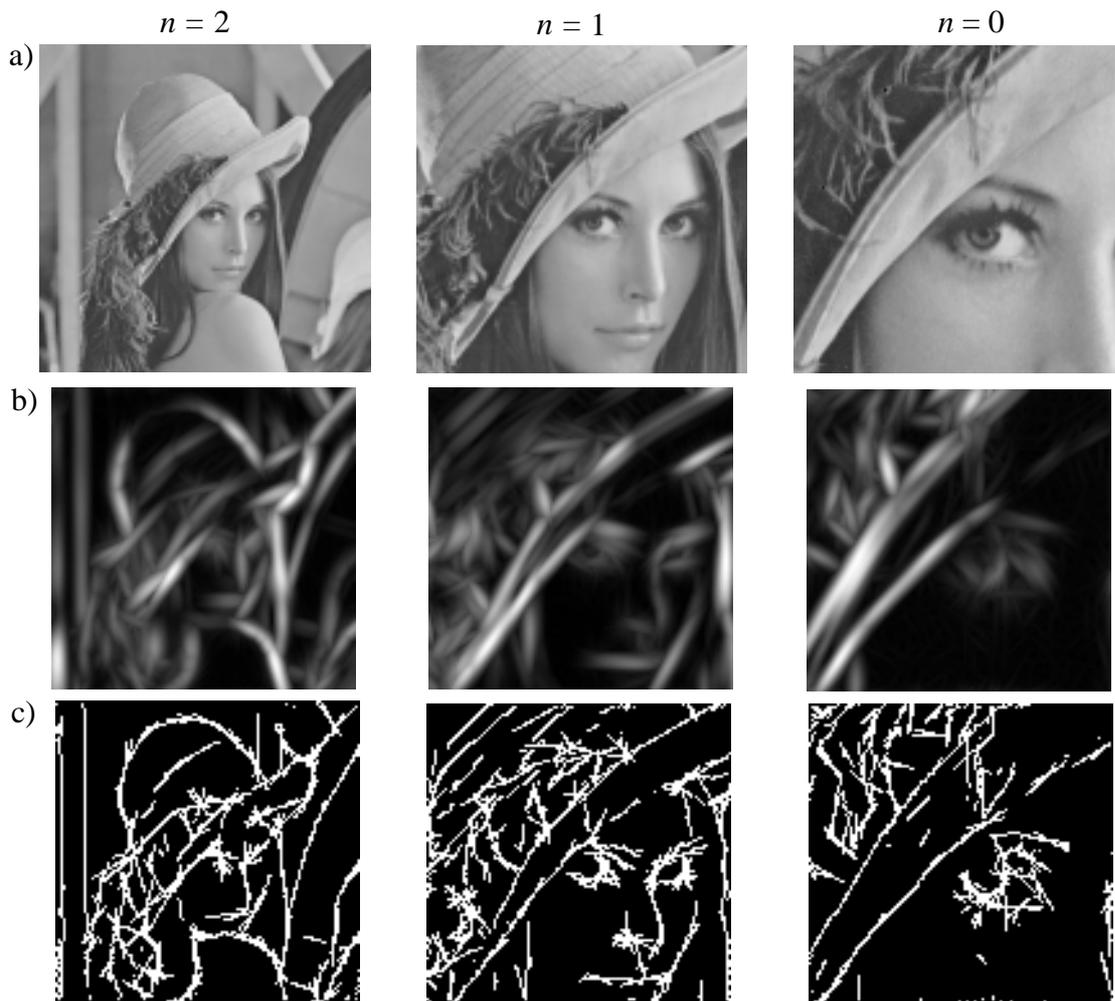


Bild 4-15: Filterung eines Grauwertbildes durch die Auflösungspyramide. a) Bildausschnitte der einzelnen Auflösungskanäle (aus [Pla72]). b) Überlagerte Beträge der Filterantworten aller Filterorientierungen. c) Überlagerte extrahierte Merkmalspunkte aller Filterorientierungen

Die Filterung der Eingangssignale im diskreten Fourierbildbereich mit endlichen Definitionsintervallen entspricht einer zyklischen Berechnung des Faltungsproduktes. Da im allgemeinen

die Eingangssignale jedoch nicht periodisch sind, treten in den Randbereichen Fehler auf, wenn in der periodischen Fortsetzung des Bildsignals Unstetigkeitsstellen vorhanden sind. Dieses Problem kann durch die Ausklammerung der Randbereiche der Filterantworten behoben werden, wobei deren Breite etwa der halben Ausdehnung der Impulsantwort entspricht.

Die Komprimierung der Filterantworten erfolgt bei der implementierten Version der Filterbank nicht explizit nach der Rücktransformation im Ortsbereich, sondern sie wird vielmehr implizit durch die Rücktransformation des jeweiligen Spektralbereiches bewirkt. Auf diese Weise kann der Berechnungsaufwand drastisch reduziert werden, da der größte Anteil der Laufzeit bei der Filterung auf die Fouriertransformationen entfällt (siehe [Tra96a]).

4.5 Das Ähnlichkeitsmaß

Ein Aspekt, der bei allen Strategien zur stereoskopischen Korrespondenzbestimmung, abgesehen von den phasenbasierten Verfahren, eine zentrale Rolle spielt, ist die Übereinstimmung zwischen den Stereobildern. Diese Übereinstimmung liefert bei vielen Ansätzen den ersten Hinweis auf mögliche Korrespondenzpaare. Die Verrechnung von Grauwerten zur Ermittlung lokaler Übereinstimmung zwischen den Grauwertbildern nach einem Korrelationsprinzip führt bei den gebietsbasierten Verfahren, im Gegensatz zu den merkmalsbasierten Verfahren, zu sehr dicht besetzten Disparitätskarten. Da sich jedoch Rauschprozesse, welche die Grauwertbilder überlagern, und andere interokuläre Differenzen bei diesen Ansätzen direkt auf das Ähnlichkeitsmaß auswirken, sind diese Verfahren entsprechend fehleranfällig.

Nachdem in den vorangegangenen Kapiteln gezeigt wurde, wie die Bilddaten für die Korrespondenzbestimmung aufbereitet werden, erfolgt in diesem Abschnitt die Beschreibung eines Verfahrens, mit dem die Ähnlichkeit zwischen den Stereokanälen ermittelt wird. Bei diesem Verfahren bildet, ähnlich wie bei den gebietsbasierten Methoden, ein lokaler Korrelationsansatz die Grundlage zur Ermittlung sehr dicht besetzter Disparitätskarten. Die Ausnutzung der Epipolareinschränkung in Form einer Rektifikation (siehe Kapitel 4.3.2) erlaubt hierbei eine effiziente Implementierung des Verfahrens. Die vorausgehende orientierungs- und frequenzselektive Filterung der Grauwertbilder durch die Gaborfilter unterdrückt einerseits gewisse Störeinflüsse und ermöglicht andererseits die Berechnung subpixelgenauer Disparitätswerte, ohne daß eine explizite Berechnung von Phasendifferenzen oder Momentanfrequenzen wie bei den phasenbasierten Verfahren erforderlich ist. Die in Kapitel 4.4.4 vorgestellte Auflösungspyramide wird zwar im weiteren zur Aufwandsreduktion verwendet, da aber der Wertebereich der Disparitäten nicht wie bei den phasenbasierten Verfahren durch die Modulationswellenlänge der Filter begrenzt wird, ist sie keine zwingende Voraussetzung für diesen Ansatz.

4.5.1 Lokale Korrelation

Wie in Kapitel 2.5.3 bereits erläutert wurde, kann die Disparität näherungsweise durch die Phasendifferenz zwischen den Filterantworten der beiden Stereokanäle berechnet werden. Die positiven Eigenschaften dieses Ansatzes, wie hohe Robustheit und dichte Disparitätskarten,

können jedoch nur bedingt ausgenutzt werden, da der Wertebereich der Disparität aufgrund der Mehrdeutigkeit der Phaseninformation durch die Modulationswellenlänge der Quadraturfilter begrenzt wird. Um diesen Nachteil zu umgehen, werden im folgenden die Antworten der in Kapitel 4.4 vorgestellten Filter zunächst in Form eines Korrelationsansatzes miteinander verrechnet. Um ein lokales Maß für die Übereinstimmung zwischen den Filterantworten des linken und rechten Stereokanals zu erhalten, wird die Korrelation durch eine Fensterfunktion räumlich gewichtet.

Seien $\tilde{r}_{mnl}(\mathbf{x})$ und $\tilde{r}_{mnr}(\mathbf{x})$ die komprimierten Antworten des Filterkanals mn im linken bzw. rechten Stereokanal und $w_{mn}(\mathbf{x})$ eine geeignete reelle Fensterfunktion, so wird das komplexe Ähnlichkeitsmaß $\rho_{mnr}(\mathbf{x}, \mathbf{d})$, das sich auf das rechte Bildkoordinatensystem bezieht, zu

$$\rho_{mnr}(\mathbf{x}, \mathbf{d}) = \frac{w_{mn}(\mathbf{x}) * \tilde{r}_{mnr}(\mathbf{x}) \tilde{r}_{mnl}^*(\mathbf{x} + \mathbf{d})}{\sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x})|^2} \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})|^2}} \quad (4-60)$$

definiert. Durch die Faltung mit der Fensterfunktion $w_{mn}(\mathbf{x})$ ergibt sich durch das Ähnlichkeitsmaß $\rho_{mnr}(\mathbf{x}, \mathbf{d})$ eine Aussage über die lokale Ähnlichkeit der rechten Filterantwort an der Koordinate \mathbf{x} zu der linken Filterantwort an der Koordinate $(\mathbf{x} + \mathbf{d})$. Der maximale Betrag des Ähnlichkeitsmaßes wird durch den Nenner in Beziehung (4-60) auf den Wert 1,0 normiert. Die in Kapitel 4.3.2 vorgestellte Rektifikation der Stereobilder bewirkt durch die horizontalen Epipolarlinien an dieser Stelle, daß die Korrelation in Beziehung (4-60) nur in Abhängigkeit der x -Koordinate durchgeführt werden muß und damit erheblich aufwandsgünstiger implementiert werden kann.

Weiterhin kann der Berechnungsaufwand durch eine geeignete Einschränkung der erfaßbaren Raumtiefe vermindert werden, indem nur ein begrenzter Disparitätsbereich zugelassen wird.

$$d_{min} \leq d_x \leq d_{max}; \quad d_y = 0 \quad (4-61)$$

Die Wahl der Fensterfunktion wird von zwei Faktoren beeinflußt. Eine ausgedehnte Fensterfunktion kann zu einer Reduktion der Mehrdeutigkeiten bei der Korrespondenzbestimmung führen, da viel Information zur Identifikation der Korrespondenzpartner ausgewertet werden kann. Mit zunehmender Ausdehnung der Fensterfunktion wird jedoch die Auflösung, mit der die Disparität ermittelt werden kann, vermindert. Das Ähnlichkeitsmaß reagiert somit empfindlicher auf Gradienten im Disparitätsverlauf, was insbesondere bei Diskontinuitäten im Disparitätsverlauf zu erheblichen Problemen führen kann.

Um eine maximale Auflösung in Richtung der y -Koordinate zu erzielen und um den Berechnungsaufwand weiter zu reduzieren, wird die Breite der Fensterfunktion in y -Richtung gleich 1

Pixel gewählt. Eine Grauwertänderung im Eingangsbild resultiert grundsätzlich in einer Filterantwort, deren räumliche Ausdehnung größer oder gleich der räumlichen Ausdehnung der Filterimpulsantwort ist. Die Fensterfunktion $w_{mn}(\mathbf{x})$ wird daher als Kompromiß zwischen den obengenannten Argumenten als eine eindimensionale Gaußfunktion gewählt, die jeweils der Einhüllenden der Filterimpulsantworten in Epipolarrichtung entspricht. Da zur Berechnung des Ähnlichkeitsmaßes die mit dem Faktor s_n abgetasteten Filterantworten dienen (vgl. Kapitel 4.4.4), wird die Fensterfunktion durch den jeweiligen Abtastfaktor entsprechend skaliert. Aufgrund der vorangegangenen Rektifikationstransformation (vgl. Kapitel 4.3.2) sind die Epipolarlinien identisch mit den Bildzeilen, so daß sich die Fensterfunktion über die Beziehungen (4-32) und (4-33) aus der entsprechend skalierten gaußförmigen Einhüllenden in Richtung der x -Koordinate ergibt.

$$w_{mn}(\mathbf{x}) = \frac{1}{2\pi\sigma_{xn}\sigma_{yn}} e^{-\frac{1}{2}x^2 s_n^2 (\sigma_{xn}^{-2}(\cos\varphi_m)^2 + \sigma_{yn}^{-2}(\sin\varphi_m)^2)} \quad (4-62)$$

Bedingt durch die Auslegung der einzelnen Filterparameter nach den Beziehungen (4-47), (4-49) und (4-54) und der Wahl des jeweiligen Abtastfaktors s_n ist die Fensterfunktion, abgesehen von einem konstanten Faktor, der jedoch keinen Einfluß auf das Ähnlichkeitsmaß hat, unabhängig von der Auflösungsebene n .

Die Bestimmung des Ähnlichkeitsmaßes nach Beziehung (4-60) beschränkt sich auf eine Messung der Übereinstimmung von Filterkanälen mit gleicher Modulationsortsfrequenz und Orientierung. Differenzen in Ortsfrequenz und Orientierung korrespondierender Elemente, die durch die unterschiedlichen Kameraperspektiven hervorgerufen werden können, werden hier nicht explizit berücksichtigt. Da jedoch diese Differenzen im allgemeinen relativ gering sind und aufgrund der Filterbandbreiten in gewissen Grenzen toleriert werden können, wird auf eine zusätzliche Berechnung der Ähnlichkeit zwischen Filterkanälen unterschiedlicher Modulationsortsfrequenz oder Orientierung im weiteren verzichtet.

4.5.2 Eigenschaften des Ähnlichkeitsmaßes

Werden die Filterantworten $\tilde{r}_{mn}(\mathbf{x})$ in einer polaren Darstellung durch Betrag $|\tilde{r}_{mn}(\mathbf{x})|$ und Phase $\phi_{mn}(\mathbf{x})$ beschrieben,

$$\tilde{r}_{mn}(\mathbf{x}) = |\tilde{r}_{mn}(\mathbf{x})| e^{j\phi_{mn}(\mathbf{x})} \quad (4-63)$$

so kann Beziehung (4-60) zu

$$\rho_{mnr}(\mathbf{x}, \mathbf{d}) = \frac{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x})| |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})| e^{j(\phi_{mnr}(\mathbf{x}) - \phi_{mnl}(\mathbf{x} + \mathbf{d}))}}{\sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x})|^2} \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})|^2}} \quad (4-64)$$

umformuliert werden. Der Real- und der Imaginärteil des Ähnlichkeitsmaßes ergeben sich somit zu:

$$\begin{aligned} \operatorname{Re}\{\rho_{mnr}(\mathbf{x}, \mathbf{d})\} &= \frac{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x})| |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})| \cos(\phi_{mnr}(\mathbf{x}) - \phi_{mnl}(\mathbf{x} + \mathbf{d}))}{\sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x})|^2} \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})|^2}} \\ \operatorname{Im}\{\rho_{mnr}(\mathbf{x}, \mathbf{d})\} &= \frac{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x})| |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})| \sin(\phi_{mnr}(\mathbf{x}) - \phi_{mnl}(\mathbf{x} + \mathbf{d}))}{\sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x})|^2} \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})|^2}} \end{aligned} \quad (4-65)$$

Bei optimaler lokaler Übereinstimmung der linken und der rechten Filterantwort ist die Phasendifferenz im Argument des Kosinus in Beziehung (4-65) gleich Null und der Realteil des Ähnlichkeitsmaßes wird maximal. Maxima im Verlauf des Realteils sind somit Kandidaten für mögliche Korrespondenzpaare. Da die Phasendifferenz der Filterantworten näherungsweise linear von der Disparität abhängt (vgl. Beziehung (2-20)), tritt bei lokalen Maxima im Realteil im Verlauf des Imaginärteils in Bezug auf die Disparität im allgemeinen ein Nulldurchgang auf. Diese wesentliche Eigenschaft des Ähnlichkeitsmaßes wird in Kapitel 4.7 ausgenutzt, um die Genauigkeit der Disparitätsmessung zu erhöhen.

Das Ähnlichkeitsmaß wird in Beziehung (4-60) auf Terme, die von den lokalen Beträgen der Filterantworten abhängen, normiert. Bedingt durch das Übertragungsverhalten der Gaborfilter liefern diese Terme eine Aussage darüber, in welchem Maß Grauwertänderungen mit geeigneter Orientierung in der durch die Fensterfunktion $w_{mn}(\mathbf{x})$ begrenzten Eingangsbildregion vorhanden sind. Um zu verhindern, daß Tiefenschätzungen an Bildkoordinaten durchgeführt werden, an denen keine signifikante Information vorhanden ist, wird das Ähnlichkeitsmaß nur an Bildkoordinaten bestimmt, an denen die Nennerterme in Beziehung (4-60) einen geeigneten Schwellwert überschreiten.

Wie sich im weiteren Verlauf der Arbeit im Zusammenhang mit der Okklusionsbehandlung noch zeigen wird, ist die Beziehung zwischen dem Ähnlichkeitsmaß $\rho_{mnr}(\mathbf{x}, \mathbf{d})$, das sich auf das rechte Bildkoordinatensystem bezieht und dem, das sich auf das linke Bildkoordinatensystem bezieht, von großer Bedeutung. Das Ähnlichkeitsmaß $\rho_{mnl}(\mathbf{x}, \mathbf{d})$ im linken Bildkoordinatensystem errechnet sich zu:

$$\rho_{mnl}(\mathbf{x}, \mathbf{d}) = \frac{w_{mn}(\mathbf{x}) * \tilde{r}_{mnl}(\mathbf{x}) \tilde{r}_{mnr}^*(\mathbf{x} + \mathbf{d})}{\sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnl}(\mathbf{x})|^2} \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x} + \mathbf{d})|^2}} \quad (4-66)$$

Wird der Disparitätsverlauf in diesem Ausdruck umgekehrt, so ergibt sich:

$$\rho_{mnl}(\mathbf{x}, -\mathbf{d}) = \frac{w_{mn}(\mathbf{x}) * \tilde{r}_{mnl}(\mathbf{x}) \tilde{r}_{mnr}^*(\mathbf{x} - \mathbf{d})}{\sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnl}(\mathbf{x})|^2} \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x} - \mathbf{d})|^2}} \quad (4-67)$$

Wird nun weiterhin die Eigenschaft des Faltungsproduktes

$$f(x + d) = g(x) * h(x + d) = g(x + d) * h(x) \quad (4-68)$$

auf Beziehung (4-67) angewendet und das Ähnlichkeitsmaß komplex konjugiert, so können die beiden Ähnlichkeitsmaße $\rho_{mnr}(\mathbf{x}, \mathbf{d})$ und $\rho_{mnl}(\mathbf{x}, \mathbf{d})$ ineinander überführt werden.

$$\rho_{mnl}^*(\mathbf{x} + \mathbf{d}, -\mathbf{d}) = \frac{w_{mn}(\mathbf{x}) * \tilde{r}_{mnl}^*(\mathbf{x} + \mathbf{d}) \tilde{r}_{mnr}(\mathbf{x})}{\sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})|^2} \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x})|^2}} = \rho_{mnr}(\mathbf{x}, \mathbf{d}) \quad (4-69)$$

Ähnlich der normierten Kreuzkorrelation ist das Ähnlichkeitsmaß invariant bezüglich einer Dämpfung einer der beiden Filterantworten durch einen reellen Faktor u . Diese Art von interokulären Differenzen kann durch einen Kontrastunterschied zwischen den Stereobildern hervorgerufen werden. Da sich diese Differenzen relativ leicht durch unterschiedliche Kameraeinstellungen ergeben können, kommt dieser Eigenschaft des Ähnlichkeitsmaßes ebenfalls eine besondere Bedeutung zu.

$$\begin{aligned} \tilde{\rho}_{mnr}(\mathbf{x}, \mathbf{d}) &= \frac{w_{mn}(\mathbf{x}) * u \tilde{r}_{mnr}(\mathbf{x}) \tilde{r}_{mnl}^*(\mathbf{x} + \mathbf{d})}{\sqrt{w_{mn}(\mathbf{x}) * |u \tilde{r}_{mnr}(\mathbf{x})|^2} \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})|^2}} \\ &= \frac{u (w_{mn}(\mathbf{x}) * \tilde{r}_{mnr}(\mathbf{x}) \tilde{r}_{mnl}^*(\mathbf{x} + \mathbf{d}))}{u \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnr}(\mathbf{x})|^2} \sqrt{w_{mn}(\mathbf{x}) * |\tilde{r}_{mnl}(\mathbf{x} + \mathbf{d})|^2}} = \rho_{mnr}(\mathbf{x}, \mathbf{d}) \end{aligned} \quad (4-70)$$

Existieren in der betrachteten Szene flächenhafte Strukturen, die eine Neigung relativ zur Stereobasis aufweisen, so unterscheiden sich die beiden Stereobilder neben der relativen Verschiebung zusätzlich durch eine Dehnung bzw. Stauchung der Intensitätsverläufe. Diese Dehnungen bzw. Stauchungen bewirken das Auftreten eines Gradienten im Verlauf der horizontalen Disparitätskomponente. Abgesehen von einigen merkmalsbasierten Ansätzen (siehe Kapitel 2.5.2) reagieren die meisten Verfahren zur Disparitätsmessung äußerst empfindlich auf diese Art von interokulären Differenzen. Eine einfache Methode zur Modellierung derartiger Unterschiede zwischen den Stereobildern wird in [JJT91] für eindimensionale Intensitätsverläufe beschrieben. In diesem Ansatz wird der Verlauf der horizontalen Disparitätskomponente durch eine lineare Funktion der Bildkoordinate x modelliert.

$$d_x(x) = c_0 + c_1x \quad (4-71)$$

Die Intensitätsverläufe i_l und i_r sind somit unter der Vernachlässigung von Beleuchtungsunterschieden zwischen den Bildern durch

$$i_l(x) = i_r(x + d_x(x)) = i_r(c_0 + (1 + c_1)x) \quad (4-72)$$

miteinander verbunden. Anstelle der relativen Verschiebungen zwischen den Intensitätsverläufen der Stereobilder muß bei einem solchen Ansatz ein Offset c_0 und eine Steigung c_1 geschätzt werden. Auf die zusätzliche Modellierung eines Gradienten im Disparitätsverlauf wird in diesem Ansatz jedoch verzichtet, da das Ähnlichkeitsmaß, wie schon erwähnt, eine gewisse Toleranz bezüglich relativer Dehnungen bzw. Stauchungen der Grauwertverläufe in den Stereobildern aufweist. Disparitätsgradienten können somit ohne eine explizite Modellierung in gewissen Grenzen erfaßt werden [Tra96d].

4.6 Korrespondenz durch Selbstorganisation

Der nahezu beliebige Disparitätsbereich, der mit dem im vorangegangenen Kapitel beschriebenen Korrelationsansatz erfaßt werden kann, wird gegenüber den phasenbasierten Verfahren durch den Nachteil erkauft, daß ähnlich wie bei den gebiets- und den merkmalsbasierten Ansätzen ein zusätzlicher Mechanismus notwendig ist, der die Mehrdeutigkeiten des Korrespondenzproblems auflöst. Obwohl Regularisierungsmethoden und stochastische Optimierungsansätze weit verbreitete Strategien zur Lösung dieses Problems sind, soll das Korrespondenzproblem in diesem Ansatz auf andere Weise gelöst werden. Den Grund hierfür bilden die zahlreichen prinzipiellen Nachteilen der obengenannten Verfahren, deren Ursprung meist in der Modellierung der Disparitätskarten zu finden ist. Ein weiterer Grund ist der enorme Berechnungsaufwand, der insbesondere bei den stochastischen Optimierungsansätzen auftritt (siehe Kapitel 2.5.4).

Um einerseits zu gewährleisten, daß die Laufzeit des Verfahrens in einem vertretbaren Rahmen bleibt, und andererseits eine ohnehin nicht mögliche exakte Modellierung der Disparitätskarten zu vermeiden, wird die Lösung des Korrespondenzproblems durch einen Selbstorganisationsansatz erzielt, der in die Gruppe der Relaxationsverfahren einzuordnen ist. Das Verfahren, das teilweise auf dem Ansatz von Reimann und Haken (siehe Kapitel 3) basiert, zeichnet sich durch eine implizite Okklusionsdetektion gegenüber den existierenden stereoskopischen Ansätzen aus. Durch diesen Mechanismus werden nur einseitig sichtbare Bereiche in der betrachteten Szene vom Zuordnungsprozeß ausgeschlossen, wodurch eine der wesentlichen Fehlerursachen in der Stereoskopie praktisch vollständig eliminiert werden kann.

Die Grundlage des im weiteren verwendeten Selbstorganisationsprozesses ist eine leicht modifizierte Version der gekoppelten Mustererkennungsgleichungen von Reimann und Haken (siehe Beziehung (3-5)).

$$\begin{aligned} \dot{\xi}(\mathbf{x}, \mathbf{d}, t) = & a\xi(\mathbf{x}, \mathbf{d}, t) - \frac{b}{P} \sum_{d' \neq d} \xi^2(\mathbf{x}, \mathbf{d}', t)\xi(\mathbf{x}, \mathbf{d}, t) - c\xi^3(\mathbf{x}, \mathbf{d}, t) \\ & + \frac{e}{Q} \sum_{\mathbf{x}' \in U} \xi(\mathbf{x}', \mathbf{d}, t)\xi(\mathbf{x}, \mathbf{d}, t) \end{aligned} \quad (4-73)$$

Der individuelle Verstärkungsterm $\lambda(\mathbf{x}, \mathbf{d})$ für jede Koordinate \mathbf{x} und Disparität \mathbf{d} in Beziehung (3-5) wird zur Vereinfachung in diesem Ansatz durch die Konstante a ersetzt und ist somit identisch für alle Variablen. Ebenso wird der Kopplungsparameter e des Umfeldes unabhängig von der Position der Variablen im Umfeld gewählt. Damit der Einfluß der Kopplungsterme nicht von der Anzahl der gekoppelten Variablen abhängt, werden die entsprechenden Parameter hier auf die Anzahl P und Q der beteiligten Variablen normiert. Die Summe $(b+c)$, die in Beziehung (3-5) als Gewichtung des dämpfenden Terms dient, wird hier zu dem Parameter b zusammengefaßt. Alle Parameter seien im weiteren positiv. Die zeitliche Dynamik des Selbstorganisationsprozesses wird ähnlich wie im Ansatz von Reimann und Haken [RH94] in zeitkontinuierlicher Form als gekoppeltes Differentialgleichungssystem beschrieben. Obwohl nur eine zeitdiskrete Formulierung direkt als Differenzgleichung numerisch auf einem Digitalrechner in Form eines Iterationsschemas implementierbar ist, wird die zeitkontinuierliche Beschreibungsform hier gewählt, da sie allgemeiner ist und die numerische Realisierung auf diese Weise unabhängig vom Prinzip der Selbstorganisation gewählt werden kann.

4.6.1 Initialisierung des Selbstorganisationsprozesses

Stimmen die Filterantworten der beiden Stereokanäle an einer Bildkoordinate bei einer bestimmten Disparität lokal überein, so ist die Phasendifferenz der Filterantworten gleich Null und der Realteil des Ähnlichkeitsmaßes nimmt den maximalen Wert 1,0 an. Der Imaginärteil weist in diesem Fall eine Nullstelle auf (vgl. Beziehung (4-65)). Anders als im Ansatz von Reimann und Haken wird hier der Realteil des in Kapitel 4.5.1 definierten Ähnlichkeitsmaßes zur

Initialisierung des Selbstorganisationsprozesses herangezogen, da dieser die Übereinstimmung der Filterantworten repräsentiert. Um dichter besetzte Disparitätskarten zu erhalten, werden zusätzlich zu den Antworten der Gaborfilter, deren Modulationsortsfrequenzvektor genau in Richtung der Epipolarlinie zeigt, auch die Filterantworten benachbarter Orientierungen zur Initialisierung verwendet (vgl. Bild 4-13). Zu diesem Zweck werden alle Variablen zum Zeitpunkt $t = 0$ mit Hilfe einer Funktion f , die von der gewichteten Überlagerung der Realteile des Ähnlichkeitsmaßes unterschiedlicher Filterorientierungen abhängt, initialisiert.

$$\xi_r(\mathbf{x}, \mathbf{d}, t = 0) = f\left(\sum_{m=-M_s}^{M_s} \mu_m \operatorname{Re}\{\rho_{mnr}(\mathbf{x}, \mathbf{d})\}\right) \quad (4-74)$$

Da der im weiteren verwendete Selbstorganisationsprozeß identisch für alle N Auflösungsebenen ist, wird auf einen zusätzlichen Index n bei den einzelnen Variablen zur Erhöhung der Übersichtlichkeit verzichtet. Die Variablenwerte beziehen sich hier zunächst auf die rechten Bildkoordinaten. Die Gewichtung μ_m der einzelnen Filterorientierungen m sei mit $\mu_m = 1/(2M_s + 1)$ identisch für alle Orientierungen. Da das dynamische Verhalten der Haken'schen Mustererkennungsgleichung für positive und negative Variablenwerte identisch ist (siehe [Hak91]), werden im folgenden nur positive Variablenwerte betrachtet. Die Funktion f dient in diesem Zusammenhang zur Abbildung der Summe in Beziehung (4-74) auf einen nicht negativen Wertebereich. Soweit nicht anders erwähnt, geschieht dies im weiteren durch eine einfache Schwellwertoperation, die negative Werte durch den Wert Null ersetzt.

$$f(x) = \begin{cases} x; & \text{für } x \geq 0 \\ 0; & \text{sonst} \end{cases} \quad (4-75)$$

Der Wertebereich der Variablen zum Zeitpunkt $t = 0$ ergibt sich unter diesen Voraussetzungen zu:

$$0 \leq \xi_r(\mathbf{x}, \mathbf{d}, t = 0) \leq 1 \quad \forall \mathbf{x}, \mathbf{d} \quad (4-76)$$

Bei allen im weiteren Verlauf der Arbeit gezeigten praktischen Beispielen werden die in Abbildung 4-13 markierten drei Filterorientierungen ($M_s = 1$) zur Initialisierung verwendet.

4.6.2 Unterdrückung von Okklusionskorrespondenzen

Ein erhebliches Problem in der Stereoskopie stellen Bereiche in der betrachteten Szene dar, die aufgrund der unterschiedlichen Perspektiven, aus denen die Stereobilder aufgenommen werden, in einem Bild zu sehen sind und im anderen nicht. Innerhalb dieser sogenannten Okklusionsbereiche ist die Disparität nicht definiert, da kein korrespondierender Bildbereich im anderen Stereokanal existiert. Bei den meisten existierenden Verfahren zur stereoskopischen Korrespondenzbestimmung wird diese Tatsache nicht weiter berücksichtigt und die Okklusionsbereiche werden mehr oder weniger zufällig Bereichen im anderen Stereobild zugeordnet. Diese Korrespondenzfehler resultieren jedoch in fehlerhaften Positionsbestimmungen bei der räumlichen Rekonstruktion des entsprechenden Bildausschnittes.

Die bislang einzige Strategie zur Behandlung dieses Problems, die in der Literatur zu finden ist, basiert auf einem bidirektionalen Zuordnungsprozeß. Die wenigen Ansätze, die das Okklusionsproblem auf diese Weise berücksichtigen, versuchen zur Detektion der Okklusionsbereiche die Stereodaten bidirektional in zwei identischen, aber weitestgehend unabhängigen Prozessen in Deckung zu bringen [GLY92, JM92, LB95]. Hierbei wird das linke Bild mit dem rechten verglichen und umgekehrt. Ergeben sich in beiden Zuordnungsversuchen für eine Bildregion nicht die gleichen Ergebnisse, so wird dieser Bereich als Okklusionsbereich deklariert. Im allgemeinen wird bei diesen Ansätzen eine binäre Okklusionskarte eingeführt, die Okklusionsbereiche im Rahmen einer Regularisierung oder Relaxation ausblendet und somit von der Optimierung ausschließt.

Die bidirektionalen Strategien basieren auf der Annahme, daß die Merkmale oder Bildbereiche in Okklusionen prinzipiell anderen Bildbereichen unähnlicher sind als diese ihren eigentlichen Korrespondenzpartnern. Da sich die Stereobilder aufgrund der verschiedenen Perspektive und anderer interokulärer Differenzen prinzipiell unterscheiden, ist diese Annahme nicht allgemein gültig. Darüber hinaus wird im allgemeinen nicht beachtet, daß die Okklusionen des einen Bildes auch Okklusionen des anderen Bildes zugeordnet werden können. Besonders aufwendig ist bei diesen Verfahren das Durchlaufen aller Verfahrensschritte, wie die Messung von Ähnlichkeiten, die Regularisierung oder die Relaxation für beide Richtungen. Der Berechnungsaufwand ist somit wesentlich höher als bei gleichen Ansätzen ohne Okklusionsbehandlung.

Um Korrespondenzfehler, die durch Okklusionen hervorgerufen werden, zu vermeiden, wird an dieser Stelle der Ansatz von Reimann und Haken um einen Mechanismus erweitert, der im Rahmen des Selbstorganisationsprozesses verhindert, daß Okklusionsbereiche anderen Bildbereichen zugeordnet werden. Unter Ausnutzung der Beziehung zwischen den Ähnlichkeitsmaßen, die in unterschiedlicher Richtung berechnet werden (siehe Gleichung (4-69)), können hierbei Okklusionen ohne einen aufwendigen bidirektionalen Ansatz zuverlässig detektiert werden.

Im Selbstorganisationsprozeß von Reimann und Haken besteht das Problem, daß die Variablen in Okklusionsbereichen, die eine Korrespondenz mit nicht okkludierten Bereichen im anderen Bild repräsentieren, nicht daran gehindert werden, den Wettbewerb zu gewinnen. Diese nicht

okkludierten Bereiche im anderen Bild besitzen jedoch korrekte Korrespondenzpartner in dem Bild, in dem die Okklusionen auftreten. Im Realteil des Ähnlichkeitsmaßes, das in umgekehrter Richtung berechnet wird, müßten somit an den entsprechenden Bildkoordinaten Korrelationsmaxima auftreten. Genau dieser Fall soll in dem nachfolgenden Beispiel verdeutlicht werden.

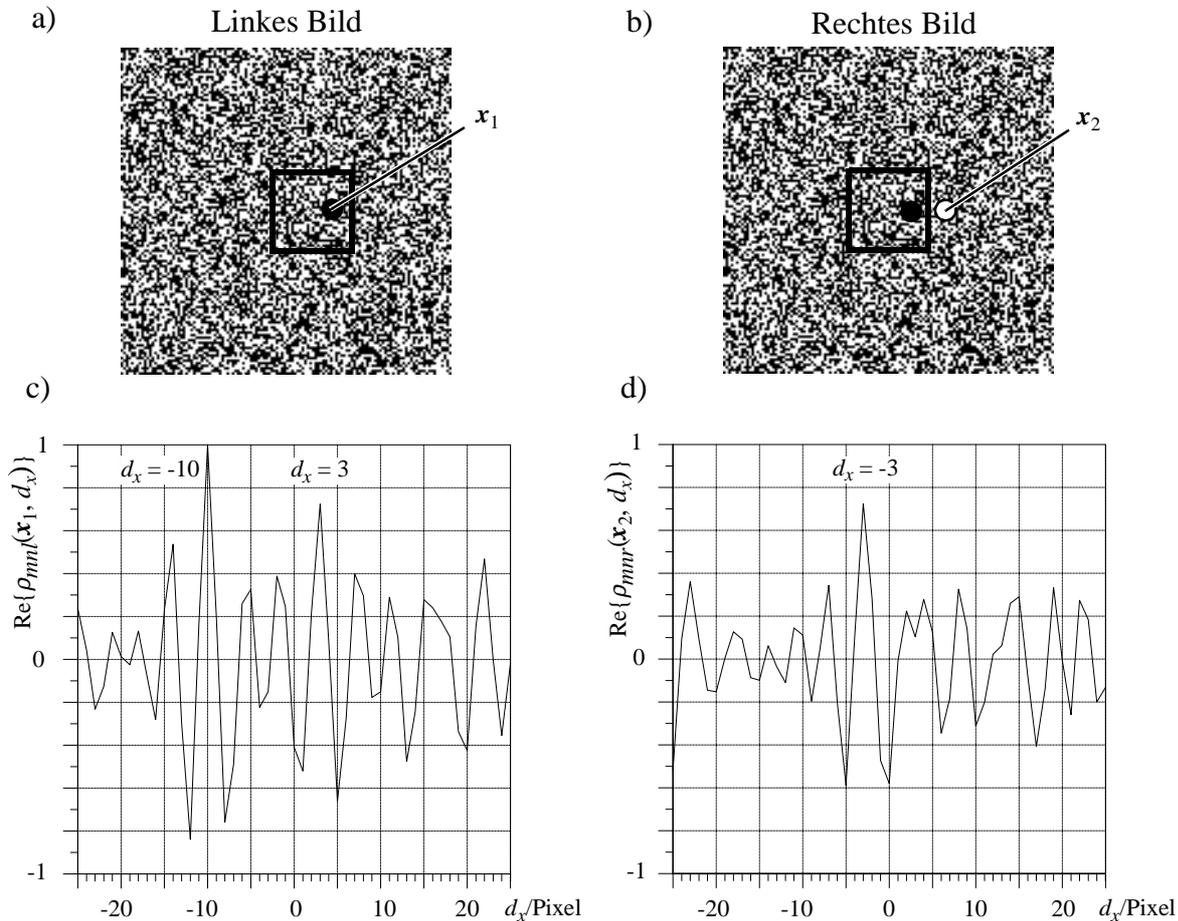


Bild 4-16: a+b) Zufallspunktstereogramm. c+d) Ähnlichkeitsmaße an den gekennzeichneten Bildpunkten im rechten und linken Bild (Erläuterung siehe Text)

Die beiden oberen Bilder in Abbildung 4-16 zeigen das linke und rechte Bild eines Zufallspunktstereogramms. Der horizontal versetzt eingefügte Bereich ($d_y = 0$) in Form eines Quadrates ist zur Orientierung schwarz umrandet. In Abbildung 4-16d ist der Verlauf des Realteils des Ähnlichkeitsmaßes in Abhängigkeit der Disparität an der Bildkoordinate, die im rechten Bild des Stereogramms als weißer Punkt gekennzeichnet ist, zu sehen. Da der korrespondierende Bildpunkt im linken Bild durch das Quadrat verdeckt wird, liegt der weiße Punkt in einem Okklusionsbereich. Das deutliche Maximum im Verlauf des Realteils des Ähnlichkeitsmaßes bei der Disparität $d_x = -3$ in Abbildung 4-16d repräsentiert die Übereinstimmung mit dem Bildbereich im linken Bild, der durch den schwarzen Punkt gekennzeichnet ist. Im Zuge eines

Selbstorganisationsprozesses ohne Okklusionsbehandlung hätte die Variable, die mit diesem Ähnlichkeitsmaß initialisiert wird, aufgrund ihres hohen Wertes gute Voraussetzungen, den Wettbewerb zu gewinnen. Der gleiche Wert ist nach Beziehung (4-69) im Realteil des Ähnlichkeitsmaßes, das sich auf das linke Bild des Zufallpunktstereogramms bezieht (vgl. Bild 4-16c), an der schwarz gekennzeichneten korrespondierenden Bildkoordinate, bei der Disparität $d_x = 3$ zu finden. Entscheidend ist jedoch an dieser Stelle, daß der korrekte Korrespondenzpartner dieses Bildpunktes, der im rechten Bild des Zufallpunktstereogramms ebenfalls schwarz gekennzeichnet ist, im Realteil des Ähnlichkeitsmaßes ρ_l als absolutes Maximum bei der Disparität $d_x = -10$ vorhanden ist.

Die grundlegende Idee besteht nun darin, zusätzlich die Variablen $\xi_l(\mathbf{x}, \mathbf{d}, t)$ in den Wettbewerbsprozeß aus Beziehung (4-73) zu integrieren, die mit dem Ähnlichkeitsmaß, das sich auf das linke Bild bezieht, initialisiert werden. Wirken nun diese Variablen ebenfalls dämpfend auf einen Variablenwert, der sich in einem okkludierten Bereich im rechten Bild befindet, so wird diese Variable den Wettbewerb verlieren, wenn eine der zusätzlichen Variablen die korrekte Korrespondenz mit einem ausreichend hohen Initialisierungswert oder einem geeignetem Umfeld U repräsentiert. Um diese Strategie zu realisieren, werden die Variablen $\xi_l(\mathbf{x}, \mathbf{d}, t)$ in den dämpfend wirkenden Term von Beziehung (4-73) eingefügt.

$$\begin{aligned} \dot{\xi}_r(\mathbf{x}, \mathbf{d}, t) &= a \xi_r(\mathbf{x}, \mathbf{d}, t) - c \xi_r^3(\mathbf{x}, \mathbf{d}, t) \\ &\quad - \frac{b}{2P} \sum_{\mathbf{d}' \neq \mathbf{d}} (\xi_r^2(\mathbf{x}, \mathbf{d}', t) + \xi_l^2(\mathbf{x} + \mathbf{d}, -\mathbf{d}', t)) \xi_r(\mathbf{x}, \mathbf{d}, t) \\ &\quad + \frac{e}{Q} \sum_{\mathbf{x}' \in U} \xi_r(\mathbf{x}', \mathbf{d}, t) \xi_r(\mathbf{x}, \mathbf{d}, t) \end{aligned} \quad (4-77)$$

Da nun die doppelte Anzahl von Variablen dämpfend auf die Variablenwerte wirkt, wird die Normierung des Dämpfungsterms entsprechend angepaßt. Die Dynamik der Variablenwerte $\xi_l(\mathbf{x}, \mathbf{d}, t)$ im linken Bildkoordinatensystem müßte analog zu Beziehung (4-77) errechnet werden, so daß eine doppelte Anzahl von Variablen und der doppelte Berechnungsaufwand notwendig wären, um diese spezielle bidirektionale Strategie zu implementieren. Da jedoch die Variablenwerte unmittelbar aus den Ähnlichkeitswerten $\rho_r(\mathbf{x}, \mathbf{d})$ hervorgehen (siehe Gleichung (4-74)), gilt nach Beziehung (4-69):

$$\xi_l(\mathbf{x}, \mathbf{d}, t) = \xi_r(\mathbf{x} + \mathbf{d}, -\mathbf{d}, t) \quad (4-78)$$

Mit Hilfe dieses Zusammenhangs können die Variablenwerte $\xi_l(\mathbf{x}, \mathbf{d}, t)$ in Gleichung (4-77) durch die entsprechenden Variablenwerte, die sich auf das rechte Bildkoordinatensystem

beziehen, ersetzt werden. Für den dämpfend wirkenden Term in Gleichung (4-77) ergibt sich somit:

$$\sum_{d' \neq d} \xi_r^2(\mathbf{x}, \mathbf{d}', t) + \xi_l^2(\mathbf{x} + \mathbf{d}, -\mathbf{d}', t) = \sum_{d' \neq d} \xi_r^2(\mathbf{x}, \mathbf{d}', t) + \xi_r^2(\mathbf{x} + \mathbf{d} - \mathbf{d}', \mathbf{d}', t) \quad (4-79)$$

Die Integration der Okklusionsbehandlung in den Selbstorganisationsprozeß kann auf diese Weise ohne einen aufwendigen bidirektionalen Prozeß erfolgen. Der doppelte Berechnungs- und Speicheraufwand entfällt somit. Im weiteren wird der Relaxationsprozeß nur auf das rechte Bildkoordinatensystem bezogen und auf den entsprechenden Index r verzichtet. Das dynamische Verhalten der Variablen wird somit durch

$$\begin{aligned} \dot{\xi}(\mathbf{x}, \mathbf{d}, t) &= a\xi(\mathbf{x}, \mathbf{d}, t) - c\xi^3(\mathbf{x}, \mathbf{d}, t) \\ &\quad - \frac{b}{2P} \sum_{d' \neq d} (\xi^2(\mathbf{x}, \mathbf{d}', t) + \xi^2(\mathbf{x} + \mathbf{d} - \mathbf{d}', \mathbf{d}', t)) \xi(\mathbf{x}, \mathbf{d}, t) \\ &\quad + \frac{e}{Q} \sum_{\mathbf{x}' \in U} \xi(\mathbf{x}', \mathbf{d}, t) \xi(\mathbf{x}, \mathbf{d}, t) \end{aligned} \quad (4-80)$$

beschrieben.

4.6.3 Unterdrückung von Okklusionspaaren

Durch den im vorangegangenen Abschnitt beschriebenen Mechanismus kann zwar verhindert werden, daß Variablen den Wettbewerb gewinnen, die eine Korrespondenz eines okkludierten Bildbereiches mit einem Bildbereich im anderen Bild repräsentieren, der seinerseits über einen korrekten Korrespondenzpartner verfügt. Jedoch kann auf diese Weise nicht ausgeschlossen werden, daß Variablen den Wettbewerb gewinnen, die eine Korrespondenz eines Okklusionsbereiches im linken Bild mit einem Okklusionsbereich im rechten Bild repräsentieren. In diesem Fall besitzt keiner der beiden Bildbereiche einen korrekten Korrespondenzpartner. Da Diskontinuitäten in der Tiefe einer betrachteten Szene grundsätzlich sowohl im linken als auch im rechten Stereobild Okklusionen erzeugen, müssen derartige Fehlkorrespondenzen verhindert werden.

In bislang bestehenden Techniken wird dieser Fall im allgemeinen nicht berücksichtigt, da die Okklusionen, die ein Objekt im linken und im rechten Bild erzeugt, grundsätzlich auf unterschiedlichen Seiten des Objektes in der Bildebene zu finden sind. Um eine Fehlkorrespondenz zwischen diesen Bildbereichen zu erzeugen, muß der zugelassene Disparitätsbereich größer sein als die Bildregion, in die das gesamte Objekt abgebildet wird. Der Disparitätsbereich muß

jedoch bei den meisten Verfahren aufgrund des Berechnungsaufwandes oder der angewendeten Technik zwangsläufig sehr klein gewählt werden, so daß Okklusionspaare bei diesen Techniken relativ selten auftreten. Da der Disparitätsbereich bei der hier zum Einsatz kommenden Strategie nicht durch das Verfahren selbst begrenzt wird und darüber hinaus große Tiefenbereiche erfaßt werden sollen, wird ein weiterer Mechanismus vorgesehen, der diese Art von Fehlkorrespondenzen unterdrückt.

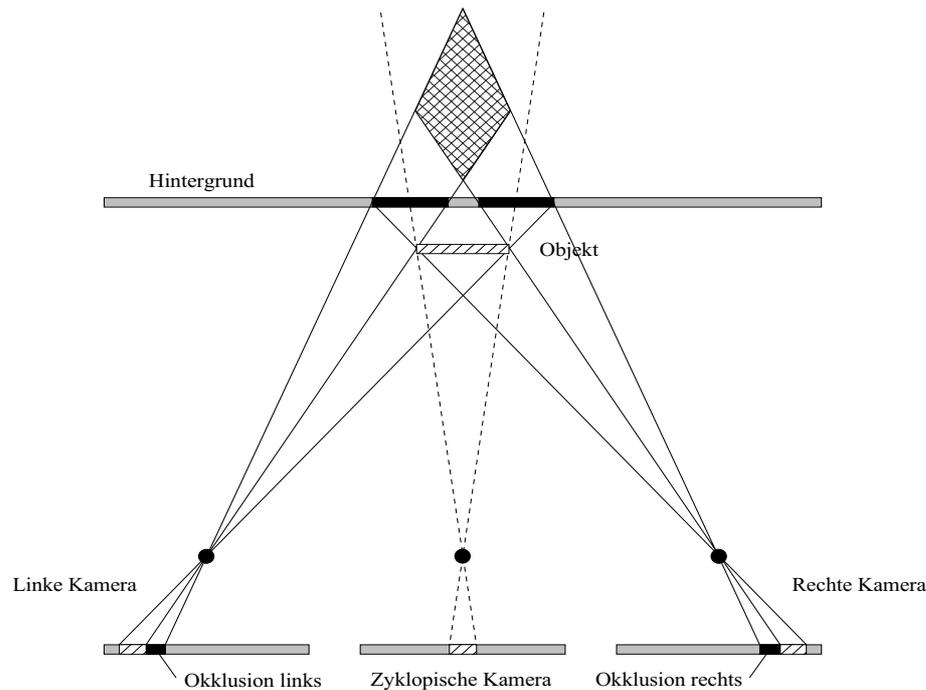


Bild 4-17: Bei unterschiedlich zum Kamerasystem entfernten räumlichen Strukturen werden Okklusionen sowohl im linken als auch im rechten Bild erzeugt (schwarze Rechtecke). Aus Sicht einer zyklopischen Kamera verdecken sich die rekonstruierte räumliche Position einer gegenseitigen Zuordnung dieser Bereiche (kariertes Bereich) und das betreffende Objekt immer gegenseitig.

Zur Vermeidung von Okklusionspaaren wird eine virtuelle zyklopische Kamera, die in der Mitte zwischen den beiden Kamerasystemen angeordnet ist, eingeführt. Aus der Perspektive einer solchen Kamera befinden sich das Objekt und die rekonstruierte räumliche Position, die sich durch die Korrespondenz der Okklusionsbereiche ergibt, direkt hintereinander und werden somit auf denselben Bildbereich in der zyklopischen Bildebene abgebildet. Die Idee besteht nun darin, den Selbstorganisationsprozeß aus Beziehung (4-80) derart zu erweitern, daß zusätzlich ein Wettbewerb zwischen allen Variablen, die in einem zyklopischen Bildkoordinatensystem die gleiche Position aufweisen würden, erzwungen wird.

Die zyklopische Bildkoordinate x_c eines Variablenwertes $\xi(x, d, t)$ an der Koordinate x im rechten Bild und der Disparität d kann mit Hilfe der Beziehungen (2-3) und (2-4) zu

$$\mathbf{x}_c = \mathbf{x} + \frac{1}{2}\mathbf{d} \quad (4-81)$$

errechnet werden. Um den oben beschriebenen Wettbewerb zu realisieren, müssen nun alle Variablen, deren Bildkoordinate $\hat{\mathbf{x}}$ und deren Disparität \mathbf{d}' die Gleichung

$$\hat{\mathbf{x}} + \frac{1}{2}\mathbf{d}' = \mathbf{x} + \frac{1}{2}\mathbf{d}, \quad \text{mit } \mathbf{d}' \neq \mathbf{d} \quad (4-82)$$

erfüllen, dämpfend auf den Variablenwert $\xi(\mathbf{x}, \mathbf{d}, t)$ wirken. Der Dämpfungsterm in Beziehung (4-80) wird zu diesem Zweck um eine zusätzliche Komponente, die alle Variablen mit der zuvor gezeigten Eigenschaft einbezieht, erweitert. Mit Beziehung (4-82) und einer entsprechend angepaßten Normierung des Dämpfungsterms ergibt sich:

$$\begin{aligned} \dot{\xi}(\mathbf{x}, \mathbf{d}, t) &= \left(a - c\xi^2(\mathbf{x}, \mathbf{d}, t) \right. \\ &\quad - \frac{b}{3P} \sum_{\mathbf{d}' \neq \mathbf{d}} \xi^2(\mathbf{x}, \mathbf{d}', t) + \xi^2(\mathbf{x} + \mathbf{d} - \mathbf{d}', \mathbf{d}', t) + \xi^2\left(\mathbf{x} + \frac{1}{2}(\mathbf{d} - \mathbf{d}'), \mathbf{d}', t\right) \\ &\quad \left. + \frac{e}{Q} \sum_{\mathbf{x}' \in U} \xi(\mathbf{x}', \mathbf{d}, t) \right) \xi(\mathbf{x}, \mathbf{d}, t) \\ &= f(\mathbf{x}, \mathbf{d}, t) \end{aligned} \quad (4-83)$$

Durch Gleichung (4-83) wird die Dynamik der einzelnen Variablen des im weiteren verwendeten Selbstorganisationsprozesses definiert. Neben der Errechnung von eindeutigen und kontinuierlichen Disparitätskarten erlaubt dieser Relaxationsansatz eine implizite Detektion von Okklusionen, da für diese im Rahmen der Selbstorganisation keine Korrespondenzen errechnet werden. Der erweiterte Wettbewerbsprozeß kann hierbei auch als eine erweiterte Eindeutigkeitseinschränkung interpretiert werden [TDH98].

4.6.4 Stabilitätsbetrachtungen

Der Wertebereich der Variablen in Beziehung (4-83) ist aufgrund der speziellen Struktur des Differentialgleichungssystems begrenzt. Ein unbegrenztes exponentielles Wachstum ist ausgeschlossen, da die verstärkend wirkenden Terme eine niedrigere Ordnung aufweisen als die Dämpfungsterme. Da sich die endgültigen Korrespondenzen, wie schon erwähnt, bei diesem Ansatz aus den stationären Lösungen der einzelnen Variablen ergeben, müssen diese unter bestimmten Voraussetzungen stabil sein. Welche Vorkehrungen zu treffen sind, damit diese

Stabilität gewährleistet werden kann, soll eine lineare Stabilitätsanalyse der stationären Lösungen klären.

Beziehung (4-83) weist ähnlich der Haken'schen Mustererkennungsgleichung (siehe Kapitel 3.1) insgesamt drei stationäre Lösungen mit $\dot{\xi}_s(\mathbf{x}, \mathbf{d}, t) = 0$ pro Bildkoordinate auf:

$$\begin{aligned}\xi_{s_1}(\mathbf{x}, \mathbf{d}) &= 0 \\ \xi_{s_{2/3}}(\mathbf{x}, \mathbf{d}) &= \pm \sqrt{\frac{\eta(\mathbf{x}, \mathbf{d})}{c}}\end{aligned}\tag{4-84}$$

mit:

$$\begin{aligned}\eta(\mathbf{x}, \mathbf{d}) &= a + \frac{e}{Q} \sum_{\mathbf{x}' \in U} \xi_s(\mathbf{x}', \mathbf{d}) \\ &\quad - \frac{b}{3P} \sum_{\mathbf{d}' \neq \mathbf{d}} \xi_s^2(\mathbf{x}, \mathbf{d}') + \xi_s^2(\mathbf{x} + \mathbf{d} - \mathbf{d}', \mathbf{d}') + \xi_s^2\left(\mathbf{x} + \frac{1}{2}(\mathbf{d} - \mathbf{d}'), \mathbf{d}'\right)\end{aligned}\tag{4-85}$$

Da der Parameter $c > 0$ gewählt wird, existieren die Lösungen $\xi_{s_{2/3}}(\mathbf{x}, \mathbf{d})$ nur dann, wenn der Term $\eta(\mathbf{x}, \mathbf{d})$ ebenfalls positiv ist. Im weiteren wird neben der Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d})$ nur die positive stationäre Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ betrachtet, da, wie noch gezeigt wird, die negative Lösung aufgrund der hier verwendeten positiven Startwerte nie angenommen wird. Um die lokale Stabilität der stationären Lösungen zu untersuchen, werden kleine Auslenkungen $\Delta\xi(\mathbf{x}, \mathbf{d}, t)$ aus der Ruhelage betrachtet:

$$\xi(\mathbf{x}, \mathbf{d}, t) = \xi_s(\mathbf{x}, \mathbf{d}) + \Delta\xi(\mathbf{x}, \mathbf{d}, t)\tag{4-86}$$

Werden alle Variablen zu einem Vektor $\xi(t)$ zusammengefaßt und wird Beziehung (4-83) um die stationäre Lösung linearisiert, so kann das Gesamtsystem durch

$$\Delta\dot{\xi}(t) = \mathbf{J}\Delta\xi(t)\tag{4-87}$$

beschrieben werden. Die Matrix \mathbf{J} stellt hierbei die Jacobi-Matrix des Differentialgleichungssystems im stationären Zustand dar. Sind die Realteile aller Eigenwerte von \mathbf{J} negativ, so ist die stationäre Lösung asymptotisch stabil [CL55]. Die Komponenten der Jacobi-Matrix, die ungleich Null sind, ergeben sich nach Beziehung (4-83) zu:

$$\begin{aligned}
\frac{\partial}{\partial \xi(x, \mathbf{d}, t)} f(x, \mathbf{d}, t)|_{\xi_s} &= \eta(x, \mathbf{d}) - 3c \xi_s^2(x, \mathbf{d}) \\
\frac{\partial}{\partial \xi(x, \mathbf{d}', t)} f(x, \mathbf{d}, t)|_{\xi_s} &= -2b \xi_s(x, \mathbf{d}') \xi_s(x, \mathbf{d}) / 3P \\
\frac{\partial}{\partial \xi(x + \mathbf{d} - \mathbf{d}', \mathbf{d}', t)} f(x, \mathbf{d}, t)|_{\xi_s} &= -2b \xi_s(x + \mathbf{d} - \mathbf{d}', \mathbf{d}') \xi_s(x, \mathbf{d}) / 3P \\
\frac{\partial}{\partial \xi\left(x + \frac{1}{2}(\mathbf{d} - \mathbf{d}'), \mathbf{d}', t\right)} f(x, \mathbf{d}, t)|_{\xi_s} &= -2b \xi_s\left(x + \frac{1}{2}(\mathbf{d} - \mathbf{d}'), \mathbf{d}'\right) \xi_s(x, \mathbf{d}) / 3P \\
\frac{\partial}{\partial \xi(x', \mathbf{d}, t)} f(x, \mathbf{d}, t)|_{\xi_s} &= \frac{e}{Q} \xi_s(x, \mathbf{d})
\end{aligned} \tag{4-88}$$

mit $\mathbf{d}' \neq \mathbf{d}$ und $x' \in U$

Um die Stabilitätseigenschaften des Systems zu ermitteln, werden nun verschiedene Konstellationen der unterschiedlichen stationären Lösungen betrachtet, die das System annehmen kann.

Zunächst soll angenommen werden, daß eine bestimmte Variable und alle auf sie dämpfend wirkenden Elemente die Lösung $\xi_{s_1}(x, \mathbf{d}) = 0$ annehmen. Die Variablen im Umfeld U der betrachteten Variablen seien dabei beliebig. Die einzigen nicht verschwindenden Elemente der Matrix \mathbf{J} ergeben sich in diesem Fall aus den Beziehungen (4-85) und (4-88) zu:

$$\frac{\partial}{\partial \xi(x, \mathbf{d}, t)} f(x, \mathbf{d}, t)|_{\xi_{s_1}} = a + \frac{e}{Q} \sum_{x' \in U} \xi_s(x', \mathbf{d}) \tag{4-89}$$

Der betreffende Unterblock der Matrix \mathbf{J} weist in diesem Fall nur auf der Hauptdiagonalen Werte ungleich Null auf. Da die Parameter a und e größer Null gewählt werden und alle Variablen durch das Ähnlichkeitsmaß aus Beziehung (4-74) mit Werten größer oder gleich Null initialisiert werden, sind alle Werte auf der Hauptdiagonalen und damit alle betreffenden Eigenwerte reell und größer Null. Die stationäre Lösung ist somit instabil. Aus diesem Grund können nicht alle Variablen, die im gegenseitigen Wettbewerb stehen, gleichzeitig die stationäre Lösung $\xi_{s_1}(x, \mathbf{d}) = 0$ annehmen.

Für den Fall, daß eine Variable an einem bestimmten Bildpunkt x_0 die stationäre Lösung $\xi_{s_1}(x_0, \mathbf{d}_0) = 0$ annimmt und alle anderen Variablen beliebige Werte aufweisen, besitzt der betreffende Unterblock der Matrix \mathbf{J} ebenfalls nur auf der Hauptdiagonalen Werte ungleich Null mit:

$$\frac{\partial}{\partial \xi(x, \mathbf{d}, t)} f(x, \mathbf{d}, t)|_{\xi_{s_1}} = \eta(x_0, \mathbf{d}_0) \tag{4-90}$$

Da $\eta(\mathbf{x}_0, \mathbf{d}_0)$ reell ist, sind auch alle Eigenwerte dieses Unterblocks reell. Ist $\eta(\mathbf{x}_0, \mathbf{d}_0)$ positiv, so ist der betreffende Eigenwert ebenfalls positiv. Die stationäre Lösung ist somit instabil und bei kleinen positiven Auslenkungen aus der Ruhelage wächst der Wert der Variable. Ist $\eta(\mathbf{x}_0, \mathbf{d}_0)$ negativ, so ist die stationäre Lösung $\xi_{s_1}(\mathbf{x}_0, \mathbf{d}_0) = 0$ asymptotisch stabil. Da durch die Initialisierung nur positive Auslenkungen aus der Ruhelage zugelassen werden und alle Eigenwerte reell sind, wird sowohl im stabilen als auch im instabilen Fall der Variablenwert $\xi_{s_1}(\mathbf{x}, \mathbf{d}) = 0$ nie unterschritten.

Die Parameter des Systems sollen nun so gewählt werden, daß die stationäre Lösung $\xi_{s_1}(\mathbf{x}_0, \mathbf{d}_0) = 0$ an einer bestimmten Bildkoordinate \mathbf{x}_0 stabil wird, wenn eine der am Wettbewerbsprozeß beteiligten Variablen die stationäre Lösung $\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1)$ annimmt und auf diese Weise den Wettbewerb gewinnt, während alle übrigen am gleichen Wettbewerb beteiligten Variablen ebenfalls die stationäre Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d}) = 0$ annehmen. Der betreffende Unterblock der Matrix \mathbf{J} weist in diesem Fall wiederum nur reelle Elemente auf der Hauptdiagonalen auf:

$$\frac{\partial}{\partial \xi(\mathbf{x}, \mathbf{d}, t)} f(\mathbf{x}, \mathbf{d}, t) \Big|_{\xi_{s_1}} = a - \frac{b}{3P} \xi_{s_2}^2(\mathbf{x}_1, \mathbf{d}_1) \quad (4-91)$$

mit $\mathbf{d}_1 \neq \mathbf{d}_0$ und $\mathbf{x}_1 \in \left\{ \mathbf{x}_0, \mathbf{x}_0 + \mathbf{d}_0 - \mathbf{d}_1, \mathbf{x}_0 + \frac{1}{2}(\mathbf{d}_0 - \mathbf{d}_1) \right\}$

Da die zugehörigen Eigenwerte die Elemente der Hauptdiagonalen sind, muß für die asymptotische Stabilität der stationären Lösung somit die Bedingung

$$a - \frac{b}{3P} \xi_{s_2}^2(\mathbf{x}_1, \mathbf{d}_1) < 0 \quad (4-92)$$

erfüllt sein. Da der Wert der stationären Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ von den Variablenwerten des betreffenden Umfeldes U abhängt (vgl. Beziehung (4-84) und (4-85)), wird im folgenden gefordert, daß die Bedingung (4-92) auch für den ungünstigsten Fall erfüllt wird. Diese Situation tritt dann ein, wenn im Umfeld U der stationären Lösung $\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1)$ alle Variablen mit der Disparität \mathbf{d}_1 die stationäre Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d})$ angenommen haben. Aus Beziehung (4-85) ergibt sich dann:

$$\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1) = \sqrt{\frac{a}{c}} \quad (4-93)$$

Die Parameter b , c und die Anzahl P der übrigen zugelassenen Disparitäten müssen nach den Beziehungen (4-92) und (4-93) mit $a > 0$ die Ungleichung

$$b > 3Pc \quad (4-94)$$

erfüllen.

Nimmt eine beliebige Variable die stationäre Lösung $\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1)$ an und alle anderen Variablen, die am gleichen Wettbewerbsprozeß beteiligt sind, nehmen die Lösung $\xi_{s_1}(\mathbf{x}_0, \mathbf{d}_0)$ an, so ist $\eta(\mathbf{x}_1, \mathbf{d}_1)$ bei den hier zugelassenen Parameterwerten immer positiv und reell. Die nicht verschwindenden Komponenten des betreffenden Unterblocks der Matrix \mathbf{J} ergeben sich für diesen Fall aus Beziehung (4-88) zu:

$$\begin{aligned} \frac{\partial}{\partial \xi(\mathbf{x}, \mathbf{d}, t)} f(\mathbf{x}, \mathbf{d}, t) \Big|_{\xi_{s_2}} &= -2\eta(\mathbf{x}_1, \mathbf{d}_1) \\ \frac{\partial}{\partial \xi(\mathbf{x}', \mathbf{d}, t)} f(\mathbf{x}, \mathbf{d}, t) \Big|_{\xi_{s_2}} &= \frac{e}{Q} \xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1) \end{aligned} \quad (4-95)$$

Eine quantitative Aussage über die Eigenwerte dieses Unterblocks der Matrix \mathbf{J} kann mit Hilfe des Satzes von Gerschgorin getroffen werden (siehe z. B. [Var62]).

Sei $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{N \times N}$ eine beliebige komplexwertige $N \times N$ Matrix, so sind die sogenannten Gerschgorin-Kreise wie folgt definiert:

$$g_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^N |a_{ij}| \right\} \quad \forall i \in [1, \dots, N] \quad (4-96)$$

Die Eigenwerte von \mathbf{A} sind dann in der Vereinigung aller Gerschgorin-Kreise g_i enthalten.

Ist \mathbf{A} strikt diagonal dominant, mit

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^N |a_{ij}| \quad \forall i \in [1, \dots, N] \quad (4-97)$$

und sind darüber hinaus alle Diagonalelemente von \mathbf{A} negative reelle Zahlen, so folgt unmittelbar aus dem Satz von Gerschgorin, daß alle Eigenwerte von \mathbf{A} negative Realteile aufweisen

[Var62]. Eine stabile stationäre Lösung $\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1)$ liegt nach Beziehung (4-95) somit immer dann vor, wenn die Ungleichung

$$2\eta(\mathbf{x}_1, \mathbf{d}_1) = 2c\xi_{s_2}^2(\mathbf{x}_1, \mathbf{d}_1) > \frac{e}{Q} \sum_{q=1}^Q \xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1) = e\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1) \quad (4-98)$$

erfüllt ist. Hieraus ergibt sich:

$$\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1) > \frac{e}{2c} \quad (4-99)$$

Um $\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1)$ nach unten abschätzen zu können und auf diese Weise weitere Aussagen über die Parameter zu erhalten, mit denen die Ungleichung (4-99) erfüllt werden kann, wird der Fall betrachtet, daß L Nachbarn im Umfeld U von $\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1)$ ebenfalls eine stationäre Lösung mit gleich hohem Wert annehmen. Die stationäre Lösung ergibt sich für diesen Fall zu:

$$\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1) = \sqrt{\frac{a + \frac{eL}{Q}\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1)}{c}} \quad (4-100)$$

Wird diese Gleichung nach $\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1)$ aufgelöst, so ergibt sich die einzig positive Lösung zu:

$$\xi_{s_2}(\mathbf{x}_1, \mathbf{d}_1) = \frac{\frac{eL}{Q} + \sqrt{\left(\frac{eL}{Q}\right)^2 + 4ac}}{2c} \quad (4-101)$$

Durch Einsetzen dieses Ausdrucks in Gleichung (4-98) ergibt sich schließlich folgende Einschränkung für die Parameter:

$$4ca > e\left(1 - 2\frac{L}{Q}\right) \quad (4-102)$$

Da alle Parameter größer Null gewählt werden, ist diese Ungleichung immer erfüllt, wenn mehr als die Hälfte der Variablen im betreffenden Umfeld U die stationäre Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$

mit mindestens gleich hohem Wert wie die betrachtete Variable annehmen. Für den Extremfall, daß alle Variablen im betreffenden Umfeld die Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d})$ annehmen, reduziert sich die Beziehung (4-102) zu:

$$4ca > e \quad (4-103)$$

Durch eine Parameterkombination, die diese Ungleichung erfüllt, würden Disparitäten zugelassen, die isoliert an einer Bildkoordinate auftreten. Aufgrund der geforderten Kontinuität der Disparitätskarten (siehe Kapitel 2.4.3) ist dieser Fall jedoch wenig sinnvoll.

4.6.5 Das Prinzip des Wettbewerbs

Die Korrespondenzpaare in den Stereobildern werden durch die stationären Lösungen $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ des Selbstorganisationsprozesses repräsentiert. Im Rahmen dieses Prozesses muß gewährleistet werden, daß nur Variablen diese Lösung annehmen, die sich entweder durch eine hohe Ähnlichkeit der korrespondierenden Filterantworten oder durch ein konformes Umfeld oder beides zugleich auszeichnen.

Die Variable $\xi(\mathbf{x}_1, \mathbf{d}_1, t)$ mit der Disparität \mathbf{d}_1 an der Bildkoordinate \mathbf{x}_1 sei die Variable mit dem größten Wert aller am gleichen Wettbewerbsprozeß beteiligten Variablen zum Zeitpunkt t .

$$\xi(\mathbf{x}_1, \mathbf{d}_1, t) > \xi(\hat{\mathbf{x}}, \hat{\mathbf{d}}, t) \quad \forall \hat{\mathbf{d}} \neq \mathbf{d}_1 \text{ und } \hat{\mathbf{x}} \in \left\{ \mathbf{x}_1, \mathbf{x}_1 + \mathbf{d}_1 - \hat{\mathbf{d}}, \mathbf{x}_1 + \frac{1}{2}(\mathbf{d}_1 - \hat{\mathbf{d}}) \right\} \quad (4-104)$$

Weiterhin soll diese Variable die größte Verstärkung durch ihr Umfeld U_1 erfahren.

$$u_1 = \sum_{\mathbf{x}' \in U_1} \xi(\mathbf{x}', \mathbf{d}_1, t) > \sum_{\mathbf{x}' \in \hat{U}} \xi(\mathbf{x}', \hat{\mathbf{d}}, t) = \hat{u} \quad (4-105)$$

Für diesen Fall soll die Variable $\xi(\mathbf{x}_1, \mathbf{d}_1, t)$ das größte Wachstum mit $\dot{\xi}(\mathbf{x}_1, \mathbf{d}_1, t) > \dot{\xi}(\hat{\mathbf{x}}, \hat{\mathbf{d}}, t)$ aufweisen. Aus Beziehung (4-83) folgt somit:

$$\begin{aligned} \xi(\mathbf{x}_1, \mathbf{d}_1, t) \left(a - \frac{b}{3P}s + \frac{b}{3P}\xi^2(\mathbf{x}_1, \mathbf{d}_1, t) - c\xi^2(\mathbf{x}_1, \mathbf{d}_1, t) + \frac{e}{Q}u_1 \right) > \\ \xi(\hat{\mathbf{x}}, \hat{\mathbf{d}}, t) \left(a - \frac{b}{3P}s + \frac{b}{3P}\xi^2(\hat{\mathbf{x}}, \hat{\mathbf{d}}, t) - c\xi^2(\hat{\mathbf{x}}, \hat{\mathbf{d}}, t) + \frac{e}{Q}\hat{u} \right) \end{aligned} \quad (4-106)$$

Hierbei ist

$$s = \sum_{\mathbf{d}} \xi^2(\mathbf{x}, \mathbf{d}', t) + \xi^2(\mathbf{x} + \mathbf{d} - \mathbf{d}', \mathbf{d}', t) + \xi^2\left(\mathbf{x} + \frac{1}{2}(\mathbf{d} - \mathbf{d}'), \mathbf{d}', t\right). \quad (4-107)$$

Unter der Voraussetzung (4-105) ergibt sich somit:

$$\left(\frac{b}{3P} - c\right) \xi^2(\mathbf{x}_1, \mathbf{d}_1, t) > \left(\frac{b}{3P} - c\right) \xi^2(\hat{\mathbf{x}}, \hat{\mathbf{d}}, t) \quad (4-108)$$

Aus Beziehung (4-104) folgt somit unmittelbar die Forderung:

$$\frac{b}{3P} > c \quad (4-109)$$

Unter den oben getroffenen Annahmen ist die Konsequenz dieser Forderung, daß die betroffene Variable schneller wächst bzw. langsamer abklingt als alle anderen Variablen, die im gleichen Wettbewerbsprozeß involviert sind. Bleibt die Voraussetzung (4-105) während der Selbstorganisation erhalten, so weist diese Variable immer den größten Wert auf und wird daher die stationäre Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ erreichen, während alle anderen Variablen die stationäre Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d})$ annehmen. Soll die Stabilität der stationären Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d})$ für den im vorangegangenen Kapitel behandelten ungünstigsten Fall gewährleistet sein (siehe Beziehung (4-93)), so ist Bedingung (4-109) automatisch erfüllt.

4.6.6 Numerische Realisierungsaspekte

Um Disparitätskarten durch den vorgestellten Selbstorganisationsprozeß mit Hilfe eines Digitalrechners ermitteln zu können, muß das zeitliche Verhalten der einzelnen Variablen durch eine numerische Integration des Differentialgleichungssystems (4-83) approximiert werden. Als eine aufwandsgünstige Realisierungsmöglichkeit soll an dieser Stelle das sogenannte Euler-Verfahren, das auch bei der Implementierung des Selbstorganisationsprozesses eingesetzt wurde, diskutiert werden. Die zeitliche Ableitung wird bei diesem Verfahren durch einen Differenzenquotient approximiert:

$$\dot{\xi}(\mathbf{x}, \mathbf{d}, t) \approx \frac{\xi(\mathbf{x}, \mathbf{d}, t_{i+1}) - \xi(\mathbf{x}, \mathbf{d}, t_i)}{\Delta t} \quad (4-110)$$

Hierbei wird $\Delta t = t_{i+1} - t_i$ als Schrittweitenparameter bezeichnet. Der Wert einer Variablen zu einem diskreten Zeitpunkt t_i läßt sich somit auf folgende Weise bestimmen:

$$\xi(\mathbf{x}, \mathbf{d}, t_{i+1}) = \xi(\mathbf{x}, \mathbf{d}, t_i) + \Delta t f(\mathbf{x}, \mathbf{d}, t_i) \quad (4-111)$$

Das Euler-Verfahren hat den Nachteil, daß der Approximationsfehler proportional zum Schrittweitenparameter Δt ist und deshalb dieser im allgemeinen relativ klein gewählt werden muß. Mit anderen numerischen Integrationsverfahren lassen sich die entsprechenden zeitkontinuierlichen Differentialgleichungen unter Umständen wesentlich genauer approximieren (siehe z. B. [Wer92]). Die Genauigkeit der Integration ist jedoch bei diesem Ansatz nicht das entscheidende Argument. Die kritischen Faktoren, die in diesem Zusammenhang für den Einsatz des Euler-Verfahrens sprechen, sind der Speicher- und Berechnungsaufwand. Das Euler-Verfahren ist ein Einschrittverfahren und benötigt im Gegensatz zu Mehrschrittverfahren daher zu Beginn der Simulation keine Anlaufberechnung. Weiterhin muß pro Iterationsschritt die Funktion $f(\mathbf{x}, \mathbf{d}, t_i)$ nur einmal berechnet und gespeichert werden, was bei den hier auftretenden Datenmengen erhebliche Auswirkungen auf die Laufzeit und den Speicherbedarf hat.

Aufgrund der rekursiven Berechnung kann es bei der numerischen Integration zur Instabilität kommen, obwohl die zugrundeliegende Differentialgleichung ein stabiles Verhalten zeigt. Bei dem Euler-Verfahren hängt der Stabilitätsbereich stark von der Wahl des Schrittweitenparameters Δt ab. Wie dieser Parameter nun zu wählen ist, ohne daß die Stabilität der stationären Lösungen gefährdet wird, soll im weiteren näher untersucht werden. Bei der Behandlung dieser Problematik wird im folgenden vorausgesetzt, daß die Parameter des Differentialgleichungssystems die Bedingungen aus Kapitel 4.6.4 und Kapitel 4.6.5 erfüllen.

Wird Beziehung (4-111) um die stationäre Lösung linearisiert, so ergibt sich mit Beziehung (4-87) für das Gesamtsystem:

$$\begin{aligned} \Delta \xi(t_{i+1}) &= \Delta \xi(t_i) + \Delta t \mathbf{J} \Delta \xi(t_i) \\ &= (\mathbf{I} + \Delta t \mathbf{J}) \Delta \xi(t_i) \end{aligned} \quad (4-112)$$

Die Eigenwerte λ_j' der Matrix $(\mathbf{I} + \Delta t \mathbf{J})$ errechnen sich zu:

$$\lambda_j' = 1 + \Delta t \lambda_j \quad \forall j, \quad (4-113)$$

wobei λ_j die entsprechenden Eigenwerte der Matrix \mathbf{J} sind. Für die Konvergenz des Euler-Verfahrens muß die Bedingung

$$|1 + \Delta t \lambda_j| < 1 \quad \forall j \quad (4-114)$$

erfüllt sein. Da der Hauptanteil der Berechnungszeit bei diesem Ansatz zur Korrespondenzbestimmung auf den Selbstorganisationsprozeß entfällt, soll die numerischen Integration möglichst zeitoptimal durchgeführt werden. Aus diesem Grund wird zur numerischen Integration ein maximaler Schrittweitenparameter gewählt, bei dem die Stabilität der stationären Lösungen noch gesichert ist. Für rein reelle negative Eigenwerte λ_j kann die maximal zulässige Schrittweite Δt_{max} durch den maximalen Eigenwertsbetrag $|\lambda|_{max}$ abgeschätzt werden.

$$\Delta t_{max} < \frac{2}{|\lambda|_{max}} \quad (4-115)$$

Die Eigenwerte der stationären Lösungen $\xi_{s_1}(\mathbf{x}, \mathbf{d}) = 0$ ergeben sich aus den Elementen auf der Hauptdiagonalen des entsprechenden Unterblocks der Matrix \mathbf{J} und sind somit reell (vgl. Beziehung (4-90)). Es läßt sich hierbei zeigen, daß unter den Parametervoraussetzungen aus Kapitel 4.6.4 und Kapitel 4.6.5 der maximale Eigenwertsbetrag dann auftritt, wenn alle Variablen im Umfeld von $\xi_{s_1}(\mathbf{x}, \mathbf{d})$ ebenfalls die stationäre Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d}) = 0$ angenommen haben und eine Variable im Wettbewerbsprozeß den maximalen Wert $\xi_{s_{2max}}(\mathbf{x}, \mathbf{d})$ aufweist. Diese Lösung wird wiederum dann angenommen, wenn alle Variablen im betreffenden Umfeld ebenfalls die maximale Lösung $\xi_{s_{2max}}(\mathbf{x}, \mathbf{d})$ annehmen und alle anderen, am gleichen Wettbewerbsprozeß beteiligten Variablen den Wert $\xi_{s_1}(\mathbf{x}, \mathbf{d}) = 0$ aufweisen. Nach den Beziehungen (4-84) und (4-85) errechnet sich diese stationäre Lösung zu:

$$\xi_{s_{2max}}(\mathbf{x}, \mathbf{d}) = \frac{e + \sqrt{e^2 + 4ac}}{2c} \quad (4-116)$$

Unter der Voraussetzung (4-92) ergibt sich somit aus Beziehung (4-115) für die maximal zulässige Schrittweite:

$$\begin{aligned} \Delta t_{max} &< \frac{2}{\frac{b}{3P} \xi_{s_{2max}}^2(\mathbf{x}, \mathbf{d}) - a} \\ \Leftrightarrow \Delta t_{max} &< \frac{12Pc^2}{be^2 + be\sqrt{e^2 + 4ac} + 2abc - 2Pac^2} \end{aligned} \quad (4-117)$$

Für eine stabile stationäre Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ zeigt Beziehung (4-95) die Form des entsprechenden Unterblocks der Jacobi-Matrix \mathbf{J} . Da in diesem Fall die Beziehung (4-98) erfüllt sein muß,

ergibt sich nach den Satz von Gerschgorin (vgl. Kapitel 4.6.4) die folgende Bedingung für den Eigenwert λ_j :

$$|\lambda_j + 2\eta(\mathbf{x}, \mathbf{d})| < e\xi_{s_2}(\mathbf{x}, \mathbf{d}) \quad (4-118)$$

Der maximale Betrag dieser Eigenwerte kann mit $\eta(\mathbf{x}, \mathbf{d}) = c\xi_{s_{2max}}^2(\mathbf{x}, \mathbf{d})$ und $|x + y| \geq ||x| - |y||$ durch

$$|\lambda|_{max} < e\xi_{s_{2max}}(\mathbf{x}, \mathbf{d}) + 2c\xi_{s_{2max}}^2(\mathbf{x}, \mathbf{d}) \quad (4-119)$$

abgeschätzt werden. Im Fall reeller Eigenwerte ergibt sich durch die Beziehungen (4-115) und (4-116) für die maximale Schrittweite:

$$\begin{aligned} \Delta t_{max} &< \frac{2}{e\xi_{s_{2max}}(\mathbf{x}, \mathbf{d}) + 2c\xi_{s_{2max}}^2(\mathbf{x}, \mathbf{d})} \\ \Leftrightarrow \Delta t_{max} &< \frac{4c}{3e^2 + 3e\sqrt{e^2 + 4ac} + 4ac} \end{aligned} \quad (4-120)$$

Da jedoch die jeweiligen Unterblöcke der Jacobi-Matrix \mathbf{J} für diese stationären Lösungen nicht immer exakt symmetrisch sind (vgl. Beziehung (4-95)), können hier komplexe Eigenwerte nicht generell ausgeschlossen werden. In der Praxis zeigte sich jedoch in allen untersuchten Fällen ein schwingungsfreies Eingleiten in die Ruhelagen $\xi_{s_2}(\mathbf{x}, \mathbf{d})$. Zur numerischen Integration werden daher Schrittweiten gewählt, die die Forderungen (4-117) und (4-120) erfüllen.

4.6.7 Parameterwahl

Das Verhalten des Selbstorganisationsprozesses wird durch die Parameter a , b , c , e und die Anzahl P bzw. Q der Variablen, die dämpfend bzw. verstärkend auf jede Variable wirken, bestimmt. Der Wertebereich der Parameter wird durch die Beziehungen (4-94), (4-102) und (4-109), welche die Stabilität der stationären Lösungen sichern und einen sinnvollen Wettbewerb zwischen den Variablen ermöglichen, eingeschränkt. Die einzelnen Forderungen, die in diesem Zusammenhang an die Parameter gestellt werden, können mit unendlich vielen Parameterkombinationen erfüllt werden. Im folgenden werden daher exemplarisch einige Parameterkombinationen, mit denen der Selbstorganisationsprozeß ein unterschiedliches Verhalten zeigt, ermittelt. Alle im weiteren gezeigten Ergebnisse wurden, soweit nicht anders erwähnt, mit diesen Parameterkombinationen berechnet.

Um die frei wählbare Zeiteinheit festzulegen, wird zunächst der Parameter b zu $b = 1,0$ festgelegt. Die übrigen Parameter werden in diesem Beispiel für einen Disparitätsbereich mit $d_{max} = -d_{min} = 15$ bestimmt. Die Anzahl der dämpfend wirkenden Variablen ergibt sich hieraus zu $P = 2d_{max} - 1 = 30$. Für den Parameter c gilt nach Beziehung (4-109) somit: $c < 0,011$. Im weiteren wird deshalb der Parameter $c = 0,01$ gewählt. Über den Parameter e wird die verstärkende Wirkung des Umfeldes und damit die Auswirkung der Kontinuitätseinschränkung (siehe Kapitel 2.4.3) beeinflußt. Weiterhin wirkt sich dieser Parameter über die Beziehung (4-102) auch auf die Stabilität der stationären Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ aus. Die Forderung (4-103), welche die Stabilität dieser Lösung bei fehlender Umfeldverstärkung sichert, ist in Bezug auf die Kontinuitätseinschränkung wenig sinnvoll. Im weiteren wird deshalb für eine stabile stationäre Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ gefordert, daß mindestens die Hälfte der Variablen im Umfeld ebenfalls diese Lösung angenommen haben. Der maximale Wert der stationären Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ wird im wesentlichen durch die nun relativ frei wählbaren Parameter a und e bestimmt (siehe Beziehung (4-116)).

Eine weitere Einschränkung an die Parameter a und e ergibt sich aus dem sehr unwahrscheinlichen Fall, daß alle Variablenwerte innerhalb eines Wettbewerbsprozesses und alle Variablen in den beteiligten Umfeldern den gleichen Wert aufweisen. Da bei gleichmäßig verteilten Variablenwerten ein Wettbewerb nicht sonderlich sinnvoll ist, wird für diesen speziellen Fall gefordert, daß keine Änderung der Variablenwerte stattfindet. Aus Beziehung (4-83) ergibt sich somit die Forderung:

$$0 = a - b\xi^2(\mathbf{x}, \mathbf{d}, t) - c\xi^2(\mathbf{x}, \mathbf{d}, t) + e\xi(\mathbf{x}, \mathbf{d}, t) \quad (4-121)$$

Da diese Forderung nur für einen bestimmten Wert von $\xi(\mathbf{x}, \mathbf{d}, t)$ erfüllt werden kann, soll sie für den Fall einer maximalen Initialisierung aller beteiligten Variablenwerte gelten. Mit $\xi(\mathbf{x}, \mathbf{d}, t = 0) = 1$ folgt somit:

$$a + e = b + c \quad (4-122)$$

Über das Verhältnis von a zu e kann der Einfluß des jeweiligen Umfeldes variiert werden. Eine starke Umfeldkopplung bewirkt eine erhöhte Gewichtung der Kontinuitätseinschränkung im Selbstorganisationsprozeß. Sie führt jedoch auch zu einer Erhöhung der stationären Variablenwerte $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ (siehe Beziehung (4-116)). Dies bewirkt wiederum, daß die maximale Schrittweite Δt_{max} zur Erhaltung der numerischen Stabilität (siehe Beziehung (4-117) und (4-120)) kleiner gewählt werden muß, was gleichbedeutend mit einer Vergrößerung der Simulationsdauer ist. In der nachfolgenden Tabelle sind exemplarisch einige Parameterkombinationen, die alle Forderungen erfüllen und die im weiteren verwendet werden, aufgeführt.

Satz	a	b	c	e	P	$\xi_{s_{2max}}$	Δt_{max}
1	1,0	1,0	0,01	0,01	30	10,51	0,86
2	0,7	1,0	0,01	0,31	30	33,11	0,062
3	0,4	1,0	0,01	0,61	30	61,64	0,017
4	0,1	1,0	0,01	0,91	30	91,10	0,0072

Tabelle 4-2: Auswahl einiger Parameterkonstellationen

4.7 Subpixelgenaue Disparitätsschätzung

Die Filterung der Bilddaten in Kapitel 4.4, die Bestimmung des Ähnlichkeitsmaßes in Kapitel 4.5 und der gesamte Selbstorganisationsprozeß in Kapitel 4.6 beziehen sich auf die diskreten Bildkoordinaten \mathbf{x} . Wie schon erwähnt, besitzen diese Koordinaten und somit auch die Disparität die normierte Einheit Pixel. In welchem Zusammenhang diese Einheit mit der Kamerageometrie steht, hängt von den einzelnen Modellparametern und dem jeweiligen Abtastfaktor (siehe Kapitel 4.4.4) ab. Die Diskretisierung der Bildkoordinaten, die durch den Bildsensor hervorgerufen wird, führt dazu, daß die stereoskopische Positionsbestimmung von Quantisierungsfehlern überlagert wird. Je nach gewähltem Abtastfaktor und Kamerageometrie kann diese Diskretisierung zu erheblichen Meßfehlern führen. Um diese Art von Meßfehlern zu reduzieren, wird die pixelgenaue Disparitätsmessung, die aus dem Selbstorganisationsprozeß hervorgeht, nachträglich durch ein einfaches Verfahren verbessert.

Wie bereits in Kapitel 4.5.2 erläutert, wird der Realteil des Ähnlichkeitsmaßes in Beziehung (4-65) maximal, wenn die Filterantworten der beiden Stereokanäle lokal übereinstimmen. Der Imaginärteil des Ähnlichkeitsmaßes weist an dieser Stelle einen Nulldurchgang auf.

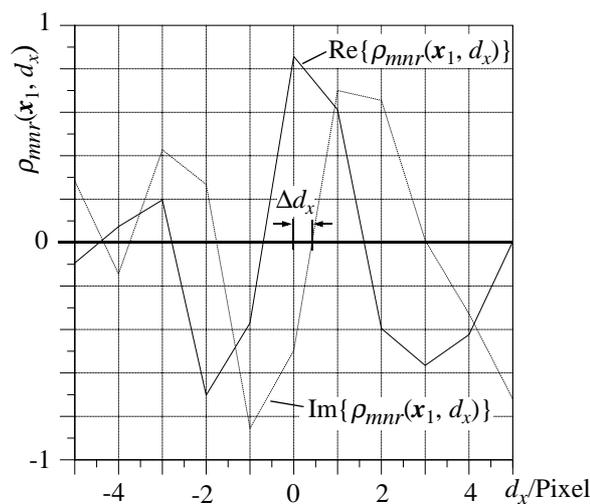


Bild 4-18: Verlauf des Real- und des Imaginärteils des Ähnlichkeitsmaßes in Abhängigkeit der Disparität an einer bestimmten Bildkoordinate \mathbf{x}_1

Um trotz des diskreten Abtastrasters subpixelgenaue Disparitätswerte zu erhalten, kann der Imaginärteil des Ähnlichkeitsmaßes auf die folgende Weise ausgenutzt werden: Der Realteil des Ähnlichkeitsmaßes dient weiterhin zur Initialisierung des in Kapitel 4.6 vorgestellten Selbstorganisationsprozesses. Die Gewinner des jeweiligen Wettbewerbs an den einzelnen Bildkoordinaten korrespondieren im allgemeinen mit lokalen Maxima im Realteil des Ähnlichkeitsmaßes, da diese bevorzugt den Wettbewerb gewinnen können (siehe Kapitel 4.6.5). Nach der Selbstorganisation werden daher die zu den Gewinnervariablen korrespondierenden Nulldurchgänge im Imaginärteil des Ähnlichkeitsmaßes durch eine lineare Interpolation ermittelt und die pixelgenaue Disparität des Gewinners um den entsprechenden Differenzwert $\Delta d_x(\mathbf{x})$ korrigiert.

$$d_{x_s}(\mathbf{x}) = d_x(\mathbf{x}) + \Delta d_x(\mathbf{x}) \quad (4-123)$$

Wie in Bild 4-18 exemplarisch zu sehen ist, tritt im Realteil des Ähnlichkeitsmaßes ein lokales Maximum bei $d_x(\mathbf{x}_1) = 0$ auf. Die eigentliche Disparität wies bei diesem Beispiel jedoch den Wert $d_x(\mathbf{x}_1) = 0,4$ auf. Wird nun der Nulldurchgang im Imaginärteil des Ähnlichkeitsmaßes in der Nähe des Maximums im Realteil nach dem oben erwähnten Verfahren zur Bestimmung der Disparität herangezogen, so ergibt sich in diesem Fall die Disparität zu $d_{x_s}(\mathbf{x}_1) = 0,415$. Wie in Kapitel 5 zu sehen ist, können die Quantisierungsfehler in der Disparitätskarte und damit auch in der Positionsbestimmung durch diesen einfachen und effizient implementierbaren Ansatz erheblich reduziert werden.

Im Gegensatz zur Initialisierung des Selbstorganisationsprozesses in Kapitel 4.6.1 werden die Imaginärteile der unterschiedlich orientierten Filterkanäle zur subpixelgenauen Disparitätsschätzung nicht überlagert. Es zeigte sich in mehreren Messungen, daß eine separate Auswertung der unterschiedlich orientierten Filterkanäle zu wesentlich geringeren Fehlern führt [Tra97]. Zur subpixelgenauen Disparitätsschätzung wird daher jeweils der Imaginärteil des Filterkanals verwendet, der an der betreffenden Bildkoordinate den größten Betrag in der Antwort aufweist.

4.8 Rekonstruktion der Weltkoordinaten

Nachdem eine gegenseitige Zuordnung der Bildpunkte in den beiden Stereokanälen erfolgt ist, können die Weltkoordinaten relativ leicht durch eine Triangulation rekonstruiert werden (vgl. Bild 2-1). Die Weltkoordinaten werden hier in Bezug auf das bereits in Kapitel 4.3.2 eingeführte Koordinatensystem ermittelt, dessen Ursprung sich in der Mitte der beiden Vergenzachsen auf der Höhe der Projektionszentren der Kamerasysteme befindet. Im idealen Fall liegt der Ursprung des Weltkoordinatensystems somit in der Mitte der Basisstrecke und die x_w -Achse hat die gleiche Richtung wie die Basisstrecke. Die y_w -Achse ist in diesem Fall kollinear mit den Vergenzachsen des Kamerasystems (vgl. Bild 4-9).

4.8.1 Räumliche Rekonstruktion unter idealen Voraussetzungen

Bedingt durch die Kamerageometrie und die Verfahren, die den Abbildungsprozeß beeinflussen, gestaltet sich die Rekonstruktion der Weltkoordinaten hier etwas komplizierter als für das einfache Beispiel aus Kapitel 2.1 (vgl. Gleichung (2-2)). Vielmehr müssen in diesem Zusammenhang alle Transformationen der einzelnen Koordinatensysteme, die zur Vereinfachung der Korrespondenzbestimmung dienen, berücksichtigt werden. Zu diesen Operationen zählen die Entzerrung in Kapitel 4.3.1, die Rektifikation in Kapitel 4.3.2 und die Unterabtastung der Filterantworten in Kapitel 4.4.4.

Zur Ermittlung der Weltkoordinaten lassen sich unter Ausnutzung der Gleichungen (4-3), (4-24), (4-25) und (4-29) zunächst die folgenden Beziehungen zwischen den x -Koordinaten des jeweiligen Stereokanals und den Weltkoordinaten angeben:

$$\begin{aligned}\tilde{x}_l - \tilde{x}_{0l} &= -f \frac{x_w}{z_w} + f \tan \beta_l - f \frac{b_s}{2z_w} \\ \tilde{x}_r - \tilde{x}_{0r} &= -f \frac{x_w}{z_w} + f \tan \beta_r + f \frac{b_s}{2z_w}\end{aligned}\tag{4-124}$$

Bezieht sich das Ähnlichkeitsmaß und damit die Disparität nach Beziehung (4-60) auf das rechte Bildkoordinatensystem, so sind korrespondierende Bildkoordinaten im linken und rechten Stereokanal durch die Disparität d_{xs} auf die folgende Weise miteinander verknüpft:

$$d_{xs}(\mathbf{x}_r) = x_l - x_r\tag{4-125}$$

Ist der Abtastfaktor s_n identisch für die beiden Stereokanäle, was im weiteren vorausgesetzt wird, so läßt sich die z_w -Koordinate eines auf die Bildkoordinate x_r abgebildeten Punktes mit den Beziehungen (4-59), (4-124) und (4-125) zu

$$z_w = \frac{-f_n b_s}{s_n d_{xs}(\mathbf{x}_r) + \tilde{x}_{0r} - \tilde{x}_{0l} + f_n (\tan \beta_r - \tan \beta_l)}\tag{4-126}$$

bestimmen. Wird die z_w -Koordinate in Beziehung (4-124) durch Gleichung (4-126) substituiert, so ergibt sich mit Beziehung (4-125) die x_w -Koordinate zu:

$$x_w = \frac{b_s(s_n x_r - \tilde{x}_{0r} - f_n \tan \beta_r)}{s_n d_{xs}(\mathbf{x}_r) + \tilde{x}_{0r} - \tilde{x}_{0l} + f_n(\tan \beta_r - \tan \beta_l)} + \frac{b_s}{2} \quad (4-127)$$

Aus den Beziehungen (4-3), (4-24), (4-25), (4-29) und (4-59) folgt für die y_w -Koordinate:

$$y_w = \frac{b_s(s_n y_r - \tilde{y}_{0r})}{s_n d_{ys}(\mathbf{x}_r) + \tilde{x}_{0r} - \tilde{x}_{0l} + f_n(\tan \beta_r - \tan \beta_l)} \quad (4-128)$$

Für jede Bildkoordinate, an der eine Korrespondenz und somit eine Disparität bestimmt werden kann, läßt sich auf diese Weise die zugehörige Raumposition rekonstruieren und eine sogenannte Tiefenkarte der betrachteten Szene erstellen.

4.8.2 Kompensation der Projektionszentrumsdrift

Die im vorangegangenen Kapitel rekonstruierten Weltkoordinaten beziehen sich auf ein Koordinatensystem, dessen Ursprung nur dann in der Mitte der Basisstrecke liegt, wenn die Projektionszentren der Kameramodelle exakt im Schnittpunkt der Bewegungsachsen positioniert sind. Da dieser Zustand mit Hilfe der in Kapitel 4.1.2 vorgestellten Mechanismen nur für einen bestimmten Kamerazustand eingestellt werden kann, verändert sich die Position des Weltkoordinatenursprungs, wenn die Kamerazustände variiert werden. Wie die Kalibrierung der Kameras zeigte, kann diese Drift insbesondere bei Veränderung der Zoomeinstellungen mehrere Zentimeter in Richtung der optischen Achse der Kameras betragen (vgl. Bild 4-7b). Bei nichtparallelen optischen Achsen führt dies dazu, daß sich die Basisstrecke, die in die Positionsbestimmung mit eingeht, verändert. Ist die Drift im linken und im rechten Kamerasystem unterschiedlich, so werden die Weltkoordinaten zusätzlich durch eine Rotation verfälscht. Darüber hinaus bewirkt diese Drift, daß sich die meßtechnisch erfaßten Vergenzwinkel von den Winkeln der optischen Achsen unterscheiden können. Neben der Rekonstruktion der Weltkoordinaten ist auch die Rektifikation der Stereobilder durch diesen Effekt betroffen. Wird diese mit falschen Vergenzwinkeln berechnet, so sind die Voraussetzungen, unter denen die Epipolareinschränkung ausgenutzt werden kann, nicht erfüllt. Da hierdurch einerseits die Korrespondenzbestimmung scheitern kann und andererseits nur eine sehr ungenaue Positionsbestimmung möglich ist, wird dieser Effekt im folgenden durch eine entsprechende Transformation der Weltkoordinaten und eine Korrektur der Vergenzwinkel kompensiert.

Da die Lageänderungen der Projektionszentren in x_k - und y_k -Richtung in Abhängigkeit der optischen Freiheitsgrade vernachlässigbar gering sind, wird im weiteren nur die Drift in Richtung der optischen Achse, die identisch mit der z_k -Achse ist, berücksichtigt.

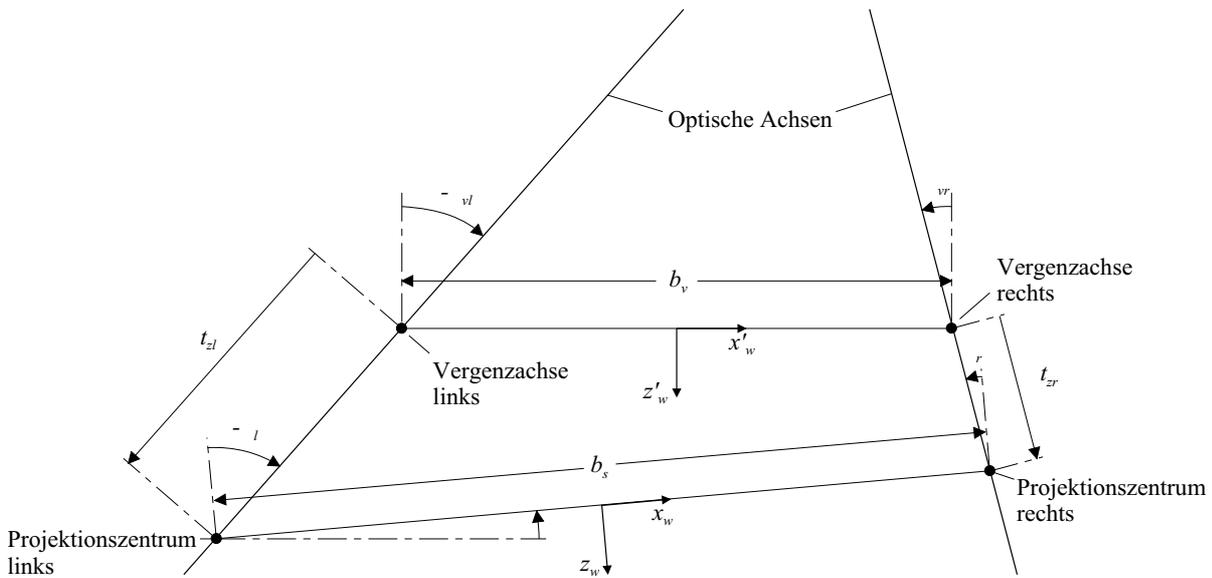


Bild 4-19: Einfluß der Projektionszentrumsdrift auf die Koordinatensysteme, die Basisstrecke und die Vergenzwinkel

Die Weltkoordinaten \mathbf{x}'_w , die sich auf ein Koordinatensystem beziehen, dessen Ursprung in der Mitte der Vergenzachsen angeordnet ist, lassen sich durch die folgende Transformation aus den in Kapitel 4.8.1 bestimmten Koordinaten \mathbf{x}_w errechnen:

$$\mathbf{x}'_w = \mathbf{R}' \mathbf{x}_w + \mathbf{t}' = \begin{bmatrix} \cos \vartheta & 0 & \sin \vartheta \\ 0 & 1 & 0 \\ -\sin \vartheta & 0 & \cos \vartheta \end{bmatrix} \mathbf{x}_w + \begin{bmatrix} t'_x \\ 0 \\ t'_z \end{bmatrix} \quad (4-129)$$

Seien Δt_{zr} und Δt_{zl} die durch die Kalibrierung ermittelten Abstände der Projektionszentren des rechten und des linken Kameramodells von der jeweiligen Vergenzachse, so läßt sich der Winkel ϑ durch

$$\vartheta = \arctan\left(\frac{\Delta t_{zl} \cos \beta_{vl} - \Delta t_{zr} \cos \beta_{vr}}{\Delta t_{zr} \sin \beta_{vr} - \Delta t_{zl} \sin \beta_{vl} + b_v}\right) \quad (4-130)$$

bestimmen. Hierbei ist b_v der Abstand der Vergenzachsen. Die korrekte Stereobasis b_s läßt sich aus der Zentrumsdrift, den Kamerawinkeln und dem Abstand der Vergenzachsen zu

$$b_s = \frac{\Delta t_{zr} \sin \beta_{vr} - \Delta t_{zl} \sin \beta_{vl} + b_v}{\cos \vartheta} \quad (4-131)$$

errechnen. Die verbleibenden Komponenten des Translationsvektors ergeben sich zu

$$t'_x = \frac{\Delta t_{zr} \sin \beta_{vr} + \Delta t_{zl} \sin \beta_{vl}}{2} \quad (4-132)$$

und

$$t'_z = \frac{\Delta t_{zr} \cos \beta_{vr} + \Delta t_{zl} \cos \beta_{vl}}{2}. \quad (4-133)$$

Die optischen Achsen der Kameras schließen mit den meßbaren Vergenzwinkeln β_v den Winkel ϑ ein. Zur korrekten Rektifikation über die Beziehungen (4-30) und (4-31) und zur korrekten Positionsrekonstruktion durch die Beziehungen (4-126) bis (4-128) müssen daher die Winkel

$$\beta = \beta_v + \vartheta \quad (4-134)$$

anstelle der gemessenen Vergenzwinkeln β_v verwendet werden.

4.9 Abschließende Betrachtungen

In den vorangegangenen Abschnitten dieses Kapitels wurden neben dem eigentlichen Problem der Korrespondenzbestimmung einige Komponenten beschrieben, ohne die eine praktische Umsetzung des Verfahrens nicht möglich wäre. Diese Aspekte sind zwar von grundlegender Bedeutung, wenn es darum geht über die Stereoskopie quantitative Positionsmessungen vorzunehmen, für das Problem der Korrespondenzbestimmung haben diese Komponenten jedoch nur einen indirekten Einfluß. In diesem Abschnitt soll daher nun unabhängig von den Systemvoraussetzungen und den geometrischen Aspekten der passiven Triangulation die prinzipiellen Grenzen des Verfahrens zur Korrespondenzbestimmung aufgezeigt werden und eine Abgrenzung zu existierenden Techniken erfolgen.

4.9.1 Verfahrensgrenzen

Um die prinzipiellen Grenzen des Verfahrens zur Korrespondenzbestimmung aufzuzeigen, werden im folgenden nur das Filterverfahren, die Bestimmung des Ähnlichkeitsmaßes und der Selbstorganisationsprozeß betrachtet. Da aufgrund des Blendenproblems in der Stereoskopie nur Strukturen bzw. Grauwertänderungen, die eine Komponente in Richtung der Epipolarlinien aufweisen, eine eindeutige Korrespondenzbestimmung erlauben, können grundsätzlich nur für die Bildkoordinaten die zugehörigen räumlichen Positionen rekonstruiert werden, an denen Grauwertänderungen dieser Art auftreten. Der Einsatz der orientierungsselektiven Bandpaßfilter in Kapitel 4.4 bewirkt in diesem Zusammenhang eine Selektion der geeigneten Grauwertänderungen. Diese Einschränkung ist von grundsätzlicher Natur und für alle binokularen Stereosysteme gleichbedeutend. Der Einsatz der Fensterfunktion bei der Berechnung des Ähnlichkeitsmaßes in Kapitel 4.5 begrenzt in gewisser Weise die Auflösung der Disparitätsbestimmung. Da aber diese Funktion gleich der Ausdehnung der unterabgetasteten Impulsantwort der verwendeten Filter in Epipolarrichtung gewählt wird, ist dieser Effekt relativ gering und wird im wesentlichen durch die Filter bzw. durch die Ebenen der Auflösungspyramide (vgl. Kapitel 4.4.4) beeinflusst. Während die Unterdrückung der Okklusionskorrespondenzen im Rahmen des Selbstorganisationsprozesses in Kapitel 4.6.2 allgemeingültig ist, gelten die Annahmen, die zur Unterdrückung der Okklusionspaare in Kapitel 4.6.3 getroffen werden, nicht uneingeschränkt. Es lassen sich in diesem Zusammenhang vielmehr Situationen konstruieren, bei denen die Relaxation nach Beziehung (4-83) unweigerlich zu falschen Korrespondenzen führt. Ein solcher Fall ist exemplarisch in der nachfolgenden Graphik dargestellt.

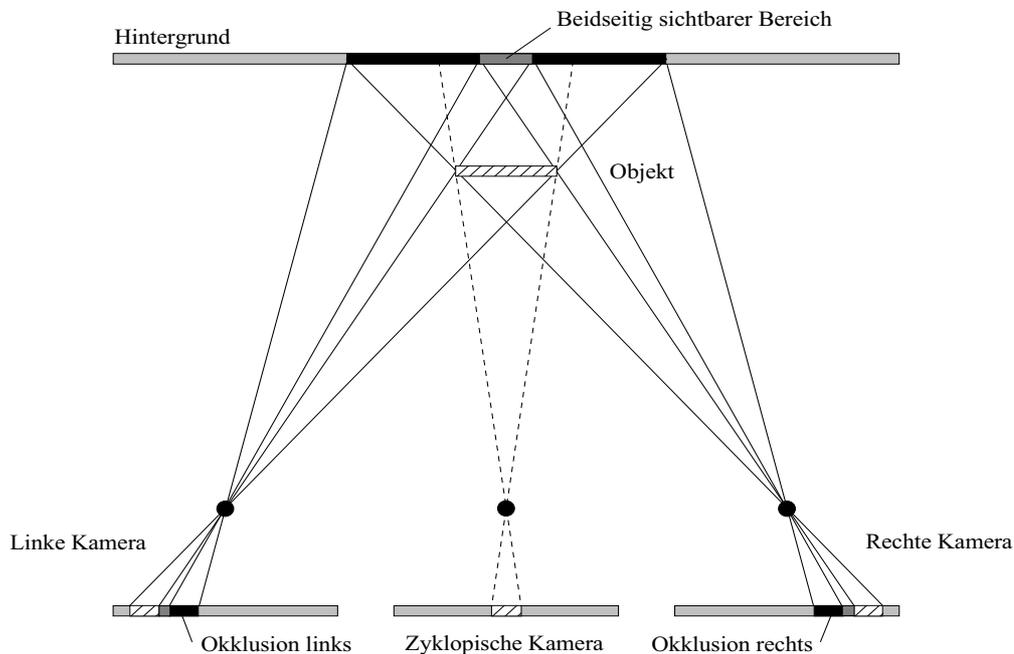


Bild 4-20: Konstellation, bei der ein Objekt und ein beidseitig sichtbarer Bereich des Hintergrundes identische zyklische Bildkoordinaten aufweisen. Eine korrekte Korrespondenzbestimmung ist aufgrund der Unterdrückung von Okklusionspaaren nicht möglich.

Ähnlich wie im Beispiel 4-17 erzeugt in Bild 4-20 ein Objekt Okklusionen im linken und im rechten Stereobild. In diesem Beispiel ist jedoch sowohl im linken als auch im rechten Bild ein Bereich vom Hintergrund zu sehen, der sich aus Sicht der zyklischen Kamera direkt hinter dem eigentlichen Objekt befindet. Da aber aufgrund von Beziehung (4-83) keine Korrespondenzen zugelassen werden, die in einem zyklischen Kamerabild die gleichen Bildkoordinaten aufweisen würden, können die räumlichen Positionen des Objektes und des betreffenden Hintergrundbereiches nicht gleichzeitig rekonstruiert werden. In diesem Fall konkurrieren im Rahmen des Selbstorganisationsprozesses jeweils zwei Variablen miteinander, die korrekte Korrespondenzen repräsentieren. Der Selbstorganisationsprozeß bewirkt hier, daß abhängig von den Initialisierungswerten eine der beiden Variablen den Wettbewerb gewinnt und somit einer der beiden konkurrierenden Bildbereiche korrekt zugeordnet wird, während für den anderen eine andere, falsche Korrespondenz errechnet wird. Der in Bild 4-20 gezeigte Fall stellt jedoch eine relativ extreme Situation dar, die in der Praxis nur selten auftritt. An dieser Stelle sei erwähnt, daß auch der Mensch aufgrund der nachgewiesenen Disparitätsgradientengrenze im visuellen System nicht in der Lage ist, diese Bereiche gleichzeitig zu fusionieren [BJ80].

Prinzipiell können bei den Unterdrückungsmechanismen zur impliziten Okklusionsbehandlung in Kapitel 4.6.2 und Kapitel 4.6.3 Probleme auftreten, wenn in Bildbereichen keine geeigneten Grauwertänderungen und damit Filterantworten vorhanden sind, durch die Okklusionszuordnungen verhindert werden könnten.

4.9.2 Einordnung des Verfahrens

Obwohl die Berechnung des Ähnlichkeitsmaßes in Kapitel 4.5 ähnlich wie bei den meisten gebietsbasierten Verfahren in Form einer lokalen Korrelation erfolgt, besteht doch ein entscheidender Unterschied zu diesen Ansätzen. Bevor es bei dem hier beschriebenen Verfahren zu einer Berechnung der Ähnlichkeit kommt, werden die Grauwertbilder zunächst orientierungs- und frequenzselektiv gefiltert, um Grauwertänderungen in Richtung der Epipolarlinien aus den Eingangsdaten zu extrahieren (siehe Kapitel 4.4). Da eine anschließende Berechnung des Ähnlichkeitsmaßes nur in Bildregionen erfolgt, in denen Grauwertänderungen mit den gewünschten Eigenschaften auftreten, kann dieses Filterverfahren entfernt als ein Interest-Operator angesehen werden (vgl. Kapitel 2.5.2). Da im Zuge der Korrespondenzbestimmung keine Merkmale aus den Grauwertbildern explizit extrahiert werden, besteht keine unmittelbare Verwandtschaft dieses Ansatzes zur Kategorie der merkmalsbasierten Verfahren. Aufgrund der zur Filterung eingesetzten Gaborfilter ist eine gewisse Übereinstimmung des hier vorgestellten Verfahrens mit den phasenbasierten Ansätzen gegeben (vgl. Kapitel 2.5.3). So können beispielsweise sehr dicht besetzte Disparitätskarten mit Subpixelgenauigkeit errechnet werden. Der wesentliche Unterschied zu dieser Kategorie von Stereoverfahren besteht jedoch darin, daß die Disparität nicht aus der Phasendifferenz der Filterantworten, sondern durch das in Kapitel 4.5 beschriebene komplexe Ähnlichkeitsmaß bestimmt wird. Der erfassbare Disparitätsbereich ist daher im Gegensatz zu den Phasendifferenzverfahren nicht durch die Modulationsfrequenz der verwendeten Filter begrenzt. Der Einsatz einer Auflösungsrampe, die

in diesem Ansatz zur Aufwandsreduktion herangezogen wird (siehe Kapitel 4.4.4), ist somit nicht zwingend erforderlich.

Der in Kapitel 4.6 beschriebene Selbstorganisationsprozeß stellt eine Erweiterung des Ansatzes von Reimann und Haken [RH94] dar und ist in die Kategorie der Relaxationsverfahren einzuordnen (vgl. Kapitel 2.5.6). Bedeutende Unterschiede zu den meisten Verfahren dieser Kategorie ergeben sich bei diesem Ansatz durch die implizite Detektion von Okklusionen und durch die Möglichkeit zu einer subpixelgenauen Disparitätsmessung. Die bislang einzige Möglichkeit, Okklusionen im Bereich der Stereoskopie zu behandeln, stellten die bidirektionalen Ansätze dar, bei denen eine Zuordnung von Bildelementen des linken Stereokanals zu Bildelementen des rechten Stereokanals erfolgt und eine weitere Zuordnung in umgekehrter Richtung [GLY92, HA89, JM92, LB95, WAH88]. Okklusionen in den Stereobildern werden hierbei durch die Differenzen, die sich zwischen den beiden Zuordnungsprozessen ergeben, explizit detektiert. Da die Korrespondenzbestimmung bei diesen Verfahren in zwei identischen, aber weitestgehend unabhängigen iterativen Prozessen erfolgt, ist der Berechnungs- und Speicher- aufwand, der im Bereich der Stereobildverarbeitung ohnehin beachtlich ist, enorm. Durch die in Kapitel 4.6.2 und Kapitel 4.6.3 beschriebenen Mechanismen erfolgt in dem hier vorgestellten Verfahren eine implizite Detektion von Okklusionen, ohne daß ein aufwendiger bidirektionaler Zuordnungsprozeß durchgeführt werden muß. Darüber hinaus können Okklusionen aufgrund des verstärkend wirkenden Umfeldes im Rahmen des Selbstorganisationsprozesses in einem begrenzten Umfang detektiert werden, selbst wenn die Okklusionen lokal eine höhere Ähnlichkeit zu anderen Bildregionen aufweisen als diese zu ihren korrekten Korrespondenzpartnern im umgekehrten Zuordnungsprozeß.

Die Unterdrückung von Korrespondenzen zwischen Okklusionen in Kapitel 4.6.3 führt dazu, daß nicht in allen Fällen Korrespondenzen errechnet werden können. Insbesondere können Objekte, die sich aus Sicht einer zyklischen Kamera direkt hintereinander befinden und in beiden Stereokanälen zu sehen sind, nicht gleichzeitig zugeordnet werden (siehe Kapitel 4.9.1). Hiermit ergibt sich eine gewisse Analogie zu der Disparitätsgradientengrenze (siehe Kapitel 2.4.5), die von einigen Verfahren eingesetzt wird, um das Korrespondenzproblem zu lösen [Fau93, LH96, PMF85, PP+85]. Während bei diesen Verfahren Korrespondenzen nur dann zugelassen werden, wenn der Disparitätsgradient einen endlichen, meist sehr kleinen Wert nicht überschreitet, ist bei dem hier vorgestellten Verfahren nur dann eine Zuordnung nicht möglich, wenn sich daraus ein unendlicher Disparitätsgradient nach der Definition (2-5) ergibt.

5 Experimentelle Ergebnisse

Während in den einzelnen Abschnitten des vorangegangenen Kapitels bereits exemplarisch einige Ergebnisse der Vorverarbeitungsstrategien und Bildaufbereitungsverfahren anhand von Filterantworten und entzerrten bzw. rektifizierten Bildern gezeigt wurden, soll in diesem Kapitel die Leistungsfähigkeit des Verfahrens zur Korrespondenzbestimmung an einigen ausgewählten Beispielen demonstriert werden. Weiterhin wird die Genauigkeit, mit der die Weltkoordinaten eines Raumpunktes durch das Gesamtsystem rekonstruiert werden können, diskutiert.

5.1 Okklusionsdetektion am Beispiel eines Zufallspunktstereogramms

Auf welche Weise sich die einzelnen Mechanismen zur Okklusionsdetektion im Rahmen des Selbstorganisationsprozesses auf die Korrespondenzbestimmung auswirken, soll am Beispiel des folgenden speziellen Zufallspunktstereogramms näher erläutert werden.

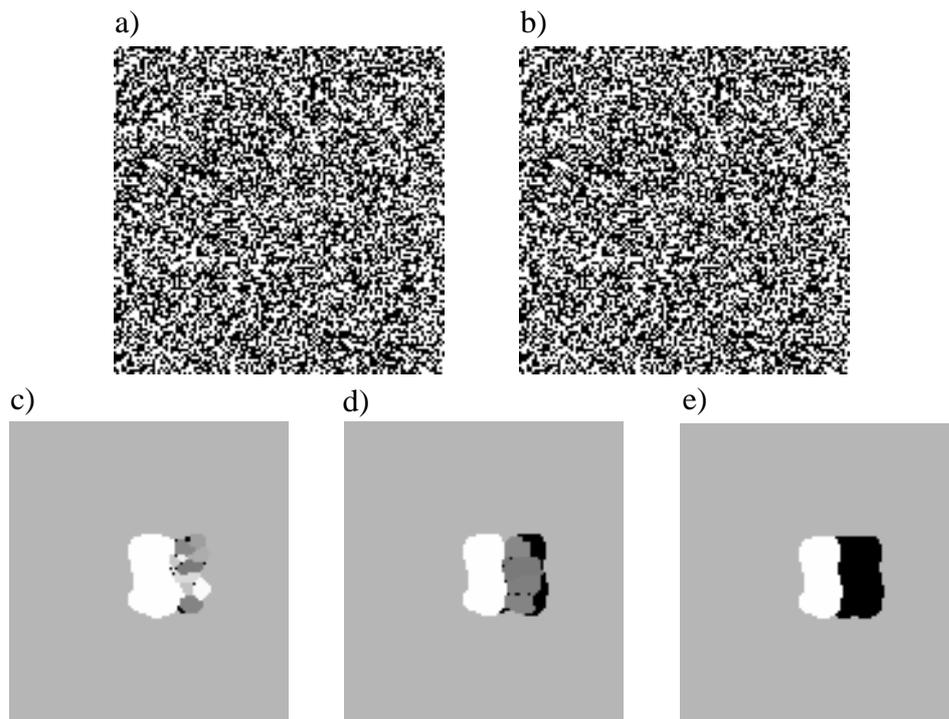


Bild 5-1: a) + b) Zufallspunktstereogramm mit versetzt eingefügtem Rechteck. Die Disparität des Rechtecks ist in diesem Beispiel gleich seiner horizontalen Ausdehnung. c)-e) Nach unterschiedlichen Verfahren errechnete Disparitätskarten (Erläuterung siehe Text)

In den Abbildungen 5-1a und 5-1b sind die zwei Bilder eines Zufallspunktstereogramms zu sehen, die sich durch einen rechteckförmigen Bereich unterscheiden, der um seine Breite ver-

setzt in die Bilder eingefügt wurde. Aufgrund der hohen Disparität des eingefügten Rechtecks tritt in diesem Beispiel ein großer Okklusionsbereich auf. Die Abbildungen 5-1c-e zeigen die zugehörigen Disparitätskarten, die sich aus unterschiedlichen Selbstorganisationsprozessen ergeben. Um die Disparität des Rechtecks erfassen zu können, wurde der Disparitätsbereich bei der Korrespondenzbestimmung gleich der Breite des verschobenen Rechtecks eingestellt. Bildbereiche, in denen alle Variablen die stationäre Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d}) = 0$ angenommen haben, wurden hier durch schwarze Regionen in den Disparitätskarten gekennzeichnet. In diesen Regionen ist somit keine Disparität definiert. An dieser Stelle sei bemerkt, daß eine sinnvolle Fusion dieser beiden Bilder dem Leser bei einer stereoskopischen Betrachtung äußerst schwer fallen wird. Der Grund hierfür ist die bereits erwähnte Disparitätsgradientengrenze im visuellen System des Menschen, die in diesem Beispiel in einigen Bildregionen überschritten wird.

Bild 5-1c zeigt eine Disparitätskarte, die durch ein Selbstorganisationsprozeß nach Beziehung (4-73) errechnet wurde. Dieser Selbstorganisationsprozeß entspricht im wesentlichen dem von Reimann und Haken vorgeschlagenen Verfahren und enthält keine Okklusionsbehandlung. Die Disparitätskarte weist rechts neben dem homogenen weißen Bereich, der die Disparität des verschobenen Rechtecks repräsentiert, zufällig verteilte Disparitätswerte auf, die Fehlkorrespondenzen des okkludierten Bildbereiches mit anderen Bildregionen im zugelassenen Disparitätsbereich darstellen.

In Bild 5-1d ist eine Disparitätskarte zu sehen, die sich aus dem Selbstorganisationsprozeß nach Beziehung (4-77) ergibt. In diesem Wettbewerbsprozeß werden Korrespondenzen zwischen Okklusionen und Regionen im anderen Bild verhindert, sofern diese einen korrekten Korrespondenzpartner aufweisen (siehe Kapitel 4.6.2). Der zulässige Disparitätsbereich wurde in diesem Beispiel jedoch so groß gewählt, daß auch Korrespondenzen zwischen den Okklusionsbereichen im rechten und linken Bild möglich sind. Da alle anderen Lösungen durch den Selbstorganisationsprozeß verhindert werden, gewinnen im Okklusionsbereich die Variablen den Wettbewerb, die gerade diese Art von Fehlkorrespondenzen repräsentieren. Aufgrund der Tatsache, daß diese Korrespondenzen auf unterschiedlichen Seiten des Objektes auftreten, ist die Disparität im Okklusionsbereich in diesem Beispiel einheitlich negativ. In der betreffenden Disparitätskarte sind die okkludierten Bereiche daher dunkler dargestellt als der Bereich, in dem das Rechteck liegt. In den schwarzen Regionen im Okklusionsbereich war das Ähnlichkeitsmaß nicht hoch genug, um Variablen, die Okklusionspaare repräsentieren, mit Werten größer Null zu initialisieren (vgl. Kapitel 4.6.1).

Die Disparitätskarte in Abbildung 5-1e zeigt das Resultat des Selbstorganisationsprozesses, der durch Beziehung (4-83) beschrieben wird. Durch den gegenseitigen Ausschluß von Korrespondenzen, die in einem zyklischen Bild die gleichen Koordinaten aufweisen, werden hier Korrespondenzen zwischen Okklusionen im rechten und im linken Stereobild verhindert. Dies führt dazu, daß im Okklusionsbereich alle Variablen die stationäre Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d}) = 0$ annehmen und diesem Bildbereich somit keine Disparitätswerte zugewiesen werden.

5.2 Dynamik der Variablen

In Abbildung 5-2 ist das dynamische Verhalten konkurrierender Variablenwerte in drei unterschiedlichen Situationen graphisch dargestellt. Die Berechnung der Variablenwerte erfolgte hierbei durch die in Kapitel 4.6.6 beschriebene numerische Integration des Differentialgleichungssystems aus Beziehung (4-83).

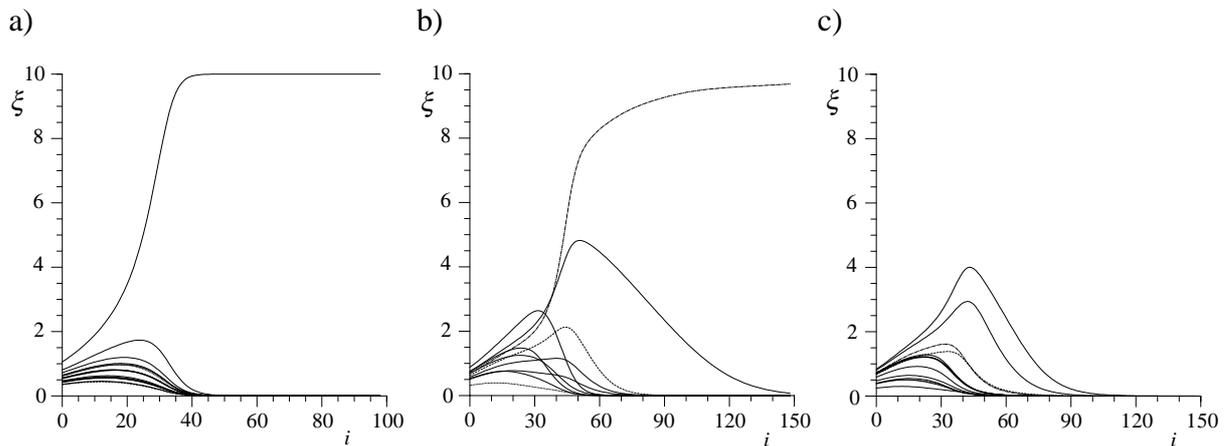


Bild 5-2: Dynamik der Variablenwerte in unterschiedlichen Situationen. a) Die Korrespondenz wird bereits durch die Initialisierung bestimmt. b) Ausgeprägter Wettbewerb im Fall von Mehrdeutigkeiten im Stereogramm. c) Okklusionsbereich

Abbildung 5-2a stellt den zeitlichen Verlauf von Variablenwerten an einem Bildpunkt dar, an dem keine Mehrdeutigkeiten im Stereogramm auftreten. In diesem Fall gewinnt die Variable den Wettbewerb, die mit dem größten Ähnlichkeitsmaß initialisiert wird, indem sie die stationäre Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ annimmt. Alle Konkurrenten dieser Variablen nehmen die stationäre Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d}) = 0$ an.

Die Grafik 5-2b zeigt eine Situation, in der Mehrdeutigkeiten im Stereogramm vorhanden sind. Der Effekt der Umfeldkopplung, durch den weitestgehend kontinuierliche Disparitätskarten erzwungen werden, bewirkt in diesem speziellen Fall, daß nicht die Variable mit dem größten Initialisierungswert den Wettbewerb gewinnt, sondern eine Variable, die einen entsprechend großen verstärkenden Einfluß aus ihrem Umfeld erhält.

In Abbildung 5-2c ist das zeitliche Verhalten der Variablenwerte im Bereich einer Okklusion zu sehen. Alle Variablen repräsentieren hier eine Fehlkorrespondenz mit entsprechenden Bildbereichen im anderen Stereokanal. Durch die in Kapitel 4.6.2 und Kapitel 4.6.3 beschriebenen Mechanismen nehmen an dieser Bildkoordinate alle Variablen die stationäre Lösung $\xi_{s_1}(\mathbf{x}, \mathbf{d}) = 0$ an, sobald eine Variable mit einem korrekten Korrespondenzpartner im umgekehrten Zuordnungsprozeß oder mit gleichen zyklischen Bildkoordinaten im jeweiligen Wettbewerbsprozeß die stationäre Lösung $\xi_{s_2}(\mathbf{x}, \mathbf{d})$ erreicht hat.

5.3 Einfluß der Umfeldkopplung

Die einzelnen Parameterkombinationen des Selbstorganisationsprozesses aus Tabelle 4-2 unterscheiden sich im wesentlichen durch das Verhältnis des Verstärkungsfaktors a zum Kopp­lungsparameter e . In welchem Maß sich die Umfeldkopplung auf die Ergebnisse, die durch den Selbstorganisationsprozeß erzielt werden, auswirkt, soll am Beispiel des folgenden stark mehrdeutigen synthetischen Stereogramms verdeutlicht werden.

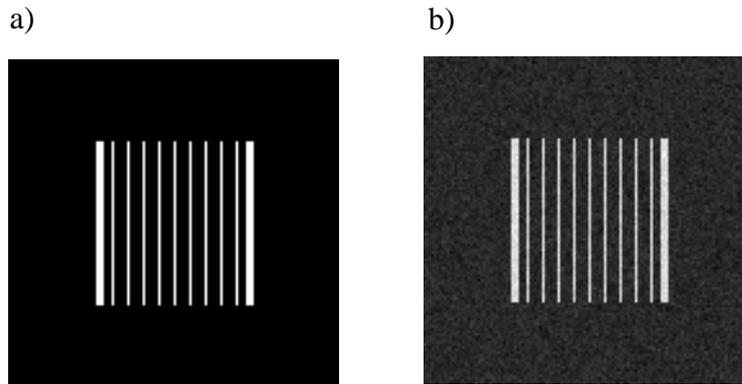


Bild 5-3: Synthetisches Stereogramm. a) Originalbild, das als Eingangsbild des linken Stereokanals dient. b) Mit einem gaußförmig verteilten Grauertrauschen ($\mu = 0$, $\sigma = 20$) überlagerte Version des Originals, das dem rechten Stereokanal zugeführt wird

Aufgrund der periodischen Strukturen des in Bild 5-3 dargestellten synthetischen Stereogramms, liefert die Messung der lokalen Ähnlichkeit keine eindeutige Aussage über die Korrespondenzen zwischen den Bildern. Erschwerend kommt in diesem Beispiel hinzu, daß sich die beiden Bilder durch einen Rauschprozeß unterscheiden, mit dem das rechte Bild überlagert wird. Eine eindeutige Zuordnung korrespondierender Bildbereiche anhand ihrer Ähnlichkeit ist nur bedingt am linken und rechten Rand des Linienmusters möglich, da hier Linien vorhanden sind, die sich durch ihre Breite von den anderen unterscheiden (vgl. Bild 5-4a).

Soll in diesem Beispiel eine korrekte gegenseitige Zuordnung aller Linien durch den Selbstorganisationsprozeß erfolgen, so kann dies nur über die Umfeldkopplung der Variablen geschehen. Ausgehend von den Rändern des Musters müssen jeweils die Variablen, welche die korrekte Disparität repräsentieren, die Variablen in ihrer unmittelbaren Nachbarschaft über die Umfeldkopplung verstärken. Dieser Effekt muß sich bis zur Mitte des Musters fortpflanzen, bevor an dieser Stelle zufällig andere Variablen den Wettbewerb gewinnen (vgl. Bild 5-5).

Die Resultate, die mit unterschiedlich starken Umfeldkopplungen bei diesem Beispiel erzielt werden können, sind in der nachfolgenden Grafik zu sehen. Der Disparitätsbereich wurde hierbei so gewählt, daß jede Linie mit maximal fünf Linien im anderen Bild verwechselt werden kann. Die Größe des verstärkend wirkenden Umfeldes wurde zu 5×5 Pixel gewählt.

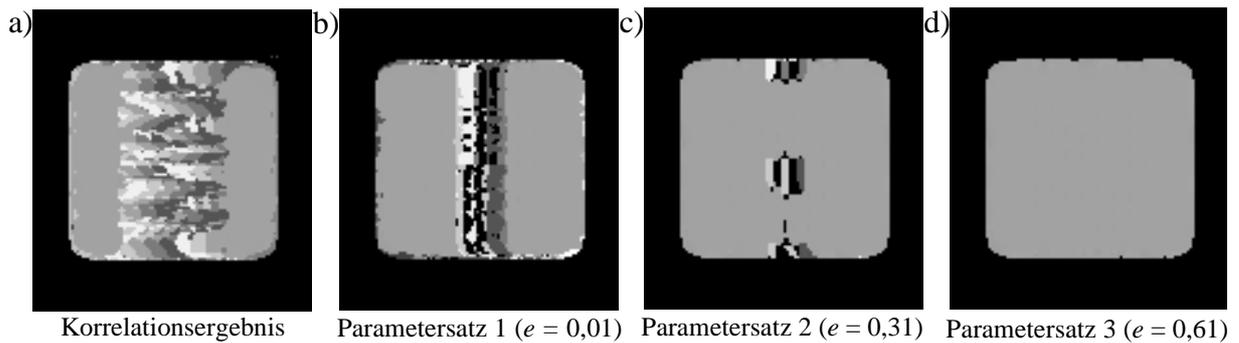


Bild 5-4: Disparitätskarten zum Stereogramm aus Bild 5-3. a) Disparitäten mit maximalem Ähnlichkeitsmaß. b,c+d) Mit unterschiedlichen Parametersätzen errechnete Disparitätskarten

Wie dieses Beispiel zeigt, können korrekte Disparitätskarten bei stark mehrdeutigen Stereogrammen nur mit einer entsprechend starken Umfeldkopplung durch den Selbstorganisationsprozeß errechnet werden. Bei zu geringem Kopplungsparameter gewinnen in der Mitte des Linienmusters Variablen den Wettbewerbsprozeß, die eine zufällige Verwechslung mit anderen Linien repräsentieren, bevor sich in diesem Bereich der verstärkende Effekt der korrekten Korrespondenzen aus dem Randbereich des Musters auswirken kann (vgl. Bild 5-4b+c).

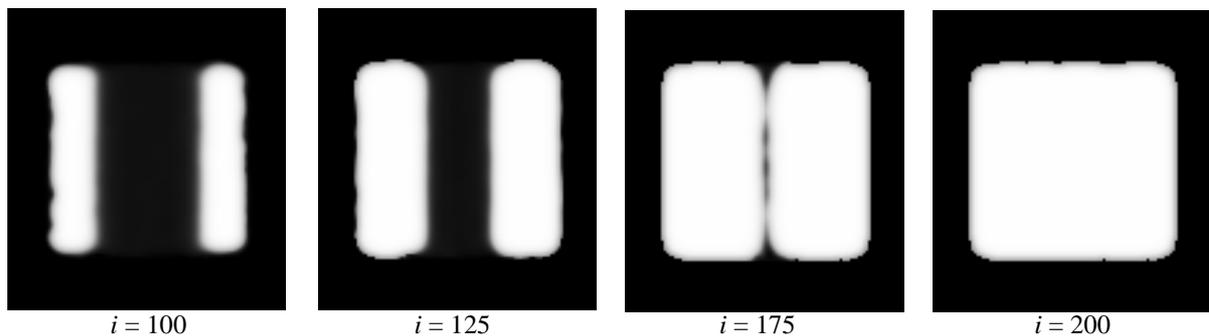


Bild 5-5: Ausbreitung der Variablenwerte, die die korrekte Disparität im Stereogramm 5-3 repräsentieren, nach unterschiedlichen Iterationsschritten des Selbstorganisationsprozesses (Parametersatz 3). Die Graustufen entsprechen hier den Werten der Variablen.

5.4 Auswirkung der subpixelgenauen Disparitätsschätzung

Die Quantisierungsfehler, die bei der Disparitätsmessung durch die diskreten Bildkoordinaten entstehen, können sich je nach Kamerageometrie enorm auf die Genauigkeit der Positionsbestimmung auswirken. Eine Möglichkeit diese Fehler zu reduzieren, bietet das in Kapitel 4.7 beschriebene Verfahren zur subpixelgenauen Disparitätsschätzung. Bevor in Kapitel 5.6 die Auswirkung dieser Methode auf die räumliche Positionsbestimmung diskutiert wird, soll an dieser Stelle der Effekt der subpixelgenauen Disparitätsschätzung zunächst an einem synthetischen Stereogramm verdeutlicht werden.

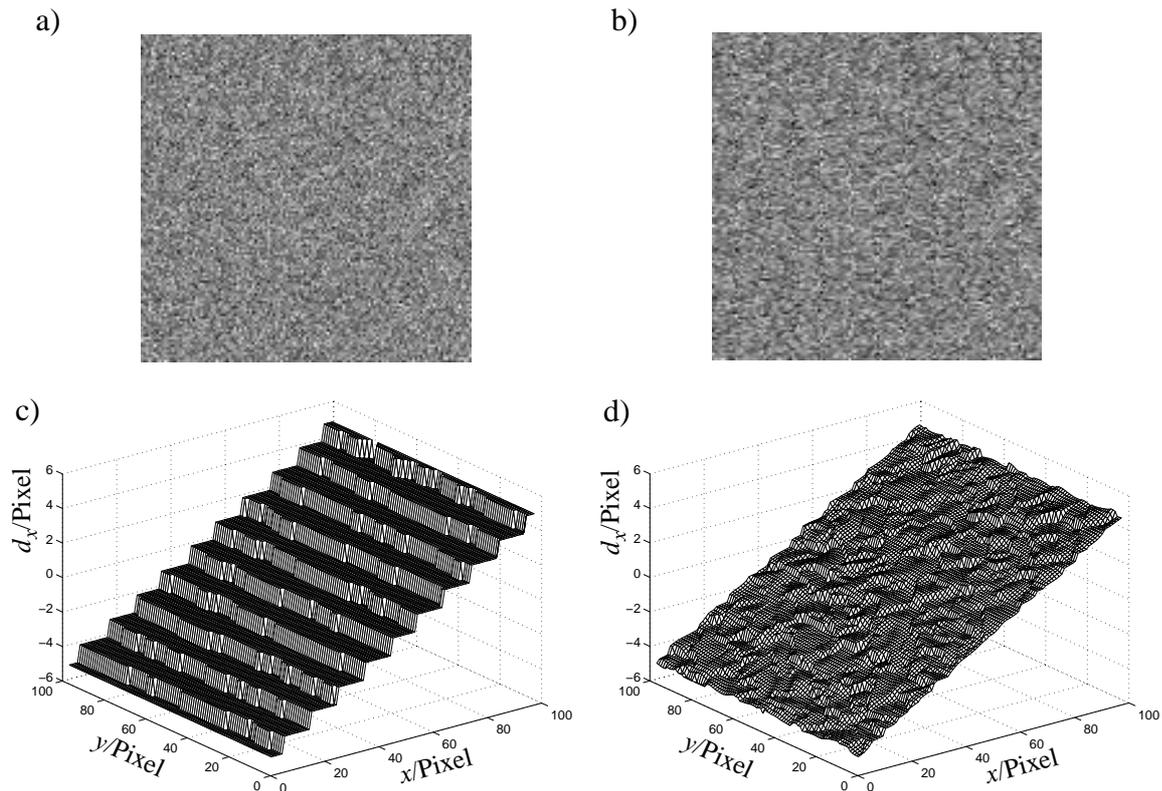


Bild 5-6: *Subpixelgenauigkeit am Beispiel eines synthetischen Stereogramms. a+b) Stereogramm, in dem beide Bilder durch ein gaußförmig verteiltes Grauertrauschen erzeugt wurden. Das rechte Bild unterscheidet sich vom linken nur durch eine horizontale Dehnung von 10%. c) Pixelgenauer Disparitätsverlauf. d) Subpixelgenauer Disparitätsverlauf*

In dem synthetischen Stereogramm, das in Bild 5-6a+b zu sehen ist, wurde ein linearer Disparitätsverlauf erzeugt, indem eines von zwei identischen Bildern horizontal gedehnt wurde. Die Bilder wurden in diesem Beispiel durch ein gaußförmig verteiltes Grauertrauschen generiert. Trotz des relativ starken Disparitätsgradienten, der in diesem Beispiel auftritt, läßt sich durch den Selbstorganisationsprozeß zunächst eine eindeutige pixelgenaue Disparitätskarte bestimmen. Die durch die diskreten Bildkoordinaten bedingte Quantisierung der Disparitätswerte ist im Disparitätsverlauf in Bild 5-6c klar zu erkennen. Die anschließende subpixelgenaue Disparitätsschätzung durch den in Kapitel 4.7 beschriebenen Prozeß ermöglicht es auf einfache Weise, die auftretenden Quantisierungsfehler erheblich zu reduzieren (siehe Bild 5-6d). Eine quantitative Aussage über die Fehler, die bei der Disparitätsmessung auftreten, kann durch die Quadratwurzel des mittleren quadratischen Fehlers (*RMS*) getroffen werden.

$$RMS = \sqrt{\frac{\sum_{i=1}^{N_m} (d_{w_i} - d_{x_i})^2}{N_m}} \quad (5-1)$$

Hierbei ist d_w jeweils der wahre Disparitätswert, d_x der gemessene Disparitätswert und N_m die Anzahl der durchgeführten Messungen. Die Fehler, die in diesem Beispiel bei pixelgenauen und bei subpixelgenauen Disparitätsmessungen in unterschiedlichen Ebenen der Auflösungs-
pyramide auftreten (siehe Kapitel 4.4.4), sind in der folgenden Tabelle aufgeführt. Die Anzahl der jeweils durchgeführten Messungen betrug hierbei etwa $N_m = 1,1 \times 10^4$.

Filtermodulati- onsortsfrequenz	Abtastfaktor s_n	RMS / Pixel pixelgenau	RMS / Pixel subpixelgenau
$\pi/2$	1	0,3038	0,1224
$\pi/4$	2	0,6193	0,2657
$\pi/8$	4	1,2551	0,5579

Tabelle 5-1: Fehler in der Disparitätsbestimmung bei unterschiedlichen Ebenen der Auflösungs-
pyramide

Wie die Tabelle zeigt, kann der Quantisierungsfehler durch die subpixelgenaue Disparitätsmes-
sung auf 40-44% des ursprünglichen Fehlers vermindert werden. Aufgrund der abnehmenden
Auflösung nimmt der Fehler mit zunehmendem Abtastfaktor s_n ebenfalls zu.

5.5 Natürliche Stereogramme

Nachdem bereits einige Ergebnisse gezeigt wurden, die das Verfahren bei speziellen syntheti-
schen Stereogrammen errechnet, soll die Leistungsfähigkeit im weiteren an einigen natürlichen
Stereogrammen veranschaulicht werden. Die Filterung der Bilddaten erfolgte bei den nachfol-
genden Beispielen mit einer Modulationsortsfrequenz von $k_0 = \pi/2$. Als Parameterkombination
wurde der Satz 3 aus Tabelle 4-2 mit einem quadratischen Umfeld von 5×5 Pixeln verwen-
det. Der horizontale Disparitätsbereich wurde jeweils auf $|d_x| \leq 15$ begrenzt.

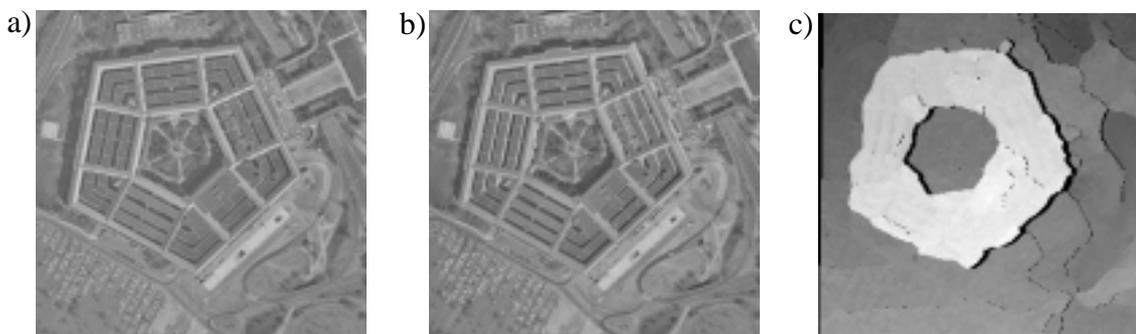


Bild 5-7: a+b) Das Pentagon-Stereogramm. c) Zugehörige Disparitätskarte

In Bild 5-7 ist das klassische Pentagon-Stereogramm mit der zugehörigen Disparitätskarte dar-
gestellt. Problematisch sind in diesem Beispiel die einzelnen Gebäudeblöcke, die durch ihre
periodische Struktur leicht zu Mehrdeutigkeiten führen können. In der Disparitätskarte hebt

sich der Gebäudekomplex klar vom Hintergrund ab. Jeweils rechts von den Wänden des Gebäudes sind in der Disparitätskarte die detektierten Okklusionen, die das Gebäude auf dem Untergrund erzeugt, als schwarz markierte Gebiete zu sehen. Aufgrund der nicht bekannten Kamerageometrie konnten in diesem wie auch im folgenden Beispiel keine quantitativen Aussagen über die Position der dargestellten Objekte getroffen werden.

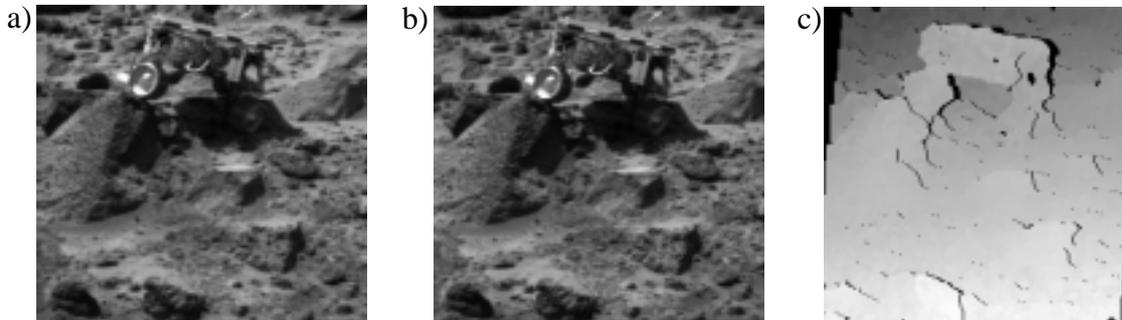


Bild 5-8: a+b) Stereogramm, das während der Pathfinder-Mission entstand. c) Zugehörige Disparitätskarte

Bild 5-8 zeigt ein Stereogramm aus der Pathfinder-Mission, in dem ein ferngesteuertes Fahrzeug auf dem Marsboden zu sehen ist. In der zugehörigen Disparitätskarte ist die nach oben zunehmende Entfernung des Marsbodens zu erkennen. Das Fahrzeug hebt sich durch seine geringere Distanz zum Kamerasystem vom Boden ab und erzeugt dadurch Okklusionen auf dem Marsboden im Hintergrund, die als schwarze Regionen an den rechten Kanten des Fahrzeugs zu sehen sind. Aufgrund der unterschiedlichen Perspektive, aus der die Bilder aufgenommen wurden, existieren auch im Randbereich der Bilder nur einseitig sichtbare Gebiete.

5.6 Meßgenauigkeit bei der Positionsbestimmung

Die Genauigkeit, mit der eine Szene stereoskopisch vermessen werden kann, ist in vielen Anwendungen eine kritische Größe. Sollen beispielsweise die Operationen eines Roboters über ein stereoskopisches Verfahren koordiniert werden (siehe Kapitel 6), so ist die Genauigkeit, mit der die räumliche Position der zu manipulierenden Objekte erfaßt werden kann, von entscheidender Bedeutung. Die Meßgenauigkeit des in der vorliegenden Arbeit beschriebenen Gesamtsystems wird im wesentlichen durch die vier folgenden Faktoren beeinflusst:

1. *Die Genauigkeit, mit der die Parameter des Kameramodells durch die Kalibrierung ermittelt werden*
2. *Die Meßgenauigkeit, mit der die Motorwinkel erfaßt werden können*
3. *Die Einstellung der optischen Freiheitsgrade (Zoom, Fokus usw.)*
4. *Die Genauigkeit der Disparitätsbestimmung*

Eine quantitative Aussage über die Auswirkung der unterschiedlichen Quantisierungsfehler auf die Positionsbestimmung liefert eine Abschätzung durch das totale Fehlerdifferential der Beziehungen (4-126) bis (4-128) (siehe [ST97]). Die kritische Größe stellt hierbei die Meßgenauigkeit in Richtung der z_w -Achse dar. Um die theoretische Fehlerabschätzung zu validieren, wurden spezielle Meßpunkte, deren Positionen im Weltkoordinatensystem bekannt waren, mit dem Gesamtsystem vermessen. Im Rahmen dieser Untersuchung wurden die Meßpunkte in unterschiedlichen Entfernungen zum Kamerasystem mit Hilfe einer Kalibrierungsvorrichtung (vgl. Bild 4-6) positioniert und durch die passive Triangulation lokalisiert.

Abbildung 5-9 zeigt den Betrag des Meßfehlers Δz_w , der sich durch das totale Fehlerdifferential ergibt, und den maximalen gemessenen Fehlerbetrag Δz_{max} , der bei pixelgenauer Disparitätsmessung auftrat. Weiterhin sind die Quadratwurzeln der mittleren quadratischen Fehler bei pixelgenauer (RMS_{zp}) und bei subpixelgenauer (RMS_{zs}) Disparitätsmessung zu sehen. Für jede Entfernung wurden durchschnittlich etwa 200 Meßpunkte ausgewertet, wobei die Vergenzwinkel des Kamerasystems mehrfach variiert wurden. Die den Messungen zugrundeliegenden Disparitätskarten wurden mit der besten Auflösungsebene ohne Unterabtastung der Filterergebnisse errechnet (vgl. Kapitel 4.4.4). Bei der Fehlerabschätzung durch das totale Fehlerdifferential wurde ein Quantisierungsfehler von einem halben Pixel in der Disparitätsbestimmung und jeweils $0,01^\circ$ bei der Erfassung der Vergenzwinkel angenommen. Letzteres entspricht der halben Winkelauflösung der verwendeten Meßwertgeber. Weiterhin wurde diese Fehlerabschätzung für eine Fixation ($d_x = 0$) der Raumpunkte in der jeweiligen Entfernung bei entsprechenden Motorwinkeln und für die jeweils kalibrierten Modellparameter, die sich durch die Fokussierung ergeben, durchgeführt.

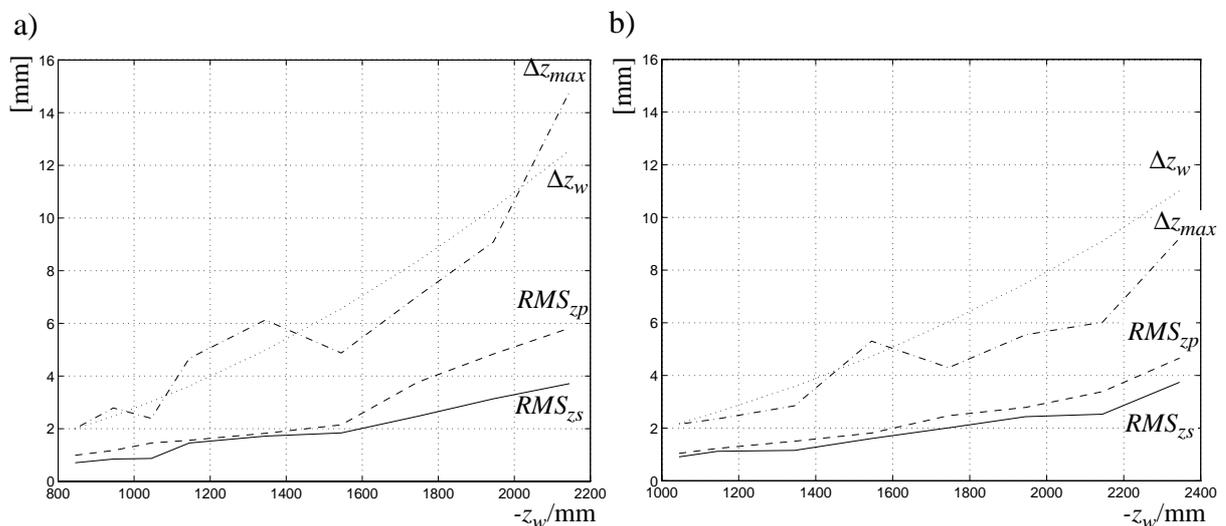


Bild 5-9: Theoretische und gemessene Fehler in z_w -Richtung in Abhängigkeit der Entfernung zum Kamerasystem. a) Fehler bei einem Weitwinkelobjektiv. b) Fehler bei einer Zoomeinstellung, die einer Erhöhung der effektiven Brennweiten gegenüber dem Weitwinkelobjektiv um etwa 75% entspricht

Die Grafiken 5-9a+b zeigen deutlich die Auswirkung der subpixelgenauen Disparitätsmessung auf die Genauigkeit der Positionsbestimmung in z_w -Richtung. So kann der mittlere quadratische Meßfehler in z_w -Richtung teilweise erheblich reduziert werden. Wie die Unterschiede zwischen den Grafiken zeigen, kann die Meßgenauigkeit darüber hinaus durch eine Vergrößerung der effektiven Brennweiten erhöht werden. Ein besonderer Vorteil der hier eingesetzten Kamerasysteme ist, daß die Brennweiten in einem sehr großen Bereich variiert werden können. Es besteht daher die Möglichkeit, die je nach Anwendung geforderten Genauigkeiten flexibel durch eine Anpassung der Kameraparameter für unterschiedliche Entfernungen einzustellen, ohne daß konstruktive Änderungen, wie beispielsweise eine Vergrößerung der Stereobasis, erforderlich sind.

Die Meßgenauigkeit des Verfahrens in x_w - und y_w -Richtung hängt in einem hohen Maß von der Genauigkeit ab, mit der die zu vermessenden Punkte in der Bildebene lokalisiert werden können. Bei der hier durchgeführten, lediglich pixelgenauen Meßpunktdetektion stieg die Quadratwurzel des mittleren quadratischen Meßfehlers mit zunehmender Entfernung zum Kamerasystem leicht an und betrug bei der größeren Brennweite zwischen 0,56 mm und 0,74 mm in x_w - und zwischen 0,4 mm und 0,76 mm in y_w -Richtung.

6 Demontage von Altfahrzeugen - ein Anwendungsbeispiel

Als eine Anwendung des in Kapitel 4 beschriebenen Verfahrens zur Tiefenrekonstruktion soll an dieser Stelle ein spezieller Prozeß im sogenannten DEMON-Projekt, in welches das Verfahren integriert wurde, näher erläutert werden. Das Ziel dieses BMBF-Verbundprojektes ist die automatische Demontage von Altfahrzeugen beliebigen Fabrikats zum Zweck der Wertstofftrennung. Im Gegensatz zur Fertigung der Kraftfahrzeuge ist bei ihrer Demontage im allgemeinen nicht mit gleichbleibenden a priori bekannten Gegebenheiten zu rechnen. Die Kraftfahrzeuge können im Rahmen ihrer Nutzung vielmehr den unterschiedlichsten Veränderungen unterliegen. Darüber hinaus unterscheiden sich die Fahrzeuge von Fabrikat zu Fabrikat oft erheblich. Zur automatischen Demontage der Kraftfahrzeuge ist aufgrund dieser Umstände ein flexibles Robotersystem notwendig, das über sensorische Fähigkeiten zur Wahrnehmung seiner Umwelt verfügt.

Eine zentrale Aufgabe stellt in diesem Prozeß die Koordination des Roboters bzw. der Demontagewerkzeuge dar. Da die räumliche Position der zu demontierenden Bestandteile nicht als bekannt vorausgesetzt werden kann, müssen folglich diese zunächst identifiziert und lokalisiert werden.

Die Identifizierung bzw. die Erkennung einzelner Objekte erfolgt im DEMON-Projekt durch das schon mehrfach erwähnte optische Mustererkennungssystem [GS+98, SH97]. Dieses Verfahren, auf das hier nicht näher eingegangen wird, analysiert die Bildinformation eines der beiden Stereokanäle des in Kapitel 4.1 beschriebenen aktiven Sehsystems. Die zu analysierenden Bilddaten werden hierbei ebenfalls den im Rahmen dieser Arbeit entwickelten Verfahren zur Bildaufbereitung und Filterung unterzogen. Aufgrund der gemeinsamen Daten, die zur Mustererkennung und zur Tiefenrekonstruktion ausgewertet werden (vgl. Bild 6-1), bietet das stereoskopische Verfahren gegenüber anderen Abstandssensoren den Vorteil, daß die resultierenden Positionsinformationen unmittelbar auf die erkannten Objekte bezogen werden können, wodurch eine aufwendige Datenfusion entfällt.

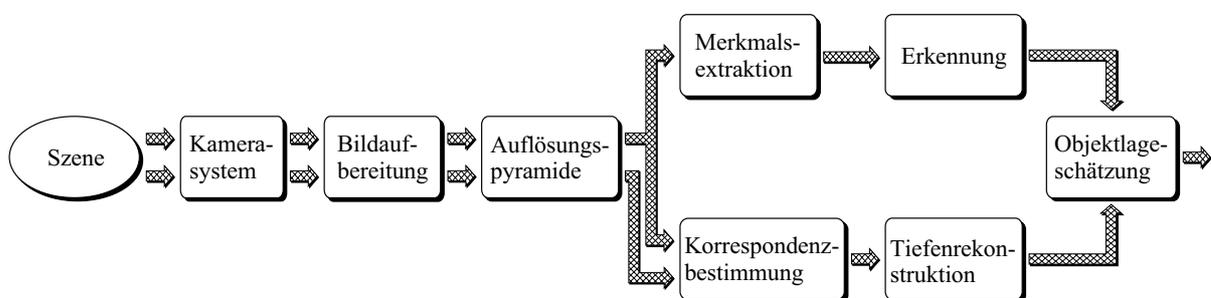


Bild 6-1: Integration der Objekterkennung und der stereoskopischen Korrespondenzbestimmung in ein Gesamtsystem zur Objektlageschätzung

Eine Station der oben beschriebenen Anwendung bildet die Demontage der Fahrzeugräder (vgl. Bild 6-2). Da in dieser Anwendung das Fahrzeug innerhalb gewisser Grenzen beliebig zum Robotersystem positioniert sein kann, müssen zunächst das Rad und anschließend die Befestigungsschrauben, die mit Hilfe des Demontageroboters gelöst bzw. herausgeschnitten werden sollen, identifiziert und räumlich lokalisiert werden.

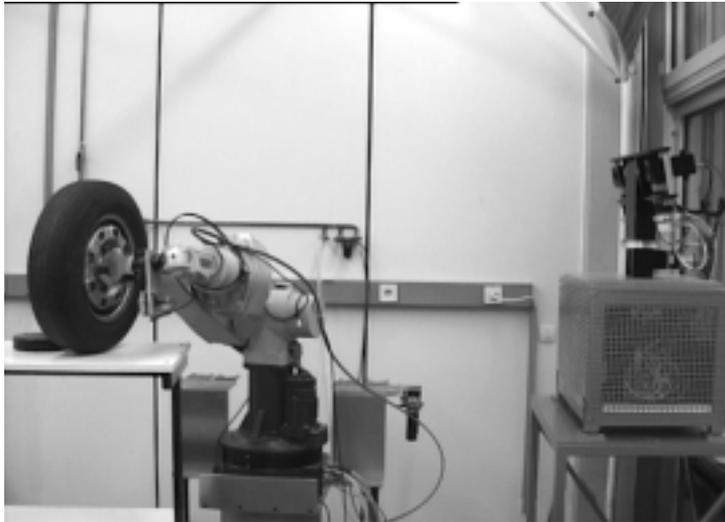


Bild 6-2: Analyse einer Demontageszene mit Hilfe eines aktiven binokularen Sehsystems zur Koordination der Roboteraktionen

Über die Kombination von unterschiedlichen Auflösungsebenen (vgl. Kapitel 4.4.4) und Zoomeinstellungen kann bei dieser Aufgabe ein Kompromiß zwischen der Größe des zu vermessenden Gebietes, der Meßgenauigkeit und dem Berechnungsaufwand erzielt werden. In den folgenden Abschnitten wird nun gezeigt, wie diese Strategie effizient in den einzelnen Schritten des Handlungsablaufes umgesetzt werden kann.

6.1 Lagerfassung des Rades

Um einen möglichst großen Raumbereich analysieren zu können, wird die Szene in einem ersten Schritt mit einer niedrigen Kamerabrennweite aufgenommen und im Rahmen der Auflösungspyramide grob abgetastet (siehe Kapitel 4.4.2). Die aus den Filterantworten gewonnenen Konturmerkmale erlauben in diesem Verarbeitungsschritt zwar noch keine Identifikation einzelner Details in der Szene, jedoch kann die Kontur der Radfelge in diesem Schritt aufwandsgünstig erfaßt und anhand der Tiefenkarte grob lokalisiert werden.

Da sich die in Bild 6-3b weiß gekennzeichneten Merkmale, anhand derer die Radfelge erkannt wird, auf die gleichen Bildkoordinaten wie die Disparitätskarte in Bild 6-3c beziehen, kann die räumliche Position dieser Merkmale durch das Auswerten der Disparitätskarte an den entsprechenden Bildkoordinaten auf sehr einfache Weise ermittelt werden (vgl. Bild 6-4).

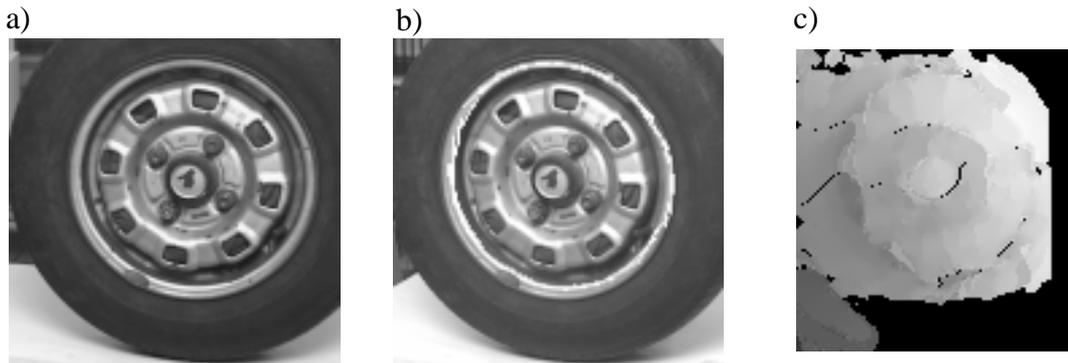


Bild 6-3: a) + b) Rektifiziertes Stereobildpaar eines Autorades. In Bild b) sind die Merkmalspunkte, die zur Erkennung der Radfelge herangezogen wurden, weiß gekennzeichnet. c) Grob aufgelöste Disparitätskarte des Stereobildpaares

Mit Hilfe der Disparitätskarte in Bild 6-3c kann die Position des Rades und auch seine Orientierung im Raum ermittelt werden. Letzterem kommt im Rahmen des Entschraubungsvorgangs eine besondere Bedeutung zu, da die Demontagewerkzeuge möglichst senkrecht zur Radebene aufgesetzt werden müssen. Die hier errechnete Disparitätskarte kann im weiteren dazu genutzt werden, das Rad durch Ausrichten der optischen Achsen zu fixieren und genauer zu vermessen.

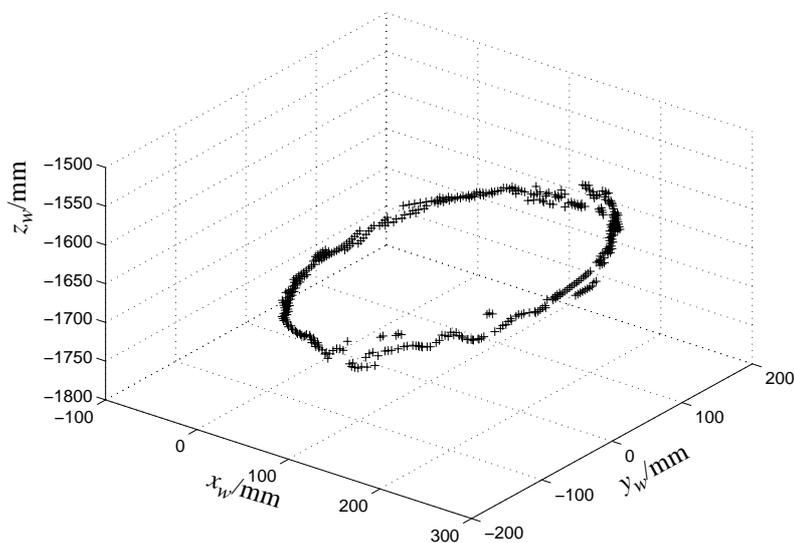


Bild 6-4: Rekonstruierte räumliche Position der Merkmalspunkte, die zur Erkennung der Radfelge dienen

6.2 Detektion des Schraubenkranzes

Durch die Fixierung des Rades mit Hilfe des aktiven Sehsystems kann der Disparitätsbereich eingeschränkt werden, da der zu vermessende Tiefenbereich durch den vorangegangenen Ver-

arbeitungsschritt bereits bekannt ist. Auf diese Weise läßt sich der Berechnungsaufwand erheblich reduzieren (vgl. Kapitel 4.5). Die Verwendung eines höheren Zoomfaktors und die Auswertung einer höher auflösenden Ebene der Auflösungspyramide erlauben es in diesem Verarbeitungsschritt bereits, die einzelnen Schrauben zu erkennen und räumlich zu lokalisieren.

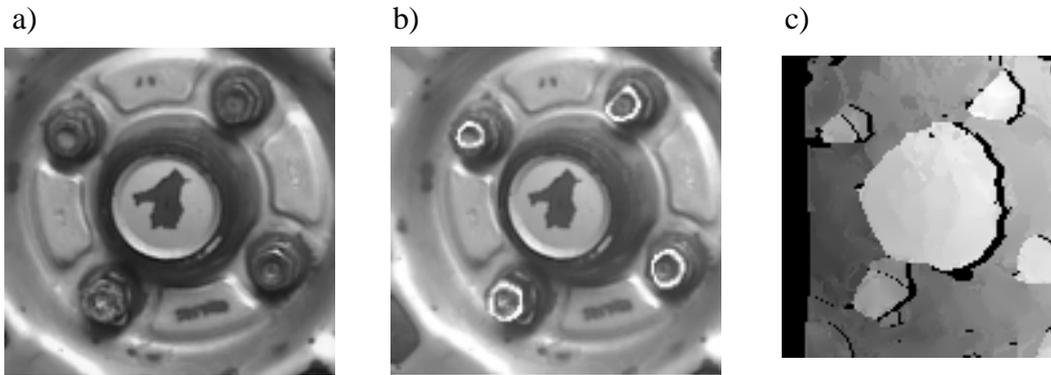


Bild 6-5: a) + b) Höher aufgelöste rektifizierte Bildausschnitte des Schraubenkranzes. In Bild b) sind die Merkmalspunkte der einzelnen Schrauben weiß markiert. c) Disparitätskarte des Bildausschnittes

Ist die Genauigkeitsanforderung entsprechend hoch oder das Objekt zu weit entfernt, so ist es möglich, daß der komplette Radkranz nicht mit einem angemessenen Berechnungsaufwand und der geforderten Genauigkeit vermessen werden kann. In diesem Fall besteht die Möglichkeit, die einzelnen Schrauben sequentiell mit höherer Genauigkeit zu vermessen.

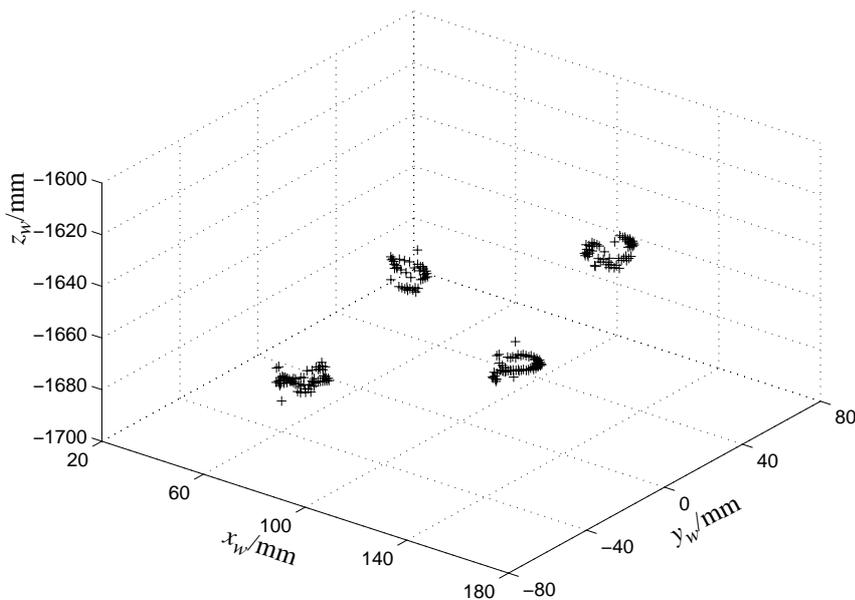


Bild 6-6: Rekonstruierte räumliche Position der Merkmalspunkte, die zur Erkennung des Schraubenkranzes dienen

6.3 Vermessung einzelner Schrauben

Anhand der im vorangegangenen Verarbeitungsschritt ermittelten Disparitätskarte besteht die Möglichkeit, einzelne Schrauben exakt zu fixieren und zu vermessen. Die höchste Genauigkeit kann hierbei durch Bildaufnahmen mit maximaler Zoomeinstellung und durch die Auswertung der nicht unterabgetasteten Auflösungsebene erzielt werden.

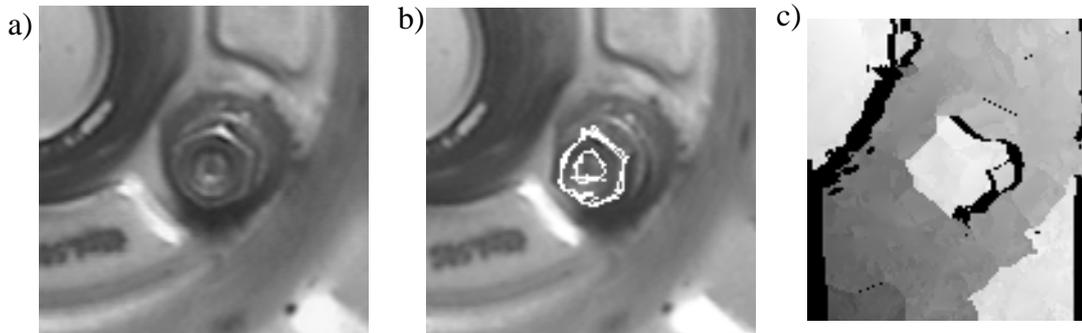


Bild 6-7: a) + b) Hochaufgelöste rektifizierte Bildausschnitte einer einzelnen Schraube. In b) sind die Merkmalspunkte der Schraube weiß gekennzeichnet. c) Disparitätskarte des Bildausschnittes

Die zuvor erwähnten Einstellungen führen dazu, daß nur sehr kleine Raumbereiche mit vertretbarem Berechnungsaufwand vermessen werden können. Für das hier diskutierte Anwendungsbeispiel bedeutet dies, daß die einzelnen Schrauben sequentiell durch jeweils erneutes Fixieren vermessen werden müssen. In diesem Zusammenhang ist eine zeitliche Verzahnung der einzelnen Positionsbestimmungen und der Aktionen des Demontageroboters denkbar.

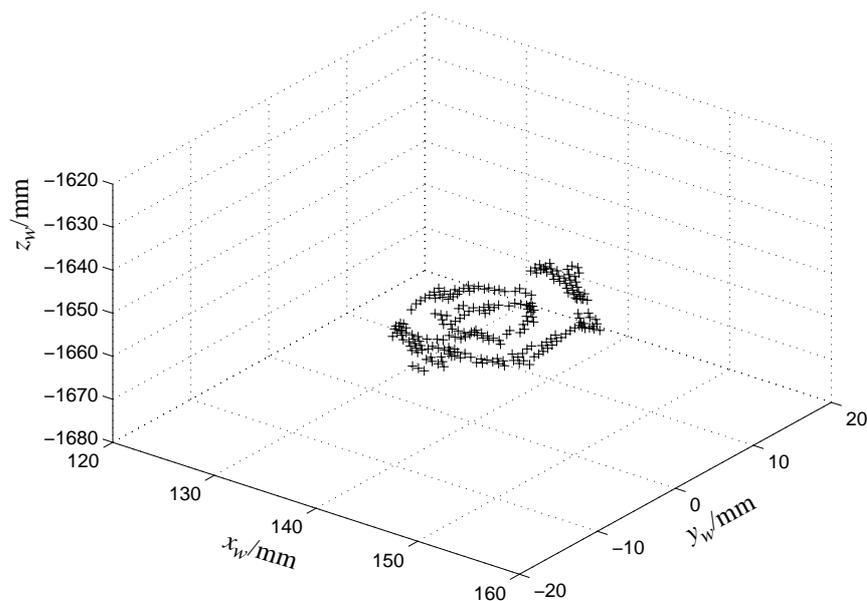


Bild 6-8: Rekonstruierte räumliche Position der Merkmalspunkte, die in Bild 6-7b weiß markiert sind

Die in den Abbildungen 6-4, 6-6 und 6-8 dargestellten rekonstruierten Weltkoordinaten beziehen sich auf das in Bild 4-19 gezeigte Koordinatensystem. Da die z_w -Achse dieses Koordinatensystems immer in der Ebene liegt, welche die beiden optischen Achsen der Kameras aufspannen, variiert seine relative Lage zu einem statischen Objekt mit Veränderungen des Neigewinkels des aktiven Sehsystems (siehe Kapitel 4.1). Während die x_w - und die z_w -Koordinaten in den obengenannten Abbildungen näherungsweise miteinander verglichen werden können, bezieht sich die y_w -Koordinate in diesen Abbildungen auf den jeweils fixierten Punkt. Da jedoch auch der Neigewinkel des Kamerasystems meßtechnisch erfaßt wird, können die rekonstruierten Weltkoordinaten in ein Koordinatensystem transformiert werden, auf das sich die Bewegungen des Demontageroboters beziehen.

6.4 Laufzeiten und spezielle Lösungen

Die Laufzeit des gesamten stereoskopischen Verfahrens kann bei der hier gezeigten Anwendung auf etwa 3,7 Sekunden eingestellt werden. Der größte Anteil der Laufzeit entfällt hierbei auf den Selbstorganisationsprozeß (etwa 3,1 Sekunden bei 150 Iterationen). Als Rechnersystem wurde bei dieser Messung eine SUN UltraSparc II mit einer Taktfrequenz von 200MHz und 128MB Arbeitsspeicher verwendet. Die Filterantwortgröße betrug jeweils 64×64 Pixel und als Parameterkombination wurde der Satz 2 aus Tabelle 4-2 mit einer Umfeldgröße von 3×3 Pixel gewählt. Der Disparitätsbereich wurde bei dieser Messung auf $|d_x| \leq 10$ beschränkt. Da alle Verfahrensschritte unter Verwendung von Multi-Threading-Strategien bis zu einem gewissen Grad parallel implementiert sind, kann durch den Einsatz einer geeigneten Mehrprozessormaschine eine Verringerung der Laufzeit erzielt werden. Innerhalb des in Bild 6-1 gezeigten Gesamtsystems können zudem Erkennungsprozeß und stereoskopische Korrespondenzbestimmung auf einfache Weise parallel implementiert werden, da diese Prozesse weitestgehend unabhängig voneinander ablaufen.

Ist der Raumbereich, in dem das zu bearbeitende Objekt frei positioniert werden kann, geeignet eingeschränkt, so kann auf ein aufwendiges aktives Sehsystem mit rotatorischen Freiheitsgraden verzichtet und stattdessen ein starres Kamerasystem eingesetzt werden. Da in diesem Fall keine Fixation möglich ist, die eine starre konzentrische Pyramidenstruktur zulassen würde, müßten die hochauflösenden Bildausschnitte selektiv aus dem Originalbild ausgeschnitten und weiterverarbeitet werden. Der Berechnungsaufwand ist bei dieser Strategie nicht wesentlich höher als bei einem aktiven Sehsystem, jedoch wird der Einsatz von unterschiedlichen Zoomfaktoren zur Erhöhung der Meßgenauigkeit bei einem Kamerasystem ohne mechanische Freiheitsgrade stark eingeschränkt.

7 Zusammenfassung

Im Rahmen dieser Arbeit konnte ein Gesamtsystem realisiert werden, das es erlaubt, die räumliche Struktur einer Szene anhand von Stereobilddaten zu rekonstruieren. Obwohl die einzelnen Komponenten, aus denen das System aufgebaut wurde, zunächst speziell für den stereoskopischen Einsatz konzipiert wurden, fanden einige der entwickelten Verfahren bereits Anwendung in Bereichen der Bildverarbeitung, die nicht in unmittelbarem Zusammenhang mit der Stereoskopie stehen. Unter diese Komponenten fällt ein aktives binokulares Sehsystem, das über die Stereoskopie hinaus in verschiedenen Objekterkennungsstrategien und Anwendungen im Bereich der Detektion bzw. Verfolgung von Bewegungen als aktives bildgebendes System eingesetzt wird. Die Filtertechnik, die im Rahmen der stereoskopische Korrespondenzbestimmung zur Extraktion der relevanten Bildstrukturen aus den Grauwertbildern dient, hat sich ebenfalls bereits in vielen Anwendungsbereichen bewährt und wird als Vorstufe zur Extraktion orientierter Merkmale in verschiedenen Mustererkennungsverfahren mit vielversprechendem Erfolg eingesetzt.

Das anwendungsübergreifend eingesetzte aktive Sehsystem zeichnet sich neben einem äußerst kostengünstigen Aufbau gegenüber den wenigen vergleichbaren Kamerasystemen durch eine hohe Positionierungsgenauigkeit und aufwendige Justierungsmöglichkeiten aus. Die sehr kompakten aber dennoch flexiblen Kameramodule erlauben es, das System über mehrere optische Freiheitsgrade an die jeweiligen Umgebungsbedingungen automatisch zu adaptieren. Durch eine aufwendige Modellierung der bildgebenden Prozesse dieser Kameras konnten die Voraussetzungen für spezielle Bildaufbereitungsverfahren, welche die Auslegung der nachfolgenden Systemkomponenten erheblich erleichterte, geschaffen werden. Diese Bildaufbereitungsverfahren umfassen eine Kompensation von bestimmten Linsenfehlern und eine Rektifikations-transformation, durch welche die Lösung des stereoskopischen Korrespondenzproblems vereinfacht werden konnte.

Die Extraktion speziell orientierter Grauwertänderungen, anhand derer eine stereoskopische Zuordnung getroffen werden kann, erfolgt durch eine orientierungs- und ortsfrequenzselektive Filterung der zuvor aufbereiteten Bilddaten. Neben einer gewissen Beleuchtungsinvarianz, die durch die Bandpaßcharakteristik der eingesetzten Gaborfilter erzielt wird, konnte der Berechnungsaufwand bei einer speziellen Auslegung der Filter durch den Einsatz einer Auflösungs-pyramide drastisch reduziert werden. Nicht zuletzt die hierdurch erzielte Laufzeiteinsparung, die sich auf alle nachfolgenden Verarbeitungsschritte auswirkt, führte zu einem anwendungs-übergreifenden Einsatz dieser Filtertechnik.

Einen ersten Hinweis auf mögliche Korrespondenzpaare im Rahmen der stereoskopischen Zuordnung liefert ein komplexes Ähnlichkeitsmaß, mit dem die lokale Übereinstimmung der zuvor gefilterten Bilddaten ermittelt wird. Aufgrund der vorangegangenen Bildaufbereitungsverfahren konnte die Berechnung dieses Ähnlichkeitsmaßes effizient durch eine nur eindimensionale lokale Korrelation realisiert werden. Durch diesen Korrelationsansatz wird der erfaßbare Disparitätsbereich im Gegensatz zu den weitverbreiteten Phasendifferenzverfahren

nicht auf die Modulationswellenlänge der verwendeten Gaborfilter beschränkt. Weiterhin ist dieses Ähnlichkeitsmaß invariant bezüglich Kontrastdifferenzen zwischen den Stereobildern und es erlaubt die Berechnung von sehr „dicht besetzten“ Tiefenkarten. Aufgrund der inhärenten Mehrdeutigkeit des Korrespondenzproblems ist jedoch eine eindeutige gegenseitige Zuordnung der Bilddaten beider Stereokanäle allein anhand deren Ähnlichkeit nicht möglich, so daß ein weiterer Mechanismus, der dieses Problem berücksichtigt, vorgesehen werden mußte.

Zur aufwandsgünstigen Lösung des Korrespondenzproblems wurde ein sogenannter Selbstorganisationsprozeß implementiert, der unter Ausnutzung bestimmten Vorwissens eine eindeutige Zuordnung korrespondierender Bildbereiche erlaubt. Im Rahmen dieses Prozesses werden mögliche Korrespondenzpaare durch Variablen repräsentiert, die durch die betreffenden Werte des Ähnlichkeitsmaßes initialisiert werden. Die Variablen sind hierbei in einen dynamischen Prozeß involviert, der durch ein nichtlineares gekoppeltes Differentialgleichungssystem beschrieben wird. Anhand der stationären Lösungen, welche die einzelnen Variablen im Zuge der Selbstorganisation annehmen, werden schließlich die Korrespondenzpaare ermittelt. Das Vorwissen, das in Form von Einschränkungen an den Lösungsraum vorliegt, wird durch eine spezielle Kopplungsstruktur in den Selbstorganisationsprozeß integriert. Eine explizite Modellierung der zu rekonstruierenden Szene ist somit überflüssig. Die dämpfend auf die einzelnen Variablen wirkenden Kopplungsstrukturen führen dazu, daß Lösungen, die eine der Einschränkungen verletzen, im Wettbewerb miteinander stehen und sich gegenseitig ausschließen. Neben der dämpfenden Kopplung innerhalb des Wettbewerbs sind Variablen, die Korrespondenzpaare mit gleicher Disparität repräsentieren, lokal verstärkend miteinander gekoppelt. Diese laterale Umfeldkopplung bewirkt, daß Mehrdeutigkeiten auch bei Störungen, die das Ähnlichkeitsmaß beeinflussen, zuverlässig aufgelöst werden.

Eine Besonderheit des hier vorgestellten Verfahrens stellt die implizite Detektion von okkludierten Bildbereichen dar, wodurch eine der Hauptfehlerursachen bei der stereoskopischen Korrespondenzbestimmung eliminiert werden konnte. Spezielle Kopplungsstrukturen innerhalb des Selbstorganisationsprozesses bewirken in diesem Zusammenhang, daß Variablen, die Korrespondenzen mit okkludierten Bildbereichen repräsentieren, im Wettbewerbsprozeß unterdrückt werden. Im Gegensatz zu den wenigen Ansätzen, in denen Okklusionen Berücksichtigung finden, erhöht sich der Berechnungs- und Speicheraufwand bei diesem Verfahren nicht wesentlich. Darüber hinaus werden aufgrund der lateralen Umfeldkopplungen Okklusionszuordnungen auch bei Störeinflüssen, die durch interokuläre Differenzen hervorgerufen werden, zuverlässig in gewissen Grenzen verhindert. Durch die ausschließlich lokalen Kopplungsstrukturen innerhalb des Selbstorganisationsprozesses besteht ein hohes Parallelisierungspotential, durch das die Laufzeit des Verfahrens bei Einsatz geeigneter Rechnerarchitekturen weiter reduziert werden kann.

Durch das verwendete Ähnlichkeitsmaß konnte in Verbindung mit der eingesetzten Filtertechnik eine einfache Möglichkeit geschaffen werden, die in den Stereobildern auftretenden Disparitäten mit Subpixelgenauigkeit zu vermessen. Die Genauigkeit der Positionsbestimmung konnte durch diese Technik teilweise erheblich verbessert werden.

Die Leistungsfähigkeit des hier vorgestellten Verfahrens zur stereoskopischen Korrespondenzbestimmung konnte anhand von einigen synthetischen wie auch natürlichen Stereogrammen demonstriert werden. Die Zuverlässigkeit, die geringe Laufzeit und nicht zuletzt die durch zahlreiche Messungen validierte Genauigkeit der Positionsbestimmung, die durch das flexible Sehsystem variabel an die jeweilige Anwendung adaptiert werden kann, waren die ausschlaggebenden Gründe für die feste Integration des Gesamtsystems in einen Demontageprozeß im Rahmen eines durch das Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie geförderten Verbundprojektes. In diesem Projekt wird das Verfahren in Verbindung mit einer Mustererkennungsstrategie dazu genutzt, die im Blickfeld des Sehsystems befindlichen Objekte anhand der Bildinformation räumlich zu lokalisieren, um auf diese Weise eine Manipulation der betreffenden Objekte durch ein Robotersystem zu ermöglichen.

8 Symbolverzeichnis

a	Gewichtungsparameter des exponentiellen Wachstums der Variablen im Selbstorganisationsprozeß
a_x, a_y	horizontaler und vertikaler Abstand der Bildsensorelemente
b	Gewichtungsparameter des dämpfend wirkenden Kopplungsterms im Selbstorganisationsprozeß
b_s	Stereobasis
b_v	Abstand der Vergenzachsen des binokularen Sehsystems
c	Gewichtungsparameter des Terms zur Begrenzung der Variablenamplituden im Selbstorganisationsprozeß
$\mathbf{d} = [d_x \ d_y]^T$	Disparität
$d_{xs}, d_{xs}(\mathbf{x})$	subpixelgenaue Disparität, bzw. Disparitätskarte in x -Richtung
e	Gewichtungsparameter der verstärkend wirkenden Umfeldkopplung im Selbstorganisationsprozeß
e_d, e_k	Kostenfunktionen
f	effektive Brennweite
f_n	Rektifikationsbrennweite
f_x, f_y	normierte effektive Brennweiten
$\mathcal{F}\{f(\mathbf{x})\} = F(\mathbf{k})$	Fouriertransformierte von $f(\mathbf{x})$
$g(\mathbf{x}), G(\mathbf{k})$	Impulsantwort, Übertragungsfunktion eines Gaborfilters
g_i	Gerschgorin-Kreise
$i(\mathbf{x}), I(\mathbf{k})$	Grauwertbild, Fouriertransformierte des Grauwertbildes
$\text{Im}\{f(x)\}$	Imaginärteil von $f(x)$
\mathbf{I}	Einheitsmatrix
j	imaginäre Einheit
\mathbf{J}	Jacobi-Matrix
$\mathbf{k} = [k_x \ k_y]^T$	Ortsfrequenz
k_0	Modulationsortsfrequenz der Gaborfilter
M	Anzahl der unterschiedlichen Filterorientierungen
M_s	Anzahl der zur Korrespondenzbestimmung herangezogenen Filterorientierungen
N	Anzahl der Ortsfrequenzbänder

\mathbf{p}	Kameraparametervektor
P	Anzahl der am Wettbewerbsprozeß beteiligten Variablen
Q	Anzahl der Variablen im verstärkend wirkenden Umfeld
$r_{mn}(\mathbf{x}), R_{mn}(\mathbf{k})$	Filterantwort, Fouriertransformierte der Filterantwort
$\tilde{r}_{mn}(\mathbf{x})$	komprimierte Filterantwort
$\operatorname{Re}\{f(x)\}$	Realteil von $f(x)$
RMS	Quadratwurzel des mittleren quadratischen Fehlers
\mathbf{R}	Rotationsmatrix
s_n	Abtastfaktor
\mathbf{S}	Parametermatrix
\mathbf{t}	Translationsvektor
U	verstärkend wirkendes Umfeld im Selbstorganisationsprozeß
\mathbf{v}	Meßfehlervektor
$w(\mathbf{x})$	Fensterfunktion
\mathbf{W}	Gewichtsmatrix
$\mathbf{x} = [x \ y]^T$	rektifizierte Bildkoordinaten
$\mathbf{x}_0 = [x_0 \ y_0]^T$	Hauptpunktkoordinaten
$\mathbf{x}_c = [x_c \ y_c]^T$	zyklopische Bildkoordinaten
$\mathbf{x}_d = [x_d \ y_d]^T$	verzerrte Bildkoordinaten
$\mathbf{x}_k = [x_k \ y_k \ z_k]^T$	Kamerakoordinaten
\mathbf{x}_m	Meßpunktkoordinatenvektor
\mathbf{x}'_m	fehlerbehafteter Meßpunktkoordinatenvektor
$\mathbf{x}_s = [x_s \ y_s]^T$	Sensorebenenkoordinaten
$\mathbf{x}_t = [x_t \ y_t]^T$	Bildkoordinaten des Distorsionszentrums
$\mathbf{x}_u = [x_u \ y_u]^T$	unverzerrte Bildkoordinaten
$\mathbf{x}_w = [x_w \ y_w \ z_w]^T$	Weltkoordinaten
β, β_v	Vergenzwinkel
$\delta(\mathbf{x})$	Dirac-Impuls
$\Delta b, \Delta b_{okt}$	relative Bandbreite
Δd_x	Disparitätskorrektur in x -Richtung
$\Delta g(\mathbf{x})$	Marr-Hildreth Operator
$\Delta\varphi$	Winkelauflösung der Filterbank

Δk	absolute Bandbreite
Δt	Schrittweite der numerischen Integration
$\phi(\mathbf{x})$	Phasenwinkel
φ_m	Filterorientierung
ϑ	Korrekturwinkel zur Kompensation der Projektionszentrumsdrift
κ	Distorsionskoeffizient
$\lambda_k, \lambda(\mathbf{x}, \mathbf{d})$	individuelle Gewichtungparameter des exponentiellen Wachstums der Variablen im Selbstorganisationsprozeß
μ_m	Gewichtung der zur Initialisierung des Selbstorganisationsprozesses verwendeten Filterorientierungen
$\rho_{mn}(\mathbf{x}, \mathbf{d})$	lokales Ähnlichkeitsmaß
$\sigma, \sigma_x, \sigma_y$	Streuungsparameter einer Gaußfunktion
Σ	Kovarianzmatrix
τ	radiale Linsendistorsion
$\xi_k, \xi(\mathbf{x}, \mathbf{d}, t)$	Ordnungsparameter, Variablen des Selbstorganisationsprozesses
$\xi_s(\mathbf{x}, \mathbf{d})$	stationäre Lösungen des Selbstorganisationsprozesses
∇d	Disparitätsgradient
∇d_g	Disparitätsgradientengrenze
$ x $	Betrag von x

9 Literaturverzeichnis

- AKM95 Aach, T.; Kaup, A.; Mester, R.: „On texture analysis: Local energy transforms versus quadrature filters.“ In *Signal Processing*, vol. 45, pp. 173-181, 1995.
- AA92 Abbot, A. L.; Ahuja, N.: „The University of Illinois Active Vision System.“ In *Proceedings of the SPIE, Intelligent Robots and Computer Vision XI*, vol. 1825, pp. 757-768, 1992.
- AK89 Alberts, J.; Kreiling, W.: „Photogrammetrisches Taschenbuch.“ Herbert Wichmann Verlag, Karlsruhe, 1989.
- AWB87 Aloimonons, J. Y.; Weiss, I.; Bandopadhyay, A.: „Active vision.“ In *Proceedings of the IEEE 1st International Conference on Computer Vision*, pp. 35-54, 1987.
- AP79 Andrews, B. A.; Pollen, D. A.: „Relationship between spatial frequency selectivity and receptive field profile of simple cells.“ In *Journal of Physiology*, vol. 287, pp. 163-176, 1979.
- AF87 Ayache, N.; Faverjon, B.: „Efficient Registration of Stereo Images by Matching Graph Descriptions of Edge Segments.“ In *International Journal of Computer Vision*, pp. 107-131, 1987.
- AH88 Ayache, N.; Hansen, C.: „Rectification of Images for Binocular and Trinocular Stereovision.“ In *Proceedings of the 9th International Conference on Pattern Recognition*, pp. 11-16, Rom, October 1988.
- AL87 Ayache, N.; Lustman, F.: „Fast and Reliable Passive Trinokular Stereovision.“ In *Proceedings of the 1st International Conference on Computer Vision*, pp. 422-427, June 1987.
- Baj85 Bajcsy, R.: „Active perception vs. passive perception.“ In *Proceedings of the 3rd IEEE Workshop on Computer Vision*, pp. 55-59, 1985.
- BB81 Baker, H. H.; Binford, T. O.: „Depth from Edge and Intensity-based Stereo.“ In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, pp. 631-636, August 1981.
- Bar89 Barnard, S. T.: „Stochastic stereo matching over scale.“ In *International Journal of Computer Vision*, vol. 2, pp. 17-32, 1989.
- BT80 Barnard, S. T.; Thompson, W. B.: „Disparity Analysis of Images.“ In *Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-2, no. 4, pp. 333-340, July 1980.

- BPT88 Bertero, M.; Poggio, T. A.; Torre, V.: „Ill-Posed Problems in Early Vision.“ In Proceedings of the IEEE, vol. 76, no. 8, pp. 869-889, August 1988.
- BCH73 Bishop, P. O.; Coombs, J. S.; Henry, G. H.: „Receptive fields of simple cells in the cat striate cortex.“ In Journal of Physiology, vol. 231, pp. 31-60, London, 1973.
- Boa92 Boashash, B.: „Estimating and interpreting the instantaneous frequency of a signal.“ In Proceedings of the IEEE, vol. 80, pp. 520-568, 1992.
- BB85 Bolles, R. C.; Baker, H. H.: „Epipolar-Plane Image Analysis: A Technique for Analyzing Motion Sequences.“ In Proceedings of the 3rd International Symposium of Robotics Research, MIT Press, Cambridge, pp. 41-48, 1985.
- Bur84 Burt, P. J.: „The pyramid as a structure for efficient computation.“ In Rosenfeld, A. (ed.): Multiresolution image processing and analysis. Springer Series in Information Sciences, vol. 12, Berlin, 1984.
- BA83 Burt, P. J.; Adelson, E. H.: „The Laplacian pyramid as a compact image code.“ In IEEE Transactions on Communication, vol. 31, pp. 532-540, 1983.
- BJ80 Burt, P.; Julesz, B.: „A Disparity Gradient Limit for Binocular Fusion.“ In Science, vol. 208, pp. 615-617, 1980.
- CR68 Campbell, F. W.; Robson, J. G.: „Applications of Fourier analysis to the visibility of gratings.“ In Journal of Physiology, vol. 197, pp. 551-556, London, 1968.
- CT90 Caprile, B; Torre, V.: „Using vanishing points for camera calibration.“ In International Journal of Computer Vision, vol. 4, pp. 127-140, April 1990.
- CR93 Chatterjee, C.; Roychowdhury, V.: „Robust camera calibration under complete lens distortion.“ In Proceedings of the SPIE: Applications of Digital Image Processing XVI, vol. 2028, pp. 322-333, 1993.
- CC92 Chang, C.; Chatterjee, S.: „A Deterministic Approach for Stereo Disparity Calculation.“ In G. Sandini (ed.): Proceedings of the 2nd European Conference on Computer Vision, Springer Verlag, Berlin, pp. 420-424, 1992.
- Chr92 Christensen, H. L.: „The AUC Robot Camera Head.“ In Proc. SPIE, Applications of Artificial Intelligence X: Machine Vision and Robotics, vol. 1708, pp. 26-33, 1992.
- CM92 Cochran, S. D.; Medioni, G.: „3-D Surface Description from Binocular Stereo.“ In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, no. 10, pp. 981-994, October 1992.

-
- CL55 Coddington, E. A.; Levinson, N.: „Theory of ordinary differential equations.“ McCraw-Hill, New York, 1955.
- CB91 Coombs, D. J.; Brown, C. M.: „Cooperative Gaze Holding in Binocular Vision.“ In IEEE Control Systems, June 1991.
- CB93 Coombs, D. J.; Brown, C. M.: „Real-Time Binocular Smooth Pursuit.“ In International Journal of Computer Vision, vol. 11, no. 2, 1993.
- Dau85 Daugman, J. G.: „Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters.“ In Journal of the Optical Society of America A, vol. 2, no. 7, pp. 1160-1169, July 1985.
- DDY79 DeValois, K. K.; DeValois, R. L.; Yund, E. W.: „Responses of striate cortical cells to grating and checkerboard patterns.“ In Journal of Physiology, vol. 291, pp. 483-505, London, 1979.
- DA89 Dhond, U. R.; Aggarwal, J. K.: „Structure from Stereo - A Review.“ In IEEE Transactions on Systems, Man, and Cybernetics, vol. 19, no. 6, pp. 1489-1510, November/December 1989.
- DP86 Drumheller, M.; Poggio, T.: „On Parallel Stereo.“ In Proceedings of the IEEE International Conference on Robotics and Automation, pp. 1439-1448, 1986.
- DH73 Duda, R. O.; Hart, P. E.: „Pattern Classification and Scene Analysis.“ Wiley, New York, 1973
- Ech90 Echigo, T.: „A camera calibration technique using three sets of parallel lines.“ In Machine Vision and Applications, vol. 3, pp. 159-167, March 1990.
- EY85 Echigo, T.; Yachida, M.: „A Fast Method for Extraction of 3-D Information using Multiple Stripes and two Cameras.“ In Proceedings of the 9th International Joint Conference on Artificial Intelligence, pp. 1127-1130, 1985.
- ERS77 Elster, K.-H.; Reinhardt, R.; Schäuble, M.: „Einführung in die nichtlineare Optimierung.“ B. G. Teubner Verlagsgesellschaft, Leipzig, 1977.
- FH84 Fang, J.-Q.; Huang, T. S.: „Solving Three-Dimensional Small-Rotation Motion Equations: Uniqueness, Algorithms, and Numerical Results.“ In Computer Vision, Graphics, and Image Processing, vol. 26, pp. 183-206, 1984.
- Fau93 Faugeras, O.: „Three-Dimensional Computer Vision.“ MIT Press, Cambridge, 1993.

- FT86 Faugeras O. D.; Toscani, G.: „Calibration Problem for Stereo.“ In Proceedings of the International Conference on Computer Vision and Pattern Recognition, pp. 15-20, June 1986.
- Fer92 Ferrier, N. J.: „The Harvard Binocular Head.“ In Proceedings of the SPIE, Applications of Artificial Intelligence X: Machine Vision and Robotics, vol. 1708, pp. 2-13, 1992.
- Fle94 Fleet, D. J.: „Disparity from Local Weighted Phase-Correlation.“ In IEEE International Conference on Systems, Man and Cybernetics, pp. 48-56, San Antonio, October 1994.
- FJ93 Fleet, D. J.; Jepson, A. D.: „Stability of Phase Information.“ In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 12, pp. 1253-1268, December 1993.
- FJJ91 Fleet, D. J.; Jepson, A. D.; Jenkin, M. R. M.: „Phase-Based Disparity Measurement.“ In Computer Vision, Graphics and Image Processing: Image Understanding, vol. 53, no. 2, pp. 198-210, March 1991.
- Gab46 Gabor, D.: „Theory of communication.“ In Journal of the IEE, vol. 93, pp. 429-457, London, 1946.
- GLY92 Geiger, D.; Ladendorf, B.; Yuille, A.: „Occlusions and Binocular Stereo.“ In G. Sandini (ed.): Proceedings of the 3rd European Conference on Computer Vision, Springer Verlag, Berlin, Heidelberg, New York, pp. 425-433, 1992.
- GG84 Geman, S.; Geman, D.: „Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images.“ In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-6, no. 6, pp. 721-741, November 1984.
- Gen88 Gennert, M. A.: „Brightness-Based Stereo Matching.“ In International Conference on Computer Vision, pp. 139-143, 1988.
- Gen79 Gennery, D. B.: „Stereo-Camera Calibration.“ In Proceedings of the DARPA Image Understanding Workshop, pp. 101-108, November 1979.
- GS+98 Götz, N.; Stemmer, R.; Trapp, R.; Drüe, S.; Hartmann, G.: „Ein aktives Stereokamerasystem zur Steuerung eines Demontageroboters.“ In Posch, S.; Ritter, H. (eds.): Proceedings in Artificial Intelligence, vol. 8, pp. 1-8, Sankt Augustin (Infix), June 1998.
- GW93 Gonzalez, R. C.; Woods, R. E.: „Digital Image Processing.“ Addison-Wesley Publishing Company, New York, 1993.

-
- GG93 Goshtasby, A.; Gruver, W. A.: „Design of a single-lens stereo camera system.“ In *Pattern Recognition*, vol. 26, no. 6, pp. 923-937, 1993.
- Gri81a Grimson, W. E. L.: „A computer implementation of a theory of human stereo vision.“ In *Philosophical Transactions of the Royal Society of London*, vol. B 292, pp. 217-253, 1981.
- Gri81b Grimson, W. E. L.: „From Images to Surfaces: A Computational Study of the Human Early Visual System.“ MIT Press, Cambridge, London, 1981.
- Gri85 Grimson, W. E. L.: „Computational experiments with a feature based stereo algorithm.“ In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-7, no. 1, pp. 17-34, January 1985.
- GST89 Grosso, E.; Sandini, G.; Tistarelli, M.: „3-D Object Reconstruction Using Stereo and Motion.“ In *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1465-1476, 1989.
- Hak79 Haken, H.: „Pattern Formation and Pattern Recognition - An Attempt at a Synthesis.“ In Haken, H. (ed.), *Pattern Formation by Dynamical Systems and Pattern Recognition*, Springer-Verlag, Berlin, 1979.
- Hak81 Haken, H.: „Erfolgsgeheimnisse der Natur.“ Deutsche Verlagsanstalt, Stuttgart, 1981.
- Hak82 Haken, H.: „Synergetik.“ Springer Verlag, Berlin, 1982.
- Hak87 Haken, H.: „Synergetic Computers for Pattern Recognition and Associative Memory.“ In Haken, H. (ed.): *Computational Systems - Natural and Artificial*, Springer-Verlag, Berlin, 1987.
- Hak91 Haken, H.: „Synergetic Computers and Cognition.“ Springer-Verlag, Berlin, 1991.
- HAL90 Hansen, C.; Ayache, N.; Lustman, F.: „High Speed Trinocular Stereo for Mobile-Robot Navigation.“ In Tou, J.; Balchen, J. (eds.): *Highly Redundant Sensing in Robot Systems*, Springer-Verlag, Berlin, Heidelberg, New York, pp. 127-146, 1990.
- HA89 Hoff, W.; Ahuja, N.: „Surfaces from Stereo: Integrating Feature Matching, Disparity Estimation, and Contour Detection.“ In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 2, February 1989.

- HSY96 Hongo, S.; Sonehara, N.; Yoroizawa, I.: „Edge-based Binocular Stereopsis Algorithm - A Matching Mechanism with Probabilistic Feedback.“ In *Neural Networks*, vol. 9, no. 3, pp. 379-395, 1996.
- HS89 Horaud, R.; Skordas, T.: „Stereo Correspondence Through Feature Grouping and Maximal Cliques.“ In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 11, pp. 1168-1180, November 1989.
- Hor86 Horn, B. K. P.: „Robot Vision.“ MIT Press, Cambridge, 1986.
- Hor79 Horst, R.: „Nichtlineare Optimierung.“ Carl Hanser Verlag, München, Wien, 1979.
- HW62 Hubel, D. H.; Wiesel, T. N.: „Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.“ In *Journal of Physiology*, vol. 160, pp. 106-154, London, 1962.
- II86 Ito, M.; Ishii, A.: „Three-View Stereo Analysis.“ In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 4, July 1986.
- Jäh93 Jähne, B.: „Digitale Bildverarbeitung.“ Springer-Verlag, Berlin, 1993.
- JJ89 Jenkin, M. R. M.; Jepson, A. D.: „The Fast Computation of Disparity from Phase Differences.“ In *Proceedings of the IEEE Computer Vision and Pattern Recognition*, pp. 398-403, 1989.
- JJ94 Jenkin, M. R. M.; Jepson, A. D.: „Recovering Local Surface Structure through Local Phase Difference Measurements.“ In *Computer Vision, Graphics and Image Processing: Image Understanding*, vol. 59, no. 1, pp. 72-93, January, 1994.
- JJT91 Jenkin, M. R. M.; Jepson, A. D.; Tsotsos, J. K.: „Techniques for Disparity Measurement.“ In *Computer Vision, Graphics and Image Processing: Image Understanding*, vol. 53, no. 1, pp. 14-30, January, 1991.
- JMT92 Jenkin, M. R. M.; Milios, E.; Tsotsos, J. K.: „TRISH: A binocular robot head with torsional eye movements.“ In *Proceedings of the SPIE, Applications of Artificial Intelligence X: Machine Vision and Robotics*, vol. 1708, pp. 36-46, 1992.
- JT94 Jenkin, M. R. M.; Tsotsos, J. K.: „Active Stereo Vision and Cyclotorsion.“ In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 806-811, June 1994.

-
- JP87a Jones, J. P.; Palmer, L. A.: „The Two-Dimensional Spatial Structure of Simple Receptive Fields in Cat Striate Cortex.“ In *Journal of Neurophysiology*, vol. 58, no. 6, pp. 1187-1211, December 1987.
- JP87b Jones, J. P.; Palmer, L. A.: „An Evaluation of the Two-Dimensional Gabor Filter Model of Simple Receptive Fields in Cat Striate Cortex.“ In *Journal of Neurophysiology*, vol. 58, no. 6, pp. 1233-1258, December 1987.
- JM92 Jones, D. G.; Malik, J.: „A computational framework for determining stereo correspondence from a set of linear spatial filters.“ In *Proceedings of the 2nd European Conference on Computer Vision*, Springer Verlag, Berlin, pp. 395-410, 1992.
- JSP87 Jones, J. P.; Stepnoski, A.; Palmer, L. A.: „The Two-Dimensional Spectral Structure of Simple Receptive Fields in Cat Striate Cortex.“ In *Journal of Neurophysiology*, vol. 58, no. 6, pp. 1212-1232, December 1987.
- Jul60 Julesz, B.: „Binokular depth perception of computer generated patterns.“ In *Bell System Technical Journal*, vol. 39, pp. 1125-1162, 1960.
- Jul71 Julesz, B.: „Foundations of cyclopean perception.“ University of Chicago Press, 1971.
- KA87 Kim, Y. C.; Aggarwal, J. K.: „Positioning Three-Dimensional Objects Using Stereo Images.“ In *IEEE Journal of Robotics and Automation*, vol. RA-3, no. 4, pp. 361-373, August 1987.
- KGV83 Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P.: „Optimization by Simulated Annealing.“ In *Science*, vol. 220, pp. 671-680, May 1983.
- KD88 Konrad, J.; Dubois, E.: „Multigrid Bayesian Estimation of Image Motion Fields Using Stochastic Relaxation.“ In *Proceedings of the 2nd International Conference on Computer Vision*, pp. 354-362, 1988.
- Kra84 Kraus, K.: „Photogrammetrie: Theorie und Praxis der Auswertesysteme.“ Ferd. Dümmler Verlag, Bonn, 1984.
- KH75 Kuglin, C.; Hines, D.: „The phase correlation image alignment method.“ In *Proceedings of the IEEE International Conference on Cybernetics and Society*, pp. 163-165, 1975.
- KB81 Kulikowski, J. J.; Bishop, P. O.: „Fourier analysis and spatial representation in the visual cortex.“ In *Experientia*, vol. 37, pp. 160-163, 1981.

- KMB82 Kulikowski, J. J.; Marcelja, S.; Bishop, P. O.: „Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex.“ In *Biological Cybernetics*, vol 43, pp. 187-198, 1982.
- LZB95 Lanser, S.; Zierl, C.; Beutlhauser, R.: „Multibildkalibrierung einer CCD-Kamera.“ In Sagerer, G.; Posch, S.; Kummert, F. (Hrsg.), *Mustererkennung 1995. Informatik aktuell*, Springer Verlag, Berlin, S. 481-491, 1995.
- Len87 Lenz, Reimar: „Linsenfehler korrigierte Eichung von Halbleiterkameras mit Standardobjektiven für hochgenaue 3D-Messungen in Echtzeit.“ In *Mustererkennung 1987*, Springer-Verlag, S. 212-216, Berlin, 1987.
- LT87 Lenz, R. K.; Tsai, R. Y.: „Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology.“ In *Proceedings of the IEEE International Conference on Robotics and Automation*, Raleigh, NC, March 1987.
- Li94 Li, M.: „Camera Calibration of the KTH Head-Eye System.“ Report from Computational Vision and Active Perception Laboratory, CVAP147, March 1994.
- LH96 Li, Z.; Hu, G.: „Analysis of Disparity Gradient Based Cooperative Stereo.“ In *IEEE Transactions on Image Processing*, vol. 5, no. 11, pp. 1493-1506, November 1996.
- LB87 Lim, H. S.; Binford, T. O.: „Stereo Correspondence: A Hierarchical Approach.“ In *Proceedings of the DARPA Image Understanding Workshop*, pp. 234-241, February 1987.
- Luk98 Lukaschek, M.: „Kalibrierung eines binokularen Kamerasystems unter Berücksichtigung der Linsendistorsion.“ Diplomarbeit, Fachgebiet Grundlagen der Elektrotechnik, Universität-GH Paderborn, April 1998.
- LB95 Luo, A.; Burkhardt, H.: „An Intensity-Based Cooperative Bidirectional Stereo Matching with Simultaneous Detection of Discontinuities and Occlusions.“ In *International Journal of Computer Vision*, vol. 15, pp. 171-188, 1995.
- MF77 Maffei, L.; Fiorentini, A.: „Spatial frequency rows in the striate visual cortex.“ In *Vision Research*, vol. 17, pp. 257-264, 1977.
- MBE95 Maki, A.; Bretzner, L.; Eklundh, J.: „Local Fourier Phase and Disparity Estimates: An Analytical Study.“ In Hlavac, Sara (eds.): *Proceedings of the 6th International Conference on Computer Analysis of Images and Patterns*, Springer-Verlag, Berlin, Heidelberg, New York, pp. 868-873, September 1995.

-
- Mar80 Marcelja, S.: „Mathematical description of the responses of simple cortical cells.“ In *Journal of the Optical Society of America*, vol. 70, no. 11, pp. 1297-1300, November 1980.
- Mar82 Marr, D.: „Vision“ Freeman, San Francisco, 1982.
- MH80 Marr, D.; Hildreth, E.: „Theory of edge detection.“ In *Proceedings of the Royal Society of London*, vol. B 207, pp. 187-217, 1980.
- MPP78 Marr, D.; Palm, G.; Poggio, T.: „Analysis of a Cooperative Stereo Algorithm.“ In *Biological Cybernetics*, vol. 28, pp. 223-239, 1978.
- MP76 Marr, D.; Poggio, T.: „Cooperative Computation of Stereo Disparity.“ In *Science*, vol. 194, pp. 283-287, 1976.
- MP79 Marr, D.; Poggio, T.: „A computational theory of human stereo vision.“ In *Proceedings of Royal Society of London*, vol. B 204, pp. 301-328, 1979.
- MMP87 Marroquin, J.; Mitter, S.; Poggio, T.: „Probabilistic Solution of Ill-Posed Problems in Computational Vision.“ In *Journal of the American Statistical Association*, vol. 32, no. 397, pp. 76-89, March 1987.
- MC+93 Masmoudi, L.; Coronado, J. L.; Trespaderne, F. M.; Peran, J. R.: „Stereo camera calibration with distortion model and using forgetting factor.“ In *Proceedings of the SPIE: Videometrics II*, vol. 2067, pp. 14-20, 1993.
- MF81 Mayhew, J. E. W.; Frisby, J. P.: „Psychophysical and Computational Studies towards a Theory of Human Stereopsis.“ In *Artificial Intelligence*, vol. 17, pp. 349-385, 1981.
- MN85 Medioni, G.; Nevatia, R.: „Segment-Based Stereo Matching.“ In *Computer Vision, Graphics, and Image Processing*, vol. 31, pp. 2-18, 1985.
- MMN89 Mohan, R.; Medioni, G.; Nevatia, R.: „Stereo Error Detection, Correction, and Evaluation.“ In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 2, pp. 113-120, February 1989.
- Mor77 Moravec, H. P.: „Towards automatic visual obstacle avoidance.“ In *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, p. 584, 1977.
- Mor93 Morozov, V. A.: „Regularization Methods for Ill-Posed Problems.“ CRC Press, London, 1993.

- MTT78 Movshon, J. A.; Thompson, I. D.; Tolhurst, D. J.: „Spatial and temporal contrast sensitivity of neurons in area 17 and 18 of the cat’s visual cortex.“ In *Journal of Physiology*, vol. 283, pp. 101-120, 1978.
- NM+92 Nomura, Y.; Michihiro, S.; Naruse, H.; Ide, A.: „Simple Calibration Algorithm for High-Distorsion-Lens Camera.“ In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-14, no. 11, pp. 1095-1099, November 1992.
- Ols92 Olsen, S. I.: „Epipolar Line Estimation.“ In G. Sandini (ed.), *Proceedings of the 2nd European Conference on Computer Vision*, Springer Verlag, Berlin, pp. 307-311, 1992.
- OP89 Olson, J. T.; Potter, R. D.: „Real-Time Vergence Control.“ In *Proceedings of the IEEE Computer Vision and Pattern Recognition*, pp. 404-409, 1989.
- OK85 Otha, Y.; Kanade T.: „Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming.“ In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-7, no. 2, pp. 139-154, March 1985.
- PE92 Pahlavan, K.; Eklundh, J.-O.: „A Head-Eye System - Analysis and Design.“ In *Computer Vision Graphics Image Processing: Image Understanding*, vol. 56, no. 1, pp. 41-56, 1992.
- PD96 Papadimitriou, D. V.; Dennis, T. J.: „Epipolar Line Estimation and Rectification for Stereo Image Pairs.“ In *IEEE Transactions on Image Processing*, vol. 5, no. 4, April 1996.
- Pap65 Papoulis, A.: „Probability, Random Variables, and Stochastic Processes.“ McGraw-Hill, New York, 1965.
- PH86 Peitikäinen, M.; Harwood, D.: „Depth from Three Camera Stereo.“ In *Proceedings of the IEEE CS Conference on Pattern Recognition*, pp. 2-8, June 1986.
- Pla72 Playboy Magazine: „Playboy’s Playmate of the Month - Miss November.“ Playboy Enterprises, Inc., vol. 11/72, Centerfold, November 1972.
- PTK85 Poggio, T. A.; Torre, V.; Koch, C.: „Computational Vision and Regularization Theory.“ In *Nature*, vol. 317, pp. 314-319, 1985.
- PMF85 Pollard, S. B.; Mayhew, J. E. W.; Frisby, J. P.: „PMF: A stereo correspondence algorithm using a disparity gradient constraint.“ In *Perception*, vol. 14, pp. 449-470, 1985.

-
- PP+85 Pollard, S. B.; Porill, J. Mayhew, J. E. W.; Frisby, J. P.: „Disparity gradient, Lipschitz continuity and computing binocular correspondences.“ In Proceedings of the 3rd International Symposium of Robotics Research, MIT Press, Cambridge, pp. 19-26, 1985.
- PLT71 Pollen, D. A.; Lee, J. R.; Taylor, J. H.: „How does the striate cortex begin the reconstruction of the visual world?“ In Science, vol. 173, pp. 74-77, New York, 1971.
- PR81 Pollen, D. A.; Ronner, S.F.: „Phase relationships between adjacent cells in the visual cortex.“ In Science, vol. 212, pp. 1409-1411, 1981.
- PA82 Posdamer, J. L.; Altschuler, M. D.: „Surface Measurement by Space-encoded Projected Beam Systems.“ In Computer Graphics and Image Processing, vol. 18, pp. 1-17, 1982.
- Pra85 Prazdny, K.: „Detection of Binocular Disparities.“ In Biological Cybernetics, vol. 52, pp. 93-99, 1985.
- RBS87 Raju, G. V. S.; Binford, T. O.; Shekhar, S.: „Stereo matching using Viterbi algorithm.“ In Proceedings of the DARPA Image Understanding Workshop, pp. 766-776, February 1987.
- Rei95 Reimann, D.: „Anwendung der Synergetik auf Korrespondenzprobleme wie die Stereoskopie.“ Dissertation, Universität Stuttgart, 1995.
- RH94 Reimann, D.; Haken, H.: „Stereovision by Self-Organisation.“ In Biological Cybernetics, vol. 71, pp. 17-26, 1994.
- RHZ76 Rosenfeld, A.; Hummel, R. A.; Zucker, S. W.: „Scene Labeling by Relaxation Operations.“ In IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-6, no. 6, pp. 420-433, June 1976.
- Sab96 Sabatini, S. P.: „Recurrent inhibition and clustered connectivity as a basis for Gabor-like receptive fields in the visual cortex.“ In Biological Cybernetics, vol. 74, pp. 189-202, 1996.
- San88 Sanger, T. D.: „Stereo Disparity Computation Using Gabor Filters.“ In Biological Cybernetics, vol. 59, pp. 405-418, 1988.
- ST97 Schneider, B.; Trapp, R.: „Untersuchung von Quantisierungsfehlern in der Stereo-Triangulation.“ Interner Bericht Nr.: NAVIS 01/97, Universität-GH Paderborn, Fachgebiet Grundlagen der Elektrotechnik, März 1997.

- SA94 Shah, S.; Aggarwal, J. K.: „A Simple Calibration Procedure for Fish-Eye (High Distorsion) Lens Camera.“ In IEEE International Conference on Robotics and Automation, no. 4, pp. 3422-3427, 1994.
- SP96 Shen, J.; Paillou, P.: „Trinocular Stereovision by Generalized Hough Transform.“ In Pattern Recognition, vol. 29, no. 10, pp. 1661-1672, 1996.
- Shi87 Shirai, Y.: „Three-Dimensional Computer Vision.“ Springer Verlag, Berlin, 1987.
- SH97 Stöhr, M.; Hartmann, G.: „Orientierungs- und skalierungsinvariante Erkennung von Objekten in komplexen Szenen.“ In Paulus, E.; Wahl, F. M. (Hrsg.): Mustererkennung 1997, Springer-Verlag, Berlin, pp. 71-78, 1997.
- Ter86 Terzopoulos, D.: „Regularization of Inverse Visual Problems Involving Discontinuities.“ In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-8, no. 4, pp. 413-424, July 1986.
- TWK87 Terzopoulos, D.; Witkin, A.; Kass, M.: „Stereo Matching as Constrained Optimization Using Scale Continuation Methods.“ In SPIE Optical and Digital Pattern Recognition, vol. 754, pp. 92-99, 1987.
- TM94 Theimer, W. M.; Mallot, H. A.: „Phase-Based Binocular Vergence Control and Depth Reconstruction Using Active Vision.“ In Computer Vision, Graphics and Image Processing: Image Understanding, vol. 60, no. 3, pp. 343-358, November 1994.
- TA77 Tikhonov, A. N.; Asenin, V. Y.: „Solutions of ill-Posed Problems.“ Winston & Sons, Washington DC, 1977.
- Tra96a Trapp, R.: „Entwurf einer Filterbank auf der Basis neurophysiologischer Erkenntnisse zur orientierungs- und frequenzspezifischen Dekomposition von Bilddaten.“ Interner Bericht Nr.: NAVIS 01/96, Universität-GH Paderborn, Fachgebiet Grundlagen der Elektrotechnik, Februar 1996.
- Tra96b Trapp, R.: „Ein flexibles binokulares Sehsystem: Konstruktion und Kalibrierung.“ Interner Bericht Nr.: NAVIS 02/96, Universität-GH Paderborn, Fachgebiet Grundlagen der Elektrotechnik, Juli 1996.
- Tra96c Trapp, R.: „Ein Verfahren zur Rektifikation von Stereobilddaten.“ Interner Bericht Nr.: NAVIS 03/96, Universität-GH Paderborn, Fachgebiet Grundlagen der Elektrotechnik, August 1996.

-
- Tra96d Trapp, R.: „Disparitätsmessung durch lokale Korrelation von orientierungs- und frequenzselektiv gefilterten Bilddaten.“ Interner Bericht Nr.: NAVIS 04/96, Universität-GH Paderborn, Fachgebiet Grundlagen der Elektrotechnik, Dezember 1996.
- Tra97 Trapp, R.: „Stereoskopische Korrespondenzbestimmung mit impliziter Okklusionsdetektion.“ Interner Bericht Nr.: NAVIS 02/97, Universität-GH Paderborn, Fachgebiet Grundlagen der Elektrotechnik, Juni 1997.
- TD96 Trapp, R.; Drüe, S.: „Ein flexibles binokulares Sehsystem: Konstruktion und Kalibrierung.“ In Mertsching, B. (ed.): Proceedings in Artificial Intelligence, vol.4, pp. 32-39, Sankt Augustin (Infix), December 1996.
- TDH98 Trapp, R.; Drüe, S.; Hartmann, G.: „Stereo Matching with Implicit Detection of Occlusions.“ In Burkhardt, H.; Neumann, B. (eds.): Proceedings of the 5th European Conference on Computer Vision, vol. II, pp. 17-33, Springer Verlag, Berlin, June 1998.
- Tsa87 Tsai, R. Y.: „A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses.“ In IEEE Journal of Robotics and Automation, vol. RA-3, no. 4, pp. 323-344, August 1987.
- Var62 Varga, R. S.: „Matrix Iterative Analyse.“ G. Forsythe (ed.), Prentice-Hall Series in Automatic Computation, New Jersey, 1962.
- VT95 Völpel, B; Theimer, W. M.: „Localization Uncertainty in Area-Based Stereo Algorithms.“ In IEEE Transactions on Systems, Man, and Cybernetics, vol. 25, no. 12, pp. 1628-1634, December 1995.
- Wah84 Wahl, F. M.: „Digitale Bildsignalverarbeitung.“ Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1984.
- WAH88 Weng, J.; Ahuja, N.; Huang, T. S.: „Two-view matching.“ In Proceedings of the 2nd International Conference on Computer Vision, pp. 64-73, 1988.
- WCH92 Weng, J.; Cohen, P.; Herniou, M.: „Camera Calibration with Distortion Models and Accuracy Evaluation.“ In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, no. 10, pp. 965-980, Oktober 1992.
- Wer92 Werner, J.: „Numerische Mathematik 2.“ Vieweg-Verlag, Braunschweig, Wiesbaden, 1992.

- WK+95 Westelius, C.; Knutson, H.; Wiklund, J.; Westin, C.: „Phase-based Disparity Estimation.“ In Crowley, L.; Christensen, H. I. (eds.): *Vision as Process*, Springer-Verlag, pp. 157-178, 1995.
- YC78 Yakimovsky, Y. Cunningham, R.: „A System for Extracting Three-Dimensional Measurements from a Stereo Pair of TV Cameras.“ In *Computer Graphics and Image Processing*, vol. 7, pp. 195-210, 1978.
- YZ88 Yan, L.; Zuxun, Z.: „Fast Implementation for Generating Epipolar Line Images with One-dimensional Resampling.“ In *Proceedings of the 16th Congress of the International Society for Photogrammetry and Remote Sensing*, pp. 511-520, Kyoto, 1988.
- YS89 Yeshurun, Y.; Schwartz, E.: „Cepstral filtering on a columnar image architecture: A fast algorithm for binocular stereo segmentation.“ In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 759-767, vol. 11, no. 7, July 1989.
- YGB91 Yuille, A.; Geiger, D.; Bülthoff, H. H.: „Stereo integration, mean field theory and psychophysics.“ In *Network*, vol. 2, pp. 423-442, 1991.

Heinz Nixdorf Institut
Interdisziplinäres Forschungszentrum für Informatik und Technik

Von Heinz Nixdorf ging die Initiative aus, das interdisziplinäre Forschungszentrum für Informatik und Technik (ZTI) zu bilden. Es ist als zentrale wissenschaftliche Einrichtung der Universität - Gesamthochschule Paderborn konzipiert. Die konstituierende Sitzung des Gründungsvorstandes fand am 23. Oktober 1987 statt. Insgesamt stellen die von Heinz Nixdorf gegründete Stiftung Westfalen, das Land Nordrhein-Westfalen und der Bund ca. 160 Millionen Mark für Einrichtung und Betrieb des Instituts zur Verfügung. Die auf 20 Jahre verteilten Mittel sind für den Ausbau der interdisziplinären Forschung und Ausbildung in Informatik, Technik und Betriebswirtschaft sowie die Einrichtung eines Graduiertenkollegs bestimmt.

Das Heinz Nixdorf Institut hat die Forschung auf anwendungsnahen Gebieten der Informatik, Technik und Wirtschaftswissenschaften als Aufgabe. Hierbei werden die einschlägigen naturwissenschaftlichen Grundlagen sowie das geistes- und sozialwissenschaftlich thematisierte Umfeld berücksichtigt. Besonderes Gewicht wird auf die wechselseitige Durchdringung traditioneller Ingenieurdisziplinen mit Verfahren der Informatik und Informationstechnik gelegt. Eine derzeitige Schwerpunktbildung kommt durch das Leitprojekt „Parallele Rechnernetzwerke in der Produktionstechnik“ zum Ausdruck, das die beiden Hauptaufgabengebiete Paralleles Rechnen und Informationsverarbeitung in der industriellen Produktion zusammen führt.

Heinz Nixdorf Institute
Interdisciplinary research centre for Computer Science and Technology

Heinz Nixdorf's initiative created an interdisciplinary research centre for Computer Science and Technology within the University of Paderborn. The constitutional meeting of the founding members of the board of trustees took place on October 23rd 19987..

The Westfahlen Foundation, founded by Heinz Nixdorf, together with the state of Nordrhein-Westfalia have placed at the disposal of the institute a total of 160 Million DM for equipment and running costs. The sum, to be spread over a piriod of 20 years, is earmarked for the development of interdisciplinary research and education in computer science, technology and economics, as well as for the establishment of a graduates' research group.

The Heinz Nixdorf Institute was established to carry out research in areas of computer science, technology and economics close to practical application. Included here is a research basis in the relevant science as well as in related areas of the humanities and social science. Particular emphasis is placed on the pervasive application of computer science and technology in traditional engineering disciplines. The present focus of research is reflected by the guideline project „Parallel Computer Networks in Production Technique“, joining the two major task fields parallel computing and information processing in industrial production.

Publikationsverzeichnis

- Bd. 1 Uta Fahrwinkel: „Methoden zur Modellierung und Analyse von Geschäftsprozessen zur Unterstützung des Business Process Reengineering“, Diss. 1995, Rechnerintegrierte Produktion, 100,00 DM, ISBN 3-931466-00-0
- Bd. 2 Dirk Hornbostel: „Methode zur Modellierung der Informationsverarbeitung in Industrieunternehmen“, Diss. 1995, Rechnerintegrierte Produktion, 100,00 DM, ISBN 3-931466-01-9
- Bd. 3 Volker Stemann: "Contention Resolution in Hashing Based Shared Memory Simulations", Diss. 1995, Theoretische Informatik, 16,50 DM, ISBN 3-931466-02-7
- Bd. 4 Norbert Ketterer: „Beschreibung von Datenaustausch eines verteilten Fertigungssteuerungssystems“, Diss. 1995, Wirtschaftsinformatik, insb. CIM, 34,00 DM, ISBN 3-931466-03-5
- Bd. 5 Thomas Hartmann: "Spezifikation und Klassifikation von Methoden zur Definition hierarchischer Abläufe", Diss. 1995, Wirtschaftsinformatik, insb. CIM, 28,50 DM, ISBN 3-931466-04-3
- Bd. 6 Alf Wachsmann: "Eine Bibliothek von Basisdiensten für Parallelrechner: Routing, Synchronisation, gemeinsamer Speicher", Diss. 1995, Theoretische Informatik, 17,50 DM, ISBN 3-931466-05-1
- Bd. 7 Jürgen Gausemeier (Hrsg.): "Die Szenariotechnik - Werkzeug für den Umgang mit einer multiplen Zukunft", 1995, Rechnerintegrierte Produktion, 88,00 DM, ISBN 3-931466-06-X
- Bd. 8 Artur Czumaj: "Parallel Algorithmic Techniques: PRAM Algorithms and PRAM Simulations", Diss. 1995, Theoretische Informatik, 23,50 DM, ISBN 3-931466-07-8
- Bd. 9 Axel Humpert: "Methodische Anforderungsverarbeitung auf Basis eines objektorientierten Anforderungsmodells", Diss. 1995, Rechnerintegrierte Produktion, 100,00 DM, ISBN 3-931466-08-6
- Bd. 10 Foued Ameer: "Space-Bounded Learning Algorithms", Diss. 1995, Theoretische Informatik, 20,00 DM, ISBN 3-931466-09-4
- Bd. 11 Matthias Paul: "Szenariobasiertes Konzipieren neuer Produkte des Maschinenbaus auf Grundlage möglicher zukünftiger Technologieentwicklungen", Diss. 1996, Rechnerintegrierte Produktion, 100,00 DM, ISBN 3-931466-10-8
- Bd. 12 Friedrich Holl: „Ordnungsmäßigkeit von Informations- und Kommunikationssystemen“, Diss. 1996, Informatik und Gesellschaft, 50,00 DM, ISBN 3-931466-11-6
- Bd. 13 Jürgen Gausemeier (Hrsg.): "First European Workshop on Global Engineering Networking - organized by GLENnet e.V.", 1996, Rechnerintegrierte Produktion, 88,00 DM, ISBN 3-931466-12-4
- Bd. 14 Kirsten Petri: "Vergleichende Untersuchung von Berechnungsmodellen zur Simulation der Dynamik von Fahrleitungs-Stromabnehmersystemen", Diss. 1996, Mechatronik und Dynamik, 100,00 DM, ISBN 3-931466-13-2
- Bd. 15 Stephan Leschka: "Fallbasiertes Störungsmanagement in flexiblen Fertigungssystemen", Diss. 1996, Rechnerintegrierte Produktion, 100,00 DM, ISBN 3-931466-14-0
- Bd. 16 Uta Schneider: "Ein formales Modell und eine Klassifikation für die Fertigungssteuerung - Ein Beitrag zur Systematisierung der Fertigungssteuerung", Diss. 1996, Wirtschaftsinformatik, insb. CIM, DM 32,00, ISBN 3-931466-15-9

Publikationsverzeichnis

- Bd. 17 Winfried Felser: "Eine Methode zur Erstellung von Fertigungssteuerungsverfahren aus Bausteinen", Diss. 1996, Wirtschaftsinformatik, insb. CIM, DM 24,00, ISBN 3-931466-16-7
- Bd. 18 Jürgen Gausemeier, Alexander Fink: „Neue Wege zur Produktentwicklung - Erfolgspotentiale der Zukunft“, 1996, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931466-17-5
- Bd. 19 Wilhelm Dangelmaier, Jürgen Gausemeier: „Fortgeschrittene Informationstechnologie in der Produktentwicklung und Fertigung“, 2. Internationales Heinz Nixdorf Symposium, 1996, DM 100,00, ISBN 3-931466-18-3
- Bd. 20 Eyke Hüllermeier: "Reasoning about Systems based on Incomplete and Uncertain Models", Diss. 1997, Wissensbasierte Systeme, DM 50,00, ISBN 3-931466-19-1
- Bd. 21 Jürgen Gausemeier: "International Symposium on Global Engineering Network - Altwerp, Belgium", 1997, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931466-20-5
- Bd. 22 Astrid Burger "Methode zum Nachweis der Wirtschaftlichkeit von Investitionen in die Rechnerintegrierte Produktion", Diss. 1997, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931466-21-3
- Bd. 23 Jürgen Gausemeier: "Entwicklung und Transfer von Entwicklungssystemen der Mechatronik - Paderborner Workshop TransMechatronik", 1997, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931466-22-1
- Bd. 24 Karl-Heinz Gerdes: "Architekturkonzeption für Fertigungsleitsysteme der flexiblen automatischen Fertigung", Diss. 1997, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931466-23-X
- Bd. 25 Peter Ebbesmeyer: "Dynamische Texturwende - Ein Verfahren zur echtzeitorientierten Bildgenerierung für Virtuelle Umgebungen technischer Objekte", Diss. 1997, Rechnerintegrierte Produktion, DM 100,00 DM, ISBN 3-931466-24-8
- Bd. 26 Gerald Frank: "Ein digitales Hardwaresystem zur echtzeitfähigen Simulation biologienaher neuronaler Netze", Diss. 1997, Grundlagen der Elektrotechnik, DM 100,00, ISBN 3-931466-25-6
- Bd. 27 Wolfgang Dittrich: "Communication and I/O Efficient Parallel Data Structures", Diss. 1997, Theoretische Informatik, DM 25,00, ISBN 3-931466-26-4
- Bd. 28 Armin Bäumker: "Communication Efficient Parallel Searching", Diss. 1997, Theoretische Informatik, DM 24,50, ISBN 3-931466-27-2
- Bd. 29 Christoph Pintaske: "System- und Schaltungstechnik neuronaler Assoziativspeicher", Diss. 1997, Schaltungstechnik, DM 75,00, ISBN 3-931466-28-0
- Bd. 30 Sören Henkel: "Ein System von Software-Entwurfsmustern für die Propagation von Ereignissen in Werkzeugen zur kooperativen Fabrikmodellierung", Diss. 1997, Wirtschaftsinformatik, insb. CIM, DM 40,00, ISBN 3-931466-29-9
- Bd. 31 Wilhelm Dangelmaier: "Vision Logistik - Logistik wandelbarer Produktionsnetze", 1997, Wirtschaftsinformatik, insb. CIM, DM 98,00, ISBN 3-931466-30-2

Publikationsverzeichnis

- Bd. 32 Dirk Brexel: "Methodische Strukturmodellierung komplexer und variatenreicher Produkte des integrativen Maschinenbaus", Diss. 1997, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931466-31-0
- Bd. 33 Axel Hahn: "Integrationsumgebung für verteilte objektorientierte Ingenieursysteme", Diss. 1997, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931466-32-9
- Bd. 34 Andreas Sabin: "Semantisches Modell zum Aufbau von Hilfsorientierungsdiensten in einem globalen Engineering Netzwerk", Diss. 1997, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931-466-3-7
- Bd. 35 Willy-B. Strothmann: "Bounded Degree Spanning Trees", Diss. 1997, Theoretische Informatik, DM 26,00, ISBN 3-931466-34-5
- Bd. 36 Wolfgang Müller, Franz-Josef Rammig: „Methoden und Beschreibungssprachen zur Modellierung und Verifikation von Schaltungen und Systemen", 1998, Praktische Informatik, DM 50,00, ISBN 3-931466-35-3
- Bd. 37 Winfrid Schneider: "Anwenderorientierte Integration von CAE-Systemen. Ein Verfahren zur Realisierung eines durchgehenden Informationsflusses entlang des Produktentwicklungsprozesses", Diss. 1997, Rechnerintegriert Produktion, DM 100,00, ISBN 3-931466-36-1
- Bd. 38 Werner Demel; Günter Schmitz (Hrsg.): "Entwicklung und Transfer von Entwicklungssystemen der Mechatronik", 1998, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931466-37-X
- Bd. 39 Rainer Grobbel; Timo Lange: "Leitfaden PPS-Systeme. Auswahl und Einführung in der Möbelindustrie", 1998 Wirtschaftsinformatik, insb. CIM, DM nicht bekannt, ISBN 3-931466-38-8
- Bd. 40 Peter Rehbein: "Tribologische Untersuchung von hochfrequent schwingenden Gleitkontakten für den Einsatz in Reibkraftschlüssigen Antrieben", 1998, Mechatronik und Dynamik, DM 100,00, ISBN 3-931466-39-6
- Bd. 41 Wilhelm Dangelmaier: „KOMNET - Kommunikationsplattform für kmU-Netzwerke", 1998, Wirtschaftsinformatik insb. CIM, ISBN 3-931466-40-X
- Bd. 42 Ferdinand Kallmeyer: „Eine Methode zur Modellierung prinzipieller Lösungen mechatronischer Systeme", Diss. 1998, Rechnerintegrierte Produktion, DM 100,00, ISBN 3-931466-41-8