

Ermittlung markanter Punkte auf den Bildern bewegter
Objekte und Berechnung einer 3D-Beschreibung auf
dieser Grundlage

Leonie Dreschler

IfI-HH-B-83/81

Oktober 1981

Zusammenfassung

Ausgehend von monokularen TV-Bildfolgen von Realwelt-Vorlagen wird als approximative volumetrische Beschreibung für bewegte starre Körper ein konvexes Polyeder abgeleitet. Gleichzeitig werden die Trajektorie und die Orientierung des bewegten Objektes relativ zur Kamera ermittelt, bis auf einen Skalierungsfaktor, der auch die Objektbeschreibung skaliert. Für zwei Straßenszenen werden Ergebnisse vorgestellt. Der Ansatz wird mit ähnlichen in der neueren Literatur veröffentlichten Verfahren verglichen und diskutiert. Außerdem wird ein Algorithmus zur Auswahl von "markanten" Punkten beschrieben, die als Meßpunkte für die 3D-Rekonstruktion verwendet werden. Auf Grund der Erfahrungen mit dem Relaxationsverfahren von Barnard und Thompson wurden an dem Relaxationsverfahren einige Änderungen vorgenommen, die hier beschrieben und diskutiert werden.

Abstract

A polyhedral approximation for the volumetric description of a moving rigid object from a real-world scene is derived, based on measurements in monocular TV-frame sequences. The trajectory and attitude of the object motion relative to the camera is simultaneously determined up to the same factor which scales the object description. Results from two street scene sequences are presented. The approach is compared to related ones reported in the recent literature. The experience with the relaxation approach of Barnard and Thompson 79+80 is described and discussed and also some modifications which were necessary in order to obtain acceptable results.

"Nächstdem, sprach ich, vergleiche dir unsere Natur in Bezug auf Bildung und Umbildung folgendem Zustande. Sieh nämlich Menschen wie in einer unterirdischen, höhlenartigen Wohnung, die einen gegen das Licht geöffneten Zugang längs der ganzen Höhle hat. In dieser seien sie von Kindheit an gefesselt an Hals und Schenkeln, so daß sie auf demselben Fleck bleiben und auch nur nach vorne hin sehen, den Kopf aber herumzudrehen der Fessel wegen nicht vermögend sind. Licht aber haben sie von einem Feuer, welches von oben und von Ferne her hinter ihnen brennt. Zwischen dem Feuer und den Gefangenen geht oben her ein Weg, längs diesem sieh eine Mauer aufgeführt wie die Schranken, welche die Gaukler vor den Zuschauern sich erbauen, über welche herüber sie ihre Kunststücke zeigen. - Ich sehe, sagte er. - Sieh nun längs dieser Mauer Menschen allerlei Geräte tragen, die über die Mauer herübertagen, und Bildsäulen und andere steinerne und hölzerne Bilder und von allerlei Arbeit; einige, wie natürlich reden dabei, andere schweigen. - Ein gar wunderliches Bild sprach er, stellst du dar und wunderliche Gefangene. - Uns ganz ähnliche, entgegnete ich. Denn zuerst, meinst du wohl, daß dergleichen Menschen von sich selbst und voneinander je etwas anderes gesehen haben als die Schatten, welche das Fenster auf die gegenüberliegende Wand der Höhle wirft? - Wie sollten sie, sprach er, wenn sie gezwungen sind, zeitlebens den Kopf unbeweglich zu halten! - Und von dem Vorübergetragenen nicht eben dieses? - Was sonst? - Wenn sie nun miteinander reden könnten, glaubst du nicht, daß sie auch pflegen würden, dieses Vorhandene zu benennen, was sie sähen? - Notwendig. - Und wie, wenn ihr Kerker auch einen Widerhall hätte von drüben her, meinst du, wenn einer der Vorübergehenden spräche, sie würden denken, etwas anderes rede als der eben vorübergehende Schatten? - Nein, beim Zeus, sagte er. - Auf keine Weise also können diese irgend etwas anderes für das Wahre halten als die Schatten jener Kunstwerke? - Ganz unmöglich. -"

(Platon, Politeia VII, 1, Das Höhlengleichnis, 514a-515d)

Gliederung

GLIEDERUNG

1.	Einleitung	9
1.1.	Übersicht	9
1.2.	Warum Beschäftigung mit Bildfolgen?	12
1.3.	Literaturübersicht zu Bildfolgen mit bewegten Objekten	19
1.4.	Aufbau der Arbeit	36
2.	Ausrüstung, Dienstprogramme und Bilddaten	37
3.	Auswahl von Objektkandidaten	44
4.	Punktefinder	50
4.1.	Merkmalsfindung und Korrespondenzproblem	50
4.2.	Meßpunkte für die 3D-Modellierung	53
4.3.	Einige bekannte Verfahren	55
4.4.	Erfahrungen mit dem Punktefinder von Moravec	64
4.5.	Ein Punktefinder auf der Basis von Krümmungsanalysen	74
4.5.1.	Strukturanalyse mittels der Gaußschen Krümmung	74
4.5.2.	Technische Probleme	84
4.6.	Kombination von Operatoren verschiedener Durchmesser	87
4.7.	Untersuchungen zur Reproduzierbarkeit	89
5.	Das Korrespondenzproblem	96
5.1.	Der Korrespondenzprozeß beim Mensch	96
5.2.	Übertragung des Wettbewerbsmodells	101
5.3.	Minimale Zuordnung (minimal mapping)	103
5.4.	Weitere Lösungsansätze	109
5.4.1.	Korrelationsverfahren	110
5.4.2.	Relaxationsverfahren	112
5.4.2.1.	Huade und Rosenfeld	113
5.4.2.2.	Das Relaxationsverfahren nach Barnard und Thompson	115

Gliederung

5.5.	Modifikationen am Relaxationsverfahren	119
5.5.1.	Symmetrie im Wettbewerb	119
5.5.2.	Das Konsistenzkriterium	121
5.5.3.	Die Anfangswahrscheinlichkeiten	123
5.5.4.	Ergebnisse	124
5.5.5.	Parameter	126
5.6.	Korrespondenzproblem und nächster Nachbar	126
5.7.	Ketten von korrespondierenden Punkten	128
6.	3D-Interpretation	131
6.1.	Übersicht	131
6.2.	Das Interpretationsverfahren von Bonde und Nagel	136
6.2.1.	Kamerakalibrierung	137
6.2.2.	Objektbezogene Parameter	139
6.2.3.	Ergebnisse	141
7.	Modellierung	143
7.1.	Übersicht über 3D-Modellierungsverfahren	143
7.2.	Die konvexe Hülle	144
7.3.	Minimale Verfahren zur Berechnung der konvexen Hülle	144
7.4.	Verdeckte Oberflächen und Schattierung	148
7.5.	Alternative Polyedermodelle zur konvexen Hülle	152
8.	Schlußbemerkungen	157
9.	Anhang: Tangententest	159
10.	Literatur	162
11.	Autorenindex	174
12.	Abbildungen	177

Abbildungen und Tabellen

Abbildungen im Text

2.1	Rechnernetz zur digitalen Bildauswertung	38
2.2	Integration von Halbbildern	43
4.1	Der Moravec-Operator	65
4.2	Krümmung einer Fläche	75
4.3	Beispiele für gekrümmte Flächen	78
4.4	Eckenstrukturen als Grauwertgebirge	79
4.5	Hauptkrümmungsrichtungen zu 4.4	79
4.6	Grenzfälle der Auflösung	82
4.7	Gerasterte Ecken und Kanten in 3*3 Auflösung	86
4.8	Messungen zur Reproduzierbarkeit	95
5.1	Das "zerbrochene Rad"	98
5.2	Statistische Unabhängigkeit von Affinitäten	104
5.3	Erwünschte und unerwünschte Zuordnungen	106
5.4	Minimale Zuordnung	107
5.5	Rotierende Kreisscheibe	108
7.1	Sternpolygontest	146
7.2	Test auf Außenfläche	150
9.1	Tangententest	159
9.2	Messung des Tangentenwinkels	160

Abbildungen im Anhang 177

I	Bildfolge 1
II	Bildfolge 2
III	Darstellung von Abb. Ia als Grauwertgebirge
IV	Ergebnis des Moravec-Operators
V	Gaußsche Krümmung
VI	Einstufiger Punktefinder
VII	Zweistufiger Punktefinder
VIII	Relaxationsverfahren
IX	Nächste-Nachbar-Zuordnung
X	Ketten von korrespondierenden Punkten
XI	3D-Rekonstruktion für das Taxi aus Abb. I
XII	3D-Rekonstruktion für das Taxi aus Abb. II
XIII	Gaußsche Krümmung in ausgewählten Bildausschnitten
XIV	Messungen zur Reproduzierbarkeit

Verzeichnis der Tabellen

1.1	Messung von Geschwindigkeit, Entdeckung oder Isolation von aktiven Bildbereichen	24
1.2	Verfolgung starrer Objekte, die sich verdeckungsfrei bewegen	26
1.3	Verfolgung von Objekten, die sich gegenseitig verdecken können oder durch den Vordergrund verdeckt werden	28
1.4	Verfolgung von nicht-starren Objekten	30
1.5	Modellierung bewegter Objekte, 3D-Rekonstruktion	32
1.6	Beschreibung von Bewegung und Interaktion bewegter Objekte	34
4.1	Ergebnis des Moravec-Operators	67
4.2	Mittlere Anzahl von markanten Punkten in ungefilterten Bildausschnitten	90
4.3	Mittlere Anzahl von markanten Punkten in Median-gefilterten Bildausschnitten	95
4.4	Wahrscheinlichkeit dafür, einen markanten Bildpunkt in n aufeinander folgenden Bildern wiederzufinden	94

Einleitung

Übersicht

1. Einleitung

1.1 Übersicht

Im Höhlengleichnis beschreibt Platon ein Gedankenexperiment, das neben erkenntnistheoretischen Problemen auch für die Verarbeitung von Bildfolgen interessante Fragen aufwirft. Welches Wissen läßt sich über bewegte Objekte gewinnen, wenn man nur ihre Projektionen beobachten kann und kein szenenspezifisches Wissen zur Verfügung hat? Genau dieses Problem wird in dieser Arbeit behandelt. Nur werden keine projizierten Objektsilhouetten untersucht, sondern TV-Bildfolgen.

Nagel hat in den beiden Artikeln /Nagel 77, Nagel 1978a/ grundsätzliche Möglichkeiten diskutiert, wie Bildfolgen von bewegten Objekten ohne Einsatz von szenenspezifischem Wissen interpretiert werden können. In dieser Arbeit werde ich ein Analysesystem für Bildfolgen vorstellen, das auf den von Nagel vorgeschlagenen Grundsätzen aufbaut. Zweck des Systems ist es, für bewegte starre Körper eine 3D-Beschreibung und ihre Bewegung relativ zur Kamera zu ermitteln. Als Vorlage dienen Realweltaufnahmen vom Straßenverkehr. Das verwendete Verfahren nutzt aus, daß durch die Bewegung eines Körpers von diesem verschiedene Ansichten entstehen, durch deren Kombination sich wiederum die räumliche Struktur erschließen läßt. Für starre Körper wird dieses im Fall der Parallelprojektion durch Ullmans "Structure from Motion"-Theorem /Ullman 79/ sichergestellt. Das Analyseproblem kann man in drei Teilaufgaben gliedern. Zunächst muß festgestellt werden, welche Bildbereiche Abbildungen von bewegten Objekten sind. Diese Objekte werden so weit wie möglich über die ganze Aufnahmesequenz verfolgt. Das zweite Problem ist, in den Objektabbildungen auffällige körperfeste Meßpunkte, hier "markante Punkte" genannt, zu finden und diese in den anderen Objektansichten

Einleitung

Übersicht

möglichst genau zu lokalisieren (Korrespondenzproblem). Die dritte Aufgabe ist das mathematische Problem, aus den Folgen von 2D-Koordinaten für die Meßpunkte die räumlichen Koordinaten für die dazugehörigen Objektpunkte und eine Hypothese für die Bewegung des Körpers relativ zum Beobachter abzuleiten. Diese Hypothesen müssen genau die gemessenen Bildkoordinaten erklären. Dazu ist es auch nötig, die Parameter zu ermitteln, die die Abbildungsverhältnisse festlegen.

Für zwei dieser Probleme existieren bereits Bausteine, die schon früher beschrieben wurden: Erstens ein Programm zur Verfolgung bewegter Objekte von Jain und Nagel /Jain and Nagel 79/ und zweitens ein Programm von Bonde und Nagel /Bonde and Nagel 79, Nagel 80a/, das für markante Punkte eines starren Körpers aus den Bildkoordinaten die räumlichen Koordinaten ermittelt, wenn das Korrespondenzproblem gelöst ist. Die Lösung des Korrespondenzproblems ist die Aufgabe des hier beschriebenen neuen Bausteins, der die Lücke zwischen den beiden vorhandenen Modulen schließt. In einer früheren Arbeit haben R. Kraasch et al. /Kraasch et al. 79a, 79b/ das Korrespondenzproblem für Laboraufnahmen von bewegten Blöcken gelöst, während in dieser Arbeit das Problem für Realweltvorlagen untersucht wird.

Der Kern des neuen Bausteins ist ein Punktefinder, der die Gaußsche Krümmung der Bildfunktion auswertet, um Ecken und isolierte Punkte zu erkennen. Bisherige lokale Punkteoperatoren haben "markante Punkte" nur nach dem Kriterium ausgewählt, daß die Grauwertumgebung solcher Punkte genug Varianz aufweist, um im nächsten Bild eindeutig wiedergefunden werden zu können. /z.B. Hannah 74, Moravec 77, 79, 80, Nevatia 76/. Im Unterschied zu anderen lokalen Punkteoperatoren wird versucht, eine weitergehende Vorstellung von der Form des Grauwertgebietes zu erhalten, indem die gefundenen Extremwerte der Gaußschen Krümmung im Zusammen-

Einleitung

Übersicht

hang mit benachbarten Extrema ausgewertet werden. So werden Kantenschnittpunkte oder isolierte Grauwertflecke als Ort von markanten Punkten gewählt. Dieses ist bei einfachen Auffälligkeitsmaßen nur mit größeren Schwierigkeiten zu erreichen. Die Korrespondenzen zum Folgebild werden durch ein Relaxationsverfahren ermittelt, das unter der Annahme, daß die Modellierungspunkte im wesentlichen auf kontinuierlichen Oberflächen liegen, die Kontinuität im optischen Fluß ausnutzt. Dieses Relaxationsverfahren wurde von Barnard und Thompson /Barnard und Thompson 79, 80/ entwickelt und von mir in einigen Punkten verbessert.

Das 3D-Interpretationsprogramm liefert als Resultat die euklidischen Koordinaten für die Modellierungspunkte in einem körperfesten Koordinatensystem. Damit man sich die relative Lage der Punkte im Raum veranschaulichen kann, wird als erstes grobes Objektmodell die konvexe Hülle dieser Punkte errechnet und auf einem Rastersichtgerät wahlweise als Linienzeichnung oder als schattierter Polyeder dargestellt. Da auch die Kameraparameter und die Bewegung des Objektes relativ zur Kamera errechnet werden, kann das Objektmodell zur Kontrolle in die einzelnen Aufnahmen der Bildfolge zurückprojiziert und den Rohdaten überlagert dargestellt werden. So ist leicht zu beurteilen, wie gut das Modellierungsergebnis ausgefallen ist.

Alle Programme wurden in PASCAL geschrieben. Die Bildverarbeitung geschieht auf einem speziell für solche Aufgaben entwickelten Kleinrechnernetzwerk, das über eine umfangreiche Video-Peripherie verfügt. Ein Ziel dieser Arbeit war es, möglichst alle szenenspezifischen Parameter, wie Schwellwerte, vom Programm automatisch wählen zu lassen, sofern sie sich nicht ganz vermeiden ließen. Wenn Parameter vom Operateur eingegeben werden müssen, soll die Wahl logisch begründet werden können, etwa durch die Geschwindigkeit des Objektes oder den Abbildungsmaßstab. Wieweit

Einleitung

Übersicht

dieser Vorsatz eingehalten werden konnte, wird im Zusammenhang mit den einzelnen Programmteilen diskutiert.

Viele der hier angesprochenen Teilprobleme sind in der Literatur schon behandelt worden (siehe Abschnitt 1.3). Jedoch kenne ich keine mit dem hier vorgestellten Bildverarbeitungssystem direkt vergleichbaren Systeme, die ausgehend von Realweltvorlagen ohne ein Modell der Objekte oder sonstiges szenenspezifisches Wissen einzusetzen, aus mit nur einer Kamera aufgenommenen TV-Bildfolgen eine 3D-Beschreibung für bewegte starre Körper erstellen können.

1.2 Warum Beschäftigung mit Bildfolgen?

Unter Bildfolgen soll in dieser Arbeit eine sinnvolle Sequenz von Einzelbildern verstanden werden. Ein Einzelbild (schwarz-weiß) läßt sich mathematisch als eine zweidimensionale Bildfunktion darstellen, die die Bildhelligkeit oder Schwärzung in Abhängigkeit von Zeile und Spalte angibt. Da Bildfunktionen in der Praxis Approximationen an das ursprüngliche Bild sind, d.h. das Bild ist gerastert und der Grauwert quantisiert, kann man von einzelnen Bildelementen reden, die ich hier Pixel nennen will. In einem Einzelbild kann man ein Pixel durch Zeile und Spalte spezifizieren, bei Bildfolgen muß noch das Bild angegeben werden, so daß die Bildfunktion drei unabhängige Parameter hat.

Diese dritte Koordinate kann eine weitere räumliche sein, sie kann die Zeit sein (bei Video-Filmen) oder irgend eine andere Physikalische Größe, die über die Folge variiert, z.B. die Wellenlänge, in der die Aufnahmen gemacht wurden, oder der Betrachtungswinkel.

Einleitung

Bildfolgen

Unter der Voraussetzung, daß die dritte Koordinate nicht die Zeit oder eine Funktion der Zeit ist und alle Bilder der Folge als gleichzeitig entstanden betrachtet werden können, ist es oft günstiger, die Einzelbilder als parallel und nicht als sequentiell angeordnet zu verstehen. Ein Beispiel dafür sind die drei Farbauszüge eines RGB-Bildes. In diesem Fall ist die Bildfunktion besser als eine 2D-Funktion beschrieben, die in Abhängigkeit von Zeile und Spalte keinen skalaren Helligkeitswert, sondern einen Helligkeitsvektor liefert. Dieser Helligkeitsvektor faßt z.B. die RGB-Werte des Pixels zusammen. Beide Darstellungsweisen sind mathematisch äquivalent. In dieser Arbeit werde ich mich auf Bildfolgen beschränken, die ein oder mehrere bewegte Objekte vor einem gut kontrastierenden, statischen Hintergrund zeigen.

Im nächsten Abschnitt möchte ich zeigen, wo die Vorteile bei der Verarbeitung von Bildfolgen gegenüber der Analyse von statischen Aufnahmen wie Einzelbildern oder Stereoaufnahmen liegen.

Ein Einzelbild ist eine gerasterte und quantisierte zweidimensionale Abbildung irgendeiner dreidimensionalen Realweltvorlage. Diese Abbildung ist nicht umkehrbar eindeutig. Ein Mensch kann jedoch in der Regel sofort erkennen, was auf einem Foto abgebildet ist. Dies liegt daran, daß er das Bild unter Zuhilfenahme von Zusatzwissen interpretiert. Oft ist er in der Lage, aus einer einzelnen Momentaufnahme Information über die Gruppierung in Objekte, über Entfernungsverhältnisse, 3D-Formen und Bewegungsverhältnisse herauszulesen.

Wenn ein Szenenanalyseprogramm dies ebenfalls leisten können sollte, müßte man ihm ein vollständiges Modell der Welt zur Verfügung stellen, die es verstehen soll. Dies ist bei den heutigen Speichern und Zugriffsmethoden für allgemeine Bild-

Einleitung

Bildfolgen

vorlagen noch nicht realisierbar. Realistischer erscheint dagegen die Beschränkung auf eine Minimalwelt, z.B. auf ein bestimmtes Fließband, auf dem nur bekannte Objekte auftauchen dürfen, oder die Verwendung von wenigen, allgemeingültigen Konzepten. Bei der Einzelbildanalyse sind je nach Analyseziel verschiedene allgemeingültige Konzepte oder Heuristiken denkbar:

1. Tiefenmessung: Im Allgemeinen nur relative Entfernung meßbar, z.B. durch Verdeckung und Abschattung, Entfernungsmarkierungen im Bild, Texturgradienten /Bajcsy and Lieberman 76/, Fluchtlinien, Deutlichkeit im Dunst oder Bildscharfe. Hat man ein Modell der im Bild vorkommenden Objekte, kann die Entfernung aus dem Verhältnis von scheinbarer Größe zu wahrer Größe oder von scheinbarer Helligkeit zu wahrer Helligkeit ermittelt werden.
2. 3D-Form: Die dreidimensionale Form kann nur ungenügend aus einer Einzelansicht ermittelt werden. Hinweise lassen sich aus der Kontur, der Kontur des Schattens, der Form von Schatten, die auf das Objekt fallen, der Position von Glanzlichtern und der Grauwertverteilung an der Oberfläche gewinnen. Nützlich ist es, wenn man ein genaues Modell der Lichtquellen und des Rückstrahlverhaltens der Objektflächen hat.
3. Bewegung: Aussagen über Bewegung sind nur durch Spezialwissen oder Verstehen der dargestellten Situation möglich, so wie ein Mensch einen Schnappschuß deuten kann; z.B. bei der Aufnahme eines fahrenden Schiffes könnte er aus den Bugwellen auf die Geschwindigkeit und die Bewegungsrichtung schließen. Ein Programm, das über kein spezialisiertes Weltmodell verfügt, könnte höchstens die

Einleitung
Bildfolgen

Bewegungsunschärfe ausnutzen.

4. Gruppierung in Objekte: Wenn man unter einem Objekt eine Materieeinheit versteht, die sich zusammenhängend als Ganzes bewegt, so hat man bei Einzelbildern so gut wie keine Möglichkeit festzustellen, welche Bildbereiche zusammenhängende Objekte darstellen, da eben die Bewegungsinformation fehlt. Hier helfen nur Objektmodelle oder eingeschränkte Weltmodelle. Für Blocksweltobjekte /Duda and Hart 75/ wurde gezeigt, wie durch konsistente Eckenmarkierung eine Zerlegung des Bildes in Einzelblöcke möglich ist. Gerade bei der Analyse von Bewegung und Gruppierung in Objekte liegen die großen Vorteile von Bildfolgen gegenüber Einzelbildern.

Ein Fortschritt gegenüber Einzelbilder-Verarbeitung ist das Arbeiten mit Stereoaufnahmen. Das Stereoverfahren basiert auf der Vermessung durch Triangulation. An den Endpunkten einer Basislinie werden zwei möglichst identische Kameras aufgestellt, deren Daten bekannt sind. Wenn beide Kameras denselben Raumpunkt abbilden, dann lassen sich aus den Bildkoordinaten die Winkel errechnen, unter denen die Kameras den Punkt sehen, und damit nach den Kongruenzsätzen der Trigonometrie die Entfernungen ermitteln. Hierbei treten einige prinzipielle Schwierigkeiten auf, die auch bei der Analyse von Bildfolgen wichtig sind.

1. Kalibrierung der Kameras: Methoden hierfür haben unter anderem Hannah /Hannah 74/, Sobel /Sobel 74/, Gennery /Gennery 79b/ und Bonde /Bonde 79/ beschrieben.

Einleitung
Bildfolgen

2. Auswahl geeigneter Punkte: Anders als bei der Feldvermessung mittels Theodoliten vermißt eine Kamera nicht nur einen Punkt zur Zeit, sondern gleich eine ganze Matrix von Punkten. Aus diesem Angebot müssen diejenigen ausgewählt werden, die in der Grauwertmatrix durch lokale, auffällige Merkmale markiert sind und für die eine Chance besteht, daß sie im zweiten Bild wiedergefunden werden können.

3. Korrespondenzproblem: Alle räumlichen Punkte, die sich auf einer gemeinsamen Geraden durch das Kamerazentrum befinden, werden von der Kamera auf denselben Punkt abgebildet. Die zweite Kamera sieht diese Gerade von der Seite, so daß die Gerade hier als Linienstück im Bild erscheint. Es kann also jeder Punkt auf dieser Linie mit dem im ersten Bild gefundenen Punkt korrespondieren. Das Korrespondenzproblem besteht nun darin, durch geeignete Heuristiken den wahrscheinlichsten Partner herauszufinden. Wenn die Stereobasis klein ist, so daß nur geringe perspektivische Verschiebungen auftreten, dann ist es üblich, mit Grauwertkorrelation oder ähnlichen Suchmethoden zu arbeiten /Hannah 74/.

Wenn die untersuchte Szene nur unbewegte Objekte enthält, kann man die Stereoanalyse auch mit nur einer Kamera durchführen, indem man die beiden Aufnahmen nicht gleichzeitig, sondern nacheinander macht, wobei man die Kamera um das Objekt bewegt. Nevatia /Nevatia 76/ hat diese Methode angewendet, weil so nicht nur zwei Aufnahmen, sondern eine ganze Serie von Aufnahmen mit sehr geringem Winkelabstand zur Verfügung stehen. Dieses sollte eine Lösung des Korrespondenzproblems erleichtern, da so auch bei großer

Einleitung
Bildfolgen

Stereobasis die Unterschiede zwischen Folgebildern klein bleiben. Dieses Verfahren wird "Motion Stereo" genannt, wohl weil Nevatia der Einfachheit halber nicht die Kamera um das Objekt, sondern das Objekt auf einem Drehteller relativ zur Kamera bewegt hat. Dies erleichtert die Bestimmung der Kameraposition.

Für das Zustandekommen verschiedener Objektansichten ist es unwichtig, ob sich das Objekt oder die Kamera bewegt hat, die Hauptsache ist, daß eine Relativbewegung stattgefunden hat. Wenn bekannt ist, wie sich Kamera und Objekt relativ zueinander bewegt haben, dann ist "Motion Stereo" nur eine Erweiterung der bekannten, statischen Stereoanalyse. Anders sieht es aus, wenn die Relativbewegung nicht bekannt ist. Dann müssen anhand der Bildfolge Objektstruktur und Bewegung gleichzeitig ermittelt werden. In dieser Arbeit wurde mit Bildfolgen gearbeitet, die dieser erweiterten Aufgabenstellung entsprechen.

Der Vorteil von Stereoaufnahmen gegenüber Einzelbildern liegt auf der Hand. Für ausgewählte Punkte auf der von beiden Kameras sichtbaren Objektfläche können exakte Entfernungsmessungen vorgenommen werden. Bei bekannter Entfernung kann auch scheinbare Größe in wahre Größe oder scheinbare Helligkeit in wahre Helligkeit umgerechnet werden. Aussagen über die Gruppierung in Objekte oder Bewegungsverhältnisse sind dagegen bei Stereoaufnahmen beinahe ebenso schwer wie für Einzelbilder, da diese ja auch nur Momentaufnahmen sind. Gennery /Gennery 79a/ versucht allerdings aus Stereomessungen eine Gruppierung in Objekte zu erreichen. Er geht von räumlichen Koordinaten für ausgewählte Oberflächenpunkte von Objekten aus, die er mittels Stereo-Aufnahmen oder Laser-Tiefenmessungen ermittelt hat. Für diese Punktkoordinaten führt er eine Ballungsanalyse durch (clustering). Die Ballungsgebiete werden als Objekte aufgefaßt und die sichtbare Oberfläche

Einleitung

Bildfolgen

durch ein Ellipsoid approximiert. Dies ist in seinem Fall nur dadurch möglich, daß er ein eingeschränktes Weltmodell verwendet. Sein Programm wurde für die Marsoberfläche entworfen, wo im wesentlichen eine Ebene mit vereinzelt Felsbrocken erwartet wird.

Die Analyse von Bildfolgen mit bewegten Objekten unterscheidet sich sehr von der Analyse von statischen Bildern, da die zeitlichen Veränderungen ein neues und vielseitiges Hilfsmittel darstellen, um die Szenen ohne szenenspezifisches Wissen zu untersuchen. Das Extrahieren von Objekten wird erleichtert, wenn sich gemeinsam und zusammenhängend bewegende Bildkomponenten als Abbild eines sich bewegenden Objekts betrachtet werden /Nagel 78a/. Außerdem können einem die verschiedenen perspektivischen Projektionen, die man von einem bewegten Objekt während einer Bildfolge zu sehen bekommt, einen Eindruck von der räumlichen Gestalt geben, und es sollte eine dreidimensionale Rekonstruktion möglich sein. Ullman hat diese Vermutung in seinem "Structure from Motion" Theorem bewiesen (weiteres siehe Kapitel 6). Dieser Beweis setzt starre Körper voraus. Umgekehrt läßt sich das Beweisverfahren auch als Test dafür verwenden, daß das beobachtete Objekt starr ist.

Ähnlich wie bei der Stereoanalyse müssen für diese Art der 3D-Analyse Ketten von korrespondierenden Punkten über die Bildfolge ermittelt werden. Die Lösung des Korrespondenzproblems ist hierbei gegenüber der Stereoanalyse dadurch erschwert, daß die in Korrespondenz zu bringenden Aufnahmen nicht gleichzeitig erstellt wurden. Sowohl die Umgebung als auch das Objekt oder die Beleuchtungsverhältnisse können sich während dessen verändert haben. Ein zu großer zeitlicher Abstand zwischen den Aufnahmen einer Bildfolge hat ähnliche Konsequenzen für das Korrespondenzproblem wie eine zu große Basislänge bei der Stereo-Analyse. Da zur Erleichterung einer Lösung des Korrespondenzproblems das

Einleitung
Bildfolgen

Intervall zwischen zwei Aufnahmen klein sein muß, enthalten die Bilder sehr viel Redundanz, die zu Konsistenzprüfungen ausgenutzt werden kann.

1.3 Literaturübersicht zu Bildfolgen mit bewegten Objekten

Bildfolgen von bewegten Objekten sind in der Literatur in vielfältiger Weise behandelt worden. Die Tabellen am Ende dieses Abschnitts geben einen Überblick über Arbeiten, die sich mit diesem Thema befassen. Bildfolgen werden bei der Überwachung von Fließbändern angewendet, zur Verkehrskontrolle, zur Auswertung von Satellitenbildern von Wolkenformationen oder landwirtschaftlichen Nutzgebieten, in der Medizin zur Organbeobachtung, Blutkörperchenkontrolle und Zellteilungskontrolle, in der Biologie zur Verhaltensforschung und Bakterienbeobachtung, bei militärischen Einsätzen zur automatischen Zielverfolgung, oder auch zur Steuerung von Robotern usw. Bildverarbeitungssysteme, bei denen der Schwerpunkt auf Problemen der symbolischen Informationsverarbeitung liegt, wie Verstehen des Bildinhaltes und der Interaktionen zwischen den Szenenkomponenten, verwenden oft keine Realweltvorlagen, sondern arbeiten mit gestellten Laboraufnahmen, Zeichentrickfilmen oder rechnergenerierten Kunstbildern.

Bei der folgenden Diskussion werde ich mich im wesentlichen auf die drei Übersichtsartikel /Martin and Aggarwal 78/, /Nagel 78b/ und /Scacchi 79/ stützen, sowie auf die Übersichten in den Dissertationen /Radig 78a/ und /O'Rourke 80a/. Der sehr ausführliche Übersichtsartikel /Nagel 80b/ diskutiert besonders den Aspekt der Anwendungen.

Einleitung Bildfolgen

Wenn man die bisherigen Arbeiten nach den Analysezielen grob unterscheidet, erhält man die Kategorien

1. Entdeckung von Bewegung
2. Verfolgung der bewegten Objekte
3. Modellierung oder Beschreibung der als bewegt erkannten Objekte oder der Bewegungsabläufe.

Eine andere mögliche Unterteilung ergibt sich aufgrund der Restriktionen für die zulässigen Objekttypen. Manche Systeme arbeiten mit einem Objektkonzept, das sehr allgemein ist und eine Vielzahl von Objekten zuläßt, während andere auf sehr spezielle Anwendungen zugeschnitten sind. Dies gilt besonders für Systeme, die mit Echtzeitanforderungen fertig werden müssen. Übliche Einschränkungen sind z.B.: Die Objekte müssen sich deutlich vom Hintergrund abheben; Die Objekte müssen in Teilen oder sogar ganz starr sein; Die Objekte müssen Kugeln, Kegel, Polyeder oder sonstige spezielle Körper sein. Die Objekte müssen zweidimensional sein, Polygonzüge sein, punktförmig sein usw.

Ähnliche Aufteilungen ergeben sich nach der Art der zulässigen Bewegungen. Manche Systeme lassen sowohl Translation als auch Rotation zu, andere nur eins von beiden, andere erlauben die Bewegung nur in einer festen Ebene oder beschränken bei der Rotation die Orientierung der Achse im Raum. Sehr wichtig ist, ob ein System zuläßt, daß bewegte Objekte zeitweilig durch den Vordergrund verdeckt werden oder sich sogar gegenseitig bei ihren Relativbewegungen verdecken oder abschatten. Interessant in diesem Zusammenhang ist die Arbeit von Aggarwal und Duda /Aggarwal and Duda 75/, die Bildfolgen von sich durcheinander bewegenden Polygonen untersucht. Die Form eines Polygons

Einleitung Bildfolgen

kann von diesem System vollständig ermittelt werden, auch wenn das betreffende Objekt in keinem Bild der Folge unverdeckt zu sehen gewesen ist, aber jeder Teil der Kontur wenigstens einmal in einem Bild zu sehen war. Tsuji et al. /Tsuji et al. 79, 80/ lassen sogar zu, daß sich nicht nur die Objekte sondern auch die Kamera bewegt, und ermitteln aus der scheinbaren Bewegung des Hintergrundes die Kamerabewegung.

Interessant ist die Beobachtung, daß ein lockerer Zusammenhang zwischen dem Objektkonzept und dem Bewegungskonzept besteht. Wenn das Objektkonzept sehr allgemein ist, dann ist in der Regel das Bewegungskonzept sehr restriktiv und umgekehrt. Aggarwal und Duda /Aggarwal and Duda 75/ lassen z.B. als Objekte nur Polygone zu, während bei der Bewegung sehr komplizierte Bedeckungsvorgänge erlaubt sind. Im Gegensatz dazu verwendet O'Rourke z.B. einen komplexen Objekttyp, aber es ist keine Verdeckung durch andere bewegte Objekte möglich, da nur ein Objekt zugelassen ist.

Ein weiteres wichtiges Kriterium ist die verwendete Bildvergleichsmethode, insbesondere auf welcher Abstraktionsebene der Bildvergleich durchgeführt wird, mehr auf der Ebene der Rohdaten oder zwischen symbolischen Beschreibungen. Radig /Radig 78a/ benutzt eine Graphdarstellung zur Bildbeschreibung, die eine ganze Hierarchie von Beschreibungsebenen enthält, so daß der Bildvergleich je nach Bedarf mehr auf signalnaher Ebene oder mehr auf symbolischer Ebene durchgeführt werden kann.

Die Zahl der zugelassenen Objekte variiert von nur einem bis zu sehr vielen. In der Regel sind umso weniger Objekte erlaubt, je komplexer sie aufgebaut sein dürfen.

Einleitung

Bildfolgen

In einigen Arbeiten liegt der Schwerpunkt nicht so sehr auf der Bildverarbeitung, sondern mehr auf der Repräsentation von Wissen und auf den Kontrollmechanismen. Hierfür sind die Arbeiten von O'Rourke und Badler /O'Rourke 80a, O'Rourke and Badler 80/ und Tsotsos /Tsotsos et al. 79, 80/ ein gutes Beispiel. Beide Arbeiten verwenden ein komplexes Objektmodell, aus dem für die Bildanalyse Hypothesen generiert werden. In einem zyklischen Kontrollfluß wird das Objektmodell mittels der Analyseergebnisse verfeinert und neue, korrigierte Hypothesen gebildet. Die Art der Hypothesenbildung und der Wissensspeicherung ist jedoch in beiden Systemen grundverschieden. O'Rourke modelliert den menschlichen Körper aus Kugelteilen und speichert zusätzliche Informationen über minimale und maximale Entfernungen, die die einzelnen Körperteile voneinander einnehmen können. Die Bildanalyse startet mit einem sehr ungenauen Modell, in dem für jeden Körperteil ein großer Unsicherheitsraum für den Ort besteht. Wenn ein Körperteil im Bild gefunden wird, werden iterativ alle Zwangsbedingungen ausgewertet, um die Lage der anderen Komponenten weiter einzuschränken. Das verbesserte Modell wird zur Vorhersage für den Ort von weiteren Körperteilen herangezogen. Die Bildanalyse selbst ist unproblematisch, da die Bildvorlagen künstlich mittels des Modells generiert werden und alle Körperteile unterschiedliche Grauwerte haben.

Tsotsos et al. /Tsotsos et al. 79, 80/ verwenden das Rahmen-Konzept (frames) von Minsky /Minsky 75/ zur Beschreibung der Objektmodelle und führen eine Wissensbasis mit erkennenden Routinen für verschiedene Bewegungskonzepte. Hypothesen über den beobachteten Bewegungsablauf werden aufgrund der bisherigen Hypothesen und der Beobachtungsdaten von allen Routinen gleichzeitig und unabhängig produziert und mit einem Konfidenzwert versehen. Die Konfidenzwerte werden in einer Art Relaxationsverfahren gegeneinander abgewogen, wobei sich ähnliche Hypothesen stützen und widersprüchliche

Einleitung

Bildfolgen

Hypothesen gegenseitig abschwächen können. Das System liefert eine verbale Beschreibung des beobachteten Bewegungsablaufs inklusive der Zeitverhältnisse. Als Eingabedaten wurden Filme von der Wandbewegung der linken Herzkammer verwendet.

Den Aspekt der Anwendungen werde ich hier vernachlässigen, da das den Rahmen dieser Arbeit sprengen würde. Eine ausführliche Übersicht ist in /Nagel 80b/ zu finden. Die Anwendungen führen nur indirekt über die mit ihnen verbundenen Anforderungen und Randbedingungen zu einer Systematik, da die Anwendungen die Entscheidungen über die verwendeten Bildverarbeitungstechniken in erster Linie beeinflussen. Wichtige Aspekte sind Rechenzeit, Robustheit, Speicherbedarf, Bedienungsfreundlichkeit, Echtzeitanforderungen oder Eignung für Kleinrechner. Ich möchte mich im Gegensatz hierzu auf die Frage konzentrieren, wie die Probleme bei der Bildfolgenbearbeitung prinzipiell und möglichst allgemeingültig zu lösen sind, unabhängig von einer speziellen Anwendung.

Die folgenden Tabellen geben eine grobe Übersicht über Arbeiten, die sich mit Bildfolgen befassen. Diese Tabellen erheben keinen Anspruch auf Vollständigkeit, sondern dienen dazu, typische Verfahren exemplarisch gegenüberzustellen. Daß in dieser Zusammenstellung viele Arbeiten aus unserer Gruppe enthalten sind, liegt einerseits daran, daß ich diese Arbeiten besonders gut kenne und andererseits daran, daß meine eigene Arbeit teilweise auf diesen Vorarbeiten aufbaut, und ich mich oft auf diese Veröffentlichungen beziehen werde. Fragezeichen in den Tabellen bedeuten, daß die Autoren zum betreffenden Punkt keine Angaben gemacht haben.

Einleitung
Literaturtabellen

Tab. 1.1: Messung von Geschwindigkeit, Entdeckung oder Isolation von aktiven Bildbereichen

Autoren	Bildvorlagen	Analyseziel
Batali and Ullman 79	synthetisch erzeugte Konfigurationen von statistischen Punkteverteilungen	lokale Geschwindigk. Objektgeschwindigkeit.
Fennema and Thompson 79	Laboraufnahmen Spielzeug	Geschwindigkeitsmessung für mehrere Objekte
Jain et al. 77 Jain and Nagel 79	Straßenszenen	Isolation von bewegten Objekten
Limb and Murphy 75	Realweltvorlagen	Geschwindigkeits-schätzung für Gesamtbild
Potter 75a	Mehrere schwarze rechteckige Objekte vor weißem Hintergrund	lokale Geschwindigk., Objektgeschwindigkeit
Potter 75b, 77	Kunstbilder und Laboraufnahmen von dreieckigen und rechteckigen Objekten	lokale Geschwindigk., Objektgeschwindigkeit

Einleitung
Literaturtabellen

Methoden	Probleme / Besonderheiten
Konvolution mit ▽ ² G Operatoren, Nullstellensuche	---
Variationen von Grauwerten, lokal	mehrere Objekte gleichzeitig auch bei Verdeckung
Auswertung von Dif- ferenzbildern 1. und 2. Ordnung FODP, SODP	Schwierigkeiten bei Verdeckung
Interbilddifferenzen, Innerbilddifferenzen	Ergebnis texturab- hängig
Pixelgeschwindigkeit durch Abstand zur nächsten Grauwert- diskontinuität	Nur für Translation geeignet keine Verdeckung
Pixelgeschwindigkeit durch Korrelation mit Kreuzmaske	Nur für Translation

Einleitung
Literaturtabellen

Tab. 1.2: Verfolgung starrer Objekte, die sich verdeckungsfrei bewegen

Autoren	Bildvorlagen	Analyseziel
Budy and Dreschler 76	Straßenszenen	Verfolgung
Dreschler and Nagel 78	Straßenszenen	Verfolgung
Gilbert et al. 78, 80	Raketen	Verfolgung Kameranachführung Modell der Rakete
Nagel 78a	Straßenszenen	Isolation von Objektkandidaten, Verfolgung
Milstein and Lazicky 78	?	bewegte Objekte im Nachführfenster halten.
Onoe and Ohba 76	Aufnahmen vom Dach ei- nes Gebäudes, Stras- senverkehr	Verkehrsflußanalyse
Wallace und Mitchell 79	synthetisch erzeugte Bilder	Objektmodell im Bild schnell finden
Wolferts 73	Luftaufnahmen von Straßenverkehr	Verfolgung von interaktiv ausge- wählten Fahrzeugen

Einleitung
Literaturtabellen

Methoden	Probleme / Besonderheiten
Korrelation mit interaktiv gewählter Maske	Nur ein Objekt, nur Translation, Anpassung der Suchmaske
Korrespondenz von Objektkandidaten durch minimalen spannenden Baum (Ballungsanalyse)	Probleme mit Differenzbild
Segmentation durch Grauerthistogramm; Trennung in Bild von Rakete und Strahl	Echtzeitverarbeitung durch Spezialelektronik
Bereichsfindung und Zuordnung von stationären Bildteilen	Nur Translation, nur ein Objekt
3 statistische Verfahren, um Objekte vom Hintergrund zu trennen	nur 1 Objekt neutraler Hintergrund erforderlich
Korrelation	Schwellwertbild durch Elektronik
Datenbank mit Fourierkoeffizienten der Kontur aus verschiedenen Blickwinkeln, Interpolation	Echtzeitanalyse erscheint möglich
Korrelation	---

Einleitung
Literaturtabellen

Tab. 1.3: Verfolgung von Objekten, die sich gegenseitig verdecken können oder durch den Vordergrund verdeckt werden

Autoren	Bildvorlagen	Analyseziel
Aggarwal and Duda 75	Konturbilder von Polygonen	Verfolgung und Ermittlung der unverdeckten Kontur
Chow and Aggarwal 77	starre Flecken beliebiger Kontur	Verfolgung
Martin and Aggarwal 78	Starre Flecken beliebiger Kontur	Verfolgung
Neumann 78a,b	Objekte auf Fließband 2D und 3D	Verfolgung, Anpassung von Modellen
Radig 78a, b, c, d	Straßenszenen Realweltvorlagen	Verfolgung eines oder mehrerer Objekte
Roach and Aggarwal 79	konvexe Polyeder, Bewegung im Raum	Verfolgung
Yachida et al. 78	Fische, Einzeller	Verfolgung, Geschwindigkeitsmessung
Gerlach 79 a, b	Straßenszenen Realweltvorlagen	Erkennen und Wiederfinden von vom Operateur markierten Objekten

Einleitung
Literaturtabellen

Methoden	Probleme / Besonderheiten
Unterscheidung der Winkel nach wahren und scheinbaren	pro Bild darf nur ein wahrer Winkel sichtbar oder neu verdeckt werden
Zuordnung durch globale Merkmale, Vorhersage durch Geschw.	Während der Verdeckung konstante Geschwindigkeit erforderlich
Konturapproximation durch Kreisbögen und Geradenstücke	keine konstante Geschwindigkeit während der Verdeckung erforderlich
Konturapproximation durch Geradenstücke, keine geschlossenen Konturen nötig	für 3D-Objekte mehrere Modelle von typischen Ansichten erforderlich
Bildrepräsentation durch Graph	Hierarchie von Abstraktionsebenen
Objektkorrespondenz durch Geschwindigkeit, relative Anordnung, Flächenkorrespondenz. Einzelobjekte durch T-Verbindungen, Tiefe durch Flächenverhältnisse	keine konstante Beschleunigung erlaubt; keine Rotation, die sichtbare Flächen verändert.
zeitliche und räumliche Grauwertdifferenzen	bei unsicheren Ergebnissen Kontrolle durch vorheriges Bild
Normierte Kreuzkorrelation, Adaption des Referenzbildes	Sensornachführung, Verfolgung auch bei Drehung oder Verdeckung durch den Vordergrund

Einleitung
Literaturtabellen

Tab. 1.4: Verfolgung von nicht-starren Objekten

Autoren	Bildvorlagen	Analyseziel
Haass and Brubaker 79	Satellitenbilder von Wolkenformationen	Verfolgung, Beschrei- bung von Größenver- änderungen und Rotation
Hernan and Jimenez 79	Turbulenzschicht zwi- schen zwei Gasströmen	Verfolgung von Gas- blasen
Tsuji et al. 79, 80	Zeichentrickfilm von Fabeltier, Strichzeich- nung	Isolation und Verfol- gung von starr bewegten Körperteilen, Kamera- bewegung.
Wolf et al. 79	Satellitenbilder von Wolkenformationen	Verfolgung, Beschrei- bung von physikalischen Eigenschaften
Yachida et al. 79	Röntgenfilm der Herz- wandbewegung,	Messung der zeitlichen Veränderungen der Herzwanddicke

Einleitung
Literaturtabellen

Methoden	Probleme / Besonderheiten
Spezielle Filter- und Korrelationstechniken	---
Anpassung von Ellipsen	Zyklische Rückkopplung zwischen Bildanalyse und Modellbildung
Bereichsfindung, Kontursegmentation, Bereichs- und Konturzuordnung	Bewegte, mitgeführte Kamera
Ballungsanalyse, Zuordnung von Ballungsgebieten	---
Konturapproximation durch Geradenstücke, Korrelation	Schnelle Bildauswertung durch Plan, spezieller Kantenfinder für vertauschte Bilder

Einleitung
Literaturtabellen

Tab. 1.5: Modellierung bewegter Objekte, 3D-Rekonstruktion

Autoren	Bildvorlagen	Analyseziel
Badler 74, 75	Strichzeichnungen von Blocksweltlandschaften	Objektbewegung, Kamerabewegung und Objektrekonstruktion; Ebenenneigung im stationären Teil
Diese Arbeit	Realwelt-Straßenszenen	Korrespondenzketten von markanten Punkten
Bonde und Nagel 79 Nagel 80a	Sequenzen von 2D-Koordinaten von Körperfesten Punkten.	3D-Koordinaten im objektbezogenen Koordinatensystem, Trajektorie
Herman et al. 79	Medizinische Daten, 3D-Daten von zeitlich variierenden Volumina	zeitliche Veränderung von Organoberflächen
Krausch et al. 79a, b	gut gegen den Hintergrund kontrastierende Polyeder	Korrespondenzen für körperfeste Punkte
Neumann 79a, 80	Ideen und Simulationsrechnungen zur Ermittlung von Ebenenneigungen aus Grauwertveränderungen	zeitlichen
Neumann 79b	einfache synthetische Testdatensätze	Gruppierung in Objekte, 3D-Rekonstruktion und Verfolgung in einem Verfahren
O'Rourke 80a O'Rourke and Badler 80	Künstliche Bilder von Menschen, jedes Körperteil anderen Grauwert.	iterative Modellverbesserung
Ullman 79	körperfeste, auffällige Punkte	Korrespondenzproblem Interpretationsproblem

Einleitung
Literaturtabellen

Methoden	Probleme / Besonderheiten
Sphärische Projektion, Fluchtpunkte für paral- lele Geraden	---
spez. Punktefinder, Relaxationsverfahren zur Korrespondenz- analyse	Bildanalyseteil für Bonde and Nagel 79
iterative Minimisierung einer Fehlerfunktion, separate Ermittlung der Kameraparameter	ebenengebundene Bewegung, Rotation nur um Achse senkrecht zur Grundebene, perspektivische Projektion
Suche im 4D-Raum nach begrenzenden Hyper- ebenen	---
Bildrepräsentation durch Segmentations- graph, Graphzuordnung	Schwierigkeiten bei der Erkennung von nicht körperfesten Punkten
Geradenparameter durch Hypothesenbildung	orthographische Projektion
modellgesteuerte Suche, Modell aus Zwangs- bedingungen für rela- tive Lage von Körper- teilen.	Modell aus Kugelteilen zyklischer Kontrollfluß
"minimal match" "Structure from Motion" Theorem	---

Einleitung
Literaturtabellen

Tab. 1.6: Beschreibung von Bewegung und Interaktion bewegter Objekte

Autoren	Bildvorlagen	Analyseziel
Futrelle and Cheng 79	Filme zur Verhaltensforschung	Verbale Beschreibung der Interaktionen zwischen vom Benutzer gekennzeichneten Objekten
Hogg 77	Realwelt-Vorlagen	Natürlich-sprachliche Beschreibung von Bewegungsabläufen.
Levine et al. 79 vgl. Ferrie et al. 80	Blutkörperchen, Zellteilungsvorgänge	Verfolgung von Blutkörperchen, Beschreibung von Zellteilungs-Vorgängen und Formveränderungen der Zellwände.
Schudy 79	Herz-Ultraschalldaten	Analytische Darstellung des Herzzinnern als Funktion der Zeit.
Tsotsos et al. 79, 80	Bewegung der linken Herzkammer (Röntgenaufnahmen)	Verstehen der Bewegung und Beschreibung durch natürliche Sprache

Einleitung
Literaturtabellen

Methoden	Probleme / Besonderheiten
Relationale Beschreibung	Speziell entwickelte Datenstruktur
Differenzbild, einfache Merkmale von Differenzbereichen ausgewertet.	Echtzeit-System
?	Auswertung einer Wissensbasis
Reihenentwicklung durch orthogonales Funktionensystem	Vorgesehen für graphische Wiedergabe der zeitlichen Veränderungen.
Wissensbasis von Bewegungskonzepten (Frames), zyklische Interaktion von Hypothesenbildung und Bildanalyse, Wettbewerb von Hypothesen im Relaxationsverfahren.	

Einleitung

Aufbau der Arbeit

1.4 Aufbau der Arbeit

Das System umfaßt folgende Bausteine:

1. Ein Programm zur Auswahl von Bildausschnitten, die ein bewegtes Objekt enthalten /Jain et al. 77, Jain and Nagel 79/. Alternativ kann diese Auswahl auch interaktiv an einem Rastersichtgerät mit Hilfe einer Rollkugel geschehen.
2. Ein Unterprogramm zur Auswahl von markanten Punkten für die 3D-Modellierung (siehe Kapitel 4). Hierfür wurden von mir zwei Algorithmen implementiert, zum einen ein Punktefinder von Moravec /Moravec 77, 79, 80/ und zum anderen ein von mir selbst entwickelter Punktefinder, der die Krümmungen im Grauwertgebirge auswertet und für Realweltvorlagen geeignet ist.
3. Ein Unterprogramm zur Lösung des Korrespondenzproblems. Dieses Unterprogramm ist eine Weiterentwicklung des Relaxationsverfahrens von Barnard und Thompson /Barnard and Thompson 79, 80/ (siehe Kapitel 5).
4. Ein von Th. Bonde entwickeltes Programm zur Kamera- kalibrierung /Bonde 79/.
5. Ein von Bonde und Nagel /Bonde 79, Bonde and Nagel 79, Nagel 80a/ entwickeltes Programm zur Berechnung der 3D-Koordinaten aus den Korrespondenzketten der markanten Punkte (siehe Kapitel 6).
6. Ein Programm zur Darstellung des 3D-Rekonstruktionsergebnisses auf einem Rastersichtgerät (siehe Kapitel 7).

Ausrüstung

2. Ausrüstung, Dienstprogramme und Bilddaten

Das in dieser Arbeit vorgestellte Bildverarbeitungssystem wurde für ein Kleinrechnernetzwerk mit umfangreicher Video-Peripherie entworfen. Dieses Netzwerk wurde im Arbeitsbereich "Kognitive Systeme" am Fachbereich Informatik der Universität Hamburg in mehrjähriger Arbeit aufgebaut und mit komfortablen Betriebssystemen für Interrechnerkommunikation und Bildverarbeitung ausgestattet. An dieser Arbeit waren bis jetzt R. Bertelsmeier, I. Heer, Dr. H. Kemen, K. Kleemann, Prof. Dr. H.-H. Nagel, Dr. B. Neumann, Dr. B. Radig und ich selbst beteiligt, unterstützt im Sekretariat durch R. Jancke. Ohne diese vorangehende Aufbau- und Entwicklungsarbeit wäre meine Arbeit in der hier vorliegenden Form nicht möglich gewesen. Daher möchte ich allen jetzigen und früheren Kollegen an dieser Stelle herzlich danken, besonders auch für die Unterstützung bei gelegentlichen Problemen mit der Apparatur oder den Systemprogrammen.

Die von mir verwendete experimentelle Ausrüstung ist seit dem Beginn dieser Arbeit in vielen Punkten weiterentwickelt und ausgebaut worden. Jedoch werde ich hier die ursprüngliche Ausbaustufe beschreiben, denn einige Entscheidungen werden verständlicher, wenn die durch die Ausrüstung gegebenen Randbedingungen bekannt sind.

Das lokale Rechnernetz besteht aus fünf Kleinrechnern vom Typ DIETZ System 621. Die Rechner sind über eine PDP11/20 untereinander und mit dem Zentralrechner des Fachbereichs verbunden, einem DEC-System-10 mit KI-10 Prozessor. (Technische Details siehe Abb.2.1)

Ausrüstung

Einer der Kleinrechner verwaltet die Video-Peripherie und besitzt zur schnellen Bildvorverarbeitung einen zweiten separaten Kernspeicher, der bei 8 bit/ Pixel Grauwertauflösung $2\frac{1}{2}$ digitisierte Vollbilder aufnehmen kann. Für diesen Rechner wurde von Dr. B. Neumann ein spezielles Betriebssystem geschrieben, das über einen digital gesteuerten Kreuzschienenverteiler die Video-Peripherie bedient und über die PDP11/20 mit den anderen Rechnern digitale Daten austauscht. Zur Peripherie gehören ein schneller Analog-Digital-Wandler, eine Farb- und zwei Schwarz-Weiß-Kameras, mehrere Filter und Monitore, zwei Analog-Plattenspeicher mit je zwei Kanälen à 600 TV-Bildern Kapazität, einem Video-Magnetband für Farbaufnahmen und einem Faksimileschreiber für Schwarz-Weiß-Bilder mit 16 Graustufen Auflösung.

Analogbilder können bis zu einem Monat problemlos auf den Analogplatten gespeichert werden; danach machen sich Alterungseffekte bemerkbar, insbesondere sinkt der Kontrast. Für die langfristige Datenhaltung müssen die Bilder in digitaler Form entweder auf den digitalen Plattenspeichern der Kleinrechner oder am Zentralrechner auf digitalen Plattenspeichern und Magnetbändern aufbewahrt werden. Zum Arbeiten mit Bildfolgen wird viel Speicherplatz benötigt. Bei 25 Bildern pro Sekunde und der von uns verwendeten Auflösung von 573 Zeilen und 512 Spalten mit 256 Graustufen müssen pro Sekunde ungefähr 7 Mega Byte (à 8 bit) gespeichert werden.

Die Speichermöglichkeiten sind für mich aus zwei Gründen wichtig: Erstens ist das System nicht in der Lage, Bildfolgen in Echtzeit zu digitisieren, ganz abgesehen von der Auswertung, und zweitens ist es günstiger, wenn verschiedene Verfahren stets anhand derselben Bildvorlagen verglichen und beurteilt werden können. Bildfolgen werden daher zunächst analog aufgezeichnet, was in Echtzeit möglich ist,

Ausrüstung

und anschließend bei Bedarf digital auf Platte oder Magnetband gespeichert. Es ist auch möglich, ohne digitale Zwischenspeicherung über den Analog-Digital-Wandler die Daten sequentiell direkt von der Analogquelle zu verarbeiten. Die zweite Vorgehensweise hat den Nachteil, daß man wegen Alterungseffekten, Digitisierungsrauschen und Zeilensynchronisationsfehlern keine genau reproduzierbaren Daten für Fehlersuche oder Vergleichsmessungen hat.

Der zweite für mich wichtige Kleinrechner verfügt über einen stark erweiterten Hauptspeicher, dessen Speicherkapazität den Befehlsadreßraum überschreitet und zum Teil nur über ein offenes Basisregister adressiert werden kann. Dieser Rechner besitzt ausserdem einen Prozessor für Gleitkommaoperationen und ein vielseitiges Farbrastersichtgerät der Firma COMTAL zur Darstellung von Graphiken, Farb- oder Schwarz-Weiß-Bildern, mit Funktions- und Falschfarbenspeichern, sowie mit einer Rollkugeleinheit. Für diesen Rechner wurde das Bildanalysesystem in der Sprache PASCAL entwickelt. Das Rastersichtgerät hat sich dabei als das wichtigste Werkzeug herausgestellt. Durch die verschiedenen Möglichkeiten zur Überlagerung von Grauwertbildern und Graphiken können Resultate unter mehreren Gesichtspunkten im Zusammenhang dargestellt, und bei Bedarf alle Zwischenschritte der Bildverarbeitung in einer geeigneten Grauwert- oder Farbverschlüsselung pixelgenau sichtbar gemacht werden.

An höheren Programmiersprachen stehen auf dem Rechner PASCAL und CONCURRENT PASCAL zur Verfügung. Die Übersetzer wurden von mehreren Studenten in Diplomarbeiten entwickelt. PASCAL hat sich als Implementationssprache für Bildverarbeitungsprogramme als sehr geeignet erwiesen. Zum einen ist es in dieser Sprache besonders leicht, größere Programme strukturiert zu entwerfen und nahezu selbstdokumentierend niederzuschreiben. Zum anderen verfügt das verwendete PASCAL-System über einen automatischen Segmentations- und

Ausrüstung

Seitenwechselmechanismus, so daß auch sehr lange Programme ausgeführt werden können, indem mit virtuellem Speicher gearbeitet wird. Eine Beschreibung des verwendeten Übersetzers ist in /Dreschler et al. 79b/ zu finden. Ein weiterer Vorteil von PASCAL ist, daß auf dem Zentralrechner ein nahezu kompatibler Übersetzer vorhanden ist, der quelltextbezogene Fehlersuchhilfen bietet. Ein großer Teil von Programmfehlern konnte so schon bei Probeläufen auf dem Zentralrechner beseitigt werden, ehe das Programm auf den Kleinrechner übertragen wurde. Dort war dann allerdings die Fehlersuche nur noch mit Hilfe von eingestreuten Kontrollausgaben oder hexadezimalen Speicher- und Registerauszügen möglich.

Nachteilig war leider, daß der PASCAL-Übersetzer für das DIETZ System 621 bei Beginn der Arbeit noch nicht fertiggestellt war und wichtige Module, wie die Kommunikation mit anderen Netzwerksrechnern, Gleitkommarechnung, externe Unterprogrammbibliotheken und die Möglichkeit zum Packen von binären Daten noch in der Entwicklung bzw. in der Testphase waren. Das Arbeiten mit diesem Übersetzer wäre sehr problematisch gewesen, wenn mich nicht die am Übersetzerbau beteiligten Diplomanden W. Benn, H. Faasch und B. Fischer stets hilfsbereit bei der Behebung von Übersetzerfehlern unterstützt hätten. Hierfür sei ihnen allen herzlich gedankt. Gleichfalls danken möchte ich B. Fischer, I. Heer, Dr. H. Kernen und Dr. B. Neumann für die Beratung beim Anschluß des Farbrastersichtgeräts an die Laufzeitunterstützungen von PASCAL und CONCURRENT PASCAL. B. Fischer und S. Shafer danke ich für die Mitarbeit bei der Entwicklung einer PASCAL Unterprogrammbibliothek.

Als Bilddaten wurden für die in dieser Arbeit vorgestellten Untersuchungen zwei Bildfolgen verwendet, die eine Strassenecke vor unserem Institut zeigen. Beide Folgen zeigen ein abbiegendes helles Taxi. Da die Fahrzeuge während des

ganzen Abbiegungsvorgangs zu sehen sein sollten und die Kamera nicht nachgeführt werden konnte, weil das verwendete Objektverfolgungsprogramm eine stationäre Kamera erfordert, sind die Bilder der Fahrzeuge recht klein. Bildfolge 1 wurde bei bedecktem Himmel aufgenommen, so daß keine starken Schlagschatten im Bild vorkommen, während Bildfolge 2 an einem sonnigen Tag entstanden ist. In Abb. I und Abb. II sind jeweils das erste und das letzte Bild der beiden Folgen zu sehen. Folge 1 umfaßt 23 Aufnahmen und Folge 2 45 Aufnahmen.

Die digitisierten Vollbilder haben ein Format von 573 Zeilen und 512 Spalten und eine Grauwertauflösung von 8 bit. Um Speicher und Rechenzeit zu sparen und die Bildgeometrie zu korrigieren, wurden die Vollbilder um den Faktor 6 komprimiert. Dabei wurden sie in Zeilenrichtung um den Faktor 3 und in Spaltenrichtung um den Faktor 2 verkleinert. Die integrierten Bilder nenne ich Geobilder und die Pixel der Geobilder Geopixel.

Da die aufgenommenen Fahrzeuge in den untersuchten Bildvorlagen sich so schnell bewegen, daß die Verschiebung zwischen dem ersten und dem zweiten Halbbild einer Aufnahme bis zu zwei Geopixel ausmachen kann, wurden bei der Integration stets nur Zeilen aus einem Halbbild verwendet. Andernfalls würden senkrecht zur Bewegungsrichtung verlaufende Objektkanten verwischt werden. Diese Integrationstechnik hat zur Folge, daß die eine Hälfte der Geopixel durch Mittelung von vier Vollbildpixeln entsteht und die andere Hälfte durch Mittelung von zwei Vollbildpixeln (Abb. 2.2). Die in Kapitel 4 erwähnten Median-gefilterten Bilder sind ebenfalls Geobilder, wobei das Filter vor der Integration auf die Vollbilder angewendet wurde. Auch bei der Filterung wurden beide Halbbilder getrennt behandelt, um die Bewegungsunschärfe nicht zu vergrößern.

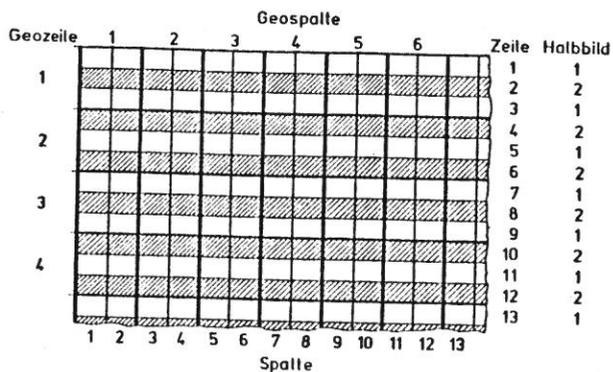


Abb. 2.2 Integration von Halbbildern

Es werden nur Zeilen des 1. Halbbildes berücksichtigt

3. Auswahl von Objektkandidaten

Die in den Kapiteln 4 und 5 beschriebenen Verfahren zur Punktefindung und Korrespondenzanalyse sind sehr rechenzeitaufwendig. Deshalb ist es vorteilhaft, wenn sie nicht auf das ganze Bild angewendet werden müssen, sondern auf die interessierenden Teilbereiche beschränkt werden können, nämlich auf die Abbildungen bewegter Objekte. Solche Teilbereiche werde ich hier als Objektkandidaten bezeichnen.

Ein weiterer wichtiger Grund dafür, zunächst nach Objektkandidaten zu suchen, ist der folgende: Für die 3D-Interpretation ist es wichtig zu wissen, wie sich die über mehrere Bilder verfolgten markanten Punkte zu Objekten gruppieren lassen. In Kapitel 6 werde ich hierauf noch näher eingehen.

Die Tabellen in Abschnitt 1.3 führen eine ganze Reihe von Verfahren an, die Objektkandidaten entdecken oder bewegte Objekte verfolgen. Nicht alle jedoch erfüllen die Randbedingungen, die ich an ein solches Verfahren stellen muß. Die Entdeckung von Objektkandidaten sollte vollautomatisch vor sich gehen, d.h. ohne Unterstützung durch einen Operateur, und ohne szenenspezifisches Wissen einzusetzen. Es sollte möglichst nur die Tatsache ausgewertet werden, daß sich durch Objektbewegungen das Bild verändert. Weiterhin müssen auch rotierende Objekte verfolgt werden können. Die Konturen der isolierten Objektkandidaten sollten möglichst genau mit denen der Objektabbildungen übereinstimmen. Diese zweite Forderung ist nicht trivial, da Rauschen und Synchronisationsfehler bei der Zeilendigitisierung ebenfalls zu Bildveränderungen führen können, die dann von den durch Objektbewegung verursachten Änderungen unterschieden werden müssen.

Auswahl von Objektkandidaten

Die Verfahren unterscheiden sich vor allem darin, wie die Bildunterschiede gemessen und interpretiert werden. Zwei grobe Kategorien sind die folgenden: Entweder werden die Bilder der Folge zunächst segmentiert und dann anhand des Segmentationsergebnisses verglichen, z.B. /Radig 78a, Kraasch et al. 79b, Tsuji et al. 79/, oder die Bilder werden sehr signalnah auf der Pixelbasis verglichen und so die Bewegung zur Segmentation eingesetzt. Bei der Segmentation durch Bewegung ist es schwer, saubere Bereichsgrenzen zu erhalten, da für ein Pixel die Entscheidung über die Objektzugehörigkeit nur aufgrund lokaler Umgebungen getroffen wird. Dieser Nachteil läßt sich ausgleichen, indem man für die provisorischen Objektkandidaten nachträglich eine Segmentierung vornimmt, die dann nicht so aufwendig ist, wie sie es für das ganze Bild wäre. Außerdem kann das Segmentationsergebnis für eine spätere Objektmodellierung nützlich sein.

Da das Verfahren von Jain et al. 77, besonders in der verbesserten Version von Blohm /Blohm 81/, befriedigende Ergebnisse liefert und außerdem auch als PASCAL-Programm auf dem Rechnernetz verfügbar ist, habe ich mich für dieses Programm entschieden. Das Programm wurde ursprünglich für das DEC-System-10 entwickelt und geht auf eine Diplomarbeit von D. Miltzer unter der Betreuung von H.-H. Nagel zurück. Das Verfahren basiert darauf, daß zeitliche Grauwertveränderungen durch ein modifiziertes statistisches Kantenkriterium nach Yakimovsky /Nagel 76/ erfaßt werden. Dieses Kriterium mißt die Wahrscheinlichkeit, daß Grauwertproben aus den beiden verschiedenen Bildern der gleichen Verteilung angehören, indem für 6×4 Testflächen (Geopixel) Mittelwert und Varianz ausgewertet werden. Das erste Bild einer Folge wird als Referenzbild gewählt und alle anderen Bilder der Folge mit diesem verglichen.

Auswahl von Objektkandidaten

Während der Auswertung werden zwei Zählmatrizen geführt, in denen für jedes Geopixel je eine Zahlvariable vorgesehen ist. Die beiden Matrizen heißen Differenzbild erster Ordnung (first order difference picture, FODP) und Differenzbild zweiter Ordnung (second order difference picture, SODP). Im Differenzbild zweiter Ordnung werden nur diejenigen Geopixel markiert, die sich im aktuellen Bild gegenüber dem Referenzbild signifikant verändert haben (Wechsel), während im Differenzbild erster Ordnung alle Wechsel vom zweiten Bild der Folge an gezählt werden. Unter den Voraussetzungen, daß das Objekt sich als starrer Körper bewegt, nicht verdeckt wird, auf der ganzen sichtbaren Oberfläche sich im Grauwert deutlich vom Hintergrund abhebt und daß es sich irgendwann innerhalb der Folge einmal um die volle Länge vom ursprünglichen Ort entfernt hat, kann aufgrund der Eintragungen in den beiden Differenzbildern die Position des Objektes im Referenzbild erkannt und das Bild des Objektes durch das des freigegebenen Hintergrundes im aktuellen Bild ersetzt werden. Auch die Objektposition im aktuellen Bild lässt sich ermitteln, ebenso Schätzwerte für Bewegungsrichtung und Geschwindigkeit.

Diese Interpretation der Differenzbilder wurde in Zusammenarbeit mit R. Jain entwickelt, der dafür eine Reihe von heuristischen Indizien mit Hilfe der Unschärfe-Logik (Fuzzy-Logic) ausgewertet hat.

Einen Teil der eingesetzten Heuristiken kann man sich an dem idealisierten Beispiel in Abb. 3.1 klarmachen, wo sich ein Objekt von konstantem Grauwert vor einem Hintergrund mit ebenfalls konstantem Grauwert bewegt.

Auswahl von Objektkandidaten

Objektposition	SODP	FODP
Bild1xxxx.	-----	-----
Bild2xxxx..1...1.1...1.
Bild3xxxx...11..11.12..12.
Bild4 ...xxxx....	...111.111.	...123.123.
Bild5 ..xxxx.....	..11111111.	..12341234.
Bild6 .xxxx.....	.1111.1111.	.123442345.

Abb. 3.1 Ideales bewegtes Objekt

Es ist deutlich erkennbar, daß im SODP Markierungen an der Vorderkante und an der Hinterkante des Objektes auftreten und daß die markierten Bereiche in der Ausdehnung der Verschiebung des Objektes gegenüber dem Referenzbild entsprechen. Wenn das Objekt sich um seine volle Länge weiterbewegt hat, verschmelzen die markierten Bereiche an der Vorder- und Hinterkante zu einem einzigen. Im Bild darauf sind wieder zwei SODP-Bereiche zu sehen, die angeben, wo sich die Abbildungen des bewegten Objektes im Referenzbild und im aktuellen Bild befinden. Im FODP fällt auf, daß die markierten Bereiche Zahlenfolgen enthalten, die entgegen der Bewegungsrichtung ansteigen. Aus diesen Folgen lassen sich Geschwindigkeit und Bewegungsrichtung, sowie Konsistenzinformation für die SODP-Bereiche ableiten.

S. und B. Schenkel /Schenkel and Schenkel 77/ haben die Auswertung des FODP gegenüber Rauschen unempfindlicher gemacht, indem sie eine Filterung eingeführt haben, die vereinzelte Markierungen löscht, wenn an dieser Stelle zu lange kein weiterer Wechsel aufgetreten ist. Außerdem wurden Ansätze für eine Datenbank implementiert, in der ein Benutzer alle wesentlichen Informationen über vorhandene Bildfolgen und Bearbeitungsergebnisse ablegen und im Dialog aktualisieren kann. Die Übertragung des Programms auf das DIETZ System

Auswahl von Objektkandidaten

621 und die Kommunikation mit dem Rechner, der die Videoperipherie steuert, wurden vorbereitet.

Diese Übertragung des Programms auf das Rechnernetzwerk wurden schließlich von B. Fischer und mir durchgeführt, nachdem der PASCAL-Übersetzer weit genug entwickelt war. Bei dieser Gelegenheit wurde auch das Farb-Raster-Sichtgerät als Ausgabemedium angeschlossen, so daß auf viel Schnelldruckerausgabe verzichtet werden konnte. Dieses umfangreiche Programm war abgesehen von der Selbstübersetzung der erste Härtestest für den Übersetzer, wobei viele Übersetzerfehler und sogar einige Fehler in der Rechnelektronik entdeckt und behoben werden konnten. Inzwischen haben H.-H. Nagel und E. Toth das Programm erheblich beschleunigt und weitere Möglichkeiten zur Messung von Bildunterschieden erprobt. Dieses war der Zustand, in dem ich das Programm zur Objektkandidatensuche übernommen habe.

Da das Programm für die Anfangsbilder der Folgen nur die SODP-Bereiche als Objektkandidaten liefern kann, so daß die Konturen nur ungefähr mit der Objektabbildung übereinstimmen, verwende ich zur Zeit die achsenparallelen umfassenden Rechtecke dieser Bereiche als Objektkandidaten. Da von S. Shafer schon Untersuchungen zur genauen Konturfindung durchgeführt worden sind und zur Zeit ein Diplomand ein verbessertes Verfahren implementiert /Blohm 81/, bin ich davon ausgegangen, daß ich dieses Problem bei meinen Untersuchungen ausklammern kann.

An dieses Programm stelle ich für die Zukunft noch die weitere Anforderung, daß es erkennen kann, ob ein Objekt verdeckt wird. Die Verdeckungsphase kann dann bei der 3D-Interpretation übergangen oder gesondert behandelt werden. Auf das Problem der Verdeckung wird in Kapitel 4 noch näher eingegangen.

Auswahl von Objektkandidaten

Für Prüfungszwecke besteht die Möglichkeit, die Objektkandidaten interaktiv am Sichtgerät auszuwählen, indem die Objektabbildungen mit der Rollkugel umfahren werden.

Punktefinder Übersicht

4. Punktefinder

4.1 Merkmalsfindung und Korrespondenzproblem

Die in diesem Kapitel vorgestellten Punktefinder müssen auch daraufhin untersucht werden, inwieweit sie das Lösen des Korrespondenzproblems unterstützen. Deshalb werde ich hier schon einige für die Diskussion des Korrespondenzproblems wichtige Begriffe einführen, obwohl das Korrespondenzproblem selbst erst in Kapitel 5 behandelt wird.

1. Bildpunkte sind durch Zeilen- und Spaltenkoordinaten angebbare Bildelemente. Im Gegensatz zu Pixeln, die nur durch diskrete Zeilen- und Spaltenwerte adressiert werden können, kann es für Bildpunkte sinnvoll sein, reellwertige Koordinaten anzugeben, z.B. interpolierte Werte.
2. Objektpunkte sind Punkte auf der Oberfläche von in der beobachteten Szene vorhandenen Objekten.
3. Zwei Bildpunkte aus zwei verschiedenen Aufnahmen einer Bildfolge oder aus der rechten und linken Aufnahme eines Stereo-Paares heißen korrespondierende Bildpunkte, wenn sie Bilder desselben Objektpunktes sind.
4. Der Differenzvektor zwischen den Bildkoordinaten zweier korrespondierender Bildpunkte heißt Verschiebung (Disparity).
5. Das Korrespondenzproblem besteht darin, korrespondierende Bildpunkte zu finden. In der Regel ist es auch wichtig, die entsprechenden Verschiebungen zu messen. Da das Korrespondenzproblem ohne einschränkende heuristische Regeln (siehe Kap. 5)

Punktefinder Übersicht

nicht eindeutig lösbar ist, heißen die von einem Programm zur Lösung des Korrespondenzproblems gewählten Paare von Bildpunkten Kandidaten für korrespondierende Bildpunkte.

Die größte Schwierigkeit beim Korrespondenzproblem besteht in der kombinatorischen Vielfalt von Lösungsmöglichkeiten. Ohne Regeln, die den Suchraum beschränken, kommt jedes Paar von Punkten aus zwei Bildern als Kandidat für ein Paar von korrespondierenden Bildpunkten in Frage.

Marr et al. /Marr and Poggio 79, Grimson and Marr '79/ diskutieren in ihrer Arbeit zwei grundsätzliche Möglichkeiten zur Suchraumbeschränkung. Beide Möglichkeiten basieren auf einer Beschränkung des Wertevorrats für die zugelassenen Verschiebungen. Zum einen könnte man eine Obergrenze für den Betrag der möglichen Verschiebungen festlegen und zum anderen könnte die Genauigkeit der Verschiebungsmessung herabgesetzt werden, indem nur noch diskrete Werte zugelassen werden. Marr et al. setzen eine Kombination beider Verfahren ein, wobei sie von folgenden Überlegungen ausgehen: Gesetzt der Fall, man hätte Aufnahmen, die mit einem Bandpaß gefiltert worden wären, der nur Signale mit einer Wellenlänge größer L durchläßt, dann könnten im Bild keine Merkmale vorkommen, deren Abstand kleiner ist als L . Wenn man gleichzeitig voraussetzen kann, daß keine Verschiebung mit einem Betrag größer als $L/2$ vorkommt, dann ist das Korrespondenzproblem eindeutig lösbar. Die Genauigkeit der Verschiebungsmessung kann iterativ verbessert werden, indem die Ergebnisse von Filtern verschiedener Kanäle kombiniert werden. Mithilfe der langwelligen Filter wird der Suchraum für die kurzwelligen Bilder beschränkt.

Punktefinder Übersicht

Marr et al. verwenden Filter mit einer Bandbreite von ein bis zwei Oktaven. Dabei stützen sie sich auf Ergebnisse von Campbell und Robson /Campbell and Robson 68/, Wilson und Giese /Wilson and Giese 77/ und Cowan /Cowan 77/, die zeigen, daß im visuellen System des Menschen Rezeptoren vorhanden sein müssen, die auf balkenförmige Bildstrukturen ansprechen. Diese Rezeptoren kommen in drei bis vier verschiedenen Größen vor und entsprechen Filtern von 1 bis 2 Oktaven Bandbreite. Genau diese Möglichkeiten des Auges wollten Marr et al. in ihrem Modell der Stereo-Wahrnehmung des Menschen simulieren.

Die Filterung in verschiedenen Bandbereichen ist nicht nur günstig für das Korrespondenzproblem, sondern erleichtert auch die Merkmalsfindung.

Marr und Hildreth /Marr and Hildreth 80/ postulieren, daß starke Kanten an einem Objekt im Bild nicht nur in einem Kanal einen Beitrag geben, sondern stets auch in den benachbarten Kanälen. Daher ist es möglich, durch Analyse des Bildsignals in mehreren getrennten Kanälen die Zuverlässigkeit von Merkmalsfindern zu verbessern. Auf die Merkmalsfindung gehe ich im Zusammenhang mit den Punktefindern im Abschnitt 4.3 ein.

Wenn es für eine bestimmte Anwendung unnötig sein sollte, alle Paare von korrespondierenden Bildpunkten zu finden, gibt es eine weitere Möglichkeit der Suchraumbeschränkung: Man bestimmt in beiden Bildern die Teilmenge derjenigen Bildpunkte, für die einen die korrespondierenden Bildpunkte interessieren. Bei der Suche nach Kandidaten für korrespondierende Bildpunkte beschränkt man sich auf die Elemente dieser Teilmengen. Diese Methode verwenden z.B. Barnard und Thompson /Barnard und Thompson 79, 80/ und Ranade und Rosenfeld /Ranade and Rosenfeld 80/. In dem hier beschriebenen Bildanalyseprogramm wurde diese zweite Möglichkeit eingesetzt, da das 3D-Rekonstruktionsprogramm von Boude und

Punktefinder Übersicht

Nagel nur für wenige Stützpunkte die Messung der Verschiebungen erfordert.

4.2 Meßpunkte für die 3D-Modellierung

Für die 3D-Modellierung sind als Meßpunkte körperfeste Objektpunkte erforderlich, deren Bildpunkte eine auffällige Umgebung haben und sich im Bild pixelgenau festlegen lassen. Die Bildpunkte müssen sich auch in verrauschten Bildern über mehrere Aufnahmen sicher verfolgen lassen. Solche Merkmale am Objekt sind Ecken, isolierte Punkte und Linienendpunkte, da sie auch im Bild als Ecken und Punkte erscheinen. Diese für die 3D-Modellierung geeigneten körperfesten Objektpunkte werde ich hier als markante Objektpunkte bezeichnen. Im Gegensatz hierzu nenne ich Bildpunkte mit auffälliger Umgebung markante Bildpunkte. Auffällige Umgebungen sind zum Beispiel helle oder dunkle Flecken, Kantenschnittpunkte und Linienendpunkte. Zwischen markanten Objektpunkten und markanten Bildpunkten muß deutlich unterschieden werden: Markante Bildpunkte sind nicht immer die Bilder markanter Objektpunkte und umgekehrt ist nicht jedes Bild eines markanten Objektpunktes ein markanter Bildpunkt. Auf Objektoberflächen können durch Verdeckung, Abschattung oder Spiegelung Grenzlinien sichtbar werden, die durch die Objektbewegung über die Oberfläche wandern können. Auch Punkte auf diesen nicht körperfesten Grenzlinien können markante Bildpunkte als Bild haben. Ein Punktefinder kann nicht entscheiden, ob ein solcher markanter Bildpunkt das Bild eines körperfesten Objektpunktes ist oder nicht, wenn er die Einzelbilder der Folge unabhängig voneinander auswertet und kein szenenspezifisches Wissen zur Verfügung steht. Daher werden markante Bildpunkte zunächst als Kandidaten für die Bilder markanter Objektpunkte betrachtet und die genauere Analyse auf später verschoben. "Kandidat

Punktefinder Übersicht

für das Bild eines markanten Objektpunktes" werde ich in Zukunft mit markanter Punkt abkürzen.

Die Trennung in körperfeste und nicht körperfeste Punkte kann frühestens dann erfolgen, wenn die Korrespondenz von Punkten untersucht wird und man die vermuteten Verschiebungen auf Konsistenz mit dem optischen Fluß /Clocksin 78, 80a, 80b/ in der Umgebung überprüfen kann.

Diese Art der Konsistenzprüfung ist allerdings nur dann möglich, wenn nur ein sehr geringer Prozentsatz an unzulässigen Punkten herausgefiltert werden muß und die Mehrzahl der Punkte körperfest ist. Zum Teil soll später das Objektverfolgungsprogramm diese Aufgabe erfüllen, da es die Objektkandidaten sauber vom Hintergrund trennen und bei Verdeckung warnen soll. Abschattung ist ähnlich problematisch wie Verdeckung, wenn scharf begrenzte Schlagschatten vorhanden sind. Daher wird die beim norddeutschen Wetter übliche diffuse Beleuchtung erwartet, so daß auch keine grellen Glanzlichter oder Spiegelungen von Punktlichtquellen vorkommen können. Glanzlichter, die dennoch entstehen, bilden sich an stark gewölbten Oberflächenstellen und sind meistens körperfest. Wenn dann noch das Objekt nicht zu sehr spiegelt und keine Löcher enthält, durch die der Hintergrund sichtbar ist, sind weitgehend alle Quellen für nicht körperfeste Punkte beseitigt.

Es bleiben noch diejenigen unerwünschten Punkte, die an konkaven Stellen des Körpers durch Selbstverdeckung oder Selbstabschattung entstehen können. Selbstabschattung entfällt wegen der Beleuchtungsbedingungen, aber Selbstverdeckung ist ein echtes Problem. Es könnte nur dadurch ausgeschlossen werden, daß nur konvexe Körper zugelassen werden, was jedoch den Anwendungsbereich dieses Programms zu stark einschränken würde. Ich fordere daher nur, daß die untersuchten Objekte näherungsweise konvex sind und nur so

Punktefinder Übersicht

viele markante Bildpunkte durch Selbstverdeckung entstehen dürfen, daß die Mehrzahl der markanten Bildpunkte Bilder markanter Objektpunkte ist. Wenn nur wenige falsche markante Punkte gefunden worden sind, sollten sie bei späteren Verarbeitungsschritten (Korrespondenzanalyse, 3D-Rekonstruktion) auffallen und von weiteren Untersuchungen ausgeschlossen werden können.

Eine Ausnahme in dieser Hinsicht bilden Szenen mit Polyedern. Roach und Aggarwal z.B. /Roach and Aggarwal 79/, die Bildfolgen mit bewegten konvexen Blöcken untersuchen, nutzen aus, daß an verdeckenden Kanten in der Regel T-Verbindungen entstehen /Duda and Hart 73, Abschnitt 12.4/, die durch eine einfache Prüfung von den weiteren Untersuchungen ausgeschlossen werden können.

Von einem Punktefinder sollte man auch erwarten, daß er Wissen über die Umgebungen der ausgewählten markanten Punkte aufbereitet, so daß die Analyse der Korrespondenzen erleichtert wird.

4.3 Einige bekannte Verfahren

In der Literatur werden im wesentlichen drei Arten von Punktefindern beschrieben: Erstens Punktefinder, die auf einer Bildsegmentation aufbauen, zweitens Punktefinder, die den Verlauf der Bildfunktion analytisch durch Approximationsfunktionen untersuchen, und drittens heuristische Verfahren.

Diejenigen Punktefinder, die ohne vorherige Bildsegmentation auskommen, verwenden in der Regel Operatoren, die lokal auf quadratische oder kreisförmige Bildausschnitte angewendet werden und lokale Auffälligkeitsmaße errechnen. Die Seitenlänge oder den Durchmesser dieser Bildausschnitte

Punktefinder Übersicht

nenne ich Operatordurchmesser. Der Operatordurchmesser ist eine wichtige Größe, um die Rauschempfindlichkeit und das Auflösungsvermögen, d.h. die Fähigkeit zum Trennen von eng benachbarten Bilddetails, zu beurteilen.

Punktefindung durch Segmentation wird besonders bei Blockswelt-Modellen eingesetzt. Ich werde hier nur auf wenige Beispiele eingehen, da für mich die lokalen Operatoren wichtiger sind, die ohne vorherige Segmentation auskommen.

Perkins und Binford /Perkins and Binford 73/ benutzen Modelle, die typische Eckenkonfigurationen beschreiben, und verwenden Heuristiken, um im segmentierten Bild den Suchraum für die Modelle einzuschränken. Ist das Suchergebnis negativ, wird lokal das Segmentationsergebnis anhand des Modells überprüft, so daß auch noch schwach kontrastierende Kanten gefunden werden können.

Kraasch et al. /Kraasch et al. 79a und 79b/ segmentieren das Bild mittels eines Bereichsfinders und approximieren Bereichsgrenzen durch Geradenstücke. Das Segmentationsergebnis wird als Graphstruktur aufbereitet. Schnittpunkte von Geraden werden in dieser Graphdarstellung durch einen Interpretierer mit Modellen von markanten Punkten verglichen.

Bullock /Bullock 76/ verwendet den Hückel-Operator, um Elementarkanten zu finden, und gruppiert diese mittels der Hough-Transformation zu Geradenstücken. Schnittpunkte von Geraden werden analytisch durch Lösen von Gleichungssystemen gefunden und die Schnittwinkel ermittelt. Das Analyseergebnis wird durch einen minimalen spannenden Baum repräsentiert, der als Knotenpunkte die Schnittpunkte der Geraden sowie Information über Zahl der Geraden und Schnittwinkel enthält. Außerdem sucht Bullock mit Hilfe einer Texturtransformation Punkte mit extremer

Punktefinder
Übersicht

Texturauffälligkeit, die ebenfalls in den minimalen spannenden Baum aufgenommen werden. Auch die Referenzmodelle sind durch einen solchen minimalen spannenden Baum beschrieben, so daß die Modellsuche durch einen Graphvergleich erfolgen kann.

Feng und Pavlidis /Feng and Pavlidis 73/ geben ein schnelles Verfahren an, um in segmentierten Bildern Punkte zu finden, an denen mindestens drei Bereiche zusammentreffen.

Die analytischen Verfahren versuchen aus Näherungsfunktionen Strukturaussagen über das Grauwertgebirge der Bildfunktion abzuleiten.

Beaudet /Beaudet 78/ entwickelt eine Taylor-Reihendarstellung der Bildfunktion und beschreibt verschiedene rotationsinvariante Operatoren, die sich aus den Ableitungen der Funktion aufbauen lassen und mit deren Hilfe sich Kanten, Ecken, Sattelpunkte und gekrümmte Linien finden lassen.

Hsu et al. /Hsu et al. 78/ berechnen aufbauend auf den Operatoren von Beaudet die Hauptkrümmungen und Hauptkrümmungsrichtungen der Bildfunktion und geben eine Gewebe-Grammatik (Web Grammar) an, mit der lokale Struktureigenschaften zu globalen Strukturaussagen verallgemeinert werden können.

Enomoto und Katayama /Enomoto and Katayama 76/ suchen im Bild Linien, auf denen Krümmungen und Gradienten bestimmte Voraussetzungen erfüllen. Diese Linien enthalten wichtige Strukturinformationen, mit deren Hilfe eine Bildrekonstruktion möglich ist. Die Schnittpunkte der Strukturlinien sind entweder Extremwerte oder Sattelpunkte.

Punktefinder Übersicht

Paton /Paton 75/ entwickelt die Bildfunktion auf kreisförmigen, radial gewichteten Stützflächen durch Legendre-Polynome. Er zeigt, wie man durch Interpretation der Koeffizienten entscheiden kann, ob lokal eine Ebene, ein Kamm, ein Tal, ein Gipfel, eine Senke, ein Sattel oder eine mehrdeutige Situation vorliegt.

Hummel /Hummel 79/ schlägt vor, für jede Art von Merkmalsuche denjenigen Satz von orthonormalen Basisfunktionen für die Bildapproximation zu wählen, der die Eigenschaften des gesuchten Merkmals am wenigsten verfälscht. Er zeigt, wie man aus einer Merkmalsbeschreibung mittels der Karhunen-Loeve-Entwicklung den optimalen Basisfunktionensatz ableiten kann. Als Beispiel dient ein dem Hückel-Operator ähnlicher Kantenfinder mit optimalen Basisfunktionen.

Dunn /Dunn 75/ beweist, daß eine bestimmte Gruppe von invarianten Transformationen Kerne besitzt, die als Kanten- oder Eckenfinder wirken.

Marr und Poggio 79 /Marr and Poggio 79/ verwenden in ihrem theoretischen Modell der menschlichen Stereowahrnehmung Kantensegmente, Linien- und Kantenendpunkte als Basiselemente für die Suche nach korrespondierenden Bildpunkten. Die Merkmalsextraktion geht wie folgt vor sich: Zunächst werden beide Bilder mithilfe von balkenförmigen Masken gefiltert, deren Querschnitt im Profil der Differenz von zwei Gaußfunktionen gleicht. Die Masken werden in vier verschiedenen Größen und zwölf verschiedenen Orientierungen angewendet, wobei die Maskengrößen die Durchlaßfrequenz des Filters bestimmen. Die Filter arbeiten als sich überlappende Bandpässe mit 1 bis 2 Oktaven Bandbreite. Da die Masken so gewählt worden sind, daß die Konvolution der Bildfunktion mit diesen Masken eine richtungsabhängige zweite Ableitung liefert, entsprechen in den gefilterten Bildern Nulldurchgänge senkrecht zur Maskenrichtung (zero crossing

Punktefinder Übersicht

segments) Kantenelementen im Originalbild. (Nullstellen der zweiten Ableitung entsprechen Extremwerten der ersten Ableitung). Zusätzlich zu solchen Kantenelementen wird nach Linien- und Kantenendpunkten gesucht. Bei der Suche nach korrespondierenden Bildelementen wird das in 4.1 beschriebene hierarchische Verfahren angewendet. Das Ergebnis wird in einem dynamischen Puffer gespeichert, der 2 1/2 D-Skizze genannt wird (siehe auch /Marr and Nishihara 78/). Die Implementation dieses hierarchischen Verfahrens wird in /Grimson and Marr 79/ beschrieben.

Marr und Hildreth /Marr and Hildreth 80/ beschreiben eine Verbesserungsmöglichkeit für die Suche nach Nulldurchgängen der zweiten Ableitung. Sie verwenden wie Marr und Poggio gaußsche Filter von 1 bis 2 Oktaven Bandbreite. Da in gefilterten, bandbreitebeschränkten Bildern bestimmte Kontinuitätsbedingungen erfüllt sind, kann anstelle der gerichteten zweiten Ableitung der Laplace-Operator verwendet werden. Die Konvolution mit Masken in verschiedenen Orientierungen kann daher durch die Konvolution mit einer einzigen ungerichteten Maske ersetzt werden. In dem gefilterten Bild $\nabla^2 G(x,y) * I(x,y)$ wird nach Nullstellen gesucht, die zu primitiven, gerichteten Bildelementen zusammengefaßt werden. Diese Bildelemente werden Nullstellensegmente genannt (zero-crossing-segments). Da diese Nullstellensegmente ebenfalls an den Stellen mit dem steilsten Gradienten liegen, verwenden Marr und Hildreth den von Marr und Poggio benutzten Namen. Marr und Hildreth geben einen Algorithmus an, wie sich die Nullstellensegmente aus verschiedenen Kanälen zu einer einzigen Bildbeschreibung zusammenfassen lassen. Diese Bildbeschreibung nennen sie "raw primal sketch".

Punktefinder Übersicht

Die Autoren sehen in diesem Ansatz im wesentlichen drei Vorteile: Erstens ist es in den bandbreitebeschränkten Bildern möglich, einfache Operatoren wie den Laplace-Operator anzuwenden, zweitens halten sie die Bildbeschreibung durch Nullstellensegmente für vollständig und reversibel, und drittens ist die Kantenfindung sicherer, wenn die Information aus mehreren Kanälen gleichzeitig ausgewertet wird. Marr und Hildreth gehen davon aus, daß Kanten, die von physikalischen Objekten herrühren, nicht nur in einem Kanal, sondern auch in den Nachbarkanälen ein Signal hervorrufen.

Kitchen und Rosenfeld /Kitchen and Rosenfeld 80/ beschreiben Experimente mit fünf verschiedenen Operatoren, die ohne vorangehende Bildsegmentation lokal Maße für die Ablenkung einer Kante oder Linie vom geradlinigen Verlauf berechnen.

Der erste vorgeschlagene Operator berechnet Gradientenrichtung und Betrag. Werden die für alle Pixel des Bildes errechneten Gradientenrichtungen als Grauwertbild dargestellt, sieht man Änderungen an denjenigen Stellen, wo im Ursprungsbild Richtungsänderungen der Kanten oder Linien sind. Der erste vorgeschlagene Operator sucht daher nach hohen Gradientenbeträgen im Gradientenrichtungsbild. Der so gefundene Operatorwert und die aller anderen von den Autoren vorgeschlagenen Operatoren, werden zusätzlich mit dem Gradientenbetrag des Ursprungsbildes multipliziert, so daß nur an Stellen mit starker Grauwertvariation hohe Operatorbeiträge auftreten können.

Eine verbesserte Version dieses Operators sucht gezielt entlang Kanten nach Richtungsänderungen. Hierzu wird wieder für jedes Pixel die Gradientenrichtung ermittelt und zwei sich gegenüberliegende Nachbarn des Pixels ausgewählt, die am besten auf einer Geraden senkrecht zur Gradientenrichtung liegen. Für diese beiden Pixel wird die Differenz in der

Punktefinder Übersicht

Gradientenrichtung errechnet. Damit erhält man für eine Kante, die durch das zentrale Pixel geht, einen Schätzwert für die Ablenkung vom geradlinigen Verlauf.

Methode 3 besteht darin, auf einem 3×3 Bildausschnitt die beiden Pixel zu suchen, die dem mittleren Pixel im Grauwert am ähnlichsten sind. Ermittelt man die Verbindungsgeraden dieser beiden Pixel mit dem mittleren Pixel, so ist der Richtungsunterschied zwischen diesen beiden Geraden ein Maß für die Richtungsänderung einer Kante, die durch das mittlere Pixel geht. Dieser Operator wird von den beiden Autoren als extrem rauschempfindlich beschrieben.

Die vierte Methode besteht darin, die Bildfunktion auf quadratischen Bildausschnitten durch eine Näherungsfunktion zu approximieren. So lassen sich die Verfahren 2 und 3 weniger rauschempfindlich realisieren, indem alle Grauwerte und Gradienten mittels der Näherungsfunktion errechnet werden. Interessanterweise führen die analytischen Versionen von Methode 2 und 3 beide auf dieselbe Formel für die Richtungsänderung.

Der fünfte untersuchte Operator ist der von Beaudet vorgeschlagene Operator DET, der in Abschnitt 4.4 noch ausführlich besprochen werden wird. Kitchen und Rosenfeld kritisieren, daß dieser Operator an sehr starken Kanten schlechte Resultate liefert.

Für alle Operatoren schlagen Kitchen und Rosenfeld vor, zur Verbesserung des Ergebnisses die Resultate von Operatoren mit verschiedenen Durchmessern pixelweise miteinander zu multiplizieren. Außerdem empfehlen sie, nur lokale Maxima des Operatorwertes zu berücksichtigen (siehe /Moravec 77, 79, 80/). Auch die Multiplikation mit dem Betrag des Gradienten kann bessere Ergebnisse bringen, wenn nur lokale Maxima des Gradienten als Faktoren zugelassen werden.

Punktefinder Übersicht

Moravec /Moravec 77, 79, 80/ berechnet auf einem quadratischen Testfeld in vier Richtungen ein Varianzmaß und akzeptiert einen Punkt dann als auffällig, wenn die kleinste der vier Zahlen ein lokales Maximum ist und einen Schwellwert überschreitet. Genaueres hierzu steht im nächsten Abschnitt.

Halé /Halé 76/ schlägt ein ganzes Sortiment von kreuzförmigen Operatoren vor, die ähnlich dem Moravec-Operator in mehreren zueinander senkrechten Richtungen die Grauwertvariation messen. Die Maße für die Variation reichen von einfacher Differenz über die Varianz bis zur Krümmung, je nach gewünschter Genauigkeit. Es wird gezeigt, wie aus den Variationsmaßen Strukturaussagen abgeleitet werden können.

Hayashi und Suto /Hayashi and Suto 77/ definieren einen ganzen Satz von elementaren Bildelementen, den sie geometrische Faktoren nennen. Unter Verwendung des Perceptronkonzepts zeigen die Autoren, daß die geometrischen Faktoren durch lokale Berechnungen ermittelt werden können.

Peucker und Douglas /Peucker and Douglas 75/ untersuchen diskrete Terrain-Elevationsdaten, indem sie die einen Punkt umgebenden acht Punkte im Kreis durchlaufen und für jeden Punkt die Höhendifferenz zum Mittelpunkt errechnen. Aus der Differenzenfolge läßt sich durch einfache logische Überlegungen feststellen, wie das Gelände in dem untersuchten Punkt geformt ist.

Hannah /Hannah 74/ benötigt für die Stereo-Analyse kleine rechteckige Bildausschnitte, die durch Kreuzkorrelation oder ähnliche Techniken im zweiten Bild des Paares sicher wiedergefunden werden können. Ein Bildausschnitt ist nur dann mit Sicherheit wiederzufinden, wenn die Korrelation ein deutliches Maximum hat. Aus diesem Grund läßt Hannah nur

Punktefinder Übersicht

diejenigen Ausschnitte im ersten Bild als Suchkandidaten zu, die ausgeprägte Autokorrelationsmaxima haben. Dazu wird die Autokorrelationsfunktion durch eine Gaußfunktion approximiert und die Form des Maximums mittels der Funktionsparameter beurteilt. Auf diese Weise werden Flächen mit wenig Struktur oder Kanten als Korrelationskandidaten vermieden. Die Autokorrelation liefert auch gleich einen Anhaltspunkt für die Akzeptanzschwelle bei der Suche im zweiten Bild. Hannahs Punktauswahlkriterium ist verwandt mit den Verfahren von Moravec und Halé.

Da die Koordinaten der gefundenen Punkte für eine 3D-Rekonstruktion eingesetzt werden sollen, ist für mich die wichtigste Anforderung an einen Punktefinder, daß er die Koordinaten von Auffälligkeiten pixelgenau und sicher ermittelt. Die zweite wichtige Bedingung ist, daß das Verfahren möglichst allgemein einsetzbar sein muß. Wenn die Art der Bildvorlagen nicht eingeschränkt werden soll, darf kein Verfahren gewählt werden, das das Bild segmentiert und anschließend Konturzüge durch Geraden, Bögen und ähnliche Strukturen approximiert. Diese Methode kann zu Näherungsfehlern führen, wenn nicht bekannt ist, welchen Gesetzmäßigkeiten die Kontur des Objektes genügt und wie sie durch die Kamera abgebildet wird. Attneaves Bild von der schlafenden Katze zeigt allerdings, daß diese Näherungsfehler für den visuellen Eindruck in der Regel nicht wichtig sind /Duda and Hart 73/. In diesem Bild wurde die Kontur zwischen den Punkten stärkster Krümmung durch Geradenstücke ersetzt. Ob die Konturpunkte stärkster Krümmung jedoch in verrauschten Bildern sicher und reproduzierbar genug bestimmbar sind, ist zweifelhaft.

Noch unsicherer erscheint mir dieses bei der iterativen Endpunktanpassung /Duda und Hart 73/. Ein Grund hierfür wird sein, daß solche Näherungsverfahren den Fehler für die Kontur als Ganzes möglichst klein halten wollen und nicht

Punktfinder Übersicht

speziell auf die Lage der Unstetigkeiten achten. Zur genauen Punktfindung benötigt man daher ein Verfahren, bei dem im Grauwertbild direkt nach Punkten gesucht wird und nicht erst nach Strukturen höherer Ordnung (z.B. Geraden). Wenn dann erst auf höherer Ebene die Lage von Punkten indirekt erschlossen wird, sind bei der Konturapproximation aufgetretene Näherungsfehler nicht mehr rückgängig zu machen. Außerdem sollte der Punktfinder möglichst gezielt nach Ecken, isolierten Punkten und Linienendpunkten suchen. Dieses setzt voraus, daß man eine Modellvorstellung von der Form des Grauwertgebirges in der Umgebung solcher Strukturen besitzt. Im Gegensatz hierzu verwendet Hannah /Hannah 74/ alle Punkte mit geeigneter Autokorrelation, ohne zu untersuchen, warum das Autokorrelationsverhalten so günstig ist.

4.4 Erfahrungen mit dem Punktfinder von Moravec

Moravec /Moravec 77, 79, 80/ verwendet zur Steuerung eines autonomen Fahrzeugs einen Punktfinder, der Ähnlichkeit mit den von Halé /Halé 76/ entwickelten Operatoren hat. Auf einem quadratischen Testfeld (4*4 bis 8*8 Pixel) werden in horizontaler, vertikaler und diagonaler Richtung jeweils die Quadrate der Grauwertdifferenzen zu den benachbarten Pixeln summiert (siehe Abb. 4.1). Damit in alle vier Summen die gleiche Anzahl von Summanden eingeht, werden für die Diagonalsummen auch außerhalb der Operatorfläche liegende Pixel berücksichtigt. Das Minimum der vier Summen wird als Auffälligkeitsmaß für das mittlere Pixel des Testfeldes genommen. Im Unterschied zum Operator von Halé werden dabei alle Pixel innerhalb des Testfeldes erfaßt.

Punktefinder
Moravec

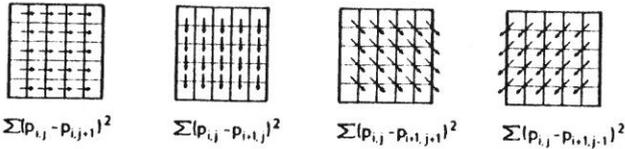


Abb. 4.1: Der Moravec-Operator

Die Idee ist, daß nur isolierte Punkte oder Ecken in alle vier Richtungen eine hohe Differenzensumme liefern, während Kanten dagegen parallel zu einer der Abtastrichtungen verlaufen und so in dieser Richtung keinen Beitrag aufweisen. Als markante Punkte werden alle lokalen Maxima des Auffälligkeitsmaßes genommen, die einen Schwellwert überschreiten.

Mit diesem Operator habe ich zu Anfang experimentiert, da er einfach zu implementieren ist und auch Barnard und Thompson /Barnard und Thompson 79, 80/ für viele verschiedenartige Realweltvorlagen über gute Ergebnisse mit diesem Operator berichtet haben. Für die von mir gewünschte Genauigkeit erwies der Operator sich jedoch nicht als ausreichend. Moravec setzt den Operator nicht auf jedes Pixel an, sondern in Abständen von jeweils einem halben Operatordurchmesser. Das bedeutet, daß sich die Operatorflächen jeweils um den halben Operatordurchmesser überlappen. Bei den von mir verwendeten 5*5 Matrizen würde eine Überlappung von 2 Pixeln bedeuten, daß der Operator nur in jeder 3. Zeile und 3. Spalte angewendet werden würde, so daß nur für jedes neunte Pixel das Auffälligkeitsmaß berechnet werden würde. Dies verkürzt die Rechenzeit und verdeckt durch geringe Auflösung einige Nachteile des Operators. Mir kommt es jedoch für die 3D-Rekonstruktion auf eine pixelgenaue Lokalisation von auffälligen Stellen an, da ich ja ohnehin

Punktefinder
Moravec

schon mit integrierten Bildern arbeite. Diese Genauigkeit wird zusätzlich durch einen Unterschied bei der Analyse von Korrespondenzen erfordert. Moravec wählt im ersten Bild einen Satz von Punkten aus und lokalisiert diese Punkte im zweiten Bild durch Korrelation. Ich dagegen suche im ersten und im folgenden Bild zunächst Punkte und versuche anschließend Kandidaten für korrespondierende Bildpunkte in den beiden Punktmengen zu finden. Daher muß ich die Punkte viel genauer auswählen als Moravec, da bei meinem Zuordnungsverfahren sich Ungenauigkeiten bei der Punktauswahl verdoppeln können, während bei Moravecs Technik Ungenauigkeiten durch sorgfältige Korrelation ausgeglichen werden können. Daher wende ich den Operator auf jedes Pixel an und speichere zur Rechenzeitverringerung alle einmal berechneten Quadrate von Grauwertdifferenzen, so daß keine Doppelrechnungen nötig sind.

Punktfinder
Moravec

Tab. 4.1 Idealisierte Bildelemente

Bildelement	Ergebnis des Moravec-Operators 5*5
*****
*****
*****
*****	..00000..
*****	..00000..
*****	..00000..
-----	..00000..
-----	..00000..
-----
-----

a. Kante in Hauptachsenrichtung

*****_
*****_
*****_
*****_	..00000..
*****_	..00000..
*****_	..00000..
*****_	..00000..
*****_	..00000..
*****_
*****_

b. Kante in Diagonalrichtung

*****
*****
*****_	..00112..
*****_	..11222..
*****_	..22222..
*****_	..22222..
*****_	..22211..
*****_
*****_

c. Kante mit 50% Steigung

Punktefinder
Moravec

```
*****
*****_
*****_
*****_
*****_
***_
**_
-----
-----
-----
.....
.....
.....00111...
.....11221...
.....11221...
.....11221...
.....21110...
.....21110...
.....
.....
.....
```

d. Kante mit 66,6 % Steigung

```
*****
*****
*****_
*****_
***_
*_
-----
-----
-----
.....
.....
.....11211...
.....11211...
.....11211...
.....11211...
.....11100...
.....
.....
.....
```

e. Kante mit 33,3 % Steigung

```
*****
*****
*****
*****
*****_
***_
-----
-----
-----
.....
.....
.....00011...
.....11111...
.....11111...
.....11111...
.....11111...
.....11111...
.....
.....
.....
```

f. Kante mit 25% Steigung

```
*****_
*****_
*****_
*****_
*****_
*****_
-----
-----
-----
.....
.....
.....00000...
.....04321...
.....03321...
.....02221...
.....01111...
.....
.....
.....
```

g. Ecke

Punktefinder
Moravec

*****
*****
*****
***	...13431...
*	...24442...
	...24442...
	...23332...
	...11111...

h. Ecke

*
*
*	...44444...
*	...48884...
*****	...48884...
*	...48884...
*	...44444...
*
*

i. Kreuz

***
***
***
*****	...32223...
*****	...22222...
*****	...22422...
*****	...22222...
***	...32223...
***
***

j. Kreuz

-----
-----
***	...22322...
***	...23332...
***	...33633...
***	...23332...
-----	...22322...
-----
-----

k. Punkt

Punktefinder
Moravec

```

-----
-----
-----
-----*-----
-----
-----
-----
-----
-----

```

```

.....
.....
...01110...
...12221...
...12221...
...12221...
...11111...
.....
.....

```

l. Punkt

```

-----
-----
-----
*****-----
-----
-----
-----
-----
-----

```

```

.....
.....
...01110...
...01111...
...01111...
...01111...
...01111...
.....
.....

```

m. Endpunkt

```

**-----
***-----
****-----
*****-----
*****-----
-----
-----
-----
-----

```

```

.....
.....
...00000...
...34321...
...23321...
...12221...
...01111...
.....
.....

```

n. Eckpunkt

```

*****-----
*****-----
*****-----
*****-----
*****-----
*****-----
*****-----
*****-----
**-----

```

```

.....
.....
...00000...
...01111...
...02221...
...01111...
...00000...
.....
.....

```

o. Ecke

Tab. 4.1 zeigt für typische idealisierte Bildelemente das Verhalten des Moravec Operators an Binärbildern für die von mir verwendeten 5×5 Fenster. Vier Dinge fallen daran auf:

1. Erstens liefert der Operator an Ecken nicht das maximale Signal, wenn er direkt über der Spitze zentriert ist, sondern dann, wenn der Operator bei beiden beteiligten Kanten gleichzeitig den größten möglichen Teil überdeckt, da so die meisten Grauwertdifferenzen verschieden von Null entstehen. Daher erzeugt der 5×5 Operator bei einer rechtwinkligen Ecke das maximale Signal um ein Pixel von der Spitze nach innen verschoben, der 7×7 Operator um zwei Pixel nach innen verschoben, usw.
2. Je größer der Operatordurchmesser ist, desto geringer ist die räumliche Auflösung. Wenn eine gesuchte Struktur um drei Pixel kleiner ist als der Operatordurchmesser, dann gibt es wenigstens vier verschiedene Positionen, bei denen die Kanten der Struktur vollständig vom Operator erfaßt werden und daher das Signal maximal und gleich ist. Je kleiner eine Struktur relativ zum Operator ist, desto größer ist der Bereich, für den der Operator den gleichen Wert liefert.
3. Symmetrische Grauwertverteilungen ergeben nicht immer ein symmetrisches Operatorsignal (siehe Tab. 4.1 l, m).
4. Am meisten stört aber die vierte Eigenschaft: Der Operator ist nicht rotationsinvariant im Ergebnis. Z.B. liefert er entlang von Kanten nur dann den Wert Null, wenn die Kante parallel zu einer der

Punktefinder
Moravec

vier Abtastrichtungen verläuft /Moravec 80/. In jeder anderen Richtung wird die Kante von allen vier Abtastrichtungen geschnitten, so daß alle vier Differenzensummen von Null verschieden werden. Das Signal ist für eine Kante zwar schwächer als für eine Ecke von gleichem Kontrast, aber da mit Schwellwerten gearbeitet wird, ist man nicht ohne weiteres in der Lage, zwischen gut kontrastierenden Kanten und schwach kontrastierenden Ecken mittels Schwellwerten zu unterscheiden. Dies erfordert eine lokale Schwellwertanpassung durch Kontrastmessung. Moravec /Moravec 80/ schlägt vor, auch noch in den Richtungen, die in einem Winkel verlaufen, der ein Vielfaches von 22.5 Grad ist, die Summen der Differenzenquadrate zu errechnen.

Außerdem ist wichtig, daß der Operator nur auf lokale Maxima anspricht. Entlang einer idealen, nicht gerasterten Kante muß das Operatorsignal überall konstant sein. Da bei gerasterten Bildfunktionen die Kanten jedoch durch Treppenfunktionen angenähert sind, kann das Operatorsignal entlang einer Kante variieren, je nachdem, wie viele Treppenstufen vom Operator zufällig überdeckt werden (Tab. 4.1c-4.1f). Abb. 1Va zeigt das Operatorsignal für den Ausschnitt um das Taxi aus Bild 1 der Folge 1 als Grauwertbild codiert. Man sieht, daß das Signal auf den Flächen des Wagens und auf der Straße sehr klein ist. Entlang kontrastreicher Kanten sind deutliche Beiträge zu sehen, während Ecken ein besonders hohes Signal geben. In Abb. 1Vb bis 1Vd ist die Lage von lokalen Extremwerten des Operatorsignals markiert, die einen Schwellwert überschreiten. Die verwendeten Schwellwerte sind 50, 100 und 500. Die Lage der gewählten Punkte ist nicht sehr zufriedenstellend. Selbst bei der niedrigsten Schwelle (50) wurde die vordere rechte Ecke vom Bug nicht als markanter Bildpunkt gefunden. Dagegen wurden viele

Punktefinder

Moravec

markante Punkte entlang den Fenster- und Karosseriekanten gewählt. Die Anzahl der auf den Kanten gefundenen Punkte kann gesenkt werden, indem der Schwellwert erhöht wird. Bei einer Schwelle von 500 werden fast nur noch Eckpunkte als markante Punkte gewählt. Allerdings liegen die gewählten Punkte nicht so genau in den Ecken, wie ein Mensch sie positionieren würde.

Ohne eine Detailanalyse der Differenzen in alle vier Richtungen, wie sie z.B. Halé vornimmt, sind lokale Maxima entlang Kanten nicht von denen an Ecken zu unterscheiden. Diese Punkte an Kanten sind aber zufällig durch die Rasterung entstanden und nicht körperfest, so daß sie für die 3D-Rekonstruktion nicht eingesetzt werden dürfen. Das Abtasten des Bildes mit einer Schrittweite von einer halben Operatorbreite unterdrückt einen großen Teil dieser unerwünschten Punkte, allerdings auf Kosten der Auflösung. Dabei werden Eckpunkte ohnehin schon nicht genau genug bestimmt, wenn man mit größerem Operatordurchmesser arbeitet. Bei kleineren Durchmessern dagegen ist der Operator zu unsicher durch Rauschempfindlichkeit.

Wenn man auf hohe Genauigkeit für die Punktkoordinaten angewiesen ist, dann bleibt nichts anderes übrig, als sorgfältig für jeden Szenentyp den Schwellwert exakt dem Bildkontrast anzupassen. Interaktiv ist diese Schwellwertwahl sicher möglich, wenn es überhaupt eine Schwelle gibt, die für das ganze Bild gleich gut geeignet ist. Für die Benutzer aber ist diese Methode auf die Dauer unzumutbar. Ein zuverlässiges automatisches Verfahren zu finden erscheint mir dagegen schwierig. Barnard und Thompson /Barnard und Thompson 79/ schlagen vor, den Schwellwert erst nachträglich festzulegen, wenn alle lokalen Maxima ermittelt und der Größe nach sortiert sind. Der Schwellwert wird so gewählt, daß die gewünschte Anzahl von Punkten übrig bleibt. Dies ist nur praktikabel, wenn man schon vorher weiß,

Punktefinder Moravec

wieviele markante Punkte zu erwarten sind. Bei einem universell einsetzbaren Szenenanalysesystem kann man ebensowenig die Zahl der zu erwartenden Punkte vorher festlegen wie den Schwellwert.

Abgesehen davon zeigen die Erfahrungen mit dem Operator von Moravec, daß es in der Regel nicht möglich ist, einen für alle Stellen im Bild gleich gut geeigneten Schwellwert zu finden. Der Schwellwert müßte lokal dem Kontrast angepaßt werden. Z.B. wäre es denkbar, den Schwellwert vom Betrag des Gradienten abhängig zu machen, so daß an kontrastreichen Stellen im Bild ein höheres Operatorsignal gefordert wird als an anderen Stellen. Eine weitere Möglichkeit ist, den Kontrast direkt anhand der vom Operator gemessenen vier Differenzsummen zu beurteilen. Aus der größten der vier Summen könnte der Schwellwert abgeleitet werden, da die größte der Summen ungefähr in Richtung der stärksten Grauwertvariation gemessen werden dürfte.

4.5 Ein Punktefinder auf der Basis von Krümmungsanalysen

4.5.1 Strukturanalyse mittels der Gaußschen Krümmung

Die Erfahrungen mit dem Operator von Moravec legten nahe, anstelle von Grauwertdifferenzen, wie Moravec sie verwendet hat, Krümmungen des Grauwertgebirges auszuwerten. Krümmungen lassen sich rotationsinvariant ermitteln und beinhalten strukturelle Aussagen über das Bild, die sich für die Lösung des Korrespondenzproblems einsetzen lassen.

Um die Beschreibung des Punktefinders zu erleichtern, möchte ich hier zunächst einige Begriffe klären. Alle Definitionen wurden von Meschkowski /Meschkowski 71/ übernommen.

Punktefinder
Gaußsche Krümmung

Die Krümmung einer ebenen, durch zweimal differenzierbare Funktionen $t \rightarrow x = x(t)$, $t \rightarrow y = y(t)$ gegebenen Kurve ist der Differentialquotient $d\theta/ds$. Dabei ist s die Bogenlänge der Kurve und θ der Winkel, den die Tangente mit der positiven Richtung der X -Achse bildet. Für Raumkurven definiert man die Krümmung durch die zweite Ableitung des Ortsvektors X nach der Bogenlänge (d^2X/ds^2).

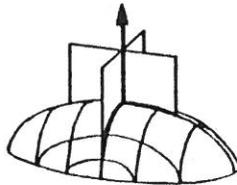


Abb. 4.2 Krümmung einer Fläche

Hauptkrümmungen (K_1 und K_2) heißen die beiden Extremwerte der Krümmung von Normalschnitten im Punkt P einer Fläche X . Normalschnitte sind dabei die durch P gehenden Flächenkurven, die durch solche Ebenen ausgeschnitten werden, die P und die Flächennormale enthalten /Abb. 4.2/.

Gaußsche Krümmung K im Punkt P einer Fläche heißt das Produkt der Hauptkrümmungen in P . $K = K_1 * K_2$.

Zu den beiden Hauptkrümmungen gehören die beiden (zueinander senkrechten) Hauptkrümmungsrichtungen. Die positivere der beiden Krümmungen werde ich positivere Krümmung nennen, die negativere negativere Krümmung.

Punktefinder Gaußsche Krümmung

Beaudet /Beaudet 78/ und Hsu /Hsu et al. 78/ geben Matrixoperatoren an, deren Konvolution mit der Bildfunktion dazu verwendet werden kann, die Koeffizienten der Taylor-Entwicklung für eine gerasterte digitale Bildfunktion um einen Bildpunkt (x, y) zu errechnen. Aus den Polynomkoeffizienten errechnen sie die Hauptkrümmungsrichtungen und die beiden Hauptkrümmungen. Die Operatoren werden für verschiedene Matrixgrößen angegeben, von 3×3 bis 8×8 .

Beaudet schlägt die Gaußsche Krümmung als Möglichkeit vor, um Sattelpunkte und Ecken zu entdecken. Nur an Stellen, wo beide Hauptkrümmungen wesentlich von Null verschieden sind, kann das Produkt von beiden groß sein. Entlang Kanten muß die Gaußsche Krümmung klein sein, da nur in Gradientenrichtung eine Grauwertvariation vorliegt und senkrecht dazu die Krümmung nahezu Null ist. Die beiden Hauptkrümmungen verhalten sich in dieser Hinsicht ähnlich wie der Operator von Moravec, nur werden sie rotationsinvariant ermittelt.

Abbildungen V zeigt am Beispiel des Wagens aus Abb. II die Gaußsche Krümmung für jedes Geopixel in Grauwertdarstellung. Dunkle Grauwerte bedeuten hohe positive Zahlen, während helle Grauwerte negative Zahlen wiedergeben. Abb. Va wurde mit einem Operatordurchmesser von 3 ermittelt, Vb mit Operatordurchmesser 5 und Vc mit Operatordurchmesser 8. Man sieht, daß die räumliche Auflösung durch die verschiedenen Operatoren mit steigendem Operatordurchmesser abnimmt.

Die Vorzeichen der Hauptkrümmungen sind eine Strukturbeschreibung der Bildfunktion. Man kann sechs typische Fälle unterscheiden: (siehe Tabelle)

Punktefinder
Gaußsche Krümmung

1. Beide Krümmungen sind negativ,
2. beide Krümmungen sind positiv,
3. beide haben verschiedene Vorzeichen,
4. beide Krümmungen sind nahe Null,
5. oder nur eine ist Null und die andere entweder positiv oder negativ.

K2 \ K1	1	-	1	0	1	+
-	1	Hügelkuppe	1	Kante oder Wendepunkt	1	Sattel
0	1	--	1	Ebene oder Wendepunkt	1	Kante oder Wendepunkt
+	1	--	1	--	1	Senke

K1 = Positivere Krümmung, K2 = Negativere Krümmung

Diese sechs Fälle entsprechen charakteristischen Verläufen der Bildfunktion. Sucht man nach Maxima der Gaußschen Krümmung, so findet man Senken oder Hügelkuppen, je nachdem, ob die beiden Hauptkrümmungen positiv oder negativ sind. Minima der Gaußschen Krümmung sind Sattelpunkte. Diese Senken und Hügelkuppen dürfen allerdings nicht mit Maxima und Minima der Bildfunktion verwechselt werden. Ein Maximum oder Minimum liegt nur dann vor, wenn auch der Gradient gleich Null ist. Die Hügelkuppen und Senken dagegen

Punktefinder
Gaußsche Krümmung

beschreiben den Kurvenverlauf unabhängig von der Lage des Koordinatensystems (Abb. 4.3).

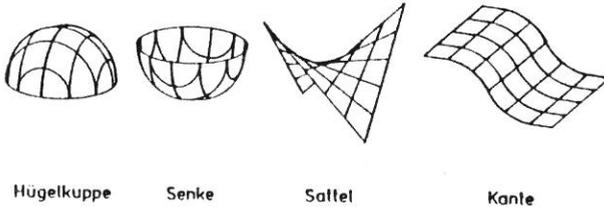


Abb. 4.3 Beispiele für gekrümmte Flächen

Abb. VIa zeigt den Betrag der Gaußschen Krümmung für das Bild des Taxis aus Abb. Ia und Abb. VIb die Lage der Extremwerte. Es sind in der Nähe von Ecken Extrema zu sehen, aber nicht genau auf dem Schnittpunkt der Kanten. Die Lage von Extremwerten der Gaußsche Krümmung gibt also Ecken ähnlich ungenau an, wie der Operator von Moravec. Dies wird besonders deutlich an den Ecken des Rückfensters. In den Innenwinkeln des Fensters befinden sich Maxima, in den Außenwinkeln Minima. Die Punkte, die ein idealer Punktefinder liefern sollte, lägen genau dazwischen.

Punktfinder
Gaußsche Krümmung

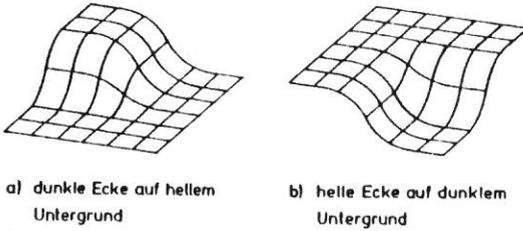
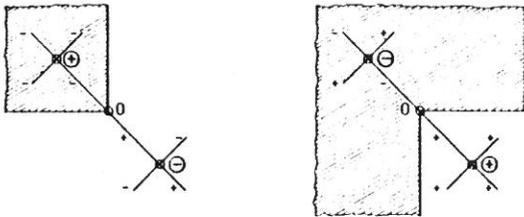


Abb. 4.4 Eckenstrukturen als Grauwertgebirge



- a) Grauwerte im Maximum
der Gaußschen Krümmung
größer als im Minimum
- b) Grauwerte im Maximum
der Gaußschen Krümmung
kleiner als im Minimum

Abb. 4.5 Hauptkrümmungsrichtungen der
Eckenstrukturen aus Abb. 4.4

- , 0 , + Vorzeichen der Hauptkrümmungen
⊖ , ⊕ Extrema der Gaußschen Krümmung

Punktefinder
Gaußsche Krümmung

Abb. 4.4 und 4.5 zeigen die Krümmungsverhältnisse an einem idealisierten Modell einer Fensterecke. Die dunkle Rückseite wird in der Bildfunktion durch ein hohes, rechteckiges Plateau wiedergegeben, das sich über der hellen Grundebene der Wagenkarosserie erhebt. Interessant ist die Schnittlinie der beiden Plateaukanten, die ungefähr S-förmig verläuft. Auf dieser Linie liegt am Oberrand des Plateaus das Maximum der Gaußschen Krümmung, denn diese Struktur entspricht ungefähr einer Hügelkuppe. Weiter unten an der Stelle des steilsten Gradienten liegt der eigentlich gesuchte Punkt, der Wendepunkt der Schnittlinie. Die Gaußsche Krümmung ist hier gleich Null. Am Fuß des Plateaus ist die Schnittlinie positiv gekrümmt, aber die Krümmung senkrecht dazu ist weiterhin negativ. Hier liegt ein Sattelpunkt und gleichzeitig das Minimum der Gaußschen Krümmung.

Ähnlich sind die Verhältnisse bei einer Innenecke, also bei einem hellen Rechteck auf dunklem Untergrund. Hier liegt das Maximum der Gaußschen Krümmung am Fußpunkt der Kanten-schnittlinie und das Minimum am Oberrand, denn der obere Rand der Ecke entspricht einem Sattel, der untere einer Senke.

Allgemein kann man feststellen (Abb. 4.5):

1. Wenn sich zwei Kanten schneiden, dann ist auf der Seite des spitzen Winkels die Gaußsche Krümmung positiv, auf der Seite des stumpfen Winkel die Gaußsche Krümmung negativ.
2. Ist der vom spitzen Winkel eingeschlossene Grauwert dunkler als die Umgebung, so sind beide Hauptkrümmungen im spitzen Winkel negativ. Auf der Seite des stumpfen Winkels ist dann die Krümmung in Richtung auf den Schnittpunkt positiv, senkrecht dazu negativ.

Punktefinder
Gaußsche Krümmung

3. Ist der vom spitzen Winkel eingeschlossene Grauwert heller als die Umgebung, so sind beide Hauptkrümmungen im spitzen Winkel positiv. Auf der dunklen Seite ist dann die Krümmung in Richtung auf den Schnittpunkt negativ, senkrecht dazu positiv.
4. Wenn die Bildfunktion nur lokal entlang den sich schneidenden Kanten gekrümmt und in der Umgebung eben ist, dann bilden sich in der Nähe des Schnittpunkts lokale Extrema der Gaußschen Krümmung, mindestens ein Maximum und ein Minimum.

Auf diesen vier Thesen baut der Punktefinder auf: Es werden für alle Pixel des gewählten Ausschnitts die Gaußsche Krümmung und deren lokale Extrema ermittelt. Alle Maxima werden als Indiz für den Innenwinkel eines Kantenschnittpunkts genommen. Innerhalb einer begrenzten Umgebung werden alle lokalen Minima daraufhin untersucht, ob sie in Einklang mit den Thesen 1 bis 3 als Gegenstück in Frage kommen. Wenn ein geeignetes Minimum gefunden wird, so ergibt sich der genaue Kantenschnittpunkt durch Interpolation der Nullstelle der Gaußschen Krümmung zwischen den beiden Extrema. Dies liefert den Krümmungsnulldpunkt und damit näherungsweise die Lage des steilsten Gradienten. Dieser Nullpunkt wird als Kandidat für einen markanten Punkt gewählt und seine Koordinaten zusammen mit zusätzlicher Information für den Korrespondenzanalysator gespeichert.

Die Paarung von Maxima und Minima ist nicht immer eindeutig. Kommen mehrere Minima für ein Maximum in Frage, so wird dasjenige ausgewählt, für das ein Konfidenzmaß am größten ist. Dieses Maß gibt an, wie gut die Hauptkrümmungsrichtung im Minimum mit der Verbindungslinie der beiden Extrema übereinstimmt. Es basiert auf der Annahme, daß im Minimum eine Hauptkrümmungsrichtung mit der Winkelhalbierenden zwischen

Punktefinder
Gaußsche Krümmung

den sich schneidenden Kanten zusammenfallen sollte (Abb. 4.5). Genauer gesagt wird der Cosinus des Winkels zwischen der Verbindungsgeraden der beiden Krümmungsextrema und derjenigen Hauptkrümmungsrichtung im Minimum errechnet, deren Krümmungskomponente sich im Vorzeichen von denen im Maximum unterscheidet. Dieser Wert wird noch mit dem Krümmungsgradienten gewichtet.

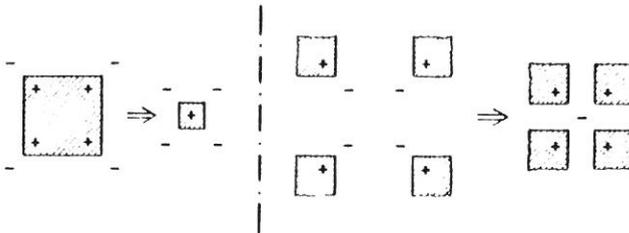


Abb. 4.6 Grenzfälle der Auflösung

Die schraffierten Gebiete entsprechen dunklen Bildbereichen. Pluszeichen geben die Lage von Maxima der Gaußschen Krümmung an, Minuszeichen die von Minima der Gaußschen Krümmung.

In Abb. XIIIa ist ein stark vergrößerter Ausschnitt (9*9 Pixel) der linken oberen Rückfensterecke des Taxis aus Abb. Ia zu sehen. In jedes Pixel sind links oben der Grauwert des Pixels und rechts die Werte der beiden Hauptkrümmungen und der Gaußschen Krümmung angegeben. Extremwerte der Gaußschen Krümmung, die den Schwellwert überschreiten, sind in Negativschrift hervorgehoben. Das mittlere Pixel ist ein markanter Punkt, da sich das Krümmungsminimum links davon und das Maximum rechts davon nach den eben besprochenen Regeln kombinieren lassen. Die eingezeichneten Geraden geben die Richtung der positiveren Hauptkrümmung an. Diese ist im Minimum der Gaußschen Krümmung genau auf das Maximum gerichtet.

Punktefinder Gaußsche Krümmung

Ein Minimum kann mehreren Maxima zugeordnet werden. Dieses ist bei kreuzähnlichen Konfigurationen nötig (z.B. Abb. 4.6). Hier überlagern sich in der Kreuzmitte vier Minima zu einem einzigen. Beim Bild des Personenwagens tritt dieser Fall an den Enden des Steges zwischen den beiden Seitenfenstern auf.

Ebenso ist es möglich, daß eine Fläche so klein wird, daß die Krümmungsmaxima in den Innenwinkeln zu einem einzigen verschmelzen und nur noch außen die Minima getrennt sind, z.B. wenn der Personenwagen sich so weit entfernt hat, daß die Seitenfenster zu klein werden. Hier hat man jetzt mehrere Möglichkeiten zur Lösung des Konflikts. Erstens könnte aus Symmetriegründen für jedes Minimum ein markanter Punkt gewählt werden, indem das Maximum mehrfach zugeordnet wird. Das wäre für den Fall des Seitenfensters eine günstige Entscheidung, da so auch noch bei geringer Auflösung alle vier Fensterecken gefunden werden können. Genausogut kann man aber das Aufeinanderfallen mehrerer Innenwinkel als punktförmige Bildstrukturen auffassen.

Bei Flecken und Glanzlichtern entstehen ähnliche Krümmungskonstellationen: Ein Maximum, das von einem Kranz Minima umgeben ist. Solche Flecken werden besser als Einzelpunkte modelliert, mit dem gewählten markanten Punkt im Zentrum, nicht als n -Eck, da man sich an der Grenze der Auflösung bewegt und Anzahl und Lage der Minima eher vom Zufall abhängt als von der Bildinformation. Die so gefundenen Eckpunkte würden wenig informativ sein und nur den Korrespondenzalgorithmus unnötig belasten. Daher wird ein Maximum, das von wenigstens drei Minima umgeben ist, als Punkt modelliert und der markante Punkt direkt auf dem Maximum gewählt. Bei dieser Lösung sind die Punkte sicher reproduzierbar, was mir für die Korrespondenzanalyse wichtiger ist als eine Einfachheit des Verfahrens durch Symmetrie. Ein Beispiel für eine solche Konstellation ist

Punktefinder Gaußsche Krümmung

das Glanzlicht in Abb. XIIIb. Dieser helle Fleck entspricht im Grauwertgebirge einem Talkessel, wobei das Maximum der Gaußschen Krümmung den tiefsten Punkt markiert. An dieser Stelle wurde der markante Punkt gewählt.

4.5.2 Technische Probleme

Der Punktefinder wird auf kontrastreichen Kanten sicherer, wenn man nicht kritiklos alle Extremwerte der Gaußschen Krümmung zuläßt, sondern verlangt, daß in einem Extremwert die Gaußsche Krümmung deutlich von Null verschieden sein soll. Dabei ist es günstig, wenn man nicht auf die Gaußsche Krümmung selbst einen Schwellwert anwendet, sondern besser auf die beiden Hauptkrümmungen, die als Faktoren eingehen. Das Produkt der beiden Hauptkrümmungen kann groß sein, auch wenn ein Faktor sehr klein ist. Z.B. kommt es entlang kontrastreicher Kanten vor, daß die Krümmung in Gradientenrichtung groß ist und senkrecht dazu die Krümmung um Null schwankt. Dies führt dazu, daß entlang einer Kante Maxima und Minima der Gaußschen Krümmung abwechseln, die nicht von Bedeutung sind. Diese unerwünschten Werte werden durch die Forderung unterdrückt, daß beide Hauptkrümmungen in einem Extremum deutlich von Null verschieden sein müssen. Diese Forderung beinhaltet eine Formanalyse; denn wenn eine der beiden Krümmungen sehr klein ist, hat man es nicht mit den gesuchten Senken, Hügelkuppen oder Sätteln zu tun, die Hinweise auf Eckpunkte sind.

Durch diese Schwelle kann es vorkommen, daß die Bilder markanter Objektpunkte, die ein menschlicher Beobachter noch als markante Punkte ausgewählt hätte, vom Punktefinder nicht mehr als markante Bildpunkte akzeptiert werden. Aber wenn die Krümmungen ohnehin so niedrig sind, daß sie unter den Schwellwert fallen, dann ist der markante Punkt auch nicht sicher reproduzierbar. In diesem Fall ist es besser, darauf

Punktefinder Gaußsche Krümmung

zu warten, daß das betreffende Objektmerkmal durch die Objektbewegung deutlicher sichtbar wird, oder andernfalls ganz auf den Punkt zu verzichten, da er zu unsicheren Messungen führen würde, die die 3-D Rekonstruktion verfälschen könnten.

Abb. VIc zeigt die mit diesem Punktefinder gewählten markanten Punkte (gleicher Bildausschnitt wie für Ergebnis des Moravec-Operators). Im Gegensatz zum Ergebnis des Moravec-Operators wurden keine markanten Punkte entlang kontrastreicher Kanten gewählt. Dieses ist ein Erfolg der genaueren Modellvorstellung von einer Ecke, die dieser Punktefinder verwendet. Die Punkte auf der hinteren Kofferraumkante werden von Schatten verursacht, die der Baum im Vordergrund auf den Wagen wirft. Im Unterschied zum Moravec-Operator konnte auch die rechte Ecke vom Bug und die rechte vordere Ecke der Fahrgastzelle als markanter Punkt erkannt werden. Die gewählten Punkte sitzen zum Teil etwas günstiger als die vom Moravec-Operator gefundenen Punkte (z.B. die beiden linken Ecken der Heckscheibe), allerdings noch immer nicht ganz so gut in den Ecken, wie ein Mensch sie wählen würde.

Das Ergebnis der Krümmungsberechnungen hängt sehr davon ab, wie der Operatordurchmesser gewählt wird. Der Operatordurchmesser bestimmt, wieviele Stützstellen für die Entwicklung der Taylor-Reihe verwendet werden. Bei Verwendung von 3×3 -Operatoren spiegelt sich im Krümmungsbild jede Grauwertschwankung wieder, d.h. auch das Rauschen. Das ist deutlich an den Binärbildern in Abb. 4.7 zu erkennen. Stufen in Kanten, die durch die Rasterung entstanden sind, können das gleiche Operatorsignal erzeugen wie Ecken. Die stumpfe Ecke (Abb. 4.7b) und die Kante (Abb. 4.7a) sind auf der vom 3×3 -Operator erfaßten Fläche völlig gleich. Die rechtwinklige Ecke (Abb. 4.7c) und die Kante unterscheiden sich nur in einem Pixel.

Punktefinder
Gaußsche Krümmung

Je größer die Operatoren sind, desto mehr unterdrücken sie die Feinheiten und geben nur noch bei großflächigen Strukturen deutliche Signale. Die räumliche Auflösung sinkt und die Rechenzeit steigt, da der Aufwand für die Konvolution proportional zur Operatorfläche ist. Abb. V zeigt die Gaußsche Krümmung bei verschiedenen Operatorgrößen.

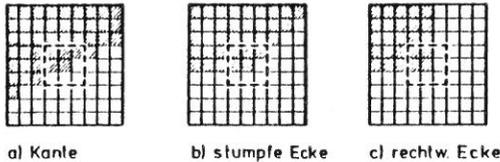


Abb. 4.7

Da die 3×3 -Operatoren zwar fein genug in der Auflösung sind, aber zu empfindlich auf Rauschen und Rasterung reagieren, dagegen die 5×5 -Operatoren unempfindlich genug sind, aber die Auflösung nicht mehr ausreicht, wird als Kompromiß ein zweistufiges Verfahren verwendet, indem die beiden Operatortypen kombiniert werden. Mit dem 5×5 -Operator werden zuerst im gewählten Ausschnitt Extremwerte der Gaußschen Krümmung gesucht. Deren Positionen werden als Anhaltspunkte dafür genommen, wo anschließend mit den 3×3 Operatoren die Extrema in feinerer Auflösung genauer lokalisiert werden müssen (Abb. VIIa). Diesen hierarchischen Punktefinder nenne ich zweistufigen Punktefinder im Unterschied zum vorher beschriebenen einstufigen Punktefinder. Abb. VIIb zeigt die mit diesem Punktefinder gefundenen markanten Punkte (gleicher Ausschnitt wie für Moravec-Operator und einstufigen Punktefinder). Außer am Steg zwischen den

Punktefinder Gaußsche Krümmung

Seitenfenstern, der wenig Kontrast bietet, sitzen alle Punkte sehr genau in den Ecken. Sogar die rechte und linke Ecke des Taxi-Zeichens konnten getrennt werden.

Dieses hierarchische Vorgehen entspricht dem zweistufigen Segmentieren mittels Plan /Yachida et al. 79, 80/. Auch Yachida et al. erstellen erst mit einem groben Segmentierer eine Planskizze, die ihnen zu entscheiden erlaubt, wo und wie ihr feiner und sehr rechenzeitintensiver Kantenfinder eingesetzt werden sollte /siehe auch Kelly 70, 71/.

4.6 Kombination von Operatoren verschiedener Durchmesser

In diesem Abschnitt werde ich einige Parallelen aufzeigen, die sich zwischen dem zweistufigen Punktefinder (5*5-Operator gefolgt von 3*3-Operator) und anderen Punktefindern mit Operatoren verschiedener Auflösung finden lassen. Auch Rosenfeld und Kitchen (siehe 4.2) haben vorgeschlagen, die Ergebnisse von Operatoren verschiedener Auflösung miteinander zu kombinieren. Allerdings hat ihr Verfahren, einfach die Ergebnisse von Operatoren mit verschiedenem Durchmesser miteinander zu multiplizieren und nur das Produkt zu bewerten, einige Nachteile. Bei dieser Methode wird wertvolle Information vernachlässigt, die man durch die Anwendung verschieden auflösender Operatoren gewonnen hat. Aus einem Produkt geht nicht mehr hervor, in welchem Verhältnis die Faktoren zueinander stehen. Genau dieses ist ja auch der Grund dafür, daß die Gaußsche Krümmung auf sehr kontrastreichen Kanten kein idealer Eckenfinder ist, obwohl die beiden Hauptkrümmungen getrennt betrachtet auch auf kontrastreichen Kanten eine gute Formbeschreibung für das Grauwertgebirge abgeben. Auch Marr und Hildreth betonen, um wieviel einfacherer und sicherer die Merkmalsfindung ist, wenn gleichzeitig in mehreren verschieden auflösenden Frequenzbereichen gearbeitet wird. Die Operatoren von

Punktefinder Gaußsche Krümmung

Beaudet erfüllen zwar nicht die Voraussetzungen, die Marr und Hildreth von einem idealen Filter fordern, aber sie lassen sich insofern damit vergleichen, als sie eine glättende Wirkung haben und die räumliche Auflösung mit wachsendem Operatordurchmesser abnimmt.

Für ebenso ungünstig wie die von Kitchen und Rosenfeld vorgeschlagene Multiplikation von Operatorergebnissen halte ich die Multiplikation von Operatorresultat und Gradientenbetrag. Zwar ist es wichtig, den Gradientenbetrag oder ein anderes geeignetes Maß für den Kontrast zu berücksichtigen, damit sichergestellt ist, daß überhaupt eine Ecke vorliegen kann. Ich halte es aber für günstiger, einfach den Gradientenbetrag mit einem niedrig gewählten Schwellwert zu vergleichen und den Punkteoperator nur noch dort anzuwenden, wo der Gradient den Schwellwert überschreitet. Diese Lösung hat zwei Vorteile: Erstens werden Operatorergebnis und Gradient getrennt ausgewertet, so daß keine Information verloren gehen kann. Zweitens wird Rechenzeit gespart, da der Punktefinder nur noch auf den Bildausschnitten angewendet wird, wo auch eine entsprechende Grauwertvariation erwartet werden kann. Diese Alternative entspricht wieder dem schon angesprochenen "Planning" von Yachida et al. . Für spätere Programmversionen plane ich ebenfalls einen einfachen und schnellen Gradientenoperator einzusetzen und den Punktefinder nur noch auf Bildausschnitten mit hohen Gradientenbeträgen einzusetzen. Gerade bei den Operatoren von Beaudet läßt sich auf diese Weise viel Rechenzeit einsparen. Die Berechnung der Gaußschen Krümmung erfordert mit den 5*5-Operatoren 75 Multiplikationen und 150 Indexberechnungen für Zugriffe auf zweidimensionale Felder, während eine grobe Abschätzung des Gradientenbetrages mit wesentlich weniger Rechenaufwand möglich ist /Davis 75/.

Punktefinder Reproduzierbarkeit

4.7 Untersuchungen zur Reproduzierbarkeit

In den vorangehenden Abschnitten wurde gezeigt, daß der zweistufige Punktefinder ganz brauchbare Kandidaten für die Bilder markanter Objektpunkte auswählt. Dieser Abschnitt ist dem Problem gewidmet, wie zuverlässig diese Punktauswahl ist, und wie sicher das Ergebnis des Punktefinders unabhängig von Rauschen und anderen Zufälligkeiten beim Aufnahme- und Digitisierungsvorgang reproduziert werden kann. Zur Reproduzierbarkeit des Ergebnisses habe ich bei keinem der in diesem Kapitel vorgestellten Punkte- und Merkmalsfinder Angaben gefunden, obwohl diese Eigenschaft sehr wichtig ist, besonders beim Verarbeiten von Bildfolgen.

Um die Reproduzierbarkeit für die von mir implementierten Punktefinder zu messen, wurde aus den Bildern der Folge 1 /Abb. 1a, 1b/ ein Ausschnitt ausgewählt, der keine bewegten Objekte enthält. Die zu untersuchenden Punktefinder wurden in 20 Bildern der Folge auf den gewählten Ausschnitt angewendet und die ausgewählten Punktmengen aus aufeinander folgenden Bildern jeweils miteinander verglichen. Da dieser Vergleich die Lösung des Korrespondenzproblems erfordert, wurde der Einfachheit halber eins der in Kapitel 5 beschriebenen Programme dafür eingesetzt. Weil im statischen Hintergrund nicht mit großen Verschiebungen zu rechnen ist, wurden solche Punktpaare als Kandidaten für korrespondierende Punkte ausgewählt, die sich gegenseitig die nächsten Nachbarn sind. Die so gemessenen Verschiebungen sind ein Anhaltspunkt dafür, wie genau die markanten Bildpunkte im jeweiligen Folgebild wiedergefunden worden sind. Diese Meßreihen wurden für drei Punktefinder durchgeführt: Mit dem Moravec-Operator, dem einstufigen Punktefinder und dem zweistufigen Punktefinder. Abb. XIV zeigt den untersuchten Ausschnitt. Die markierten Stellen zeigen, wo mit dem zweistufigen Punktefinder im ersten Bild von Folge 1 (Punkte) und im zweiten Bild (offene Quadrate)

Punktefinder
Reproduzierbarkeit

markante Bildpunkte gefunden wurden. Gefüllte Quadrate markieren Stellen, an denen in beiden Bildern am gleichen Ort ein markanter Punkt gefunden wurde.

Die folgende Tabelle gibt für alle drei Punktefinder die in 20 Bildern im Mittel gefundene Punkteanzahl und die gemessenen Varianzen und Streuungen an.

Punktefinder	I Moravec	I einstufig	I zweistufig
Schwellwerte	100	4	1. Stufe 4 2. Stufe 10
Mittlere Punktezahl	71.1	53.7	54.6
Empirische Varianz	13.3	3.6	11.7
Empirische Streuung	3.6	1.9	3.4
Empirische Streuung %	5.1	3.5	6.3

Tabelle 4.2: Mittlere Anzahl von markanten Bildpunkten in ungefilterten Bildausschnitten

Zwei Besonderheiten fallen an diesem Ergebnis auf: Zum einen liefert der Moravec-Operator wesentlich mehr markante Punkte als die beiden anderen Punktefinder, und zweitens ist bei allen drei Punktefindern eine deutliche Schwankung in der Zahl der gefundenen Punkte vorhanden, obwohl der untersuchte Ausschnitt nur statischen Hintergrund zeigt. Die kleinste Streuung in der Punktzahl zeigt der einstufige Punktefinder.

Punktefinder Reproduzierbarkeit

Eine wenig variierende Punktzahl bedeutet nicht automatisch, daß die Punkte immer wieder am selben Ort gewählt werden. Wie oft jeder einzelne Punkt in Folgebild wieder als markanter Bildpunkt gewählt wird, wurde mit dem oben erwähnten Korrespondenzprogramm untersucht. In Abb. 4.8a (am Ende dieses Kapitels) ist in einem Histogramm zu sehen, wie oft im Mittel für einen markanten Bildpunkt im Folgebild ein korrespondierender Punkt gefunden werden konnte. Am unteren Ende der Säulen des Histogramms ist die Streuung des jeweiligen Meßwerts markiert. Alle Werte sind in Prozenten der mittleren Punktzahl angegeben. Ähnlich wie bei der Varianz der Punktzahl zeigt der einstufige Punktefinder wieder die besten Werte, wobei der Moravec-Operator ebenfalls wieder ein wenig besser abschneidet als der zweistufige Punktefinder. Wenn man allerdings die Streuung der Meßwerte berücksichtigt, dann ist die Reproduzierbarkeit vom Moravec-Operator und dem zweistufigen Punktefinder ungefähr vergleichbar, während der einstufige Punktefinder ein wenig besser ist.

Die Histogramme in Abb. 4.8c zeigen die Häufigkeitsverteilung der gemessenen Verschiebungen für die drei Punktefinder. Ungefähr die Hälfte der Punkte wird am gleichen Ort wiedergefunden wie im Bild davor, während der Rest 1 bis 2 Pixel versetzt oder überhaupt nicht wiedergefunden wird. Auch diese Histogramme fallen wieder für den einstufigen Punktefinder am günstigsten aus.

Die geringe Rauschempfindlichkeit des einstufigen Punktefinders wird verständlich, wenn man bedenkt, daß zur Berechnung der Krümmungen die Bildfunktion durch ein Polynom 2. Grades approximiert wird und alle Krümmungswerte anhand dieser Näherungsfunktion bestimmt werden. Da in die Näherungsfunktion 25 Stützstellen eingehen, stören einzelne Ausreißer nur wenig, während bei den 3×3 -Operatoren des zweistufigen Punktefinders schon ein einzelner Ausreißer zu wesentlich

Punktefinder Reproduzierbarkeit

verschiedenen Krümmungswerten führen kann (siehe Abb. 4.7). Auch der Moravec-Operator ist rauschempfindlich, da er ja gerade Varianzen der Bildfunktion in verschiedenen Richtungen auswertet. Vermutlich läßt sich auch der Unterschied in der Reproduzierbarkeit zwischen dem Moravec-Operator und dem zweistufigen Punktefinder durch den unterschiedlichen Operatordurchmesser erklären.

Aus diesen Überlegungen folgt, daß sich durch eine Verringerung des Rauschens, z.B. durch Filterung, die Sicherheit der Punktefinder erhöhen ließe. Für den zweistufigen Punktefinder lohnt sich eine Glättung der Bilder aber nur dann, wenn die Bilder nicht so stark geglättet werden, daß die gute räumliche Auflösung des Punktefinders nicht mehr wirksam werden kann. Andernfalls wäre es günstiger, von vornherein mit dem einstufigen Punktefinder zu arbeiten. Das Medianfilter ist für diesen Zweck ideal, da bei dieser Art der Filterung scharfe Kontraste kaum verwischt werden. Vergleicht man die vom zweistufigen Punktefinder auf gefilterten (Abb. VIIb) und auf ungefilterten Bildern (Abb. VIIc) gewählten markanten Punkte, sieht man, daß in der Auflösung kaum ein Unterschied zu sehen ist. Die Reproduzierbarkeit der markanten Punkte wird beim Moravec-Operator und beim zweistufigen Punktefinder um ungefähr 6% besser, während sie beim einstufigen Punktefinder nur um ca 4% wächst. Dieses Ergebnis (Abb. 4.8b, d) bestätigt die Annahmen über die Rauschempfindlichkeit der Punktefinder. Eine Nebenwirkung der Filterung ist, daß sich die Anzahl der im Mittel gefundenen Punkte verringert. Besonders beim Moravec-Operator ist dieser Effekt auffällig. Nach der Filterung sind die Zahl der vom Moravec-Operator und die von den anderen beiden Punktefindern gewählten markanten Punkte vergleichbar.

Punktefinder
Reproduzierbarkeit

Punktefinder	I Moravec	I einstufig	1 zweistufig
Schwellwerte	1	1	1
	1	100	4
	1	1	1
Mittlere Punkt- zahl	1	1	1
	1	48.8	41.9
	1	1	1
Empirische Varianz	1	1	1
	1	6.6	3.5
	1	1	1
Empirische Streuung	1	1	1
	1	2.6	1.9
	1	1	1
Empirische Streuung %	1	1	1
	1	5.3	4.5

Tabelle 4.3: Mittlere Anzahl von markanten Bildpunkten
in Median-gefilterten Bildern

Anhand der Prozentangaben über die von einem Bild zum nächsten zugeordneten Punkte läßt sich abschätzen, wie groß die Wahrscheinlichkeit dafür ist, für einen im ersten Bild einer Folge gewählten markanten Punkt in n aufeinander folgenden Bildern einen korrespondierenden Punkt zu finden. Unter der Annahme, daß die Wahrscheinlichkeit, einen markanten Punkt im Folgebild wiederzufinden, für alle Punkte gleich und statistisch unabhängig ist, ergibt sich die folgende Tabelle:

Punktefinder
Reproduzierbarkeit

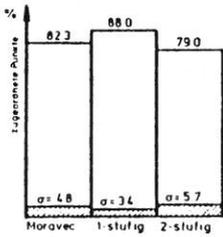
n	1 Moravec- 1 Operator	1 Einstufiger 1 Punktefinder	1 Zweistufiger 1 Punktefinder
2	0.88	0.92	0.85
3	0.77	0.84	0.72
4	0.68	0.78	0.62
5	0.60	0.72	0.52
10	0.32	0.47	0.23
20	0.09	0.21	0.05

Tab. 4.4: Wahrscheinlichkeit dafür, einen markanten Bildpunkt in n aufeinander folgenden Bildern zu finden

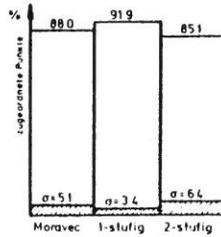
Die Ergebnisse dieser einfachen Modellrechnung sind für die Verarbeitung von Bildfolgen nicht sehr ermutigend. Selbst für Punkte aus dem statischen Hintergrund ist die Wahrscheinlichkeit klein, daß sie über eine längere Folge immer wieder gefunden werden. Die Messung der gefundenen Kettenlängen gibt aber ein viel günstigeres Ergebnis. Im Widerspruch zum Ergebnis der Modellrechnung wurden mit allen drei Punktefindern für mehr als die Hälfte der markanten Bildpunkte aus dem ersten Bild der untersuchten Folge in allen 19 weiteren Bildern ein korrespondierender Punkt gefunden. Dies zeigt, daß die Reproduktionswahrscheinlichkeiten nicht für alle markanten Punkte gleich sein können, oder daß die Annahme der statistischen Unabhängigkeit unzutreffend ist. Ein großer Teil der markanten Bildpunkte läßt sich sehr sicher reproduzieren, während andere Punkte weniger sicher reproduziert werden können; vermutlich diejenigen, bei denen die Operatorwerte in der Nähe der Schwellwerte liegen. Ein vermuteter Zusammenhang zwischen der Reproduzierbarkeit und der Höhe der Operatorwerte wurde aus Zeitgründen noch nicht näher untersucht.

Punktefinder Reproduzierbarkeit

Mittelwert u Streuung der Anzahl der im jew. Folgebild zugeordneten Punkte in % der mittleren Punktezahl

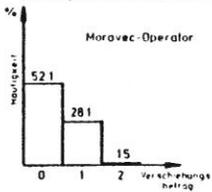


a) ungefilterte Bilder

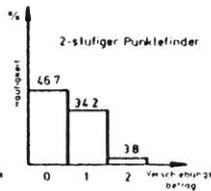
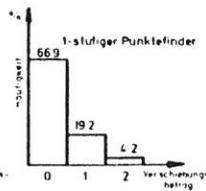
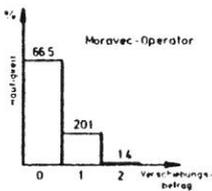
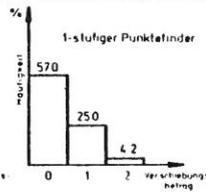


b) Median-gefilterte Bilder

Häufigkeitsverteilung der Verschiebungsbeträge (% der mittl. Punktezahl)



c) ungefilterte Bilder



d) Median-gefilterte Bilder

Abb. 4.8 Messungen zur Reproduzierbarkeit

Vergleich von Moravec-Operator, ein- und zweistufigem Punktefinder auf einer Folge von 20 Bildausschnitten aus dem statischen Hintergrund

5. Das Korrespondenzproblem

5.1 Der Korrespondenzprozeß beim Menschen

Ullman /Ullman 79/ hat in seinem Buch eine Reihe von Experimenten beschrieben, die untersuchen, wie der Mensch das Korrespondenzproblem löst. Bei diesen Experimenten wurden Versuchspersonen Folgen von Momentaufnahmen vorgeführt, die bewegte Punktwolken zeigen. Dabei wurde untersucht, unter welchen Bedingungen die Personen die richtigen Korrespondenzen erkannten und die dargestellte räumliche Bewegung korrekt erfassen konnten.

Eine wichtige Frage ist, welches die Basiselemente sind, die das menschliche Auge für die Korrespondenz in Beziehung setzt. Es sind alle Stufen zwischen zwei Extremen denkbar:

1. Die Korrespondenzen könnten direkt durch Helligkeitsvergleiche ermittelt werden, noch bevor irgend eine Art von Merkmalsextraktion erfolgt wäre.
2. Als anderes Extrem könnte der Mensch alle typischen Ansichten eines Objektes lernen und in einer Art Datenbank speichern, so daß er die Korrespondenzen dadurch ermitteln könnte, daß er erkannte Objekte zueinander in Beziehung bringt.

Ullman vermutet, daß beim Menschen die Zuordnung auf der Basis von einfachen Bildelementen, wie Kanten, Liniensegmenten und Punkten erfolgt. Er führt zwar keinen endgültigen Beweis, aber er macht seine Annahme plausibel, indem er die technischen Probleme der einzelnen Lösungsmöglichkeiten gegeneinander abwägt und Ergebnisse von psychologischen Untersuchungen anführt, die seine Hypothese stützen. Interessant ist, daß die von Ullman vermuteten

Korrespondenzproblem Wettbewerbsmodell

Basiselemente ungefähr der "primal sketch" von Marr /Marr 76/ entsprechen, von der Marr annimmt, daß sie als erster Schritt vom visuellen System bei der Bildanalyse gebildet wird und als Baustein für Konstrukte höherer Ordnung dient. Bullock /Bullock 76/ und Kovalevsky /Kovalevsky 79/ kommen allein durch Praktikabilitätsüberlegungen zu der Feststellung, daß für ein komplexes und flexibles Bildverarbeitungssystem die Ebene der einfachen, lokalen Merkmale, wie Kanten, Eckpunkte und Texturmerkmale die günstigste ist.

Ullman erwartet, daß der Mensch für diese Basiselemente ein Wahrscheinlichkeitsmaß für die Korrespondenz gespeichert hat. Dieses Maß nennt Ullman Affinität. Für zusammengesetzte Konstrukte wird die Affinität durch Kombination der Einzelaffinitäten ermittelt. Daß der Mensch nicht Objekte von Bild zu Bild zueinander in Beziehung setzt, zeigt deutlich das Experiment mit dem zerbrochenen Rad. Der Mensch sieht bei geeigneter Rotationsfrequenz den mittleren Teil des Rades in Abb 5.1 in entgegengesetzter Richtung zum äußeren und inneren Teil rotieren. Diese optische Täuschung läßt sich mit der Annahme erklären, daß auch der Mensch bewegte Bilder als Folgen von Einzelbildern verarbeitet. Wenn für Teile der Radspeichen im Folgebild jeweils der Speichenteil mit der geringsten Verschiebung als Kandidat für ein korrespondierendes Bildelement gewählt wird, kann es vorkommen, daß Teile von ganzen Speichen mit Teilen von zerbrochenen Speichen in Korrespondenz gebracht werden. Dieser Effekt hängt davon ab, um welchen Winkel das Rad sich zwischen den zwei Einzelbildern gedreht hat.

Ullman fand heraus, daß für das menschliche Auge die Länge des zeitlichen Intervalls zwischen zwei Momentaufnahmen die wahrgenommenen Korrespondenzen wesentlich bestimmt, ebenso daß die gleiche räumliche Anordnung am Rande des Bildfeldes anders bewertet werden kann als im Zentrum, und daß korre-

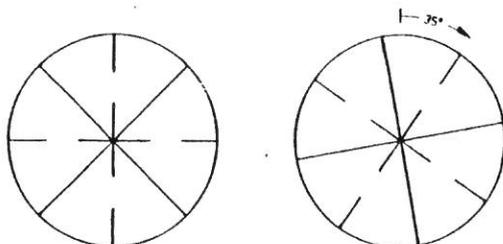


Abb. 5.1 „das zerbrochene Rad“

spondierende Punktpaare nicht isoliert betrachtet werden können, sondern daß der Kontext wichtig ist. Er schlägt vor, den Korrespondenzmechanismus des Menschen durch einen Wettbewerbsprozeß zu modellieren. Dieser Prozeß wird durch drei Konzepte bestimmt:

1. Affinität (Affinity): Für je zwei Basiselemente aus aufeinanderfolgenden Bildern ist die Affinität ein Maß für die Güte einer möglichen Zuordnung. Die Affinität hängt von vielen verschiedenen Dingen ab, besonders von der Entfernung der Basiselemente. Zusätzlich sind eine Reihe von Ähnlichkeitsmaßen wichtig, wie Helligkeit, Orientierung und Größe. Zum Rande des Gesichtsfeldes hin nehmen die Affinitäten bei sonst gleichen Parametern zu, und außerdem sind die Unterschiede zwischen den Affinitäten hier im Vergleich zum Zentrum weniger deutlich. Auch längere Zeitintervalle zwischen den Momentaufnahmen machen die Affinitäten einheitlicher.

Korrespondenzproblem
Wettbewerbsmodell

Ullman hat experimentell nachgewiesen, daß Gleichmäßigkeit der Bewegung kein Kriterium für die wahrgenommenen Korrespondenzen ist. Dies steht im Widerspruch zu den bisherigen Annahmen vieler Psychologen.

Für die Affinitäten sind vor allem Entfernungen wichtig. Dabei sind drei Entfernungsbegriffe zu unterscheiden, erstens die echte räumliche 3D-Entfernung, zweitens die dem Auge sichtbare 2D-Entfernung und drittens die interpretierte Entfernung, d.h. die durch irgendwelche Tiefenhinweise, wie Größe oder Perspektive, wahrgenommene Entfernung. Ullman hat experimentell die Annahme untermauert, daß nur die 2D-Entfernungen für die Affinitäten ausgewertet werden.

2. Teilungswettbewerb (Split Competition): Für ein Basiselement im ersten Bild besteht Teilungswettbewerb zwischen den Affinitäten zu den möglichen Zuordnungskandidaten im zweiten Bild.
3. Verschmelzungswettbewerb (Merge Competition): Mehrere Basiselemente im ersten Bild stehen mit ihren Affinitäten im Wettbewerb um den gleichen möglichen Zuordnungskandidaten im zweiten Bild.

Teilweise ist dieser Wettbewerbsprozeß im Relaxationsverfahren von Barnard und Thompson /Barnard und Thompson 79, 80/ realisiert, wenn auch unter anderem Gesichtspunkt.

Der Begriff Affinität wurde auch in anderen Arbeiten für Ähnlichkeitsmaße benutzt. Freuder /Freuder 76/ hat in seinem hierarchischen Bereichverschmelzer das Ähnlichkeitsmaß für benachbarte Bereiche Affinität genannt. Als

Korrespondenzproblem Wettbewerbsmodell

Verschmelzungsbaum wurde bei diesem Verfahren der minimale spannende Baum gebildet, wobei stets diejenigen benachbarten Bereiche verschmolzen wurden, die die größte Affinität aufwiesen. Im Unterschied zu Ullman läßt sich Freuders Affinitätsmaß lokal ohne Interaktion mit benachbarten Affinitäten berechnen.

Dreschler und Nagel /Dreschler 77, Dreschler and Nagel 78/ haben bei der Verfolgung von bewegten Objekten das Korrespondenzproblem für auffällige Bildbereiche durch eine Ballungsanalyse zu lösen versucht. Nachdem jeder auffallende Bildbereich durch einen normierten Merkmalsvektor beschrieben war, wurde für alle Bereiche aus allen Bildern gleichzeitig der minimale spannende Baum ermittelt. Dieses zeigt starke Analogien zum Bereichverschmelzer von Freuder. Als Affinitätsmaß wurde der euklidische Abstand im Merkmalsraum verwendet. Die Erwartung war, daß auffällige Bereiche, die Abbildungen von bewegten Objekten entsprechen, im Merkmalsraum langgestreckte Ballungsgebiete bilden, da sich die Merkmale systematisch mit der Zeit verändern. Diese Bereiche sollten daher im minimalen spannenden Baum durch Ketten in zeitlicher Reihenfolge aufgefüdelt werden. Mit diesem Verfahren konnte ein PKW über mehr als 90 TV-Aufnahmen (wovon nur jede zweite ausgewertet wurde) verfolgt werden, wobei er eine Drehung von über 90 Grad ausführte. Probleme bei diesem Verfahren waren erstens die geeignete Normierung des Merkmalsraums und zweitens das Ermitteln von auffälligen Bereichen, da das einfache Differenzbildverfahren sich als nicht sicher genug erwies. Auch bei diesem Verfahren sind die Affinitäten unabhängig voneinander berechenbar, wenn man von der Normierung des Merkmalsraums absieht, bei der eine Statistik über Mittelwert und Streuung aller Merkmale ermittelt wird.

5.2 Übertragung des Wettbewerbsmodells

Das visuelle System des Menschen ist ein vielseitiges Instrument, das sehr allgemein einsetzbar ist und selten Fehler macht. Deshalb ist es erwägenswert, ob man nicht Bildverarbeitungssysteme für Maschinen ähnlich aufbauen sollte, da die Evolution erfahrungsgemäß für viele Probleme die beste Lösung herausgefunden hat. Allerdings ist zu bedenken, ob nicht einige Fehler, die das menschliche Auge macht (optische Täuschungen), daher rühren, daß das visuelle System des Menschen zu allgemein ausgelegt ist. Der Mensch kann nach Ullman z.B. drei verschiedene Arten von Bewegung verstehen, Bewegung fester Körper, Verwandlung (der Frosch im Märchen verwandelt sich in einen Prinzen) und Ersetzung (ein Gegenstand wird plötzlich durch einen anderen ersetzt). Das hier beschriebene Bildanalyse-system braucht nur die Bewegung als fester Körper zu kennen; das menschliche Auge wäre für diese Aufgabe unnötig flexibel.

Ullman hat ein Verfahren entwickelt, mit dessen Hilfe man nachprüfen kann, ob ein Satz von Korrespondenzen über mehrere Bilder sich als räumliche Bewegung eines starren Körpers interpretieren läßt (Genauerer siehe Kap. 6). Wenn es nicht durch die kombinatorische Vielfalt und durch die mangelnde Genauigkeit von Punktefindern illusorisch wäre, könnte man das Korrespondenzproblem für feste Körper durch systematisches Probieren lösen. Da dies am Aufwand scheitern muß, erscheint mir tatsächlich Ullmans Wettbewerbsverfahren als Heuristik am geeignetsten. Jedoch braucht man nicht alle Automaten nachzubilden, die zu optischen Täuschungen führen können.

Optische Täuschungen können z.B. dann auftreten, wenn sich ein periodisches Muster so bewegt, daß von einem Bild der Folge bis zum nächsten das Muster mit sich selbst nahezu zur Deckung kommt. Das Rad in Abb. 5.1 kann sich entweder um 35

Korrespondenzproblem Wettbewerbsmodell

Grad nach rechts oder um 55 Grad nach links bewegt haben. Das menschliche visuelle System entscheidet sich automatisch für die Lösung mit dem kleineren Drehwinkel. Ein gutes Bildanalyseprogramm könnte erkennen, daß es mehrere gleichwertige Lösungen gibt und die Entscheidung auf einen späteren Zeitpunkt verschieben. Die optische Täuschung mit dem zerbrochenen Rad kann von einem Analyseprogramm vermieden werden, wenn es weiß, welche Bildkomponenten sich zusammenhängend als starrer Körper bewegen sollten. Die Beschränkung auf starre Körper erlaubt es auch, auszunutzen, daß in der Bewegung eine gewisse Kontinuität vorhanden sein sollte, weil starre Körper nicht beliebig stark beschleunigt oder gebremst werden können. Das menschliche Auge verzichtet auf dieses Kriterium, wie Ullman gezeigt hat. Läßt man außerdem nur undurchsichtige Körper zu, kann man auch den Verlauf des optischen Flusses an der Objektoberfläche auswerten, denn dieser ist außer an räumlichen Diskontinuitäten lokal stets ungefähr parallel und von gleichem Betrag. Barnard und Thompson /Barnard and Thompson 79, 80/ nutzen bei ihrem Relaxationsverfahren besonders diese Eigenschaft aus.

Interessant sind diejenigen Automaten im visuellen System, die Affinitäten für die gleiche Anordnung von Basiselementen am Rande des Blickfeldes anders bewerten als im Zentrum. Für das menschliche Auge erscheint dies aus zwei Gründen eine sinnvolle Einrichtung: Erstens ist die Netzhaut nicht homogen aufgebaut, so daß Messungen am Rande anders gewichtet werden müssen als in der Mitte, weil Empfindlichkeit und Auflösung nicht überall gleich sind, und zweitens sorgt der Mensch durch Kopf- und Augenbewegungen in der Regel dafür, daß die wichtigen Objekte in der Mitte des Bildfeldes zu sehen sind. Für eine Maschine wären solche Automaten dann erwägenswert, wenn auch sie mit Sensoren ausgestattet wäre, die lokal unterschiedliche Eigenschaften haben, z.B. Randunschärfe, Verzeichnungen usw., oder wenn Kameraorientierung oder Vario-Objektiv automatisch

Korrespondenzproblem Wettbewerbsmodell

nachgeführt werden könnten. Die Abhängigkeiten der Affinitäten von der zeitlichen Auflösung dagegen erscheint mir grundsätzlich sinnvoll, da die Korrespondenzen sich um so zuverlässiger ermitteln lassen, je ähnlicher sich die aufeinanderfolgenden Bilder noch sind.

5.3 Minimale Zuordnung (minimal mapping)

Im Kapitel 3 seines Buches /Ullman 79/ beschreibt Ullman das Korrespondenzproblem als Optimierungsaufgabe aus statistischer Sicht. Er erläutert das Problem an einem einfachen Beispiel. Gegeben seien zwei Aufnahmen mit je n Bildelementen und gesucht sei die beste bijektive Abbildung der Elemente aus der ersten Aufnahme auf die der zweiten. Es sind $n!$ solcher Abbildungen möglich. Ohne Ausnutzung von Zusatzwissen, das die kombinatorische Vielfalt einschränkt, ist die Fehlerwahrscheinlichkeit sehr hoch. Bei der Stereoanalyse dienen in der Regel Kameramodelle als solche Einschränkungen. Ullman führt eine ganze Reihe von sehr allgemeinen Bedingungen an, die auf der Ebene der elementaren Zuordnungen anwendbar sind. Diese Einschränkungen sollen dazu dienen, aus den $n!$ möglichen Lösungen diejenigen mit der größten Wahrscheinlichkeit herauszufinden.

Um die Gesamtwahrscheinlichkeit für eine Zuordnung beurteilen zu können, muß man wissen, durch welche statistischen Abhängigkeiten sich die Wahrscheinlichkeiten der elementaren Zuordnungen gegenseitig beeinflussen. Ein Beispiel für solche Abhängigkeit zeigt Abbildung 5.2. Die offenen Kästchen stellen Punkte aus dem ersten Bild der Folge dar, die geschlossenen Kästchen Punkte aus dem zweiten Bild. Punkte aus dem ersten Bild haben Nummern, die mit X beginnen; Punktnummern für das zweite Bild beginnen mit Y .

Korrespondenzproblem
 Minimale Zuordnung

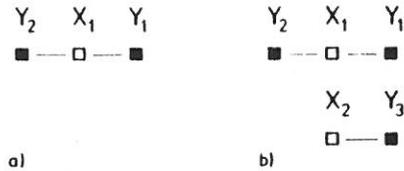


Abb. 5.2 statistische Unabhängigkeit
 von Affinitäten

In Abb. 5.2b könnte man annehmen, daß die Korrespondenz X1-Y1 hier wahrscheinlicher ist als in Abb. 5.2 a, da das Paar X2-Y3 einen optischer Fluß in diese Richtung nahelegt. Da Ullman in sein statistisches Modell nur solche Abhängigkeiten eingehen lassen wollte, die auch das menschliche Auge auswertet, hat er untersucht, welche Korrespondenzen der Mensch bei der oben beschriebenen Konfiguration wahrnimmt. Er hat keine Hinweise darauf gefunden, daß das parallele Paar X2-Y3 den Konflikt zwischen X1-Y1 und X1-Y2 lösen hilft. Ullman kommt daher zu dem Schluß, daß man die Bildelemente als sich unabhängig bewegend betrachten kann. Für den Korrespondenzprozeß ist zwar Wettbewerb zwischen den einzelnen Zuordnungen erforderlich, aber es werden keine verstärkenden Interaktionen zwischen den Affinitäten ähnlicher Verschiebungsrichtungen benötigt. Ullman zeigt später, daß solche Interaktionen redundant wären. Dieses ist ein konträrer Standpunkt zum Relaxationsverfahren von Barnard und Thompson /Barnard and Thompson 79, 80/, das gerade auf solchen verstärkenden Interaktionen beruht.

Korrespondenzproblem Minimale Zuordnung

Die Hypothese der statistischen Unabhängigkeit erleichtert die Berechnung der wahrscheinlichsten Zuordnung, da man so die "Maximum-Likelihood"-Methode anwenden kann. Sei v die Geschwindigkeit, mit der ein Bildelement sich von einer Aufnahme zur nächsten verschoben hat und $p(v)$ die dazugehörige Wahrscheinlichkeitsdichte. Hat das Bild n Elemente, so findet man die sicherste Lösung, indem man das Produkt $\prod p(v)$ maximiert über alle legalen Zuordnungen oder besser die Summe $\sum q(v)$ minimiert, mit $q(v) = -\log p(v)$. Die Funktion q kann man als Kostenfunktion auffassen und die Optimierungsaufgabe so verstehen, daß man die Lösung mit der geringsten Kostensumme sucht.

Ullman zeigt, wie dieses Optimierungsproblem mit einem Netzwerk von lokalen Prozessoren gelöst werden kann. Jeder Prozessor bearbeitet einen Bildpunkt und steht nur mit den Prozessoren benachbarter Bildelemente in Verbindung. Die Programme für diese Prozessoren können als lineare Programme formuliert werden.

Bisher wurde nur über eineindeutige Zuordnungen gesprochen. Da es jedoch nicht immer garantiert ist, daß die beiden beteiligten Aufnahmen gleich viele Bildelemente enthalten, bleibt die Frage, ob Mehrfachzuordnungen möglich sein oder lieber überzählige Bildelemente vernachlässigt werden sollten. Ullman läßt nur solche Zuordnungen zu, die alle Elemente erfassen. Ich halte dieses nicht für günstig, da so falsche Paarungen erzwungen werden, wenn für ein Bildelement der Partner im anderen Bild nicht sichtbar ist. Dieses kann vorkommen, wenn ein Bildelement neu hinzukommt, aus dem Bildbereich verschwindet oder verdeckt wird. Selbst im Falle, daß beide Bilder gleich viele Elemente enthalten, kann es falsch sein, eine deckende Zuordnung zu erzwingen, da zufällig gerade so viele Elemente neu hinzugekommen sein können wie unsichtbar geworden sind. Bei meinem Problem halte ich es für günstiger, lieber ein Bildelement gar nicht

Korrespondenzproblem Minimale Zuordnung

zuzuordnen, als es falsch zuzuordnen, da ich für die Korrespondenzketten von Bild zu Bild noch keine Konsistenzprüfung vorgesehen habe.

Ullmans Forderung nach deckenden Zuordnungen ist verständlich, da sein Optimierungsverfahren sonst die Lösung mit leerer Zuordnung als die kostengünstigste liefern würde.

Allerdings sind nicht alle deckenden Zuordnungen erwünscht. Wenn z.B. alle Elemente aus dem ersten Bild demselben Element aus dem zweiten Bild zugeordnet werden, sind viele überflüssige Paarungen nötig, um auch die restlichen Elemente aus dem zweiten Bild zu erfassen. Erwünscht sind vielmehr diejenigen Kombinationen, bei denen man mit einem Minimum an Paarungen vollständige Deckung erreicht. Im Idealfall sind das gerade so viele Paare, wie das Bild mit der größeren Elementzahl Elemente enthält.



Abb. 5.3 erwünschte (a) und
unerwünschte (b) Zuordnungen

Die Zahl der Paarungen läßt sich durch eine geeignete Kostenfunktion klein halten: Die Kostenfunktion wird um einen konstanten Summanden erweitert, der für jede Paarung gleich ist. Damit wird eine Lösung umso teurer, je mehr Paarungen sie enthält, denn der konstante Term bestraft jede Paarung gleichmäßig, unabhängig davon, ob sie ansonsten kostengünstiger wäre oder nicht. So können sich deckende

Korrespondenzproblem
 Minimale Zuordnung

Zuordnungen, die mit wenigen Verbindungen auskommen, gegenüber denjenigen durchsetzen, die viele Paarungen erfordern, aber dafür im variablen Kostenanteil günstiger gewesen wären.

Ullman hat eine ganze Reihe von Eigenschaften herausgefunden, die die Kostenfunktion haben müßte, wenn man damit das menschliche Auge simulieren wollte. Viele dieser Eigenschaften erreicht man durch eine äußerst einfache Kostenfunktion. Als Kosten für eine Paarung wird einfach der Abstand zwischen den beiden Partnern gewählt. Es ist erstaunlich, wie günstige Eigenschaften diese einfache Funktion hat. Sie verhindert, daß Paarungen über Kreuz auftreten können (Abb. 5.4) und sorgt dafür, daß möglichst parallele Paarungen gefunden werden. Im Extremfall allerdings ist der vorher diskutierte konstante Zusatzterm erforderlich, um die Zahl der Verbindungen klein zu halten. Dieses Verfahren nennt Ullman "minimale Zuordnung" (Minimal Mapping).

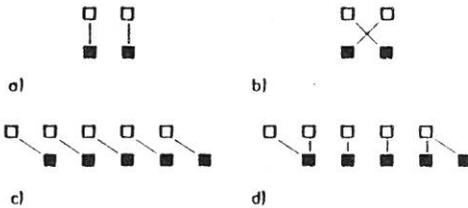


Abb 5.4 minimale Zuordnung

Lösung (a) ist kostengünstiger als (b), (c) günstiger als (d)

Korrespondenzproblem
Minimale Zuordnung

Die Eigenschaften der minimalen Zuordnung haben ihn in der Annahme bestärkt, daß keine verstärkenden Interaktionen zwischen elementaren Zuordnungen erforderlich sind, sondern daß diese redundant wären. Auch ohne die explizite Verstärkung von parallelen, kreuzungsfreien Paarungen werden genau diese vom minimalen Zuordnungsverfahren gefunden. Ich glaube allerdings nicht, daß sich diese Erfahrung auf den Fall der nicht deckenden Lösungen übertragen läßt, da hier Kriterien benötigt werden, die entscheiden, welche Verbindungen im Konfliktfall günstiger sind.

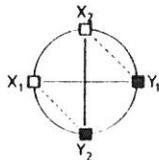


Abb 55 rotierende Kreisscheibe

In manchen Situationen verfolgt das minimale Zuordnungsverfahren genau die falsche Strategie. Besonders bei Rotationsbewegungen ohne Translationsanteil ist die Heuristik der kreuzungsfreien Zuordnung ungeeignet. Das wird am Beispiel der rotierende Scheibe in Abb. 5.5 deutlich. Wenn die Lösung mit den sich kreuzenden Verschiebungsvektoren gewählt wird, entspricht das einer Drehung der Scheibe um 180° , während die kreuzungsfreie Lösung einer reinen Translation entspricht.

5.4 Weitere Lösungsansätze

In einer ganzen Reihe von Arbeiten wird das Korrespondenzproblem auf der Ebene von komplexen Strukturbeschreibungen zu lösen versucht. Interessant sind z.B. die Graphvergleichsmethoden von Radig /Radig 78a, 78b, Radig et al. 80/, Kraasch et al. /Kraasch et al. 79a, 79b/ und Bullock /Bullock 76/. Auch das hierarchische Verfahren von Tsuji et al. /Tsuji et al. 79, 80/ und das Verfahren von Price und Reddy /Price and Reddy 79/ sind bemerkenswert. Alle diese Verfahren gehen von einer Segmentation des Bildes aus und verwenden Bildbeschreibungen mit komplexen Merkmalen. Da für das von mir verfolgte Konzept Verfahren wichtig sind, die ohne Segmentation auskommen, werde ich auf diese Arbeiten nicht näher eingehen, sondern mich auf Verfahren beschränken, die lokale Bildmerkmale vergleichen.

Ein großer Teil der Lösungsansätze für das Korrespondenzproblem wurde für die Analyse von Stereobildern entwickelt. Allen diesen Arbeiten ist die folgende Konfliktsituation gemeinsam: Je größer der Abstand zwischen den Kameras ist, desto größere Tiefen kann man genau messen, denn eine größere Stereobasis vergrößert die Parallaxe und damit die relative Genauigkeit der Triangulation. Andererseits wird die Ermittlung der Korrespondenzen umso schwieriger, je größer die Parallaxe und damit die perspektivischen Unterschiede zwischen den Bildern werden. Außerdem wird bei größerem Kameraabstand auch der von beiden Kameras gleichzeitig erfaßte Raumwinkel kleiner, so daß nur für einen kleineren Bereich Tiefenmessungen vorgenommen werden können. Hier muß zwischen zwei gegeneinander wirkende Fehlerquellen ein gangbarer Mittelweg gefunden werden. Der Lösungsvorschlag von Marr und Poggio /Marr and Poggio 79/ wurde schon in Kapitel 4.1 angesprochen. Wichtig sind drei Anforderungen:

Korrespondenzproblem Korrelationsverfahren

1. Die Korrespondenzen müssen richtig sein.
2. Sie müssen genau sein .
3. Die kombinatorische Vielfalt muß eingeschränkt werden.

5.4.1 Korrelationsverfahren

Eine grundlegende Arbeit stammt von Hannah /Hannah '74/. Hannah unterscheidet zwischen "Matching" und "Mapping". Mit "Mapping" ist jede allgemeine Transformation gemeint, die Teilbereiche aus einem Bild in das folgende Bild überführt. "Matching" ist der Spezialfall, daß kleine Teilbereiche des ersten Bildes nur durch Translation in das zweite Bild abgebildet werden können. Hannah und die meisten anderen Autoren arbeiten mit einer so kleinen Stereobasis, daß es zulässig ist, ein "Matching"-Verfahren zu verwenden. Dieses ermöglicht direkte Grauwertvergleiche als Ähnlichkeitsmaß einzusetzen. Hannah wählt im ersten Bild kleine Grauwertmasken um die markanten Punkte aus. Für diese Masken sucht sie im zweiten Bild durch Korrelation oder andere statistische Ähnlichkeitsmaße (Siehe Kap. 4) nach ähnlichen Ausschnitten. Aus der Autokorrelation der Maske wird ein Schwellwert für die Güte des Ähnlichkeitswertes abgeleitet.

Die Suche mit der Korrelation oder anderen Grauwertvergleichsmethoden ist nur dann angebracht, wenn die Suchmasken nicht für schwer abzuschätzende Veränderungen von Perspektive und Entfernung korrigiert werden müssen. Solange man selbst Kontrolle über die Kamerabasislänge hat, kann man die maximal möglichen perspektivischen Verzerrungen abschätzen und die Basislänge geeignet wählen. Wenn man dagegen bewegte Objekte beobachtet, kann man über

Korrespondenzproblem
Korrelationsverfahren

Verzerrungen wenig aussagen, da man keinen Einfluß auf die Objektbewegungen hat (es sei denn, man kennt Schätzwerte für die Geschwindigkeit). Ein Fahrzeug z.B., das mit der im Stadtverkehr zugelassenen Höchstgeschwindigkeit von 50 km/h fährt, legt in einer fünfundzwanzigstel Sekunde ungefähr einen halben Meter zurück, das ist ein nicht zu vernachlässigender Teil der Gesamtlänge. Wenn das Fahrzeug nur nahe genug ist, dann sind die perspektivischen Änderungen beträchtlich, selbst bei einer reinen Translationsbewegung /Budy und Dreschler 76/.

Interessant sind Hannahs Methoden zur Suchraumeinschränkung. Als erstes werden einige Stützpunkte gewählt, die durch den Verlauf ihrer Autokorrelation eine sichere und genaue Identifikation versprechen. Für diese Punkte wird mittels eines Kameramodells diejenige Linie ermittelt, auf der sie im zweiten Bild liegen müssen. In der Umgebung dieser Geraden wird mittels einer Maske durch Korrelation oder ähnliche Methoden der beste Zuordnungskandidat gesucht. Kommen mehrere Punkte in Frage, dann wird die Entscheidung verschoben. Für jede sicher gefundene Zuordnung wird in einem rekursiven Verfahren jeder Nachbar überprüft, ob er sich mit dem entsprechenden Nachbarn des Zuordnungspunktes paaren läßt. Auf diese Weise wachsen ohne weitere Suche aus Einzelzuordnungen ganze Zuordnungsbereiche, wobei zum Teil auch Mehrdeutigkeiten geklärt werden können. Der Suchaufwand läßt sich noch mehr einschränken, wenn man ausnutzt, daß die Korrelation in der Nähe des Maximums in der Regel monoton ansteigt. Man kann zunächst mit einem gröberen Raster die Nähe des Maximums suchen und dann durch ein Gradientensuchverfahren die genaue Lage ermitteln.

Ebenfalls mit statistischen Grauwertvergleichen und einem Kameramodell arbeiten Moravec /Moravec 77, 79, 80/ und Nevatia /Nevatia 76/. Moravecs Verfahren unterscheidet sich von Hannahs im wesentlichen nur durch den Punktefinder,

Korrespondenzproblem Korrelationsverfahren

während Nevatias auch für große Stereobasislinien anwendbar ist. Nevatia verwendet nicht nur ein Stereopaar sondern eine ganze Folge von Aufnahmen, die mit sehr kleinem Winkelabstand entstanden sind. Die Korrespondenzen über die ganze Basis ergeben sich durch Konkatenation von kleinen Schritten. Die Frage hierbei ist, ob die Genauigkeit, die man für die kleinen Stereowinkel gewinnt, nicht durch Fehlerfortpflanzung wieder aufgehoben wird, da eine einzige falsche Korrespondenz in einer Kette die ganze Kette fehlerhaft macht, wenn keine Konsistenzprüfungen vorgenommen werden. Dieses Verfahren von Nevatia ist unter dem Namen "Motion stereo" bekannt.

5.4.2 Relaxationsverfahren

Relaxationsverfahren sind iterative, parallele Algorithmen, die sich dazu eignen, durch lokale Nachbarschaftsuntersuchungen die Knoten eines Graphs konsistent zu färben. Im allgemeinen Fall sind eine endliche Menge von Objekten und eine endliche Menge von Klassen gegeben, wobei nach der Klassenzugehörigkeit der Objekte gefragt wird. Eine Randbedingung ist, daß die Klassenzugehörigkeiten von benachbarten Objekten bestimmte, problemabhängige Konsistenzbedingungen erfüllen müssen. Die Objekte können als Knoten eines Graphen aufgefaßt werden, dessen Kanten die Nachbarschaftsrelation darstellen. Jedem Objekt wird zu Beginn des Relaxationsverfahrens eine Liste von Marken (label) zugeordnet. Diese Markenliste enthält die Namen der möglichen Klassen, denen das Objekt angehören kann. Zu jeder Marke wird eine Anfangsschätzung für die Wahrscheinlichkeit ermittelt, daß sie für das jeweilige Objekt die richtige Klasse angibt. Die Anfangswahrscheinlichkeiten werden iterativ verbessert, indem für jede Marke ein Maß für die Kompatibilität zu den Markenlisten der benachbarten Objekte ausgewertet wird. Die meisten Unterschiede zwischen

Korrespondenzproblem Relaxationsverfahren

den gängigen Relaxationsverfahren bestehen bezüglich dieser Konsistenzmessung.

Relaxationsverfahren werden schon recht lange in der Bildverarbeitung und auch in der Spracherkennung eingesetzt /Rosenfeld et al. 76, Rosenfeld 77/, aber die Untersuchung ihrer mathematischen Eigenschaften, wie Konvergenzverhalten, ist noch nicht abgeschlossen. Pavlidis /Pavlidis 77/ vergleicht ein allgemeines Relaxationsverfahren mit einem Tiefpaß, der bei jedem Iterationsschritt einen Teil der Anfangsinformation wegfiltiert. Daher können Relaxationsverfahren zu einem Fixpunkt konvergieren, der keine geeignete Lösung des gestellten Problems darstellt. Es ist wichtig, das Relaxationsverfahren rechtzeitig zu beenden. Haralick et al. /Haralick et al. 80/ diskutieren für den Spezialfall von nur zwei Klassen, wie der erreichte Fixpunkt von den Kompatibilitätskoeffizienten abhängt und wie stabil er ist. Zucker et al. /Zucker et al. 79/ untersuchen, wie sich die Äquivalenz zweier Relaxationsverfahren beweisen läßt. Hummel und Zucker /Hummel and Zucker 80/ diskutieren den Zusammenhang zwischen der Variationsrechnung und allgemeinen Relaxationsverfahren. Richards et al. /Richards et al. 80/, Helland /Helland 80/ und Peleg /Peleg 80/ schlagen Kontrollstrategien vor, mit denen die Konvergenzgeschwindigkeit verbessert werden /Helland 80/ oder ein Informationsverlust durch zu langes Iterieren verhindert werden kann.

5.4.2.1 Ranade und Rosenfeld

Ranade und Rosenfeld /Ranade and Rosenfeld 80/ haben ein Relaxationsverfahren entwickelt, mit dem diejenige Translation ermittelt werden kann, für die zwei Punktwolken am besten zur Deckung gebracht werden können. Dieses Verfahren ist für den Fall gedacht, daß ein großer zusammenhängender

Korrespondenzproblem Relaxationsverfahren

Bildbereich als ganzes starr verschoben worden ist. Das ist ein typischer Anwendungsfall für die Suche mit Korrelationsstechniken. Der Korrelationsaufwand ist mindestens proportional zur Fläche des gesuchten Ausschnitts. In dem von Ranade und Rosenfeld vorgeschlagenen Verfahren werden einige wenige markante Punkte im gesuchten Ausschnitt ausgewählt und mit markanten Punkten im Suchbild zu kombinieren versucht. Der Aufwand hierfür ist proportional zur vierten Potenz der Punktezahl. Bei großen Bildern mit nur wenigen ausgewählten Punkten ist dieses Verfahren der Korrelation vorzuziehen.

Das Relaxationsverfahren selbst basiert auf einer einfachen Idee. Für jede mögliche Kombination eines Punktes P aus dem ersten Bild mit einem Punkt Q aus dem zweiten Bild wird ein Gütemaß der Deckung errechnet. In dieses Gütemaß geht eine Fehlerfunktion ein, die für alle Punkte aus dem ersten Bild mißt, wie dicht sie bei der gerade betrachteten Verschiebung P-Q mit einem Punkt aus dem zweiten Bild zur Deckung kommen. In den nächsten Iterationsschritten werden die gleichen Güteberechnungen wiederholt, nur werden bei den Fehlerberechnungen die vorher ermittelten Gütewerte berücksichtigt. Die Autoren geben an, daß sich in der Regel schon nach drei bis vier Iterationsschritten ein stabiler Zustand einstellt, bei dem sich die Gütewerte für die gesuchte Translation deutlich von den anderen Werten abheben.

Dieses Verfahren ermittelt den besten Verschiebungsvektor für den betrachteten Ausschnitt als Ganzes, wobei kleinere Verzerrungen nicht stören. Auch bei einer leichten Rotation wird diejenige Translation ermittelt, für die der Fehler am kleinsten ist; die Rotation wird allerdings nicht erkannt.

Korrespondenzproblem Relaxationsverfahren

Für mich ist dieses Verfahren in der vorgeschlagenen Form nicht verwendbar, da die von mir untersuchten Objekte rotieren und perspektivische Veränderungen auftreten können. Außerdem liefert das Verfahren nur einen globalen Verschiebungsvektor und nicht die genauen Einzelpunktkorrespondenzen, die ich als Eingangsdaten für die 3D-Rekonstruktion benötige. Ein weiterer Nachteil ist, daß das Relaxationsverfahren keine Ähnlichkeitsmaße für die markanten Punkte auswertet, so daß z.B. auch schwarze mit weißen Punkten kombiniert werden könnten. Dieser Nachteil ließe sich allerdings durch einfache Modifikationen beheben. Es ist aber zu überlegen, ob man dieses Verfahren nicht dazu einsetzen könnte, um einen Schätzwert für den Translationsanteil der Bewegung zu erhalten. Nach dieser Grobanalyse könnte ein einfaches heuristisches Verfahren wie das in 5.7 beschriebene "Nächster-Nachbar-Verfahren" die korrespondierenden Einzelpunkte finden. Dieses Vorgehen ist aber nur dann sinnvoll, wenn der Rotationsanteil der Bewegung klein ist gegen den Translationsanteil.

5.4.2.2 Das Relaxationsverfahren nach Barnard und Thompson

Barnard und Thompson /Barnard and Thompson 79, 80/ bauen ihren Lösungsvorschlag für das Korrespondenzproblem auf drei Prinzipien auf: Unterscheidbarkeit, Ähnlichkeit und Konsistenz.

1. Unterscheidbarkeit (discreteness): Im Bild müssen markante Bildpunkte erkennbar sein.
2. Ähnlichkeit (similarity): Für kleine Umgebungen von markanten Bildpunkten müssen Ähnlichkeitsmessungen möglich sein.

Korrespondenzproblem
Relaxationsverfahren

3. Konsistenz (consistency): Die vermuteten Verschiebungen gegenüber dem nächsten Bild sollen lokal ähnlich sein. Dies basiert auf der Annahme, daß der optische Fluß auf kontinuierlichen Oberflächen sich ebenfalls kontinuierlich ändert.

Diese drei Prinzipien werden in einem Relaxationsverfahren eingesetzt, das Ähnlichkeiten mit dem Verfahren von Ranade und Rosenfeld hat, aber in einigen Punkten leistungsfähiger ist. Ebenso wie Ranade und Rosenfeld verringern Barnard und Thompson die kombinatorische Vielfalt, indem sie sich auf ausgewählte markante Punkte beschränken. Dies ist mit dem Unterscheidbarkeitskriterium gemeint. Die wesentlichen Vorteile gegenüber dem Verfahren von Ranade und Rosenfeld sind die folgenden: Erstens werden die Startwerte für die Iterationen mit Ähnlichkeitsmaßen für die beiden kombinierten Punkte initialisiert, so daß unwahrscheinliche Kombinationen von vornherein ausgeschlossen werden können, und zweitens ermittelt dieses Verfahren die Translationswerte lokal, d.h. für jeden Punkt aus dem ersten Bild wird der am besten geeignete Punkt im zweiten Bild gesucht, so daß auch Rotationen erfaßt werden können. Für den ersten Vorteil wird das Ähnlichkeitsprinzip ausgenutzt und für den zweiten Vorteil das Konsistenzprinzip. Ein großer Nachteil des Verfahrens ist, daß sehr viele Parameter benötigt werden.

Das Verfahren arbeitet in groben Zügen wie folgt: Die Objekte, für die im Verlauf des Relaxationsverfahrens die Klassenzugehörigkeiten ermittelt werden, sind in diesem Fall markante Bildpunkte aus dem ersten Bild. Die gesuchten Klassen sind die Verschiebungen. Als erstes wird eine Entfernung R' festgelegt, die die Maximalzahl von Pixeln angibt, um die ein Punkt sich in horizontaler oder vertikaler Richtung verschoben haben darf. Damit beschränkt sich die Menge der möglichen korrespondierenden Bildpunkte im

Korrespondenzproblem
Relaxationsverfahren

zweiten Bild auf die markanten Bildpunkte innerhalb der durch R' vorgegebenen Umgebung. Für jeden Punkt aus dem ersten Bild wird eine Liste von Marken (l) erstellt, die für jeden möglichen Verschiebungsvektor die Komponenten (l_x, l_y) enthält. Außerdem ist eine Marke l^* vorgesehen, die für undefinierte Verschiebung steht. Für jede Marke (l) wird eine Anfangsschätzung $w(l)$ für ihre Wahrscheinlichkeit ermittelt, die aus der Grauwertähnlichkeit in der 5×5 -Umgebung der kombinierten Punkte errechnet wird. Dieses Maß ist so normiert, daß es gleich 1 ist, wenn die Grauwertumgebungen identisch sind und gegen Null geht, wenn sie stärker differieren. Ein Parameter C steuert, wie stark ein Unterschied im Grauwert in das Maß eingeht.

$$w(l) = 1 / (1 + C * s(l)), \quad l \neq l^*$$

$s(l)$ = Summe der quadrierten Grauwertdifferenzen der Testflächen um die Punkte (x,y) und $(x+l_x, y+l_y)$

Für die Marke l^* wird die Anfangsschätzung aus dem größten vorkommenden Wert abgeleitet, indem man annimmt, daß die Wahrscheinlichkeit dafür, daß keine Zuordnung möglich ist, ungefähr $1 - (\text{größter Wert})$ ist. Um zu Wahrscheinlichkeitsmaßen $p(l)$ zu kommen, wird für jeden Punkt die Summe der Anfangsschätzungen $w(l)$ aller Marken l auf die Summe 1 normiert. Die von Barnard und Thompson verwendeten Wahrscheinlichkeiten sind im mathematischen Sinne strenggenommen keine Wahrscheinlichkeiten sondern heuristische Konfidenzmaße, mit denen wie mit Wahrscheinlichkeiten gerechnet wird. Ich werde diese Konfidenzmaße dennoch weiterhin Wahrscheinlichkeiten nennen, um dieselbe Terminologie wie Barnard und Thompson zu verwenden.

Korrespondenzproblem
Relaxationsverfahren

Die Anfangswahrscheinlichkeiten $p(1)$ werden iterativ verbessert, indem das Konsistenzkriterium ausgewertet wird. Dazu wird eine weitere Entfernung R'' gewählt, die festlegt, bis zu welcher Entfernung Punkte als benachbart betrachtet werden. Innerhalb dieser Nachbarschaft sollen sich die Wahrscheinlichkeiten $p(1)$ von konsistenten Marken l während des Iterationsvorganges gegenseitig erhöhen können. Marken sind konsistent, wenn die entsprechenden Verschiebungen ungefähr gleichen Betrag und gleiche Richtung haben. Barnard und Thompson betrachten zwei Marken l' und l'' als konsistent, wenn der Betrag des Differenzvektors kleiner ist als eine Schwelle T , wobei der Betrag nach der Maximumsnorm gemessen wird. Bei jedem Iterationsschritt $k+1$ wird die alte Wahrscheinlichkeit $p_k(1)$ einer Marke l dadurch verändert, daß eine Summe $q_k(1)$ ermittelt wird, in der für alle konsistenten Marken von benachbarten Punkten die Wahrscheinlichkeiten aufsummiert werden. $q_k(1)$ soll für alle Marken l hoch sein, für die es viele benachbarte ähnliche Marken mit hoher Wahrscheinlichkeit gibt. Die neuen Schätzungen $w_{k+1}(1)$ ergeben sich nach folgenden Formeln:

$$w_{k+1}(1) = p_k(1) * (A + B * q_k(1)), \quad l \neq l^*$$

$$w_{k+1}(l^*) = p_k(l^*)$$

Die neuen Wahrscheinlichkeiten $p_{k+1}(1)$ ergeben sich durch Normieren der $w_{k+1}(1)$ auf die Summe 1. Nur bei diesem Normierungsschritt können Wahrscheinlichkeiten für eine Marke sinken. Eine Marke, die in der Liste keinen Konkurrenten hat, kann während der Iteration ihren Wahrscheinlichkeitswert nur erhöhen. In diesem Fall hängt es nur von der Anfangswahrscheinlichkeit ab, ob für den betreffenden Punkt mit nur einer möglichen Verschiebung eine Zuordnung getroffen wird oder nicht. Die Konfigurationen in der Umgebung haben nur einen verstärkenden Einfluß, können aber die Wahrscheinlichkeit nicht abwerten. Die Parameter A und B steuern, wie stark der Wert von q die Wahrscheinlichkeit

Korrespondenzproblem Relaxationsverfahren

einer Marke beeinflusst. Dieses ist nötig, da die Wahrscheinlichkeit einer Marke nicht zu schnell sinken darf, denn es können sich manchmal erst nach einer Reihe von Iterationsschritten Werte von weiter entfernten Punktpaaren durch das Netz von Punkten hindurch durchsetzen und den Wert noch positiv beeinflussen.

Die Iterationen werden nach einer fest vorgegebenen Zahl von 10 Schritten abgebrochen. Für jeden Punkt wird in der Markenliste nachgesehen, welche Marke die größte Wahrscheinlichkeit hat. Ist nur eine Marke mit großer Wahrscheinlichkeit vorhanden, so wird der Punkt als zugeordnet betrachtet; sind mehrere Marken mit ähnlich hoher Wahrscheinlichkeit vorhanden, ist der Fall mehrdeutig und es wird keine Zuordnung vorgenommen, ebenso wenn alle Marken eine niedrige Wahrscheinlichkeit haben.

Das Relaxationsverfahren erinnert stark an Ullmans Wettbewerbsmodell: Die Wahrscheinlichkeiten kann man mit Ullmans Affinitäten vergleichen. Die Wahrscheinlichkeiten stehen zueinander im Teilungswettbewerb, da für markante Punkte aus dem ersten Bild die Wahrscheinlichkeiten aller möglichen Verschiebungen miteinander konkurrieren.

5.5 Modifikationen am Relaxationsverfahren

5.5.1 Symmetrie im Wettbewerb

Beim Arbeiten mit dem Relaxationsverfahren von Barnard und Thompson haben sich einige Effekte als störend herausgestellt. Zunächst einmal fällt auf, daß die Ergebnisse nicht unabhängig von der Zuordnungsrichtung sind. Wenn man Punkte aus dem ersten Bild dem zweiten zuordnet, ist das Ergebnis nicht unbedingt das gleiche, als wenn man die Zuordnung umgekehrt vornehmen würde. Dies liegt an einer

Korrespondenzproblem Relaxationsverfahren

Unsymmetrie des Verfahrens. In Ullmans Terminologie ausgedrückt, ist nur einer der beiden Wettbewerbsmechanismen realisiert, nämlich der Teilungswettbewerb. Das Gegenstück dazu, der Verschmelzungswettbewerb, wurde nicht berücksichtigt. Dieses hat neben der Richtungsabhängigkeit des Korrespondenzergebnisses noch einen weiteren Nachteil zur Folge: Es können mehrdeutige Situationen auftreten, wenn für zwei oder mehr Punkte aus dem ersten Bild der gleiche Punkt im zweiten Bild als Kandidat mit der größten Wahrscheinlichkeit ermittelt wurde. Ein solcher Konflikt ist nicht durch den Vergleich der Wahrscheinlichkeiten zu entscheiden, da diese nur bezüglich der gemeinsamen Ausgangspunkte normiert werden und daher bezüglich der Zielpunkte nicht vergleichbar sind.

Ich habe das Verfahren daher so verändert, daß für jedes Paar von Zuordnungskandidaten zwei Wahrscheinlichkeitswerte geführt werden, einer, der bezüglich des Ausgangspunktes für den Teilungswettbewerb normiert wird, und einer, der bezüglich des Zielpunktes für den Verschmelzungswettbewerb normiert wird. Bei der Iteration wird jetzt anstelle der alten Wahrscheinlichkeiten der Mittelwert der beiden neuen Wahrscheinlichkeitswerte verwendet. So werden schon in der Iterationsphase beide Wettbewerbsarten berücksichtigt und nicht erst nach Beendigung der Iteration bei der Entscheidung über die endgültigen Zuordnungen. Durch diese Erweiterung ist das Relaxationsverfahren richtungsunabhängig geworden. Darüber hinaus können einige vorher mehrdeutige Situationen geklärt werden, in denen vorher anhand des Iterationsergebnisses keine Entscheidung möglich war.

5.5.2 Das Konsistenzkriterium

Ein weiterer unbefriedigender Aspekt ist, daß Barnard und Thompson das Konsistenzkriterium nicht vollständig ausnutzen. Innerhalb der Nachbarschaft können Marken von hoher Wahrscheinlichkeit die Wahrscheinlichkeiten von konsistenten Marken erhöhen, aber sie können nicht die Wahrscheinlichkeiten von inkonsistenten Marken erniedrigen. Wahrscheinlichkeiten können daher nur durch den Wettbewerb sinken, indem konkurrierende Werte stärker erhöht werden und bei der anschließenden Normierung der weniger stark gestiegene Wert sinkt. Es gibt Situationen, in denen dieser Mechanismus ungenügend ist. Wenn die Konkurrenz fehlt, weil ein Punkt nur einen einzigen Zuordnungskandidaten hat, kann sich seine Wahrscheinlichkeit nur noch erhöhen. Dies ist selbst dann der Fall, wenn in der Umgebung nur inkonsistente Zuordnungen vorhanden sind, die gegen die Paarung sprechen. Durch die Hinzunahme des Verschmelzungswettbewerbs wurde dieser Nachteil etwas gemildert, denn oft sind bei einem Punktepaar für den einen Partner zwar keine Zuordnungsalternativen vorhanden, aber für den zweiten wohl. Für eine solche Konstellation kann sich wenigstens die mittlere Wahrscheinlichkeit durch den zweiten Wettbewerbsmechanismus erniedrigen.

Wenn allerdings beide Partner keine Zuordnungsalternativen besitzen, reichen auch die bisherigen Modifikationen nicht aus, um die Korrespondenzwahrscheinlichkeit abzuschwächen. Dieses läßt sich nur erreichen, indem das Konsistenzkriterium besser ausgenutzt wird, als es bisher getan wurde. Der Korrekturterm in der Iterationsformel muß so umgestaltet werden, daß die Wahrscheinlichkeit für das Korrespondieren eines Punktepaares durch nicht konsistente benachbarte Paare gesenkt und durch konsistente erhöht werden kann. Wenn die Umgebung eines Punktepaares leer ist, oder sich konsistente und nicht konsistente Paare die Waage halten, darf kein

Korrespondenzproblem
Relaxationsverfahren

korrigierender Beitrag erfolgen. In solch einem Fall sollte nur das Ähnlichkeitsmaß entscheiden, das in der Anfangswahrscheinlichkeit enthalten ist.

Aufgrund dieser Überlegungen habe ich die Iterationsformel in folgender Weise verändert. Für jedes mögliche Zuordnungspaar werden innerhalb der R'' -Umgebung zwei Korrektursummen ermittelt, einmal die Wahrscheinlichkeitssumme der konsistenten Paare (S_k) und einmal die Wahrscheinlichkeitssumme der inkonsistenten Paare (S_i). Die a-priori Wahrscheinlichkeiten dafür, daß ein zufällig gewählter Verschiebungsvektor zu einer gegebenen Verschiebung konsistent oder inkonsistent ist, können verschieden voneinander sein. Deshalb werden beide Summen mit dem jeweiligen Reziprokwert der Wahrscheinlichkeit gewichtet, daß eine beliebige legale Zuordnung konsistent (w_k) beziehungsweise inkonsistent ist (w_i). Diese a-priori Wahrscheinlichkeiten für konsistente und inkonsistente Verschiebungen lassen sich daraus errechnen, daß es innerhalb der R' -Umgebung $(2 \cdot R' + 1)^2$ mögliche Verschiebungen gibt, wovon maximal $(2 \cdot T + 1)^2$ konsistent sind (zur Verschiebung (R', R') gibt es nur $(T + 1)^2$ konsistente Verschiebungen). Dabei ist T die größte Differenz, die zwei Verschiebungsvektoren in einer Koordinate haben dürfen, wenn sie noch konsistent sein sollen. Bei der Ermittlung der Wahrscheinlichkeiten wurde statistische Unabhängigkeit vorausgesetzt. Die neue Iterationsformel lautet:

$$w_{k+1}(l) = p_k(l) * F * (1 + (S_k/w_k - S_i/w_i) / (S_k/w_k + S_i/w_i))$$

Der von den Wahrscheinlichkeitssummen abhängige Quotient kann nur zwischen den Werten 1 und -1 schwanken. Wenn der Summenanteil der konsistenten Nachbarn dominiert, geht der Quotient gegen 1 und die Affinität wird erhöht, während ein

Korrespondenzproblem Relaxationsverfahren

Dominieren der inkonsistenten Nachbarn den Quotienten g gegen -1 gehen läßt und die Wahrscheinlichkeit abnimmt.

Ungeklärt ist noch die Rolle des konstanten Faktors F . Dieser Faktor wird in dem Fall wichtig, wo die Umgebung eines Punktepaars leer ist. Ohne diesen Faktor würde über alle Iterationen hinweg die alte Wahrscheinlichkeit erhalten bleiben. Da aber die Startwerte in der Regel zu niedrig sind, um am Ende akzeptiert werden zu können, muß sich die Wahrscheinlichkeit während der Iteration steigern. Der Faktor F ist so zu wählen, daß relativ hohe Anfangswerte nach Beendigung der Iterationen die Akzeptanzschwelle überschreiten, die niedrigen sie aber nicht erreichen können. Da es schwer ist, F im voraus festzulegen, wird er vom Programm ermittelt, wenn die Startwahrscheinlichkeiten festliegen. F wird so ausgerechnet, daß Werte, die größer sind als der Mittelwert der Startwahrscheinlichkeiten, nach der vorgegebenen Iterationszahl den Wert 1 erreichen. Zehn Iterationsschritte haben sich experimentell als ausreichend erwiesen, um einen stabilen Zustand zu erreichen.

5.5.3 Die Anfangswahrscheinlichkeiten

Der dritte wesentliche Unterschied gegenüber der ursprünglichen Version liegt in der Berechnung der Startwahrscheinlichkeiten. Bei meinem Verfahren wird mehr vom Punktefinder übernommenes Wissen ausgewertet, da der neu entwickelte Punktefinder im Gegensatz zu dem von Barnard und Thompson eingesetzten Moravec-Operator eine Strukturbeschreibung der gefundenen Punkte liefert. Bei der Initialisierung des Relaxationsverfahrens lasse ich von vornherein nur solche Paarungen zu, bei denen die Vorzeichen der Hauptkrümmungen im Maximum der Gaußschen Krümmung für beide Partner die gleichen sind. Dieser Test ist gleichbedeutend mit der Prüfung, ob das zum Punkt gehörige Maximum einer

Korrespondenzproblem Relaxationsverfahren

Hügelkuppe oder einer Senke entspricht, d.h. ob es sich um dunkle Ecken auf hellem Grund oder um helle Ecken auf dunklem Grund handelt. Diese Vorauswahl beschleunigt die Iteration auf zweifache Weise: Erstens sind weniger mögliche Paarungen zu überprüfen und zweitens enthalten die Nachbarschaften weniger Kandidaten, so daß die Summenbildung beschleunigt wird.

Bei der Berechnung der Anfangswahrscheinlichkeiten habe ich den Parameter C nicht mehr variabel gehalten, sondern auf $1/25$ festgelegt. Bei dieser Wahl ist das Ähnlichkeitsmaß genau die mittlere quadratische Grauwertabweichung (Maskengröße 5×5). Zusätzlich multipliziere ich die Startwerte mit den vom Punktefinder übernommenen Konfidenzmaßen der beiden beteiligten Punkte (siehe Abschnitt 4.5.1), so daß in der Anfangsphase der Iteration diejenigen Zuordnungen im Vorteil sind, bei denen die beteiligten Punkte besonders sicher gefunden wurden.

5.5.4 Ergebnisse

Abb. VIII zeigt die Arbeitsweise des Relaxationsverfahrens anhand von markanten Punkten aus dem ersten und dritten Bild von Folge 1. In Abb. VIIIa ist der untersuchte Ausschnitt aus dem dritten Bild der Folge zu sehen. Die in den beiden Bildern gewählten Punkte sind als offene Quadrate (drittes Bild) und als Punkte (erstes Bild) in das Bild des Taxis eingetragen. Je nach Untergrundhelligkeit wurden diese Markierungen in schwarz oder weiß eingetragen, so daß sie sich deutlich von der Umgebung abheben. Die Abb. VIIIb bis VIIIe zeigen die Zwischenergebnisse des Relaxationsverfahrens. Alle Punkte innerhalb der R' -Nachbarschaft, die als Kandidaten für korrespondierende Punkte in Frage kommen, sind durch Striche verbunden dargestellt. Die Striche sind

Korrespondenzproblem Ergebnisse

umso dunkler, je größer die Korrespondenzwahrscheinlichkeit ist. Abb. VIIIb zeigt die aufgrund der Grauwertähnlichkeiten in den Umgebungen der Punkte geschätzten Anfangswahrscheinlichkeiten und die folgenden Abbildungen den 2., 4. und 10. Iterationsschritt.

Es ist schon bei der zweiten Iteration zu sehen, daß die Wahrscheinlichkeiten derjenigen Verschiebungen steigen, die mit der Verschiebungsrichtung des Objektbildes übereinstimmen. Einige Kandidaten mit sehr hoher Anfangswahrscheinlichkeit dagegen sinken in der Wahrscheinlichkeit, da ihre Verschiebungen nicht parallel zu denjenigen in ihrer Nachbarschaft sind.

In der oberen rechten Ecke liegen die markanten Punkte, die im parkenden VW im Hintergrund gefunden wurden. Hier konnten nicht alle Korrespondenzkonflikte gelöst werden, da diese Punkte sich nicht konsistent mit den Punkten aus dem Taxi verschoben haben.

Bei den markanten Punkten an den Seitenfenstern des Taxis fällt auf, daß sich der Konflikt zwischen den parallelen und den sich kreuzenden Verschiebungen nicht lösen läßt. Die Ursache hierfür ist, daß bei der von mir gewählten Konsistenzschwelle von $T=3$ die sich kreuzenden Verschiebungen mit den parallelen konsistent sind. Um die sich kreuzenden Verschiebungen mit der Bewegungsrichtung inkonsistent zu machen, könnte man zwar die Schwelle erniedrigen, aber das würde das Problem nur verlagern. Es erscheint mir günstiger, die Schwelle T durch ein kontinuierliches Konsistenzmaß zu ersetzen, das die Ähnlichkeit zwischen den Verschiebungen berücksichtigt.

Korrespondenzproblem Ergebnisse

5.5.5 Parameter

Bei meiner Version des Relaxationsverfahrens sind im Vergleich zum Ansatz von Barnard und Thompson /Barnard and Thompson 79, 80/ drei Parameter nicht mehr eingesetzt worden (A, B und C). Diese drei gingen in die Relaxationsformel und Anfangswahrscheinlichkeit ein und wurden von mir durch errechenbare Größen, wie Pixelzahl und Wahrscheinlichkeit einer Marke ersetzt. Geblieben sind leider noch folgende fünf Parameter: Die Umgebungsradien R' und R'' , die Konsistenzschwelle T , die Akzeptanzschwelle für die Endwahrscheinlichkeiten und die Zahl der Iterationen. Die ersten drei Parameter hängen nur von der erwarteten Objektbewegung und dem Abbildungsmaßstab ab und können daher auf keinen Fall vom Programm ermittelt werden, da diese Größen erst mit Hilfe des Programms gefunden werden sollen. Allerdings ist es denkbar, vom Objektverfolgungsprogramm Schätzwerte für die Objektgröße und die Geschwindigkeit zu übernehmen. Der vierte Parameter bestimmt, wie sicher die getroffenen Entscheidungen sein sollen. Der Parameter, der die Zahl der Iterationen vorgibt, sollte durch einen Algorithmus ersetzt werden, der durch Auswerten der Zwischenergebnisse einen günstigen Abbruchzeitpunkt bestimmt, z.B. /Peleg 80/. Eine feste Zahl von Iterationsschritten ist ungünstig, da dann aus Sicherheitsgründen unnötig oft iteriert werden muß, obwohl der Rechenzeitaufwand für einen Iterationsschritt groß ist.

5.6 Korrespondenzproblem und nächster Nachbar

Das Korrespondenzproblem läßt sich durch eine einfache Heuristik lösen, wenn die markanten Bildpunkte in allen Bildern einer Folge untereinander große Abstände haben und die Verschiebungen von Bild zu Bild im Vergleich dazu gering

Korrespondenzproblem

Nächster Nachbar

sind. Wenn alle markanten Bildpunkte untereinander einen wenigstens doppelt so großen Abstand haben wie die größte vorkommende Verschiebung, dann ist der nächste Nachbar der beste Kandidat für einen korrespondierenden Bildpunkt. Dieses ist ein ähnlicher Sachverhalt wie in der Codierungstheorie der Zusammenhang zwischen Fehlerkorrigierbarkeit und Hammingdistanz. Weil dieses einfache Verfahren problematisch ist, wenn in beiden Bildern Punkte vorkommen, für die es im anderen Bild keinen korrespondierenden Punkt gibt, habe ich das Verfahren so implementiert, daß nur Punkte kombiniert werden dürfen, die sich gegenseitig der nächste Nachbar sind. Die Entfernung der beiden Punkte darf einen vorgegeben Schwellwert nicht überschreiten, der je nach maximal erwarteter Verschiebung festgelegt wird. Die Suche nach dem nächsten Nachbarn erfordert wesentlich weniger Rechenzeit als das Relaxationsverfahren von Barnard und Thompson, da die Iterationsschritte entfallen. Um aber mit diesem einfachen Verfahren arbeiten zu können, müssen die Verschiebungen klein sein gegen den Abstand der markanten Bildpunkte. Marr und Poggio /Marr und Poggio 79/ erreichen dieses durch eine Filterung der Bilder, wobei sie die Bilder in mehreren Kanälen mit verschiedenen großer Auflösung untersuchen. Ich habe den zweiten Weg gewählt und die Beträge der Verschiebungen vermindert, indem ich den zeitlichen Abstand zwischen den Einzelbildern einer Folge vermindert habe. Durch diese Methode wird allerdings die Zahl der Bilder erhöht, in denen markante Punkte gesucht werden müssen, so daß es eine Grenze gibt, ab der es nicht mehr ökonomisch ist, das Relaxationsverfahren durch die schnellere "Nächste-Nachbar-Zuordnung" zu ersetzen, weil dann der Rechenaufwand für den Punktefinder zu groß wird. Das kleinste mögliche Zeitintervall zwischen zwei Folgebildern ist bei unserer Apparatur das Aufnahmeintervall zwischen den beiden Halbbildern eines Video-Vollbildes.

Korrespondenzproblem Nächster Nachbar

Abb. IX zeigt Zuordnungsergebnisse für mit dem einstufigen Punktfinder im Bild des Taxis aus Bildfolge 1 gewählte markante Punkte. In Abb. IX b sind die Zuordnungen für Punkte aus dem ersten und zweiten Halbbild des ersten Bildes der Folge zu sehen, während Abb. IX d die Zuordnungen für Punkte aus dem ersten und zweiten Bild zeigt. Bei der Halbbild-Zuordnung ist das Ergebnis nur dort unbefriedigend, wo der Punktfinder nicht reproduzierbar genug gearbeitet hat. Bei einem Abstand von einem Vollbild dagegen treten auch schon an einigen Stellen Schwierigkeiten durch eine zu große Punktdichte auf. Dies ist z.B. am Steg zwischen den beiden Seitenfenstern der Fall. In solchen Situationen ist das Relaxationsverfahren oder Ullmans "Minimale Zuordnung" besser geeignet. Andererseits gibt es Situationen, in denen das Relaxationsverfahren ungeeignet ist, weil der Abstand zwischen den markanten Punkten so groß ist, daß die Kontinuitätsvoraussetzung nicht erfüllt ist. Dieses ist bei Szenen mit Objekten der Fall, bei denen die markanten Objektpunkte sehr weit auseinanderliegen. Nächste-Nachbar-Zuordnung und Relaxationsverfahren sind also zwei Verfahren, die gerade für sich gegenseitig ausschließende Anwendungssituationen geeignet sind. Das eine Verfahren erfordert eine hohe Punktdichte, das andere eine niedrige. Da bei Realwelt-Vorlagen immer damit gerechnet werden muß, daß die Dichte der gefundenen Punkte lokal unterschiedlich ist, erscheint es wünschenswert, einen Algorithmus zu haben, der lokal die Punktabstände analysiert und das geeignete Verfahren einsetzt.

5.7 Ketten von korrespondierenden Punkten

Wenn für alle Bilder einer Folge eine Lösung des Korrespondenzproblems gefunden worden ist, wird nach markanten Bildpunkten gesucht, die sich in mehreren aufeinander folgenden Bildern zuordnen ließen. Die korrespondierenden

Korrespondenzproblem Ketten

Punkte werden nach aufsteigender Bildnummer sortiert miteinander verkettet und als Eingabedaten für das 3D-Rekonstruktionsprogramm verwendet. Um das Rekonstruktionsprogramm nicht unnötig mit Daten zu belasten, werden nur solche Ketten berücksichtigt, die sich wenigstens über vier Bilder erstrecken. In Abb. X sind solche Ketten dem letzten Bild einer Teilfolge (Bild 1 bis Bild 15) von Bildfolge 1 überlagert worden, wobei die Koordinaten der verketteten Punkte durch Linien verbunden sind. Ketten über die ganze Bildfolge würden sich so stark gegenseitig verdecken, daß auf der Abbildung nichts mehr zu erkennen wäre.

Von den Daten, die das Objektverfolgungsprogramm liefert, wurden bisher nur die umfassenden Rechtecke der FODP-Bereiche als Objektkandidaten verwendet. Wenn in diesen Rechtecken viel auffällig strukturierter Hintergrund zu sehen ist, z.B. parkende Wagen, kann es vorkommen, daß so viele markante Bildpunkte im statischen Hintergrund gewählt werden, daß die 3D-Rekonstruktion problematisch wird. Deshalb hat der Benutzer im Moment noch zwei Möglichkeiten, in die Auswahl der Ketten für die 3D-Rekonstruktion interaktiv einzugreifen.

Die eine Möglichkeit ist, daß der Benutzer in einem Dialog am Rastersichtgerät entscheiden kann, welche Ketten zu Punkten aus dem Hintergrund gehören und vernachlässigt werden sollen. Zweitens kann der Benutzer aber auch mit Hilfe der Rollkugel die Objektkandidaten von Hand auswählen und dabei möglichst genau die Kontur des bewegten Objekts umfahren. Diese zweite Möglichkeit ist die günstigere, da sie erstens die Rechenzeit für den Punktefinder verringert und zweitens auch das Relaxationsverfahren sicherer arbeitet. Wenn die Objektgrenzen genau vorgegeben werden, können keine markanten Bildpunkte im Hintergrund gewählt werden, deren zur Objektbewegung inkonsistente Verschie-

Korrespondenzproblem Ketten

bungen das Relaxationsverfahren behindern. Wenn die Arbeiten von Blohm /Blohm 81/ abgeschlossen sind, kann auch dieser letzte interaktive Programmteil entfallen.

Im Zusammenhang mit der Analyse dieser Ketten gibt es noch eine Reihe von offenen Problemen, die noch gelöst werden müssen. Z.B. fehlt bisher noch die Möglichkeit, Lücken zu ergänzen und kurze Ketten zu verbinden, wenn eine Kette abgebrochen werden mußte, weil für den letzten Punkt kein korrespondierender Punkt gefunden werden konnte. Ebenso wünschenswert erscheint ein Verfahren, um innerhalb der Ketten Fehler zu erkennen und zu korrigieren, die bei der Lösung des Korrespondenzproblems gemacht worden sind. Die Kette, die an der linken oberen Ecke der Heckscheibe ansetzt, enthält z.B. eine Fehlzuordnung am Ende dieser Kette (Abb. X). Solche krassen Fehler sollten mit einfachen Konsistenzprüfungen zu erkennen sein.

Ob es auch möglich ist, allein durch Vergleich der Ketten zu erkennen, ob die Punkte einer Kette zum stationären Hintergrund oder zu einem bewegten Objekt gehören, ist eine schwierige Frage. Bei rotierenden Körpern z.B. erscheinen die Punkte in der Nähe der Rotationsachse stationär, obwohl sie zum bewegten Objekt gehören. Daher reicht es nicht aus, diejenigen Ketten auszusondern, deren Punkte alle angenähert die gleichen Koordinaten haben.

Da noch so viele Verbesserungsmöglichkeiten bei der Analyse der Ketten vorhanden sind, erscheinen die guten Ergebnisse der 3D-Rekonstruktion umso bemerkenswerter (siehe Kap. 6).

3D-Interpretation Übersicht

6. 3D-Interpretation

6.1 Übersicht

Eine Folge von Einzelbildern läßt noch keine eindeutige 3D-Interpretation für die dargestellten Objekte zu, wenn man keine einschränkenden Randbedingungen berücksichtigt. Dies liegt an der fehlenden Tiefeninformation. Für jeden Bildpunkt weiß man nur, auf welcher Raumgeraden durch das optische Zentrum sich das Original befinden muß, aber nicht in welcher Entfernung von der Kamera. Die weiteren Bilder der Folge können die Mehrdeutigkeit nur klären helfen, wenn man Zwangsbedingungen kennt, die die Bewegungsmöglichkeiten für Objektpunkte einschränken. Eine mögliche Zwangsbedingung ist unter anderem, daß alle im Bild befindlichen Objekte sich als starre Körper bewegen, anders ausgedrückt, daß alle Elemente eines Körpers zueinander stets die gleiche relative Anordnung behalten. Diese Randbedingung wurde in allen mir bekannten Arbeiten zur 3D-Analyse eingesetzt.

Einige Autoren gehen sogar noch über diese Einschränkung hinaus, indem sie nur bestimmte Bewegungsformen zulassen oder verlangen, daß die bewegten Objekte Polyeder sind.

Badler /Badler 74/ arbeitet mit Strichzeichnungen von geradlinig begrenzten Objekten und verwendet zur Abbildung die sphärische Projektion. Die Bilder sehen ungefähr wie mit einem Fischaugen-Objektiv aufgenommen aus. Bei dieser Projektion werden parallele Geraden in Großkreise abgebildet, die sich alle in denselben zwei Polen schneiden. Diese beiden Pole entsprechen dem Fluchtpunkt der Geradenschar bei der perspektivischen Darstellung. Aus der Lage der Pole läßt sich eindeutig die Orientierung der Geradenschar im Raum errechnen. Bei einer rechtwinklig begrenzten Fläche reicht es aus, die Schnittpunkte der projizierten begrenzenden Großkreise zu ermitteln, um die

3D-Interpretation Übersicht

Orientierung der Ebene im Raum festzustellen. Bei einer Bildfolge kann man aus den Daten von korrespondierenden Geraden die Trajektorien von bewegten Objekten und die des Beobachters ermitteln.

Neumann /Neumann 79b/ arbeitet ebenfalls mit durch Geradenstücke begrenzten Körpern. Im Gegensatz zu Badler verwendet er die Parallelprojektion. Bei der Parallelprojektion läßt sich keine eindeutige Lösung finden, da sich das Bild eines bewegten Objektes nicht von seinem an der Bildebene gespiegelten Abbild unterscheiden läßt, bei dem der Rotationsanteil der Bewegung in entgegengesetzter Richtung erfolgt. Außerdem läßt sich die Entfernung von der Kamera nicht ermitteln. Ein Vorteil der Parallelprojektion ist die Einfachheit der Gleichungen. Außerdem ist sie bei größeren Entfernungen nahezu mit der Zentralprojektion identisch. Neumanns Verfahren ermöglicht es, in einem einzigen Interpretationsvorgang drei Aufgaben gleichzeitig zu lösen:

1. Im Bild gefundene Geraden werden zu Objekten gruppiert.
2. Die Korrespondenzen zu den Folgebildern werden ermittelt.
3. Für alle Objekte werden die räumlichen Parameter der begrenzenden Geraden und die Trajektorien errechnet.

Diese Aufgaben werden durch eine Tiefensuche gelöst, bei der Hypothesen über die Zusammengehörigkeit von Geraden aufgestellt und die jeweils am meisten Erfolg versprechenden verfolgt werden. Die Tiefensuche ist günstig, da sie z.B. erlaubt, schnell den statischen Hintergrund zu erkennen und

3D-Interpretation

Übersicht

von der weiteren Untersuchung auszuschließen. Neumann hat für seine Modellierung Geradenstücke gewählt, da sie sich im Bild auf mehr Pixel stützen als einzelne markante Punkte, und so ihre Parameter relativ genau gefunden werden können.

Für mich sind die Arbeiten von Ullman /Ullman 79/ wichtiger, da er das Interpretationsproblem für Einzelpunkte behandelt hat. In seinem "Structure from Motion Theorem" macht er eine Aussage über die Eindeutigkeit der Struktur bei Bildfolgen von bewegten festen Körpern in Parallelprojektion:

"Given three distinct orthographic views of four non-coplanar points in a rigid configuration, the structure and motion compatible with the three views are uniquely determined".

Diese Eindeutigkeitsaussage gilt, wie schon oben erwähnt, nur bis auf eine Spiegelung an der Bildebene und die absolute Entfernung von der Kamera.

Die Zentralprojektion, die perspektivische Effekte mit berücksichtigt, hat Ullman nur für einen eingeschränkten Bewegungstyp untersucht: Rotation um eine senkrechte Achse, gefolgt von einer beliebigen Translation. In diesem Fall genügen drei Punkte in zwei Ansichten, um eine Gleichung, genannt Polargleichung, zu lösen, die den Rotationswinkel θ liefert. Diese Gleichung ist vierten Grades, so daß man maximal vier Lösungen erhält. Oft genügt es, einen weiteren Punkt hinzuzunehmen, um ein eindeutiges Ergebnis zu erhalten (siehe auch /Meiri 80/, /Nagel 81/, /Nagel und Neumann 81/). Kennt man den Rotationswinkel, dann sind alle anderen Bewegungsparameter und die objektbezogenen Punktkoordinaten berechenbar. Die absolute Entfernung des Objekts von der Kamera kann nicht ermittelt werden, da in der Zentralprojektion kleine nahe Objekte nicht von großen entfernten zu

3D-Interpretation

Übersicht

unterscheiden sind. Allerdings sind die perspektivischen Effekte oft gering, so daß die Lösungen unsicher sind.

Ullman hat bei verschiedenen psychologischen Experimenten festgestellt, daß der Mensch nicht in der Lage ist, perspektivische Effekte bei bewegten Objekten in räumliche Struktur umzusetzen, sondern daß er das Gesehene nach der Parallelprojektion interpretiert. Andererseits gewinnt der Mensch räumliche Information, die sich nicht allein durch orthographische Auswertung erklären läßt. Als Näherung hierfür hat Ullman ein Interpretationsverfahren vorgeschlagen, das er polar-parallel nennt, und das beide Projektionstechniken vereint. Lokale Gruppen von benachbarten Punkten werden nach den Gleichungen für Parallelprojektion ausgewertet, was eine gute Näherung für die Zentralprojektion ist. Global werden die Lösungen nach den Gesetzmäßigkeiten der Perspektive kombiniert, so daß spiegelbildliche Mehrfachlösungen vermieden werden können.

Ullman zeigt unter den beiden Oberbegriffen "Structure from Motion" und "Motion from Structure", daß das Verstehen von Struktur und Bewegungsvorgängen eng gekoppelt sind. Kennt man die Struktur, so folgt daraus oft schon die Bewegung und umgekehrt. Daraus folgt aber auch, daß falsche Annahmen über die Struktur zu einer fehlerhaften Interpretation der Bewegung führen können. Dieser Zusammenhang ist die Ursache für viele optische Täuschungen, denen das menschliche Auge erliegt.

Durch das Beschränken auf starre Körper sind nicht mehr alle denkbaren Bewegungsvorgänge interpretierbar. Der Mensch kann im wesentlichen drei Arten von Bildveränderungen verstehen: Bewegung, Ersetzung und Verwandlung /Ullman 79/. Veränderung durch Bewegung bedeutet, daß ein Körper aus verschiedenen Blickwinkeln oder Entfernungen gesehen wird; Ersetzung bedeutet das plötzliche Austauschen eines Körpers

3D-Interpretation Übersicht

gegen einen zweiten, deutlich verschiedenen, wie bei einem Zaubertrick, während Verwandlung eine allmähliche physikalische Veränderung eines Körpers meint, z.B. eine sich öffnende Blüte oder ein schmelzender Schneemann.

Die Ersetzung läßt sich überhaupt nicht als starrer Körper in Bewegung modellieren, während es bei der Verwandlung in bestimmten Fällen Möglichkeiten gibt. Zum einen sind viele nicht starre Körper im wesentlichen aus mehreren gelenkig verbundenen starren Teilen zusammengesetzt. Dies gilt z.B. für fast alle Lebewesen mit Innen- oder Außenskelett (Glieder- und Wirbeltiere) und auch für viele Maschinen. Diese Objekte könnten in Teilen als starre Körper behandelt werden, wobei solche Lösungsansätze wie die von Neumann /Neumann 79b/, Asada et al. /Asada et al. 80/ und Webb and Aggarwal /Webb and Aggarwal 80/ interessant sind, die automatisch die Objektbestandteile in Gruppen von sich starr bewegenden Einheiten unterteilen. Eine weitere Möglichkeit wäre Ullmans Prüfmethode, mit deren Hilfe man jeweils vier Punkte in drei verschiedenen orthographischen Ansichten daraufhin untersuchen kann, ob sie zu einem bewegten starren Körper gehören.

Kommt diese Möglichkeit nicht in Frage, weil der Körper sich amorph verändert, kann man untersuchen, ob es möglich ist, den Körper wenigstens über kurze Zeiträume als angenähert starr zu betrachten und für jeweils kurze Intervalle eine neue, zeitlich variable 3D-Interpretation zu erzeugen. Dies setzt allerdings voraus, daß die Bildveränderungen durch Bewegung wesentlich die durch Verwandlung überwiegen, da sonst die Ergebnisse zu unzuverlässig werden. Der Extremfall sind stationäre, sich verändernde Objekte, wo man sicher die Grenze des Verfahrens erreicht hat. Hier sehe ich nur noch die Möglichkeit, die Struktur unabhängig von der Bewegung festzustellen, durch Stereoaufnahmen, Laser-Tiefenmessung, Konturanalyse, Schattenwurf usw.

3D-Interpretation Übersicht

Meiri /Meiri 80/ hat untersucht, wieviele Punkte in wievielen Ansichten bei den verschiedenen Projektionstechniken jeweils nötig sind, um die relative Anordnung der Punkte herauszufinden.

Eine alternative Möglichkeit zu den hier besprochenen Verfahren ist, 3D-Information mittels des "optischen Flusses" abzuleiten, z.B. /Clocksin 78, 80a, 80b/, /Prazdny 79, 80/ und /Williams 80/. Auf diese Arbeiten werde ich hier nicht näher eingehen.

6.2 Das Interpretationsverfahren von Bonde und Nagel

Bonde und Nagel /Bonde 79, Bonde and Nagel 79, Nagel 80a/ haben das Interpretationsproblem für folgende Realwelt-Situation gelöst: Ein starrer, beliebig geformter Körper führt in einer Ebene eine Translationsbewegung aus, wobei er um eine senkrecht auf der Grundebene stehende Achse rotiert. An der Objektoberfläche müssen einige körperfeste, markante Punkte identifizierbar sein, für die das Korrespondenzproblem als lösbar vorausgesetzt wird. Es wird mit einer stationären TV-Kamera und Folgen von Schwarz-Weiß-Aufnahmen gearbeitet. Für das Verfahren ist es nicht zwingend, daß die Kamera stationär ist, da es nur auf Relativbewegungen ankommt. Wichtig ist die stationäre Kamera nur, um die bewegten Objekte mit dem bisher verwendeten Ansatz /Jain and Nagel 79/ vom Hintergrund zu trennen.

Zur Interpretation werden drei verschiedene Koordinatensysteme eingesetzt:

1. Das Weltkoordinatensystem: Die XY-Ebene dieses Systems ist die Ebene, in der das Objekt sich bewegt. Alle Koordinatenangaben für den stationären Hintergrund beziehen sich auf dieses System.

3D-Interpretation

Bonde und Nagel

2. Das objektbezogene Koordinatensystem: Alle Koordinaten für Objektmodellpunkte werden in diesem System angegeben. Dieses System ist körperfest mit dem Objekt verbunden. Der Ursprung ist ein frei gewählter Referenzpunkt im Objekt, während die XY-Ebene mit der des Weltkoordinatensystems zusammenfällt.

3. Das kamerabezogene Koordinatensystem: In diesem System dient als Ursprung der Schnittpunkt der optischen Achse mit der Abtastfläche der Kamera, während die XZ-Ebene durch die Abtastfläche gebildet wird. Die X-Achse verläuft parallel zur Bild-Horizontalen. Dabei wird davon ausgegangen, daß die Abtastfläche genau senkrecht zur optischen Achse justiert wurde. Dieses Koordinatensystem wird für Bildkoordinaten verwendet.

6.2.1 Kamerakalibrierung

Im Gegensatz zu den im Abschnitt 6.1. erwähnten Verfahren setzen Bonde und Nagel perspektivische Projektion voraus. Bei den von ihnen untersuchten Straßenszenen kann sich die Entfernung der beobachteten Objekte von der Kamera so stark ändern, daß Parallelprojektion nicht als Näherungslösung für die Abbildung durch die Kameraoptik in Frage kommt. Auch der polar-parallele Lösungsansatz ist nicht geeignet, weil es nicht garantiert werden kann, daß die gewählten markanten Punkte dicht genug liegen, um lokal Parallelprojektion annehmen zu können. Bonde und Nagel sind in der Lage, die Objektgröße und die Entfernung von der Kamera absolut zu ermitteln, da sie die Höhe der Kamera über der Grundebene und die Orientierung der optischen Achse der Kamera im Raum bestimmen. Als Kameramodell verwenden sie das der

Lochkamera. Dieser Ansatz erfordert, daß eine ganze Reihe von festen Kameraparametern bekannt sein müssen, die den Abbildungsvorgang beeinflussen. Diese Parameter sind

1. die Transformationsparameter für den Übergang vom Weltkoordinatensystem zum kamerabezogenen Koordinatensystem; das sind drei Ortskoordinaten (wovon nur die Kamerahöhe über der Grundebene benötigt wird) und drei Orientierungswinkel für die Kamera (Drehung, Kippung und Neigung);
2. die Objektivbrennweite (Vario-Objektiv),
3. die Zeilen- und Spaltenauflösung auf der Abtastfläche (Pixel/mm)
4. und die Verschiebung des Mittelpunktes der Abtastfläche gegen den Durchstoßpunkt der optischen Achse.

Die Blendenöffnung dagegen wird nicht benötigt, da die Bildhelligkeit und die Tiefenschärfe nicht ausgewertet werden.

Alle oben erwähnten Parameter sind unabhängig von dem beobachteten Objekt und dessen Bewegung und können daher in einem separaten Kalibrierungsschritt ermittelt werden. Hierzu werden im statischen Hintergrund geeignete Kalibrierungspunkte ausgewählt, und die räumlichen Abstände zwischen den Punkten vermessen. Diese Punkte sollten die folgenden drei Bedingungen erfüllen:

3D-Interpretation Bonde und Nagel

1. Die Punkte sollten möglichst einfach vermeßbar sein.
2. Sie müssen im Bild deutlich als markante Punkte unterscheidbar sein und
3. sie sollten gleichmäßig über das Bild verteilt sein.

Für die ausgewählten Kalibrierungspunkte werden im Weltkoordinatensystem möglichst viele relative Abstände vermessen und für die entsprechenden Bildpunkte die Zeilen- und Spaltenkoordinaten ermittelt. Das Auszählen der Pixel ist am einfachsten am Farb-Rastersichtgerät in Verbindung mit der Rollkugeleinheit. Nach der Vermessung wird mit Hilfe der Kalibrierungsdaten eine quadratische Fehlerfunktion ausgewertet und iterativ minimisiert.

6.2.2 Objektbezogene Parameter

Die objektabhängigen Daten, nämlich die 3D-Koordinaten der markanten Punkte im objektbezogenen Koordinatensystem und die Relativbewegung von Kamera und Objekt, werden in einem zweiten Verfahren errechnet. Hierbei gehen nur die vorher ermittelten festen Kameraparameter und die Bildkoordinatenfolgen der markanten Objektpunkte in das Verfahren ein. Ebenso wie bei der Kamerakalibrierung wird wieder eine Fehlerfunktion iterativ minimiert, in der die Fehlerquadrate von gemessenen Bildkoordinaten gegen die aus den geschätzten Parametern errechneten Koordinaten aufsummiert werden.

5D-Interpretation Bonde und Nagel

Bei der Auswahl des Minimierungsverfahrens waren im wesentlichen drei Gesichtspunkte zu berücksichtigen: Die Startwerte sind sehr ungenau; es sind sehr viele freie Parameter (mehr als zehn) gegeneinander zu optimieren, und das Verfahren sollte schnell sein und sowohl feine als auch grobe Schrittweiten zulassen. Aus diesen Gründen haben Bonde und Nagel die Davidon-Fletcher-Powell-Methode gewählt, da diese als schnelles und zuverlässiges Verfahren gilt /Himmelblau 72/.

Das getrennte Kalibrieren von bewegungsabhängigen und bewegungsunabhängigen Parametern hat sowohl Vorteile als auch Nachteile. Vorteilhaft ist, daß so bei jedem Kalibrierungsschritt die kombinatorische Vielfalt sinkt, da weniger freie Parameter vorhanden sind. Das Verfahren wird dadurch schneller und sicherer. Außerdem können die festen Kameraparameter unabhängig von eventuellen Ungenauigkeiten oder Fehlern bei der Korrespondenzanalyse so genau ermittelt werden, wie es die Auswahl und Vermessung der Kalibrierungspunkte zuläßt. Ein Nachteil ist, daß das Kalibrierungsverfahren für die Kamera nicht ohne absolute Abstandsmessungen an den Objekten im statischen Hintergrund auskommt, so daß es streng genommen kein Fernvermessungsverfahren (remote sensing) ist, auch wenn die Abstandsmessungen notfalls an unzugänglichen Stellen durch trigonometrische Methoden vorgenommen werden könnten. Andererseits wäre es ohne absolute Längenmessungen bei der Kalibrierung nicht möglich, die Kamerabrennweite und Höhe über der Grundebene, und damit den Abbildungsmaßstab festzustellen, so daß alle Längenangaben wie bei Ullmans Zentralprojektion nur relativ erfolgen könnten. Auf die Kalibrierung der Kamera kann verzichtet werden, wenn die absoluten Abstandsmessungen nicht erforderlich sind. In der Diplomarbeit von Bonde /Bonde 79, Kap. 4 und 5/ ist eine ausführliche Abschätzung der Einflüsse von ungenau ermittelten Kameraparametern oder ungenau isolierten

markanten Punkten auf das Interpretationsergebnis zu finden. Dort wird auch das Konvergenzverhalten der Minimierungsfunktion in Abhängigkeit von den gewählten Startwerten diskutiert, so daß diese Untersuchungen hier nicht noch einmal aufgeführt zu werden brauchen.

6.2.5 Ergebnisse

Abb. X1a und X11a zeigen die Genauigkeit der 3D-Rekonstruktion für die beiden Taxis aus den Bildfolgen 1 und 2. Die offenen Quadrate geben die Lage der im ersten Bild der Folge gewählten markanten Punkte an, die bei der Rekonstruktion berücksichtigt worden sind. Die kleinen Punkte markieren die Stellen im Bild, auf die die Punkte des errechneten 3D-Modells abgebildet werden würden, wenn man das Modell anstelle des Objektes durch die Kamera abbilden würde. Dabei wird das Modell in derjenigen Entfernung und Orientierung relativ zur Kamera angenommen, die die Modellrechnung für die Position des Objektes während der ersten Aufnahme ergeben hat. Für viele Punkte fallen die Bilder der rückprojizierten Modellpunkte und die ursprünglich gewählten markanten Punkte genau aufeinander. Wo dieses nicht der Fall ist, wurden zusammengehörige Punkte durch Linien verbunden. An der Länge dieser Linien kann man ablesen, wie gut das errechnete Modell die gemessenen Punktkoordinaten-Folgen erklären kann. Die Näherungsfehler bewegen sich in der Größenordnung von wenigen Geopixeln.

Die Abbildungen zeigen auch die Bilder von einigen Modellpunkten, zu denen es keine korrespondierenden markanten Bildpunkte gibt. Diese Punkte wurden aufgrund von Ketten von markanten Punkten errechnet, die erst in späteren Bildern der Folge gefunden worden sind. In Abb. X11a kann man einige Stellen finden, an denen mehrere Bilder von

3D-Interpretation
Bonde und Nagel

Modellpunkten dicht beieinander liegen. Hierbei handelt es sich vermutlich um Modellpunkte für denselben markanten Objektpunkt, dessen Bild nicht über die ganze Folge ununterbrochen verfolgt werden konnte, und daher Anlaß zu mehr als einer Korrespondenzkette gegeben hat. Solche Punkte sollten eventuell durch eine Ballungsanalyse als Modell desselben Objektpunktes zu erkennen sein.

Anhand dieser Rückprojektionen kann man sich nur schwer klar machen, wie die Punkte des Modells im Raum relativ zueinander liegen. Das nächste Kapitel zeigt, wie die relative Anordnung der Punkte im Raum sichtbar gemacht werden kann.

Modellierung
Übersicht

7. Modellierung

7.1 Übersicht über 3D-Modellierungsverfahren

Für die 3D-Formbeschreibung von Volumina sind eine ganze Reihe von Verfahren bekannt, z.B. aus der Medizin, wo aus Computertomogrammen die Form vom Herz oder anderen Organen ermittelt wird. Soroka und Bajcsy /Soroka and Bajcsy 78/ beschreiben Organinnenräume durch Stapelung von generalisierten Zylindern. Schudy /Schudy 79/ wählt zur schnellen Wiedergabe von Organbewegungen auf einem Graphik-sichtgerät eine analytische Darstellung. Er approximiert die Herzzinnenfläche durch eine Reihenentwicklung, die auch den Parameter der Zeit berücksichtigt.

Gennery /Gennery 79a/ untersucht die 3D-Modellierung von Felsblöcken zur Steuerung eines autonomen Fahrzeugs. Da er von Felsblöcken, die als Hindernis auftauchen, nur auf der sichtbaren Seite 3D-Tiefenmessungen gewinnen kann, schätzt er den Verlauf der Oberfläche, indem er annimmt, daß die Form im wesentlichen elliptisch ist. An die auf der sichtbaren Seite gemessenen Punkte wird einfach ein Näherungsellipsoid angepaßt. Da das Modell nur dafür benötigt wird, den Kurs des Fahrzeugs ungefähr vor auszuplanen, kommt es nicht auf eine exakte Modellierung der Hindernisse an.

Andere Arten der Modellierung erlauben es, Körper zu beschreiben, die aus gelenkig miteinander verbundenen starren Teilen bestehen, z.B. die Pfeifenreinigerfiguren von Marr und Nishihara /Marr and Nishihara 76/ oder der Kugelmann von O'Rourke /O'Rourke 80a/.

All diese Verfahren sind aber nur sinnvoll, wenn man so viele Punkte an der Objektoberfläche gemessen hat, daß man die vollständige Oberfläche durch Interpolation zwischen benachbarten Punkten ziemlich gut annähern kann. Oder man

Modellierung Übersicht

muß schon ein grobes Modell des Objektes kennen, das durch wenige genaue Messungen verfeinert werden kann, z.B. durch "constraint propagation" /O'Rourke 80a/.

7.2 Die konvexe Hülle

Die markanten Objektpunkte, für die bei meinem Verfahren Messungen gewonnen werden, sind im wesentlichen die Eckpunkte und einige ausgezeichnete Punkte auf der Oberfläche. Diese Daten allein lassen noch keine Rekonstruktion des Oberflächenverlaufs zu. Eine Möglichkeit, die räumliche Anordnung für einige der rekonstruierten Meßpunkte darzustellen, ist die dreidimensionale konvexe Hülle (das ist der kleinste konvexe Körper, der ein gegebenes Volumen umschließt). Die konvexe Hülle für eine Punktwolke ist ein Polyeder.

Solange ein Körper nur wenige Konkavitäten aufweist, ist die konvexe Hülle eine gute Annäherung an die Form. Im allgemeinen gilt zwar, daß zu jedem Körper eine eindeutige konvexe Hülle existiert, aber umgekehrt gehören zu jeder konvexen Hülle unendlich viele Körper, die die gleiche Hülle haben.

7.3 Minimale Verfahren zur Berechnung der konvexen Hülle

Zur Berechnung der konvexen Hülle von Punktwolken gibt es eine ganze Reihe von mehr oder weniger schnellen Verfahren, von denen einige nur für zwei Dimensionen, andere auch für drei oder beliebig viele Dimensionen anwendbar sind. Ein einfacher Algorithmus für dreidimensionale konvexe Hüllen wurde von Appel und Will /Appel and Will 76/ entwickelt. Dieses Verfahren prüft, ob eine durch zwei Punkte gegebene

Modellierung Konvexe Hülle

Kante zur konvexen Hülle gehört oder nicht. Wenn ein Polyeder auf eine Ebene senkrecht zur zu prüfenden Kante projiziert wird, dann ist das Bild der Kante ein Punkt. Gehört diese Kante zur konvexen Hülle des Körpers, so liegt ihr Bild außerhalb oder auf dem Rand der ebenen konvexen Hülle aller anderen projizierten Punkte des Körpers. Es muß also geprüft werden, ob ein Punkt in die konvexe Hülle der Projektion fällt oder nicht.

Für diesen Test schlagen Appel und Will einen Sternpolygontest vor. Die euklidischen Koordinaten der projizierten Punkte werden in Polarkoordinaten umgerechnet. Das Bild der untersuchten Kante dient als Ursprung für das Koordinatensystem. Die Nullrichtung kann willkürlich gewählt werden. Anschließend werden die übrigen Punkte nach aufsteigendem Polarwinkel sortiert. Diese sortierte Punkfolge kann als Kantenzug verstanden werden und wird Sternpolygon genannt. Die Frage, ob das Sternpolygon den Ursprung umschließt, ist gleichbedeutend mit der Frage, ob der Ursprung innerhalb der konvexen Hülle der Punkte des Polygons liegt.

Für jedes Kantenstück des Sternpolygons wird in aufsteigender Reihenfolge ermittelt, wie der Drehsinn des Dreiecks bestehend aus den Endpunkten der Kante und dem Ursprung gerichtet ist. Eine Änderung des Drehsinns zeigt, daß der Ursprung außerhalb des Sternpolygons liegen muß. Der Drehsinn wird aus der Determinante der Koordinaten der drei beteiligten Punkte errechnet. Abb. 7.1 zeigt ein solches Sternpolygon einmal mit dem Ursprung innerhalb und einmal außerhalb der konvexen Hülle.

Dieses Verfahren ist günstig, wenn man nur wenige Punkte zu untersuchen hat. Bei mehr als dreißig bis vierzig Punkten kann die Rechenzeit je nach Geschwindigkeit des Rechners schon in die Größenordnung von Stunden gehen. Die Autoren

Modellierung
Konvexe Hülle

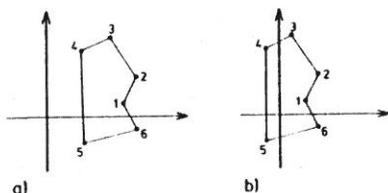


Abb. 7.1 Sternpolygontest

selbst geben nur an, daß der Rechenaufwand ähnlich komplex wie bei einem Eliminationsverfahren für verdeckte Linien (hidden line) ist.

Ich möchte hier für Vergleiche eine grobe Abschätzung des Aufwands geben: Für eine Wolke von N Punkten gibt es $N*(N-1)/2$ mögliche Kanten, die überprüft werden müssen. Jeder Kantentest erfordert eine Projektion aller N Punkte auf eine Ebene senkrecht zur Kante (genauer nur $N-1$ Punkte, da die beiden Endpunkte der Kante das gleiche Bild haben). Dies geschieht am bequemsten durch zweimalige Rotation des Koordinatensystems, so daß eine Achse parallel zur fraglichen Kante wird. Hierzu müssen eine $3*3$ Rotationsmatrix ermittelt und $N-2$ Vektoren mit der Matrix multipliziert werden. Der Rechenaufwand bisher sieht so aus:

$$N*(N-1)/2 * (\text{Projektionsmatrix errechnen} + \\ (N-2)* \text{Matrixmultiplikation} + \text{Sternpolygontest})$$

Modellierung Konvexe Hülle

Älter als das Verfahren von Appel und Will ist der Algorithmus von Chand und Kapur /Chand and Kapur 70/. Dieser Algorithmus ist deshalb bemerkenswert, weil er für alle Dimensionen geeignet ist und die Komplexität nur $O(N^2)$ beträgt. Das Verfahren basiert auf dem "Geschenk Einwickeln" (Gift Wrapping), das heißt, dem Körper wird die konvexe Hülle angepaßt, indem man von außen Tangentialebenen heranklappt, bis sie den Körper berühren.

Kurz nach der Veröffentlichung von Appel und Will erschien mit der Arbeit von Preparata und Hong /Preparata and Hong 77/ das bisher schnellste Verfahren für zwei und drei Dimensionen, das mit $O(N \log N)$ Schritten auskommt. Die Punkte werden zunächst nach einer Koordinate sortiert und dann in einem Teil- und Mischverfahren in kleine Gruppen unterteilt, für die separat die konvexe Hülle ermittelt wird. Kompliziert ist das hierarchische Verschmelzen von den elementaren konvexen Hüllen zu größeren Einheiten. Das anfängliche Sortieren ist dabei wichtig, weil dadurch garantiert wird, daß der Durchschnitt aller Teilhüllen stets leer ist.

7.4 Verdeckte Oberflächen und Schattierung

Die Darstellung eines Polyeders als Linienzeichnung wird um einiges übersichtlicher, wenn nur die in der gewählten Ansicht unverdeckten Linien gezeichnet werden. Bei konkaven Körpern ist dies ein nicht triviales Problem, da durch Selbstverdeckung Flächen und Kanten nur teilweise sichtbar sein können. Bei konvexen Körpern dagegen ist Selbstverdeckung unmöglich. Kanten und Flächen sind daher entweder vollständig oder gar nicht sichtbar. Um festzustellen, ob ein Flächenelement sichtbar ist, braucht man nur zu überprüfen, ob es um mehr als 90° gegen die Beobachtungsrichtung geneigt ist. Dies läßt sich leicht anhand des Skalarproduktes von Flächennormale und dem Einheitsvektor in Richtung zum Beobachterstandort ermitteln.

Modellierung Konvexe Hülle

Leider liefert das Verfahren von Appel und Will nur die Kanten der konvexen Hülle, aber nicht die Kantenzüge, die die Oberflächenfacetten begrenzen. Appel und Will schlagen das folgende Verfahren vor, um die Flächen zu finden: Jedes Oberflächenelement wird durch einen geschlossenen Polygonzug begrenzt, dessen Kanten koplanar sind. Diese Eigenschaft wird ausgenutzt, indem nach ebenen Kantenzyklen gesucht wird. Es wird irgend ein Eckpunkt der konvexen Hülle ausgewählt und dann zwei von diesem Punkt ausgehende Kanten gesucht. Diese beiden Kanten spannen eine Ebene auf. Wenn alle Punkte der konvexen Hülle auf der gleichen Seite der Ebene oder in der Ebene liegen, handelt es sich um eine Außenfläche. Es werden solange Kantenpaare gewählt und nach dieser Methode überprüft, bis ein Paar gefunden ist, das zu einer Außenfläche gehört. Der so begonnenen Polygonzug wird vervollständigt, indem ausgehend vom Endpunkt nach einer weiteren koplanaren Kante gesucht wird. Dies wird solange fortgesetzt, bis der Startpunkt wieder erreicht und der Polygonzug geschlossen ist. Auf die gleiche Weise wird nach weiteren Flächenelementen gesucht. Wenn alle Kanten der konvexen Hülle genau zweimal verwendet worden sind, bricht das Verfahren ab.

Bei der Implementation habe ich das Verfahren ein wenig vereinfacht. Um zu prüfen, ob die beiden Startkanten für einen Kantenzklus zu einer Begrenzungsfläche gehören, kann man einen schnelleren Test verwenden. Man braucht nicht alle Punkte der konvexen Hülle daraufhin zu untersuchen, ob sie auf derselben Seite der durch die beiden Kanten aufgespannten Ebene liegen. Es reicht aus, die Endpunkte der übrigen vom gemeinsamen Ausgangspunkt der beiden Startkanten wegführenden Kanten zu überprüfen.

Modellierung
Konvexe Hülle

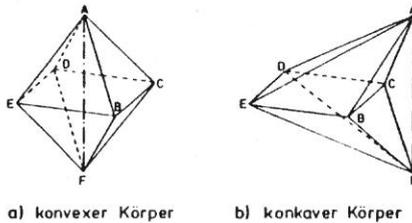


Abb 72

Dies kann man sich leicht an der Abb. 7.2 klarmachen. Die Ebene ABC ist eine Außenfläche des Oktaeders und die Punkte D und E liegen offensichtlich auf der gleichen Seite dieser Ebene. Für den Punkt F gibt es zwei Möglichkeiten: Entweder er liegt auf derselben Seite wie D und E; dann ist die Fläche ABC eine Außenfläche, oder er liegt auf der anderen Seite. Daraus würde aber folgen, daß die Ebene ABC den Körper schneidet, und daß es eine weitere von A ausgehende Kante geben muß, die auf der anderen Seite der Ebene liegt als die Verbindungslinien AD und AE und über die der Punkt F von A aus durch einen Kantenzug erreicht werden kann. Andernfalls würde die Verbindungsgerade A-F teilweise außerhalb des Körpers verlaufen, was durch die Definition der Konvexität ausgeschlossen ist. Es sollen aber nach Voraussetzung alle von A aus direkt erreichbaren Punkte auf der gleichen Seite der geprüften Ebene ABC liegen, was ebenfalls zum Widerspruch führt. Daraus folgt, daß ein konvexer Polyeder unmöglich ist, bei dem alle von einem Punkt A aus direkt über Kanten erreichbaren Punkte auf derselben Seite einer den Körper begrenzenden Ebene ABC liegen, andere Körperpunkte aber auf der anderen Seite der Ebene.

Modellierung Konvexe Hülle

In Abb. X1b ist die konvexe Hülle der in Abb. X1a gezeigten Modellpunkte dem Bild des Taxis überlagert dargestellt. Es wurden nur die sichtbaren Linien der konvexen Hülle eingetragen. Die verdeckten Linien sind in Abb. X1c zu sehen. Man sieht, daß die konvexe Hülle sich im Wesentlichen an den Ecken von Bug und Heck und am Taxizeichen auf dem Dach orientiert. Die konkaven Stellen des Wagens können selbstverständlich mit diesem einfachen Modell nicht wiedergegeben werden.

Abb. X1b zeigt die entsprechende Darstellung für den Wagen aus Bildfolge 2. An dieser Abbildung fällt auf, daß die rechte Ecke vom Heck nicht vom Modell erfaßt wird. Das liegt daran, daß diese Ecke immer im Schatten der Fahrgastzelle lag und das Bild der Ecke daher nie genug Kontrast aufwies, um als markanter Bildpunkt gewählt zu werden. Eine andere Schwachstelle dieses Rekonstruktionsergebnisses ist die errechnete Lage des Taxizeichens. Die Abweichung läßt sich dadurch erklären, daß der Punktefinder regelmäßig auch die Spitze des Schattens vom Taxizeichen als markanten Punkt gewählt hat. Der Schatten ist aber nicht körperfest, da der Wagen auch eine Drehung ausführt. Das Rekonstruktionsergebnis für die anderen Punkte zeigt, daß das Verfahren von Bonde und Nagel robust genug ist, um auch einzelne nicht körperfeste Punkte zu verkraften.

Wenn alle Oberflächenelemente eines Polyeders mit ihren Normalenvektoren bekannt sind, kann der Körper auch mit verschiedenen schattierten Oberflächen dargestellt werden. Dies gibt besser den räumlichen Eindruck wieder als eine Linienzeichnung der Kanten. Zur Schattierung wurde der Einfachheit halber angenommen, daß der Körper ein lambertscher Strahler ist. Die Helligkeit einer Fläche hängt dann nur vom Winkel zwischen Flächennormaler und

Modellierung Konvexe Hülle

Einfallrichtung des Lichtes ab, nicht vom Ort des Beobachters. Im ersten Beleuchtungsversuch habe ich paralleles Licht gewählt, das senkrecht zur Grundebene einfällt, in der das Objekt sich bewegt. Damit sollte die Beleuchtung durch das von oben kommende Himmelslicht angenähert simuliert werden. Diese Beleuchtung hat allerdings den Nachteil, daß einige Flächenelemente, die dem Beobachter sichtbar sind, von der Lichtquelle abgewandt sind und daher schwarz abgebildet werden. Diese Flächen sind dann vom Hintergrund nicht mehr zu trennen (Abb. XI e). Der optische Eindruck ist günstiger, wenn der Körper genau aus der Beobachtungsrichtung beleuchtet wird, da dann alle sichtbaren Flächen Licht erhalten. Diese Beleuchtung verstärkt auch den plastischen Eindruck, da die Flächenhelligkeit der Neigung zum Beobachter entspricht (Abb. XI d, XI c). Abb. XI f zeigt das Modell für das Taxi aus Bildfolge 1 zum Vergleich noch einmal in der Orientierung, wie sie für das letzte Bild der Folge errechnet wurde.

7.5 Alternative Polyedermodelle zur konvexen Hülle

Die konvexe Hülle hat als Darstellungsform für eine Punktwolke den offensichtlichen Nachteil, daß einige der Punkte, die auf der Oberfläche liegen sollen, nicht auf der Oberfläche der konvexen Hülle liegen und in deren Innern versteckt werden. J. O'Rourke /O'Rourke 80b/ hat im Sommer 1980 während seines Aufenthalts an der Universität Hamburg eine Reihe von Möglichkeiten untersucht, um bessere Darstellungen zu finden. Im Unterschied zur konvexen Hülle sollten alle Modellierungspunkte als Oberflächenpunkte des Modells verwendet werden. Das verbesserte Modell sollte aber alle Vorteile der konvexen Hülle bewahren und mit vertretbarem Aufwand berechenbar sein. Zwei der vielen von

Modellierung

Weitere Polyedermodelle

O'Rourke untersuchten Möglichkeiten sind besonders interessant:

Das eine ist eine leichte Modifikation des Algorithmus von Preparata und Hong, der in einem Teil- und Mischverfahren die Punktewolke in Teilwolken zerlegt und die konvexe Hülle rekursiv aufbaut, indem die konvexen Hüllen von benachbarten Teilgruppen verschmolzen werden. Die Frage ist, ob der Teil- und Mischvorgang nicht so gesteuert werden kann, daß die sich dabei bildenden Zwischenstufen eine sinnvolle Zerlegung des modellierten Objektes darstellen. Z.B. wäre es wünschenswert, für einen T-förmigen Block ein Modell zu erhalten, das aus zwei (konvexen) Quadern besteht.

Die Zwischenergebnisse hängen von zwei Bedingungen ab. Die eine ist die Lage der Achse, entlang der die Punktewolke in Untergruppen zerlegt wird, und die andere ist die Mischreihenfolge. Preparata und Hong haben als Achse der Einfachheit halber eine Koordinatenachse gewählt. Das ist ungünstig, wenn die Zwischenergebnisse lageinvariant sein sollen. Die Achsenrichtung muß dann einen Bezug zur Orientierung der Punktewolke haben. Es bieten sich z.B. die Hauptachsen der Punktewolke an; wird die längste Hauptachse gewählt, erhält man als Resultat eine Reihung von konvexen Blöcken in Längsrichtung des modellierten Körpers. Die Mischreihenfolge könnte so gewählt werden, daß stets die nächsten Nachbarn vereinigt werden. So bilden sich möglichst kompakte Untereinheiten.

Ein weiteres Kriterium wäre es, beim Mischen darauf zu achten, daß keine Punkte zu inneren Punkten werden. Dies ist mit dem Problem verwandt, für ein Polyeder diejenige Zerlegung in konvexe Unterkörper zu finden, die aus den wenigsten Elementen besteht und deren Vereinigung das Polyeder ist. Dies Problem ist für den dreidimensionalen Fall noch nicht gelöst.

Modellierung

Weitere Polyedermodelle

Die zweite interessante Möglichkeit ist die Modellierung durch Polyedertypen, die wie die konvexe Hülle (in den meisten Fällen) eindeutig sind, aber alle Punkte der Wolke als Eckpunkte enthalten. Die vorgeschlagenen Polyedertypen bezeichnet O'Rourke als Minieder (minihedron) und Maxieder (maxihedron). Das Minieder ist das zu einer Punktwolke gehörige Polyeder kleinster Oberfläche, das genau die Punkte der Wolke als Eckpunkte enthält. Analog dazu ist das Maxieder das zu einer Punktwolke gehörige Polyeder größten Volumens, das genau die Punkte der Wolke als Eckpunkte enthält. Im Spezialfall, wo die konvexe Hülle keine Punkte im Innern verbirgt, sind Minieder, Maxieder und konvexe Hülle identisch. Im Gegensatz zur konvexen Hülle sind Minieder und Maxieder nicht unbedingt eindeutig.

Minieder und Maxieder enthalten konkave Stellen, so daß keine Modellierungspunkte mehr verdeckt werden. Die Konkavitäten werden so gewählt, daß die Abweichungen von der konvexen Hülle entweder nach der Oberfläche oder nach dem Volumen gemessen minimal sind.

Die Konstruktion dieser Körper ist jedoch enorm aufwendig. Das Problem ist verwandt mit dem "Problem des Handlungsreisenden". J. O'Rourke schlägt verschiedene Heuristiken vor, mit deren Hilfe man das Minieder durch Suche konstruieren kann, jedoch garantieren die Heuristiken nicht, daß man die Lösung schneller als mit der systematischen Suche findet. Dagegen gibt es verschiedene Möglichkeiten, um in kurzer Zeit gute Näherungslösungen zu finden. Ein Vorschlag ist, von der konvexen Hülle auszugehen und iterativ die Oberfläche lokal nach Innen einzubeulen, bis alle Innenpunkte auf der Oberfläche liegen. Dabei ist darauf zu achten, daß die Einbuchtungen so gewählt werden, daß der Oberflächenzuwachs jeweils der kleinstmögliche ist.

Modellierung

Weitere Polyedermodelle

Da Minieder und Maxieder auch konkave Stellen aufweisen, sehen sie realistischer aus als die konvexe Hülle. Allerdings ist es ein Trugschluß anzunehmen, daß diese Modelle dem ursprünglichen Objekt dadurch ähnlicher sind. Es ist sehr wohl möglich, daß das Minieder an Stellen konkav ist, wo das Objekt konvex ist und umgekehrt. So kann ein völlig falscher Eindruck vom Objekt entstehen, der gefährlicher ist als bei der konvexen Hülle. Wenn man dagegen so viele Modellierungspunkte gewonnen hat, daß die Oberflächen durch lineare Interpolation zwischen benachbarten Punkten rekonstruiert werden können, dann ist das Minieder ein sehr günstiges Modell. In der Regel jedoch werden die Modellierungspunkte nicht dicht genug liegen.

Zwar halte ich es für günstig, wie vorgeschlagen die konvexe Hülle so zu modifizieren, daß die versteckten Punkte mit in das Modell aufgenommen werden. Allerdings würde ich beim Einbau der nötigen konkaven Stellen keine numerischen Kriterien wie minimale Oberfläche oder maximales Volumen einsetzen, die in keinem Zusammenhang mit dem modellierten Objekt stehen. Spätestens zu diesem Zeitpunkt halte ich es für besser, statt zu raten, sich am Bild des Objektes zu orientieren. Da das Kameramodell bekannt ist, kann das errechnete Modell ins Bild zurückprojiziert und mit der Objektabbildung verglichen werden. So ist es möglich, die Korrekturen am Modell von den beobachteten Unterschieden abhängig zu machen. Wenn das Objektverfolgungsprogramm weit genug entwickelt ist, wird es hilfreiche Informationen zur Verfügung stellen. In allen Bildern der Folge wird die genaue Objektkontur bekannt sein, die dann mit dem projizierten Umriß des Modells verglichen werden kann. Aus den Unterschieden läßt sich schließen, wo im Modell Konkavitäten eingefügt werden müssen. Wenn das Objekt sich während der Aufnahme weit genug gedreht hat, so daß mehrere typische Ansichten zur Verfügung stehen, kann das Objekt so schon recht wirklichkeitsnah werden. Auch eine Segmentation

Modellierung

Weitere Polyedermodelle

des Objektbildes könnte hilfreich sein. Wenn das Objekt im wesentlichen konvex ist, sollten sich einige Kanten oder Flächen der konvexen Hülle im Segmentationsergebnis identifizieren lassen. Die so verifizierten Modellbestandteile dürfen dann nicht mehr modifiziert werden.

Mit welchen Problemen bei dieser bildorientierten Modellverbesserung zu rechnen ist, habe ich noch nicht untersucht. Mein Vorschlag erscheint mir aber aussichtsreicher als die rein mathematischen Verfahren, da auf diese Weise aus der Bildfolge entnommenes Wissen über das Objekt eingesetzt wird, um aus der Vielzahl von möglichen Polyedermodellen das am besten geeignete auszuwählen.

8. Schlussbemerkungen

Die in dieser Arbeit vorgestellten Ergebnisse haben gezeigt, daß es möglich ist, allein durch Beobachtung der verschiedenen Ansichten von bewegten Objekten eine angenäherte 3D-Beschreibung zu erhalten. Diese kann als Ausgangsbasis für die Ermittlung einer genaueren Beschreibung dienen. Das hier beschriebene Verfahren kommt ohne eine Segmentation des Bildes aus. Die Koordinaten von markanten Punkten sind die einzigen Merkmale des Objektbildes, die für die 3D-Rekonstruktion ausgewertet werden. Nur zur Auswahl der Objektkandidaten wird Bildsegmentation (durch Bewegung) eingesetzt. Das Verfahren wurde an zwei Realwelt-Bildfolgen erprobt. Die beiden neu entwickelten Punktefinder auf der Basis der Gaußschen Krümmung haben sich in der Auswahl von markanten Punkten einfachen statistischen Auffälligkeitsmaßen, wie dem Moravec-Operator, überlegen erwiesen. In der Reproduzierbarkeit sind sie mindestens gleichwertig.

Es gibt noch eine Reihe von Teilproblemen, bei denen der vorgestellte Ansatz verbessert oder erweitert werden könnte. Der entwickelte zweistufige Punktefinder wählt markante Punkte zwar genau an Ecken oder isolierten Punkten, aber noch fehlt die Möglichkeit, zu entscheiden, welche der markanten Bildpunkte wirklich Bilder von markanten Objektpunkten sind. Außerdem wird die in den Zuordnungsketten enthaltene Information noch nicht vollständig ausgewertet, da noch keine Konsistenzprüfung für die einzelnen Kettenglieder vorgenommen wird, um Zuordnungsfehler nachträglich zu erkennen oder eventuell sogar zu korrigieren. Auch die Wahl von markanten Bildpunkten könnte sicherer gemacht werden, indem zumindest für die Objektkandidaten eine grobe Bildsegmentation vorgenommen wird und nur solche Punkte gewählt werden, die auch im globalen Zusammenhang auffällig sind. Auf diese

Schlußbemerkungen

Weise könnte die Reproduzierbarkeit der Punkteauswahl noch mehr verbessert werden. Man hat aber immer noch keine Garantie dafür, daß die gewählten Punkte Bilder körperfester Punkte sind. Eine zuverlässigere Punkteauswahl erleichtert auch die Lösung des Korrespondenzproblems. Um körperfeste Objektpunkte zu finden, könnte es nützlich sein, Teile der Bildfolge erneut auszuwerten, wobei die Analyse durch das vorläufige Objektmodell gesteuert wird.

Ein weiteres offenes Problem ist das einer verbesserten Auswahl von Objektkandidaten. Bei der 3D-Rekonstruktion wäre es wünschenswert, von der ebenengebundenen Bewegung loszukommen (siehe /Roach and Aggarwal 80/, /Nagel 81/ und /Nagel and Neumann 81/).

Das Relaxationsverfahren von Barnard und Thompson zeigt auch nach den vorgenommenen Verbesserungen nicht vollständig befriedigende Ergebnisse, so daß hier noch Alternativen untersucht werden müssen.

Das Problem der Verdeckung wurde bisher ebenfalls ausgeklammert. Dieses Problem tritt sowohl bei der Auswahl von Objektkandidaten als auch beim Korrespondenzproblem auf.

Obwohl es noch die erwähnten Probleme und Verbesserungsmöglichkeiten gibt, sind die bisherigen Ergebnisse recht befriedigend und ermutigen zu weiteren Untersuchungen in der bisher verfolgten Richtung.

9. Anhang: Tangententest

Die Rechenzeit für das Verfahren von Appel und Will /Appel und Will '76/ habe ich verringert, indem ich den Sternpolygontest durch ein einfacheres Verfahren ersetzt habe, das nicht auf die Sortierung der Punkte angewiesen ist. Meine Idee war die folgende: Wenn man die von Appel und Will vorgeschlagene Polarkoordinatendarstellung beibehält und vom Ursprung aus zwei Tangenten an die Wolke der anderen Punktbilder legt, dann ist der Ursprung außerhalb der konvexen Hülle der Punktwolke, wenn der Winkel zwischen den Tangenten $\leq 180^\circ$ ist, andernfalls ist er innerhalb. Abb. 9.1 zeigt diese Verhältnisse an einem Beispiel.

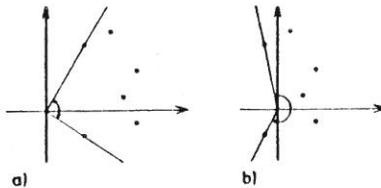


Abb. 9.1 Tangententest

Der Winkel zwischen den Tangenten läßt sich leicht aus vier Extremwerten der Polarkoordinaten ermitteln, ohne daß die Punkte sortiert werden müssen. Es werden der kleinste und der größte vorkommende Wert im Bereich 0° bis 180° (Min180, Max180) benötigt und ebenso der kleinste und größte Winkel im Bereich 180° bis 360° (Min360, Max360).

Anhang
Tangententest

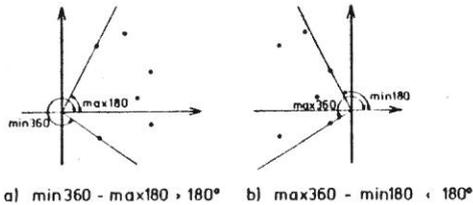


Abb. 9.2 Messung des Tangentenwinkels

Wenn $\text{Min}360 - \text{Max}180$ größer 180° ist, dann ist der Winkel zwischen den Tangenten kleiner als 180° , ebenso wenn $\text{Max}360 - \text{Min}180$ kleiner als 180° ist (Abb. 9.2). Liegen alle Punkte nur im Intervall 0° bis 180° oder nur im Intervall 180° bis 360° , dann ist der Winkel zwischen den Tangenten ebenfalls kleiner als 180° . Durch diese Prüfmöglichkeit reduziert sich der Aufwand auf $N-2$ Polarkoordinatentransformationen und auf das gleichzeitige Ermitteln von vier Extremwerten. Das Sortieren und Vergleichen des Orientierungssinns entfällt. Der Gesamtaufwand ist daher nicht mehr proportional zu $N^3 \cdot \log N$, sondern nur noch zu N^3 . Bei sehr vielen Punkten ist das eine deutliche Rechenzeiterparnis, da auch die Einzeloperationen weniger rechenzeitaufwendig sind.

Man könnte die Rechenzeit noch weiter reduzieren, indem man nicht systematisch alle möglichen $N \cdot (N-1) / 2$ Kanten auf Konvexität testet, sondern in einem rekursiven Verfahren alle Endpunkte von Kanten, die schon als zur konvexen Hülle gehörig erkannt wurden, in einem Kellerspeicher ablegt, und nur noch solche Kanten prüft, die von konvexen Punkten ausgehen. Die Rechenzeiterparnis hängt sehr davon ab, wieviele Kanten man überprüfen muß, ehe man einen Startpunkt

Anhang
Tangententest

gefunden hat. Es wird allerdings eine gute Heuristik benötigt, die es erlaubt, sicher und schnell einen Anfangspunkt zu finden. Eventuell könnte es nützen, die möglichen Kanten nach der Länge zu sortieren und mit den kürzeren zu beginnen, da diese wahrscheinlicher zur konvexen Hülle gehören als die langen, die diagonal durch den Körper laufen. Dieses rekursive Verfahren wurde wegen des Problems mit dem Startpunkt noch nicht erprobt.

Literatur

10. Literatur

Aggarwal and Duda 75

Computer Analysis of Moving Polygonal Images,
J.K. Aggarwal and R.O. Duda, IEEE Trans. Computers C-24
(1975) 966-976

Appel and Will 76

Determining the Three-Dimensional Convex Hull of a
Polyhedron, A. Appel and P.M. Will, IBM J. Research and
Development 20 (1976) 590-601

Asada et al. 80

Three Dimensional Motion Interpretation for the Sequence
of Line Drawings, M. Asada, M. Yachida, and S. Tsuji,
ICPR-80, pp. 1266-1273

Badler 74

Three-Dimensional Motion from Two-Dimensional Picture
Sequences, N.I. Badler, IJCPR-74 Aug. 13-15, 1974
Copenhagen, pp. 157-161

Badler 75

Temporal Scene Analysis: Conceptual Descriptions of
Object Movements, N.I. Badler, Technical Report No. 80,
February 1975, Dept. Computer Science, University of
Toronto

Bajcsy and Lieberman 76

Texture Gradient as a Depth Cue, R. Bajcsy and
L.I. Lieberman, Computer Graphics and Image Processing 5
(1976) 52-67

Barnard and Thompson 79

Disparity Analysis of Images, S.T. Barnard and
W.B. Thompson, Technical Report 79-1 (January 1979),
Computer Science Department, University of Minnesota,
Minneapolis/MN

Barnard and Thompson 80

Disparity Analysis of Images, S.T. Barnard and
W.B. Thompson, IEEE Trans. Pattern Analysis and Machine
Intelligence PAMI-2 (1980) 333-340

Batali and Ullman 79

Motion Detection and Analysis, J. Batali and S. Ullman,
Proc. Image Understanding Workshop, pp. 69-73,
L.S. Baumann (ed.), Los Angeles/CA, Nov. 7-8, 1979,
Science Applications, Inc., Arlington/VA 22209

Beaudet 78

Rotationally Invariant Image Operators, P.R. Beaudet,
IJCPR-78, pp. 579-583

Blohm 81

Klassifizierung der Bilder bewegter Objekte und
Auswertung von Bildfolgen auf einem Rechnerverbund,
W. Blohm, Diplomarbeit, Fachbereich Informatik,
Universität Hamburg 1981

Bonde 79

Untersuchungen zur dreidimensionalen Modellierung
bewegter Objekte durch Analyse von Formveränderungen der
Objektbilder in TV-Aufnahmefolgen, T. Bonde,

Literatur

- Diplomarbeit (Januar 1979) Fachbereich Informatik,
Universität Hamburg
- Bonde and Nagel 79
Deriving a 3-D Description of a Moving Rigid Object from
Monocular TV-Frame Sequences, T. Bonde and H.-H. Nagel,
WCATV1-79, pp. 44-45
- Budy and Dreschler 76
Untersuchungen zur Identifikation eines bewegten
Objektes (PKW) in der Video-Bildsequenz einer
Straßenszene, B. Budy and L. Dreschler, Mitteilung
IfI-III-M-53/76, Institut für Informatik, Universität
Hamburg, 1976
- Bullock 76
Finding Structure in Outdoor Scenes, B.L. Bullock, in:
Pattern Recognition and Artificial Intelligence,
C.H. Chen (ed.), Academic Press, New York 1976,
pp. 61-85
- Campbell and Robson 68
Application of Fourier Analysis to the Visibility of
Gratings, F.W. Campbell and J.G. Robson, J. Physiol.,
Lond. 197 (1968) 551-556
- Chand and Kapur 70
An Algorithm for Convex Polytopes, Chand and Kapur,
JACM-17(1970) 78-86
- Chow and Aggarwal 77
Computer Analysis of Planar Curvilinear Moving Images,
C. K. Chow and J. K. Aggarwal; IEEE Trans. Computers
C-26 (1977) 179-185
- Clocksın 78
Determining the Orientation of Surfaces from Optical
Flow, W.F. Clocksın, Proc. AISB/GI-78 on Artificial
Intelligence, Hamburg, July 18-20, 1978, pp. 93-102
- Clocksın 80a
The Effect of Motion Contrast on Surface Slant and Edge
Detection, W.F. Clocksın, Proc. AISB-80 Conference on
Artificial Intelligence, St. Hardy (ed.), Amsterdam,
July 1-4, 1980
- Clocksın 80b
Perception of Surface Slant and Edge Labels from Optical
Flow: A Computational Approach, W.F. Clocksın,
Perception 9 (1980) 253-269
- Cowan 77
Some Remarks on Channel Bandwidth for Visual Contrast
Detection, J.D. Cowan, Neurosci. Res. Prog. Bull. 15
(1977) 492-517
- Davis 75
A Survey of Edge Detection Techniques, L.S. Davis,
Computer Graphics and Image Processing 4 (1975) 248-270

Literatur

Dreschler 77

Isolierung sich bewegender Objekte in TV-Bildfolgen unter Ausnutzung der Formkonstanz bei variabler Objektbildgröße, L. Dreschler, Diplomarbeit, Institut für Informatik, Univ. Hamburg, August 1977

Dreschler and Nagel 78,

Using "Affinity" for Extracting Images of Moving Objects from TV-Frame Sequences, L. Dreschler and H.-H. Nagel, IfI-HH-B-44/78 (February 1978), Fachbereich Informatik, Universität Hamburg

Dreschler et al. 79a

Experience with a Minicomputer Network for the Analysis of TV-Frame Sequences, L. Dreschler, B. Fischer, Ma. Mittelstein, and H.-H. Nagel, WCATVI-79, pp. 111-112

Dreschler et al. 79b

Ein Pascal-Compiler für das DIETZ-System 621x2, L. Dreschler, B. Fischer, Ma. Mittelstein, and H.-H. Nagel, PASCAL- 2.Tagung Kaiserslautern, H.-W. Wippermann (ed.), Kaiserslautern, February, 16-17, 1979, Berichte des German Chapter of the ACM, 1, B.G. Teubner Stuttgart 1979

Duda and Hart 73

Pattern Classification and Scene Analysis, R.O. Duda and P.E. Hart, John Wiley & Sons, New York, 1973

Dunn 75

Group Averaged Linear Functions that Detect Corners and Edges, J.C. Dunn, IEEE Trans. Computers C-24 (1975) 1191-1201

Enomoto and Katayama 76

Structure Lines of Images, H. Enomoto and T. Katayama, IJCPR-76, pp. 811-815

Feng and Pavlidis 73

Finding "Vertices" in a Picture, H.Y. Feng and T. Pavlidis Computer Graphics and Image Processing 2 (1973) 103-117

Fennema and Thompson 79

Velocity Determination in Scenes Containing Several Moving Objects, C.L. Fennema and W.B. Thompson, Computer Graphics and Image Processing 9 (1979) 301-315

Ferrie et al. 80

Cell Tracking: A Modelling and Minimization Approach, F.P. Ferrie, M.D. Levine, and S.W. Zucker, Technical Report 80-4R (March 1980), Dept. Electrical Engineering, McGill University Montreal/Quebec, Canada

Freuder 76

Affinity: A Relative Approach to Region Finding, E. C. Freuder, Computer Graphics and Image Processing 5 (1976) 254

Literatur

- Futrelle and Cheng 79
Data Abstractions for the Description of Complex Ongoing Activities - the Design of a Moving Image Analysis System, R.P. Futrelle and W.K. Cheng, WCATV1-79, pp. 131-133
- Gennery 79a
Object Detection and Measurement Using Stereo Vision, D.B. Gennery, IJCAI-79, pp. 320-327, see, too: IUW April 1980, pp. 161-167
- Gennery 79b
Stereo-Camera Calibration, D.B. Gennery, Proc. IUW November 1979, pp. 101-107
- Gerlach 79a
Digitale Bildfolgenauswertung zum Wiederfinden von Objekten in natürlicher Umgebung, H. Gerlach, Dissertation, Universität Karlsruhe, Juli 1979
- Gerlach 79b
Digitale Bildfolgenauswertung zum Wiederfinden von Objekten in natürlicher Umgebung, H. Gerlach, in: Angewandte Szenenanalyse, J.P. Foith (ed.), Informatik Fachberichte 20, pp. 199-207, Springer Verlag, Berlin-Heidelberg-New York 1979
- Gilbert et al. 78
A Real-Time Video Tracking System Using Image Processing Techniques, A.L. Gilbert, M.K. Giles, G.M. Flachs, R.B. Rogers, and Y. Hsun U, IJ CPR-78 Nov. 7-10, 1978 Kyoto/Japan, pp. 1111-1115
- Gilbert et al. 80
A Real-Time Video Tracking System, A.L. Gilbert, M.K. Giles, G.M. Flachs, R.B. Rogers, and Y. Hsun U, IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-2 (1980) 47-56
- Grimson and Marr 79
A Computer Implementation of a Theory of Human Stereo Vision, W.E.L. Grimson and D. Marr, Proc. Image Understanding Workshop, pp. 41-47, L.S. Baumann (ed.), Palo Alto/CA, April 24-25, 1979, Science Applications, Inc., Arlington/VA 22209
- Haass and Brubaker 79
Estimation of Lateral and Rotational Cloud Displacement from Satellite Pictures, U. Haass and T.A. Brubaker, WCATV1-79, pp. 103-104
- Halé 76
Detection of Elementary Features in a Picture by Non Linear Local Numerical Processing, J.A.G. Halé, IJ CPR-76, pp. 764-768
- Hannah 74
Computer Matching of Areas in Stereo Images, M.J. Hannah, Ph.D. Thesis, Memo AIM 239 (July 1974), Stanford University, Stanford/CA

Literatur

- Haralick et al. 80
Compatibilities and the Fixed Points of Arithmetic Relaxation Processes, R.M. Haralick, J.L. Mohammed, and S.W. Zucker, Computer Graphics and Image Processing 13 (1980) 242-256
- Hayashi and Suto 77
A Feature Extraction Algorithm Based on Local Picture Computation, Y. Hayashi and Y. Suto, IEEE Trans. Systems, Man and Cybernetics SMC-7 (1977) 743-749
- Helland 80
Convergence Properties of Two Label Relaxation, A.W. Helland, Proc. IUW April 1980, pp. 176-181
- Herman et al. 79
Detection of Changing Boundaries in Two and Three Dimensions, G.T. Herman, S.N. Srihari, and J.K. Udupa, WCATV1-79, pp. 14-16
- Hernan and Jimenez 79
Automatic Analysis of Movies in Fluid Mechanics, M.A. Hernan and J. Jimenez, WCATV1-79, pp. 134-135
- Himmelblau 72
Applied Nonlinear Programming, D.M. Himmelblau, McGraw-Hill Book Company, New York 72
- Hogg 77
A Methodology for Real Time Scene Analysis, D.C. Hogg, IJCAI-77, 627
- Hsu et al. 78
Web Representation of Image Data, S. Hsu, J.L. Mundy and P.R. Beaudet, IJCPR-78, pp.675-680
- Hummel 79
Feature Detection Using Basic Functions, R.A. Hummel, Computer Graphics and Image Processing 9 (1979) 40-55
- Hummel and Zucker 80
On the Foundations of Relaxation Labeling Processes, R.A. Hummel and S.D. Zucker, ICPR-80, pp. 50-53, and TR 80-7 (July 1980) Computer Vision and Graphics Laboratory, Department of Electrical Engineering, McGill University, Montreal/Quebec, Canada
- Jain and Nagel 79
On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes, R. Jain and H.-H. Nagel, IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-1 (1979) 206-214
- Jain et al. 77
Separating Non-Stationary from Stationary Scene Components in a Sequence of Real World TV-Images, R. Jain, D. Militzer, and H.-H. Nagel, IJCAI-77, 612-618 and IfI-III-B-32/77 (March 1977), Institut für Informatik, Universität Hamburg,
- Kelly 70
Visual Identification of People by Computer, M.D. Kelly, Memo AI-130, Comp.Sc.Dept. Stanford Univ. 1970

Literatur

- Kelly 71
Edge Detection in Pictures by Computer Using Planning
M.D. Kelly, Machine Intelligence 6 (B. Meltzer,
D. Michie, eds.), Edinburgh University Press, 1971,
pp. 379-409
- Kitchen and Rosenfeld 80
Gray-Level Corner Detection, L. Kitchen and
A. Rosenfeld, TR-887 (April 1980), Computer Science
Center, University of Maryland, College Park/MD
- Kovalevsky 79
Local versus Global Decisions in Image Recognition,
V.A. Kovalevsky, Proc. IEEE 67 (1979) 745-752
- Kraasch et al. 79a
Automatic Three-Dimensional Description of Simple Moving
Objects, R. Kraasch, B. Radig, and W. Zach, WCATV1-79,
pp. 42-43
- Kraasch et al. 79b
Automatische Dreidimensionale Beschreibung bewegter
Gegenstände. R. Kraasch, B. Radig, W. Zach. in:
Angewandte Szenenanalyse. J.P. Foith (ed.) Informatik
Fachberichte 20, pp. 208-215. Springer Verlag,
Berlin-Heidelberg-New York 1979
- Levine et al. 79
Cell Movements: Its Characterization and Analysis,
M.D. Levine, Y.M. Youssef, and F. Ferrie, WCATV1-79,
pp. 93-95
- Limb and Murphy 75
Estimating the Velocity of Moving Images in Television
Signals, J.O. Limb and J.A. Murphy, Computer Graphics
and Image Processing 4 (1975) 311-327
- Marr 76
Early Processing of Visual Information, D. Marr,
Phil. Trans. Royal Society, Series B, 275 (1976) 483-524
- Marr and Hildreth 80
Theory of Edge Detection, D. Marr and E. Hildreth,
Proc. R. Soc. Lond. B 207 (1980) 187-217
- Marr and Nishihara 76
Representation and Recognition of the Spatial
Organisation of Three-Dimensional Shapes, MIT Artificial
Intelligence Laboratory, A.I. Memo 372, October 1976
- Marr and Nishihara 78
Representation and Recognition of the Spatial
Organization of Three Dimensional Shapes, D. Marr,
H.K. Nishihara, Proc. Royal Society B.200 (1978) 269-294
- Marr and Poggio 79
A Computational Theory of Human Stereo Vision, D. Marr
and T. Poggio, Proc. Royal Society London B 204 (1979)
301-308
- Martin and Aggarwal 78
Survey: Dynamic Scene Analysis, W.N. Martin and
J.K. Aggarwal, Computer Graphics and Image Processing 7
(1978) 356-374

Literatur

Meiri 80

On Monocular Perception of 3-D Moving Objects, A. Zvi Meiri, IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-2 (1980) 582-583

Meschkowski 71

Mathematisches Begriffswörterbuch, H. Meschkowski, B.I. Hochschultaschenbücher 99/a/b, Bibliographisches Institut Mannheim 1971

Milstein and Lazicky 78,

Statistical Tests for Image Tracking, L.B. Milstein and T. Lazicky, Computer Graphics and Image Processing 7 (1978) 413-424

Minsky 75

A Framework for Representing Knowledge, M. Minsky, in: The Psychology of Computer Vision, Winston (ed.) McGraw Hill New York 1975

Moravec 77

Towards Automatic Visual Obstacle Avoidance, H.P. Moravec, IJCAI-77, p. 584

Moravec 79

Visual Mapping by a Robot Rover, H.P. Moravec, IJCAI-79, pp. 598-600

Moravec 80

Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover, H.P. Moravec, Ph.D. Thesis, Department of Computer Science, Stanford University, available as CMU-RI-TR-3 (September 1980), Robotics Institute, Carnegie-Mellon University Pittsburgh/PA

Nagel 76

Experiences with Yakimovsky's Algorithm for Boundary and Object Detection in Real World Images: H.-H. Nagel, IJCPR-76, 753-758

Nagel 77

Analysing Sequences of TV-Frames: System Design Considerations, H.-H. Nagel, IJCAI-77, p.626, IfI-III-B-33/77 (March 1977) Fachbereich Informatik, Universität Hamburg

Nagel 78a

Formation of an Object Concept by Analysis of Systematic Time Variations in the Optically Perceptible Environment, H.-H. Nagel, Computer Graphics and Image Processing 7 (1978) 149-194

Nagel 78b

Analysis Techniques for Image Sequences, H.-H. Nagel, IJCPR-78, pp. 186-211

Nagel 80a

From Digital Picture Processing to Image Analysis, H.-H. Nagel, Proc. International Conference on Image Analysis and Processing, Pavia/ Italy, October 22-24, 1980

Literatur

Nagel 80b

Image Sequence Analysis: What can we learn from Applications? H.-H. Nagel, FBI-HH-M-79/80 (September 1980), Fachbereich Informatik, Universität Hamburg/Germany to appear in: T.S. Huang (ed.) Image Sequence Analysis, Springer Verlag Berlin-Heidelberg-New York 1981

Nagel 81

On the Derivation of 3D Rigid Point Configurations from Image Sequences, H.-H. Nagel, IEEE Conference on Pattern Recognition and Image Processing, Dallas/TX, August 3-5, 1981

Nagel and Neumann 81

On 3D Reconstruction from two Perspective Views, H.-H. Nagel and B. Neumann, IJCAI-81, pp. 661-663

Neumann 78a

Best-First Interpretation of Imperfect Object Contours, B. Neumann, AISB/GI Conference on Artificial Intelligence, July 18-20, 1978, Hamburg, pp. 222-228

Neumann 78b

Interpretation of Imperfect Object Contours for Identification and Tracking, B. Neumann, IJ CPR-78 Nov. 7-10, 1978 Kyoto/Japan, pp. 691-693

Neumann 79a

Exploiting Image Formation Knowledge for Motion Analysis, B. Neumann, WCATVI-79, pp. 34-35

Neumann 79b

Räumliche Analyse von Bildsequenzen mithilfe korrespondierender Kanten, B. Neumann, in: Angewandte Szenenanalyse, J.P. Foith (ed.), Informatik Fachberichte 20, pp. 216-221, Springer Verlag, Berlin-Heidelberg-New York 1979

Neumann 80

Exploiting Image Formation Knowledge for Motion Analysis, B. Neumann, IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-2 (1980) 550-554

Nevatia 76

Depth Measurement by Motion Stereo, R. Nevatia, Computer Graphics and Image Processing 5 (1976) 203-214

O'Rourke 80a

Image Analysis of Human Motion, J. O'Rourke, Ph.D. Dissertation (May 1980), The Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia/PA

O'Rourke 80b

Polyhedral Object Models from 3D-Points, J. O'Rourke, IFl-HH-M-77/80 Fachbereich Informatik Universität Hamburg 1980

O'Rourke and Badler 80

Model-Based Image Analysis of Human Motion Using Constraint Propagation, J. O'Rourke and N.I. Badler, IEEE Pattern Analysis and Machine Intelligence PAMI-2 (1980) 522-536

Literatur

- Onoe and Ohba 76
Digital Image Analysis of Traffic Flow, M. Onoe and K. Ohba, IJCPR-76, 803-808
- Paton 75
Picture Description Using Legendre Polynomials, K. Paton, Computer Graphics and Image Processing 4 (1975) 40-54
- Pavlidis 77
Structural Pattern Recognition, T. Pavlidis, Springer Series in Electrophysics, Vol. 1, G. Ecker, W. Engl, L.B. Felsen, K.S. Fu, T.S. Huang (ed.) Springer-Verlag Berlin Heidelberg New York 1977
- Peleg 80
Monitoring Relaxation Algorithms Using Labeling Evaluations, S. Peleg, ICPR-80, pp. 54-57
- Perkins and Binford 73,
A Corner Finder for Visual Feedback, W.A. Perkins and T.O. Binford, Computer Graphics and Image Processing 2 (1973) 355-374
- Peucker and Douglas 75
Detection of Surface-Specific Points by Local Parallel Processing of Discrete Terrain Elevation Data, T.K. Peucker and D.H. Douglas, Computer Graphics and Image Processing 4 (1975) 375-387
- Potter 75a
Velocity as a Cue to Segmentation, J. Potter, IEEE Trans. Systems, Man and Cybernetics, SCM-5 (1975) 390-394
- Potter 75b
Scene Segmentation by Velocity Measurements Obtained with a Cross-Shaped Template, J. Potter, IJCAI-75, 803-810, Tbilisi, Georgia/USSR, September 3-8, 1975
- Potter 77
Scene Segmentation Using Motion Information, J.L. Potter, Computer Graphics and Image Processing 6 (1977) 558-581
- Prazdny 79
Motion and Structure from Optical Flow, K. Prazdny, IJCAI-79, pp. 702-704
- Prazdny 80,
Egomotion and Relative Depth Map from Optical Flow, K. Prazdny, Biological Cybernetics 36 (1980) 87-102
- Preparata and Hong 77
Convex Hulls of Finite Sets of Points in Two and Three Dimensions, F. P. Preparata and S. J. Hong, CACM-20 (1977) 87-93
- Price and Reddy 79
Matching Segments of Images, K. Price and R. Reddy, IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-1 (1979) 110-116

Literatur

- Radig 78a
Auswertung von digitisierten Fernsehbildern zur Beschreibung bewegter Objekte, B. Radig, Dissertation, Fachbereich Informatik (März 1978), Universität Hamburg
- Radig 78b
Parameterized Region Extracting for the Description of Moving Objects, B. Radig, AISB/GI Conference on Artificial Intelligence, July 18-20, 1978, Hamburg, pp. 263-272
- Radig 78c
Description of Moving Objects Based on Parameterized Region Extracting, B. Radig, IJCPR-78 Nov. 7-10, 1978 Kyoto/Japan, pp. 723-725
- Radig 78d
Parametrisierte Bereichsfindung in digitisierten Fernsehbildern als Grundlage für die Beschreibung bewegter Objekte, B. Radig, DAGM Symposium Bildverarbeitung und Mustererkennung, E. Triendl (ed.), Informatik Fachberichte vol 17, pp. 155-163, Springer Verlag Berlin-Heidelberg-New York 1978
- Radig et al. 80
Matching Symbolic Descriptions for 3-D Reconstruction of Simple Moving Objects, B. Radig, R. Kraasch, and W. Zach, ICPR-80, pp. 1081-1084
- Ranade and Rosenfeld 80
Point Pattern Matching by Relaxation, S. Ranade and A. Rosenfeld, Pattern Recognition 12 (1980) 269-275
- Richards et al. 80
Overcoming Accuracy Deterioration in Pixel Relaxation Labeling, J.A. Richards, D.A. Landgrebe, and P.H. Swain, ICPR-80, pp. 61-65
- Roach and Aggarwal 79
Computer Tracking of Objects Moving in Space, J. Roach and J.K. Aggarwal, IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-1 (1979) 127-135
- Roach and Aggarwal 80
Determining the Movement of Objects from a Sequence of Images, J.W. Roach and J.K. Aggarwal, IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-2 (1980) 554-562
- Rosenfeld 77
Iterative Methods in Image Analysis, A. Rosenfeld, Proc. IEEE Conf. Pattern Recognition and Image Proc., Troy, N.Y., 1977
- Rosenfeld et al. 76
Scene Labelling by Relaxation Operations, A. Rosenfeld, R.A. Hummel, S.W. Zucker, IEEE Trans. Systems, Man, and Cybernetics SMC-6 (1976) 420-433

Literatur

- Scacchi 79
Visual Motion Perception by Intelligent Systems,
W. Scacchi, IEEE Conference on Pattern Recognition and
Image Processing, Chicago/IL, August 6-8, 1979,
pp. 646-652
- Schenkel and Schenkel 77
Teilimplementation eines Systems zur Analyse von
Fernsehbildfolgen, B. Schenkel and S. Schenkel,
Diplomarbeit (1977), Fachbereich Informatik der
Universität Hamburg
- Schudy 79
Towards an Anatomical Model of Heart Motion as seen in
Cardiac Ultrasound Data, R. B. Schudy, WCATVI-79,
pp. 87-89
- Sobel 74
On Calibrating Computer Controlled Cameras for
Perceiving 3D-Scenes, I. Sobel, Artificial Intelligence
5 (1974) 185-198
- Soroka and Bajcsy 78,
A Program for Describing Complex Three-Dimensional
Objects Using Generalized Cylinders as Primitives,
B.I. Soroka and R.K. Bajcsy, IEEE Conference on Pattern
Recognition and Image Processing, Chicago/IL, May
31-June 2, 1978, pp. 331-339
- Tsotsos et al. 79
ALVEN: A Study on Motion Understanding by Computer,
J.K. Tsotsos, J. Mylopoulos, H.D. Covvey, and
S.W. Zucker, IJCAI-79, pp. 890-892
- Tsotsos et al. 80
A Framework for Visual Motion Understanding,
J.K. Tsotsos, J. Mylopoulos, H.D. Covvey, and
S.W. Zucker, IEEE Trans. Pattern Analysis and Machine
Intelligence PAMI-2 (1980) 563-573
- Tsuji et al. 79
Three Dimensional Movement Analysis of Dynamic Line
Images, S. Tsuji, M. Osada, and M. Yachida, IJCAI-79,
pp. 896-901
- Tsuji et al. 80
Tracking and Segmentation of Moving Objects in Dynamic
Line Images, S. Tsuji, M. Osada, and M. Yachida, IEEE
Trans. Pattern Analysis and Machine Intelligence PAMI-2
(1980) 516-522
- Ullman 79
The Interpretation of Visual Motion, S. Ullman, The MIT
Press, Cambridge/Mass., 1979
- Wallace and Mitchell 79
Real-Time Analysis of Three-Dimensional Movement Using
Fourier Descriptors, T.P. Wallace and O.R. Mitchell,
WCATVI-79, pp. 32-33
- Webb and Aggarwal 80
Observing Jointed Objects, J.A. Webb and J.K. Aggarwal,
ICPR-80, pp. 1246-1250

Literatur

Williams 80

Depth from Camera Motion in a Real World Scene,
T.D. Williams, IEEE Trans. Pattern Analysis and Machine
Intelligence PAMI-2 (1980) 511-516

Wilson and Giese 77

Threshold Visibility of Frequency Gradient Patterns,
H.R. Wilson and S.C. Giese, Vision Res. 17 (1977)
1177-1190

Wolf et al. 79

An Automatic Method for Determining Cloud Motions from
Pictures Taken by Geosynchronous Weather Satellites,
D.E. Wolf, D.J. Hall, and R.M. Endlich, WCATV1-79,
pp. 101-102

Wolferts 73

Ein interaktives Verfahren zur teilautomatischen
Auswertung von Luftbildern für Verkehrsanalysen,
K. Wolferts, Proc. NTG/GI Fachtagung Cognitive Verfahren
und Systeme, p. 349, Th. Einsele, W. Giloi, H.-H. Nagel
(eds.), Hamburg, April 11-13, 1973, Lecture Notes in
Economics and Mathematical Systems vol. 83, Springer
Verlag Berlin-Heidelberg-New York 1973

Yachida et al. 78

Automatic Motion Analysis System of Moving Objects from
the Records of Natural Processes, M. Yachida, M. Asada,
and S. Tsuji, IJCP-78 Nov. 7-10, 1978 Kyoto/Japan,
pp. 726-730

Yachida et al. 79

Efficient Analysis of Noisy Dynamic Pictures Using Plan,
M. Yachida, M. Ikeda, and S. Tsuji WCATV1-79, pp. 90-92

Yachida et al. 80

A Plan-Guided Analysis of Cineangiograms for Measurement
of Dynamic Behaviour of Heart Wall, M. Yachida,
M. Ikeda, S. Tsuji, IEEE Trans. Pattern Analysis and
Machine Intelligence PAMI-2 (1980) 537-543

Yachida et al. 81

Automatic Analysis of Moving Images, M. Yachida,
M. Asada, and S. Tsuji, IEEE Trans. Pattern Analysis and
Machine Intelligence PAMI-3 (1981) 12-20

Zucker et al. 79

Continuous Relaxation and Local Maxima Selection:
Conditions for Equivalence, S.W. Zucker, Y.G. Leclerc,
and J.L. Mohammed, IJCAI-79, pp. 1014-1016

INDEX

Aggarwal	19-21, 28, 55, 135, 158
Appel	144-145, 149, 159
Asada	135
Badler	22, 32, 131
Bajcsy	14, 143
Barnard	3, 11, 36, 52, 65, 73, 99, 102, 104, 115, 121, 126
Batali	24
Beaudet	57, 75
Binford	56
Blohm	45, 130
Bonde	10, 15, 32, 36, 136, 140
Brubaker	30
Budy	26, 111
Bullock	56, 97, 109
Campbell	52
Chand	148
Cheng	34
Chow	28
Clocksini	54, 136
Cowan	52
Davis	88
Douglas	62
Dreschler	26, 32, 41, 100, 111
Duda	15, 20-21, 28, 55, 63
Dunn	58
Enomoto	57
Feng	57
Fennema	24
Ferrie	34
Freuder	99
Futrelle	34
Gennery	15, 17, 143
Gerlach	28
Giese	52
Gilbert	26
Grimson	51, 59
Haass	30
Halé	62, 64
Hannah	10, 15-16, 62, 64, 110
Haralick	113
Hart	15, 55, 63
Hayashi	62
Helland	113

Herman	32
Hernan	30
Hildreth	52, 59
Himmelblau	140
Hogg	34
Hong	148
Hsu	57, 75
Hummel	58, 113
Jain	10, 24, 36, 45, 136
Jimenez	30
Kapur	148
Katayama	57
Kelly	87
Kitchen	60
Kovalevsky	97
Kraasch	10, 32, 45, 56, 109
Lazicky	26
Levine	34
Lieberman	14
Limb	24
Marr	51-52, 58-59, 97, 109, 127, 143
Martin	19, 28
Meiri	133, 135
Meschkowski	74
Milstein	26
Minsky	22
Mitchell	26
Moravec	10, 36, 61-62, 64, 72, 111
Murphy	24
Nagel	9-10, 18-19, 23-24, 26, 32, 36, 45, 100, 133, 136, 158
Neumann	28, 32, 132-133, 135, 158
Nevatia	10, 16, 111
Nishihara	59, 143
O'Rourke	19, 22, 32, 143, 152
Ohba	26
Onoe	26
Paton	57
Pavlidis	57, 113
Peleg	113, 126
Perkins	56
Peucker	62
Platon	4, 9
Poggio	51, 58, 109, 127
Potter	24
Prazdny	136

Preparata	148
Price	109
Radig	19, 21, 28, 45, 109
Ranade	52, 113
Reddy	109
Richards	113
Roach	28, 55, 158
Robson	52
Rosenfeld	52, 60, 113
Scacchi	19
Schenkel	47
Schudy	34, 143
Sobel	15
Soroka	143
Suto	62
Thompson	3, 11, 24, 36, 52, 65, 73, 99, 102, 104, 115, 121, 126
Tsotsos	22, 34
Tsuji	21, 30, 45, 109
Ullman	9, 24, 32, 96, 103, 133-134
Wallace	26
Webb	135
Will	144-145, 149, 159
Williams	136
Wilson	52
Wolf	30
Wolferts	26
Yachida	28, 30, 87
Yakimovsky	45
Zucker	113



a) erstes verarbeitetes Bild

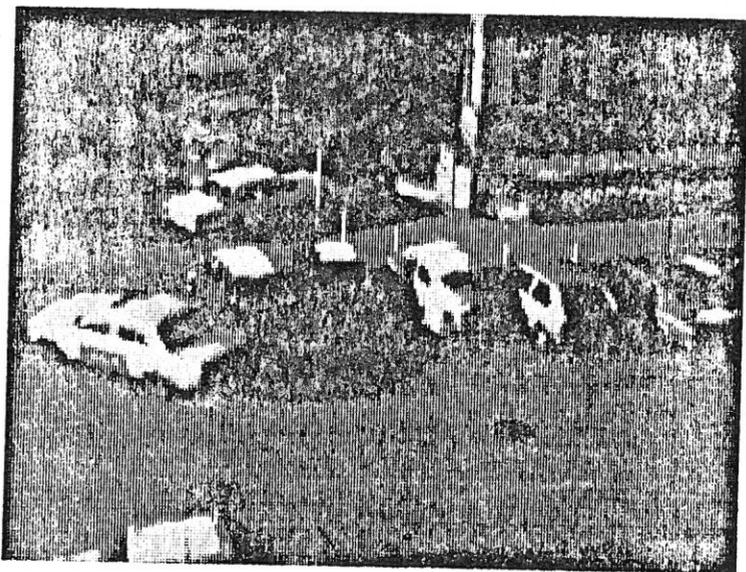


b) letztes verarbeitetes Bild

Abb. 1 Bildfolge 1

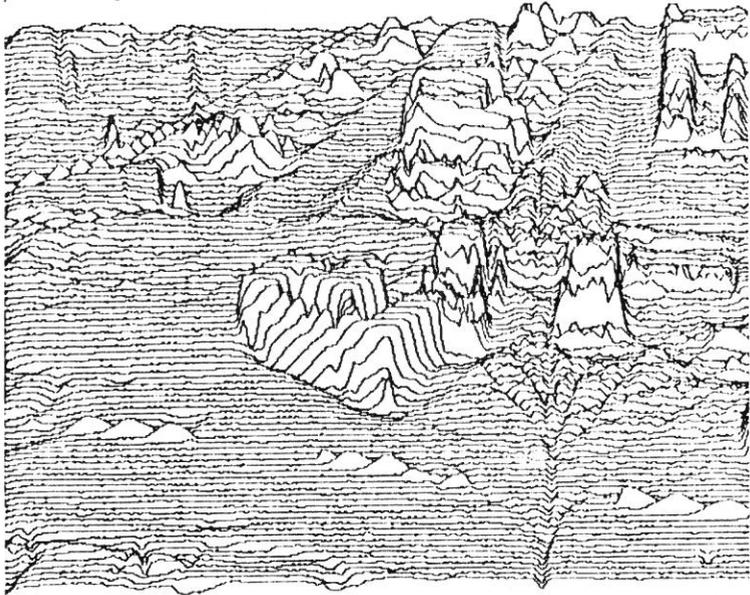


a) erstes verarbeitetes Bild



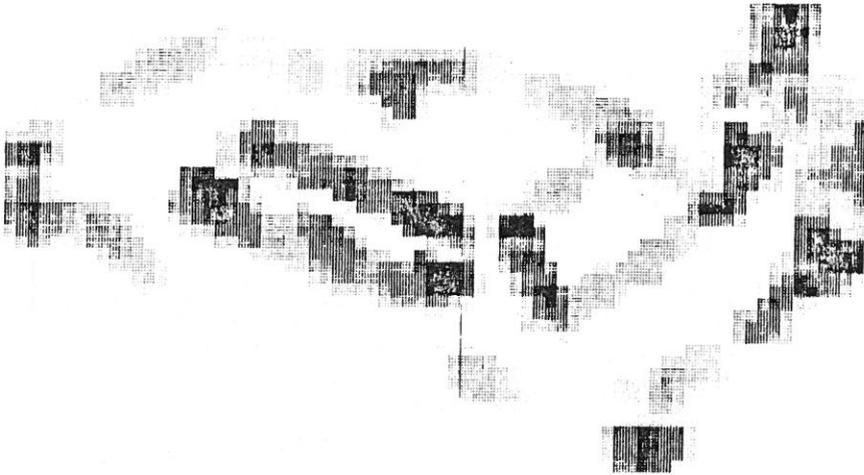
b) letztes verarbeitetes Bild

Abb. II Bildfolge 2

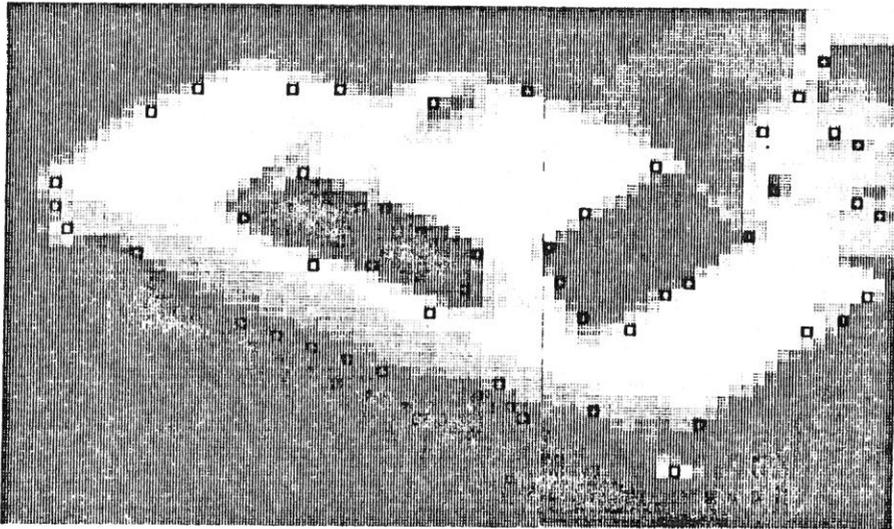


Pseudo-3D U211 of TAXI 28-SEP-1979 11:54:49

Abb. III Darstellung von Abb. I a als Grauwertgebirge



a) Operatorresultat für jedes Pixel durch einen Grauwert codiert (dunkel $\hat{=}$ hoher Wert)



b) ausgewählte Modellierungspunkte (Schwellwert: 50)

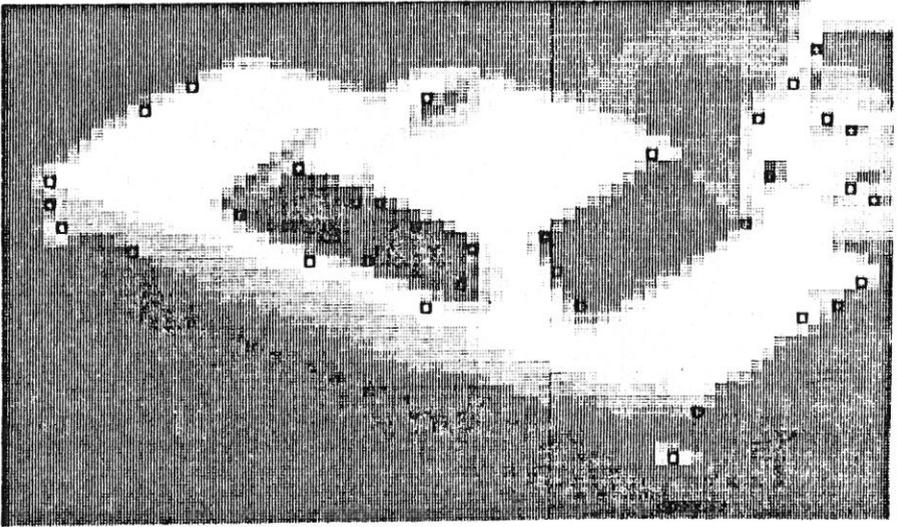
Abb. IV Ergebnis des Moravec-Operators

Ausschnitt um das Bild des Taxis aus Folge 1 (Abb. I)

schwarze Quadrate markieren lokale Maxima, die den Schwellwert überschreiten

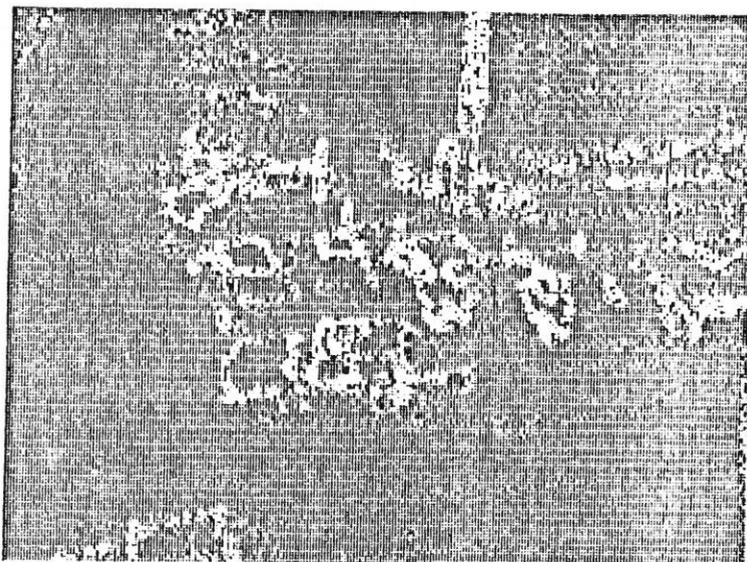


c) ausgewählte Modellierungspunkte (Schwellwert: 100)

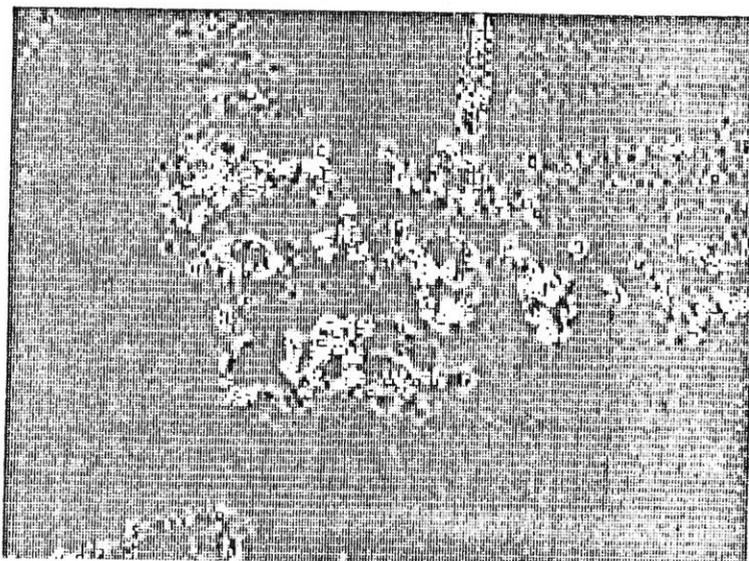


d) ausgewählte Modellierungspunkte (Schwellwert: 500)

Abb. IV Ergebnis des Moravec-Operators



a) Operatordurchmesser 3

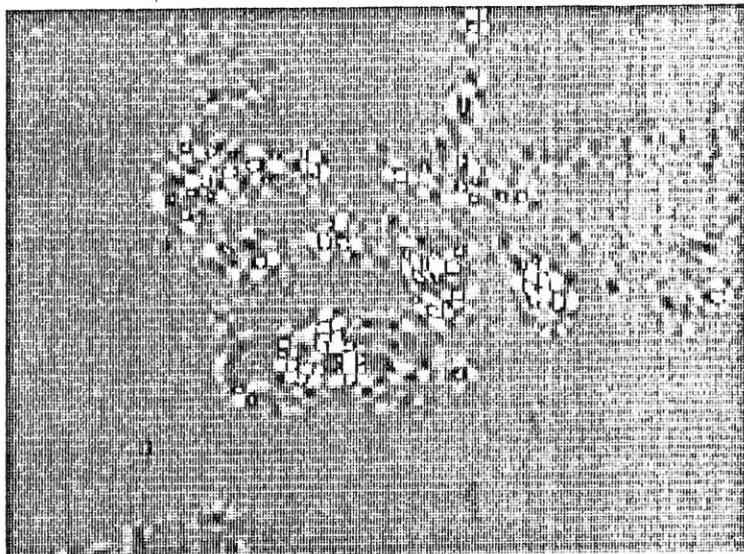


b) Operatordurchmesser 5

Abb. V Gaußsche Krümmung

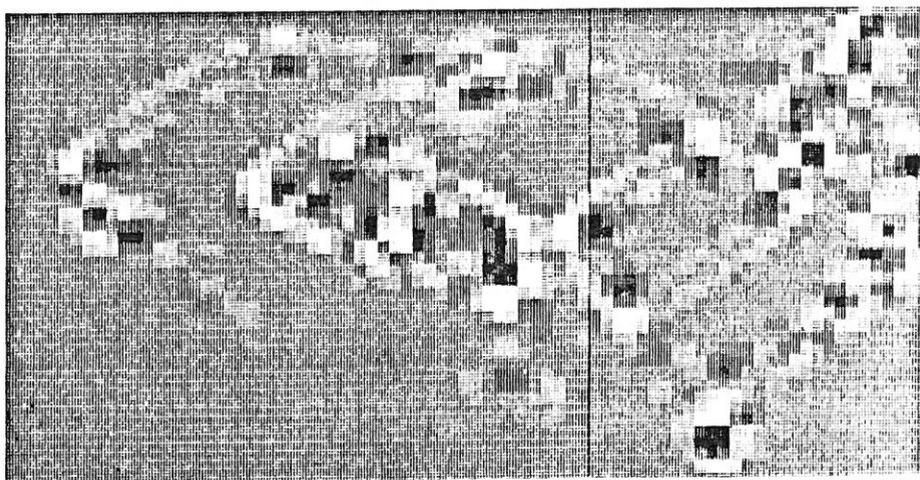
Beispiel aus Abb II a

Darstellung durch Grauwerte: hell $\hat{=}$ negative Werte, dunkel $\hat{=}$ positive Werte



c) Operatordurchmesser 7

Abb. V Gaußsche Krümmung



a) Gaußsche Krümmungen in Grauwertdarstellung (hell $\hat{=}$ neg Werte, dunkel $\hat{=}$ pos Werte)

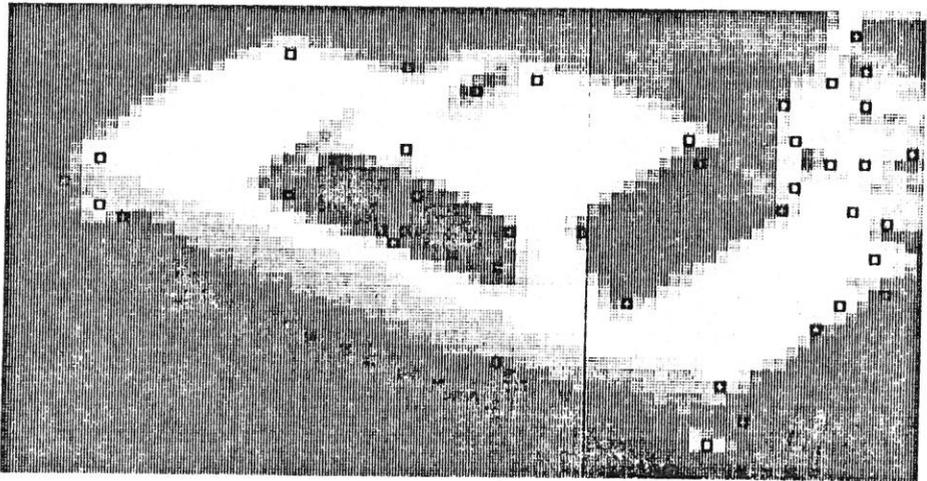
Abb. VI Einstufiger Punktefinder

Operatordurchmesser 5

Bildausschnitt wie in Abb. IV

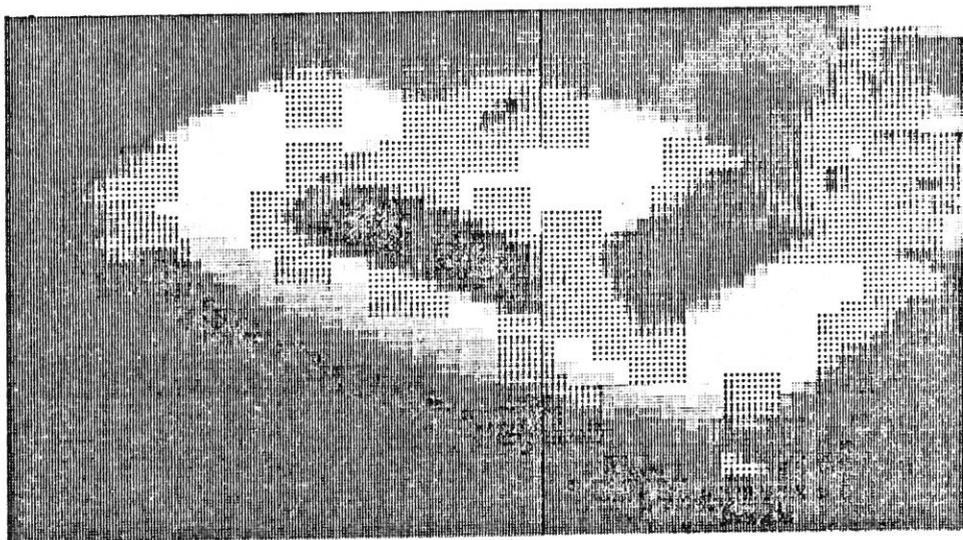


b) Lage der Extremwerte der Gaußschen Krümmung (-: lok. Minima, +: lok. Maxima)

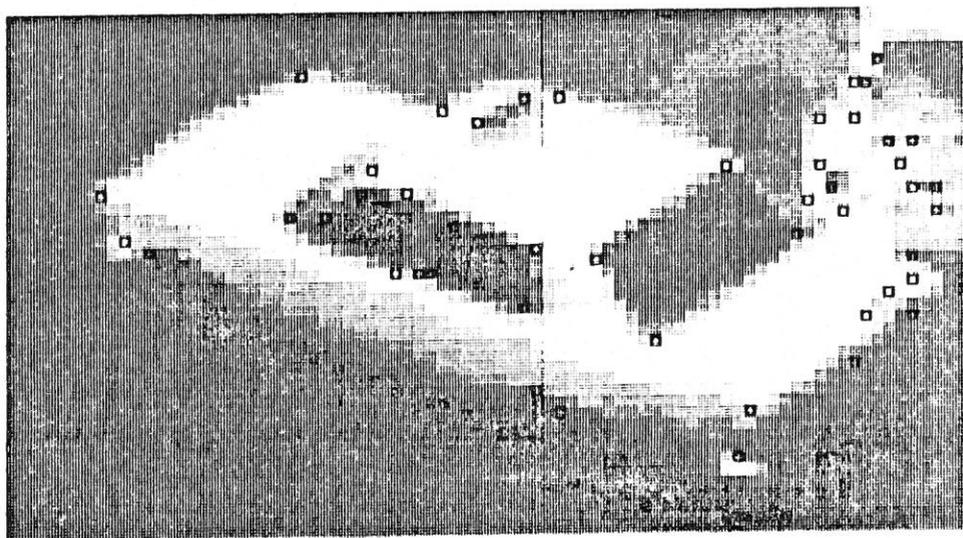


c) Lage der gefundenen markanten Bildpunkte

Abb. VI Einstufiger Punktefinder

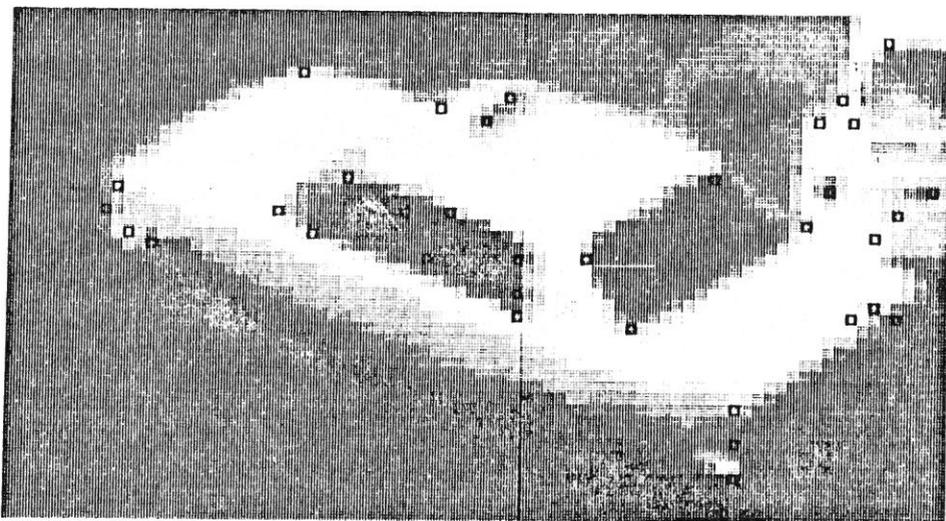


a) Füllung: das dunkle Raster markiert die Bereiche, die mit dem 3×3 -Operator untersucht werden



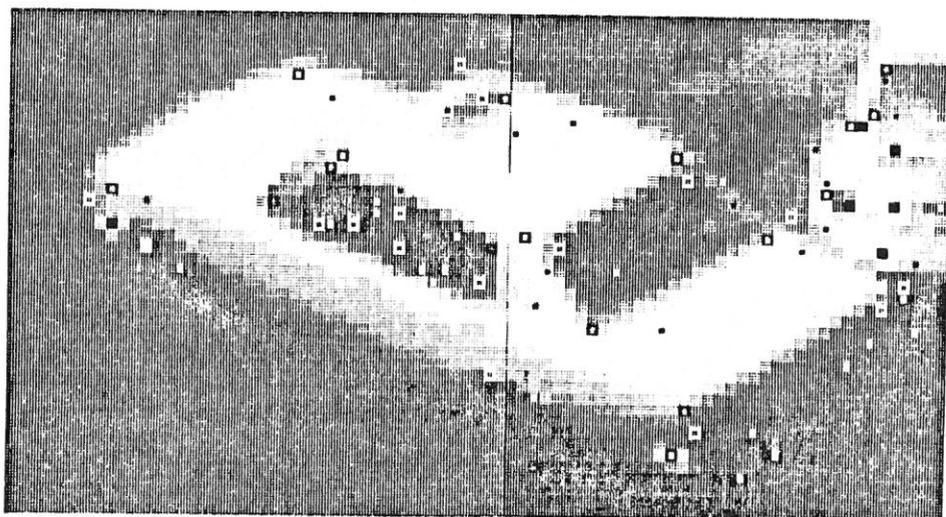
b) markante Bildpunkte im ungefüllten Bild

Abb. VII Zweistufiger Punktfinder
Operatordurchmesser 5, gefolgt von Operatordurchmesser 3
Bildausschnitt wie in Abb IV



c) markante Bildpunkte im Median-gefilterten Bild

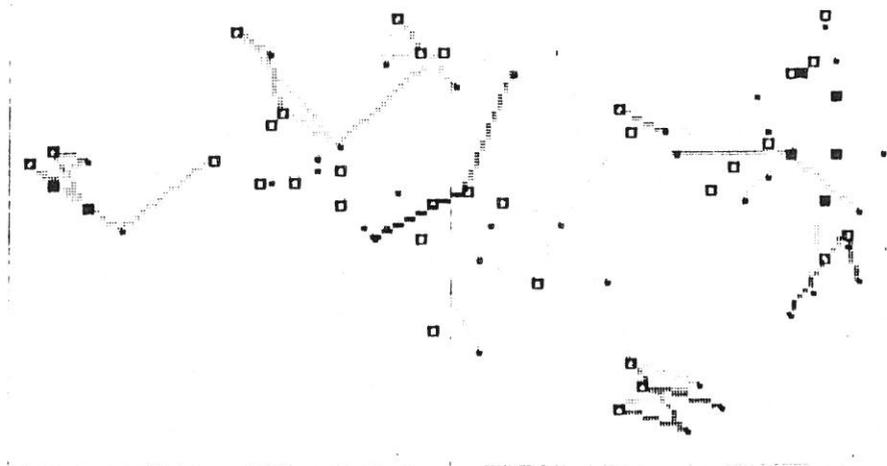
Abb. VII Zweistufiger Punktefinder



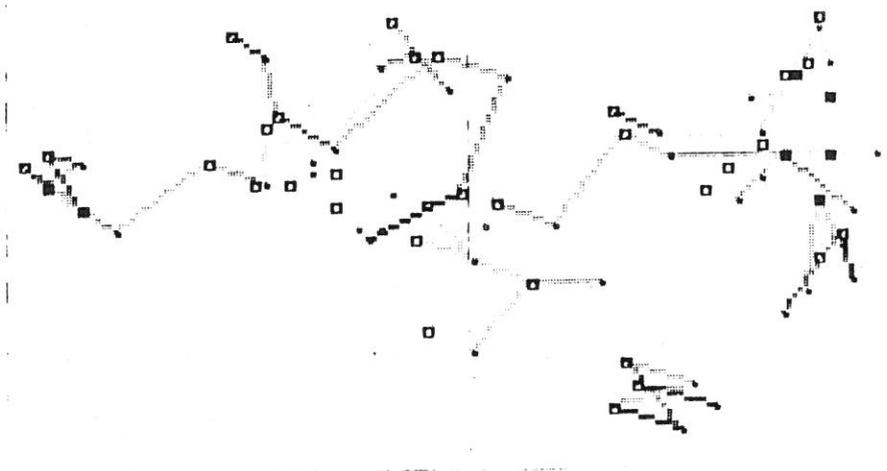
a) Lage der markanten Punkte (Bild 1: Punkte, Bild 3: Quadrate)

Abb. VIII Zwischenergebnisse des Relaxationsverfahrens

markante Punkte aus dem ersten und dritten Bild von Folge 1; die Wahrscheinlichkeiten für die Korrespondenz zwischen Punkten wird durch die Grauwerte der Verbindungslinien in Abb. b) - e) wiedergegeben (dunkel $\hat{=}$ hohe Wahrscheinlichkeit)



b) Anfangswahrscheinlichkeiten

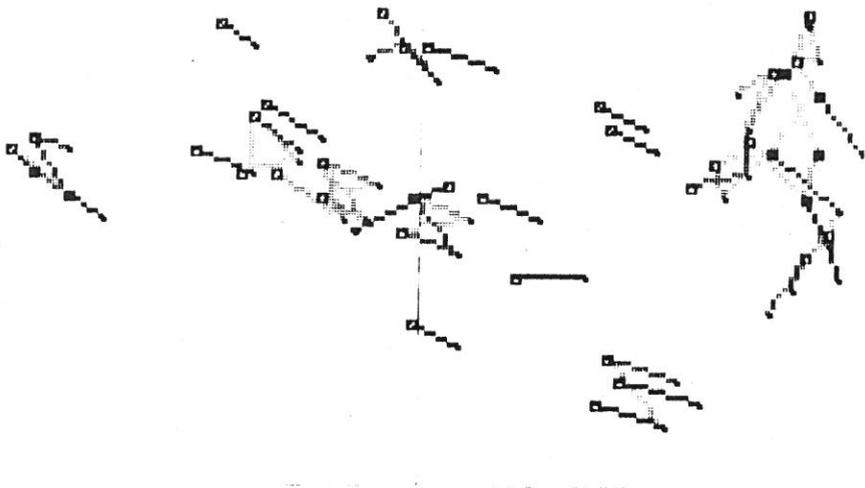


c) nach 2 Iterationen

Abb. VIII Zwischenergebnisse des Relaxationsverfahrens

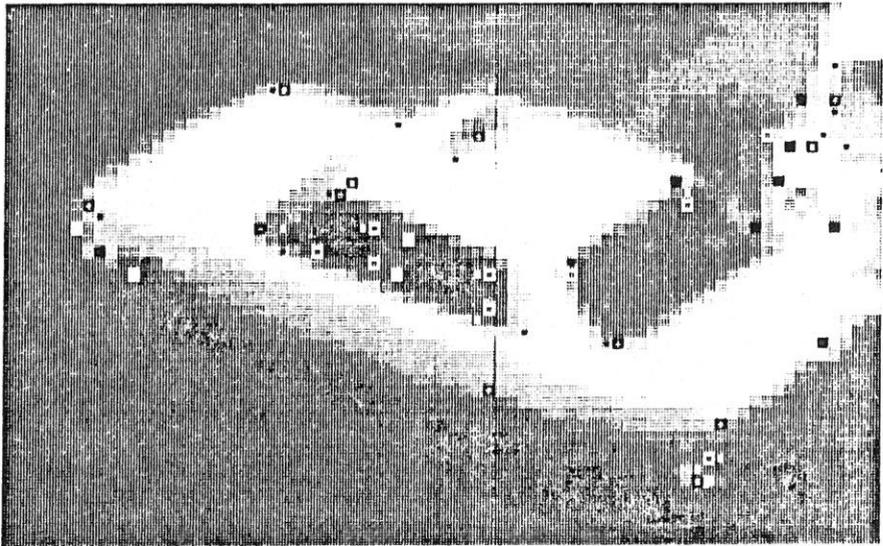


d) nach 4 Iterationen



e) nach 10 Iterationen

Abb. VIII Zwischenergebnisse des Relaxationsverfahrens



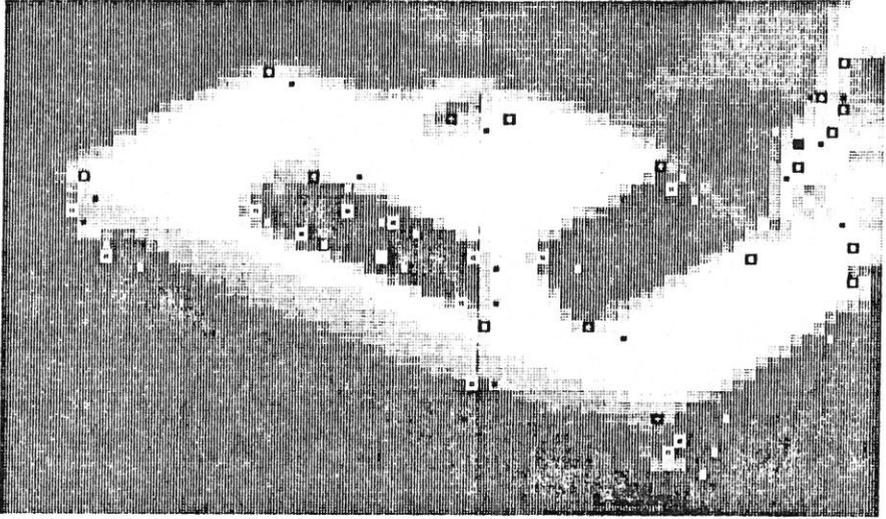
a) markante Bildpunkte aus dem ersten (Punkte) und zweiten (Quadrate) Halbbild von Bild 1 der Folge 1



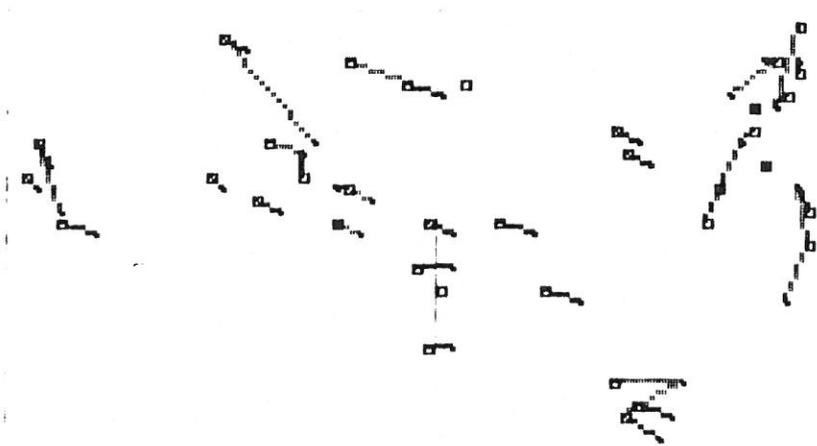
b) Zuordnungsergebnis für die Punkte aus Abb. a)

Abb. IX Nächste-Nachbar-Zuordnung

In Abb. b) und d) sind die markanten Punkte durch mittelgraue Linien mit ihren jeweils nächsten Nachbarn im anderen Bild verbunden; schwarze Linien verbinden die Punkte, die sich gegenseitig nächste Nachbarn sind und deshalb einander zugeordnet werden



c) markante Bildpunkte aus Bild 1 (Punkte) und Bild 2 (Quadrate) von Folge 1



d) Zuordnungsergebnisse für die Punkte aus Abb. c)

Abb. IX Nächste-Nachbar-Zuordnung

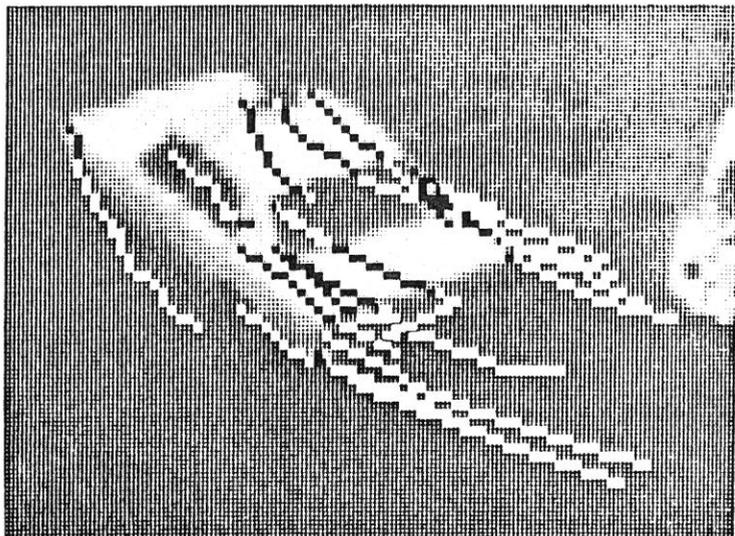
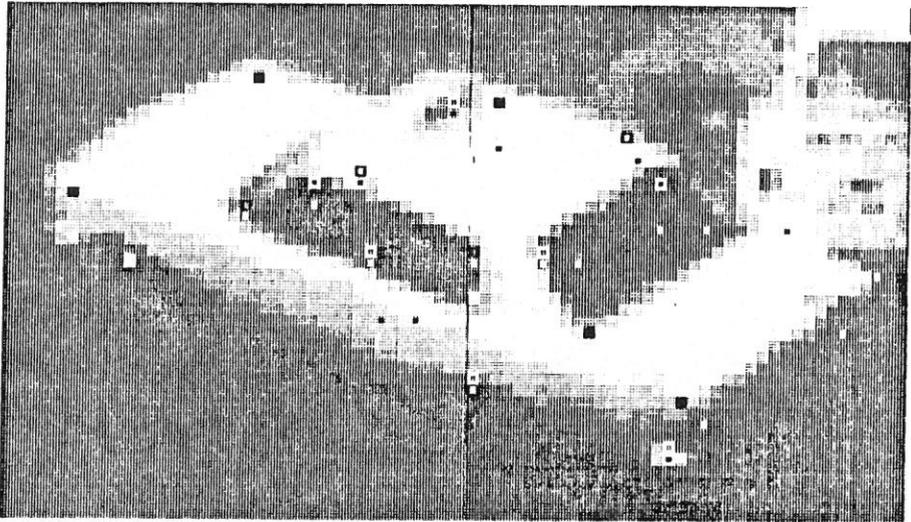
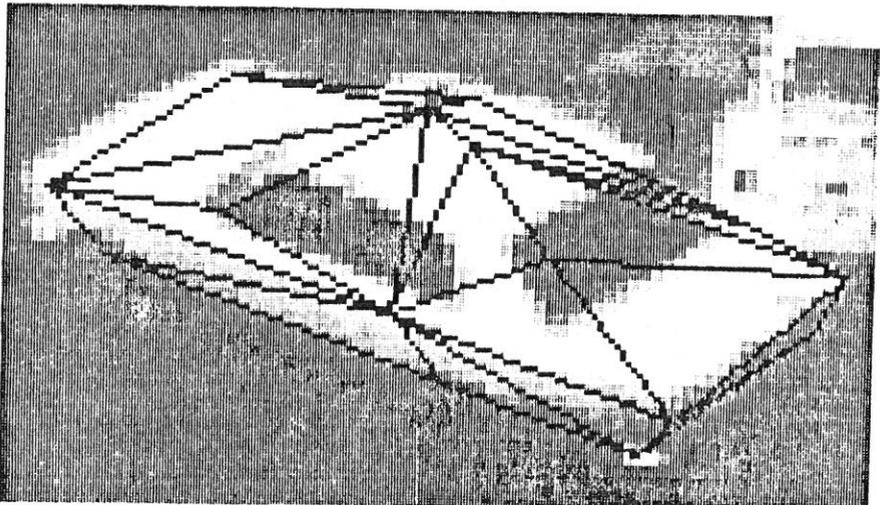


Abb. X Ketten von korrespondierenden Punkten
Bild 1 bis 14 aus Folge 1

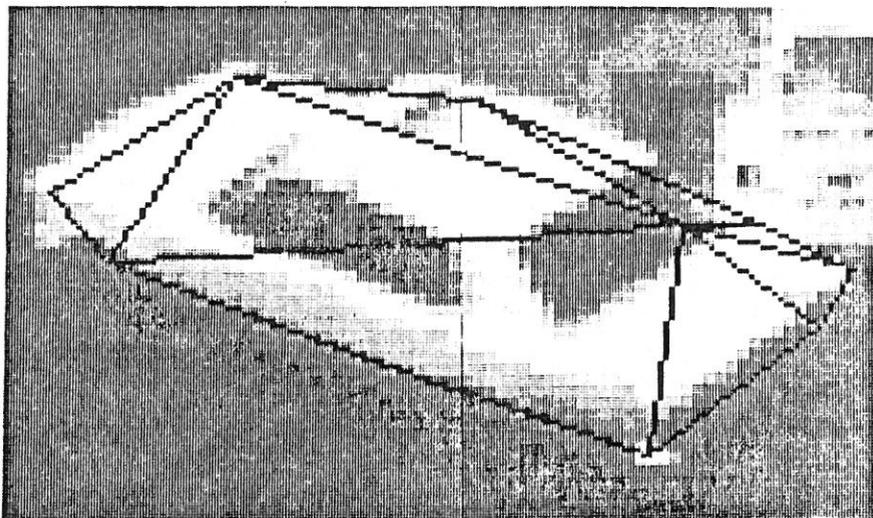


- a) Genauigkeit der 3D-Rekonstruktion
 Quadrate: markante Punkte aus Bild 1
 der Folge 1 (s. Abb. oben)
 Punkte: in das Bild rückprojizierte
 Punkte des 3D-Modells
 Vollquadrate ergeben sich, wo urspr.
 u. rückproj. Punkt zur Deckung kommen,
 während die anderen Paare in der Abb.
 rechts durch Linien verbunden sind

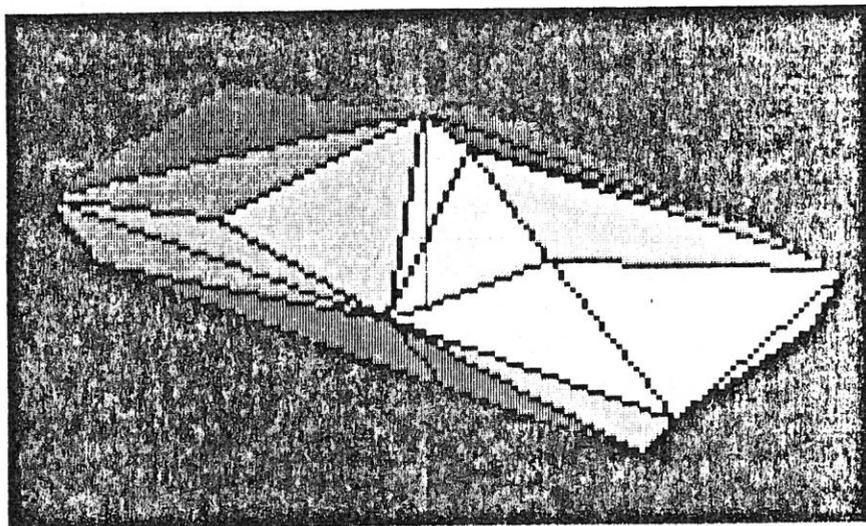


- b) sichtbare Linien der konvexen Hülle

Abb. XI 3D-Rekonstruktion für das Taxi aus Bildfolge 1

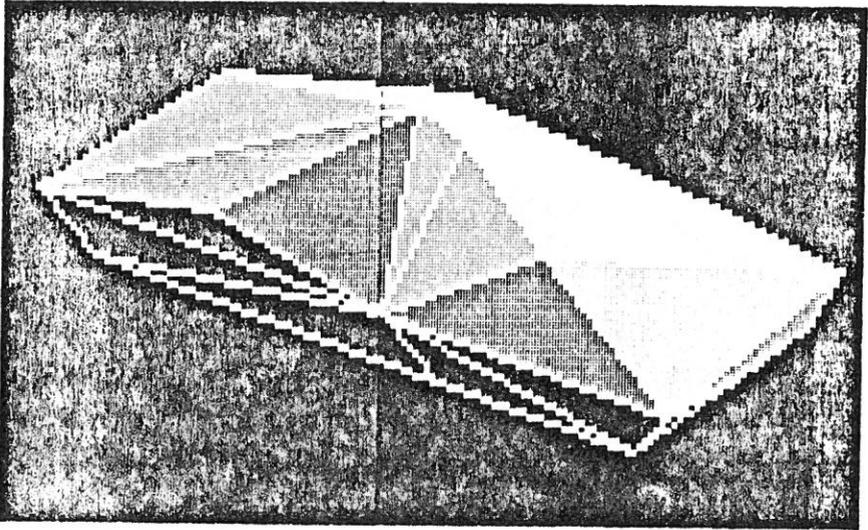


c) verdeckte Linien der konvexen Hülle

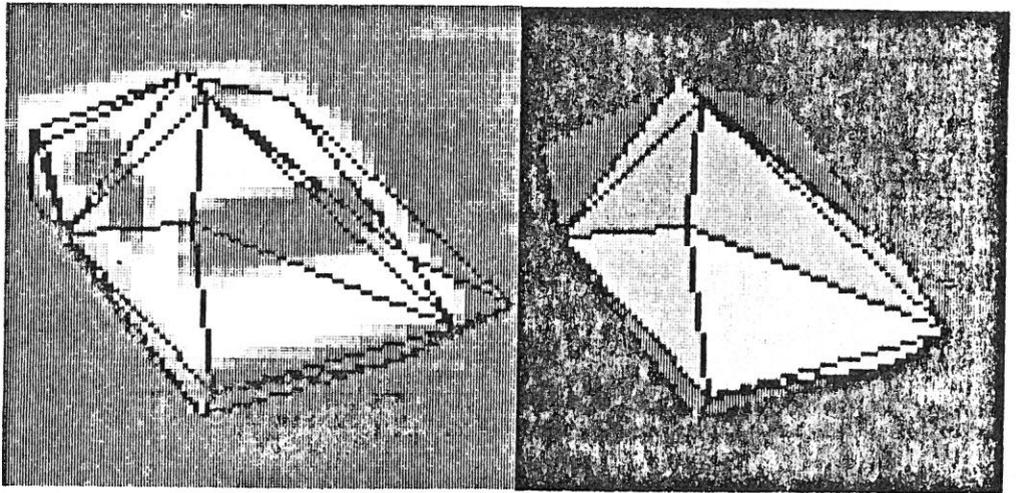


d) konvexe Hülle schattiert als lamberscher Strahler bei Beleuchtung aus der Betrachtungsrichtung;
räumliche Orientierung für das erste Bild der Folge

Abb. XI 3D-Rekonstruktion für das Taxi aus Bildfolge 1

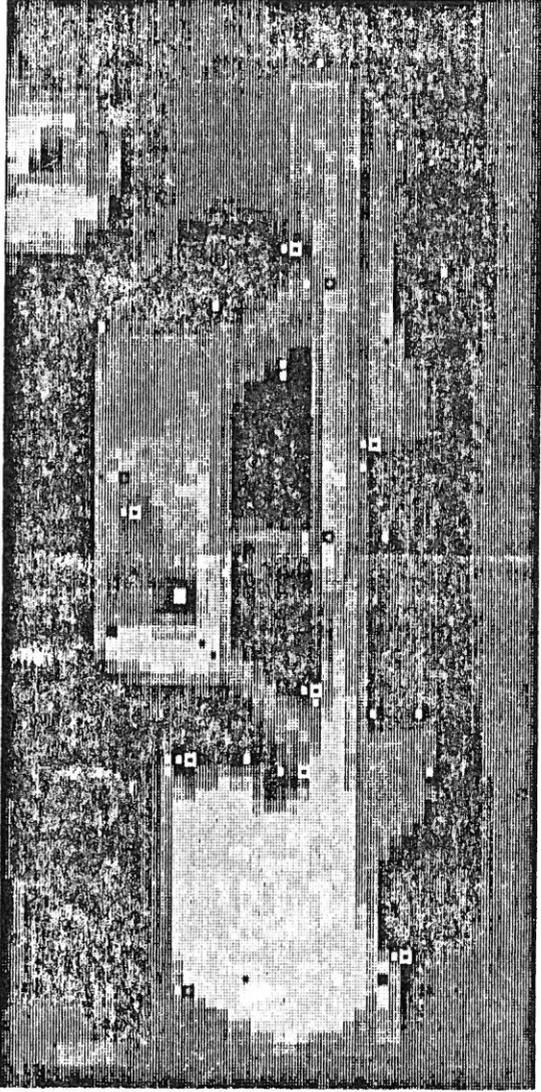


e) wie Abb. d), jedoch Beleuchtung von oben



f) wie Abb. b) und d), jedoch räumliche Orientierung für das letzte Bild der Folge

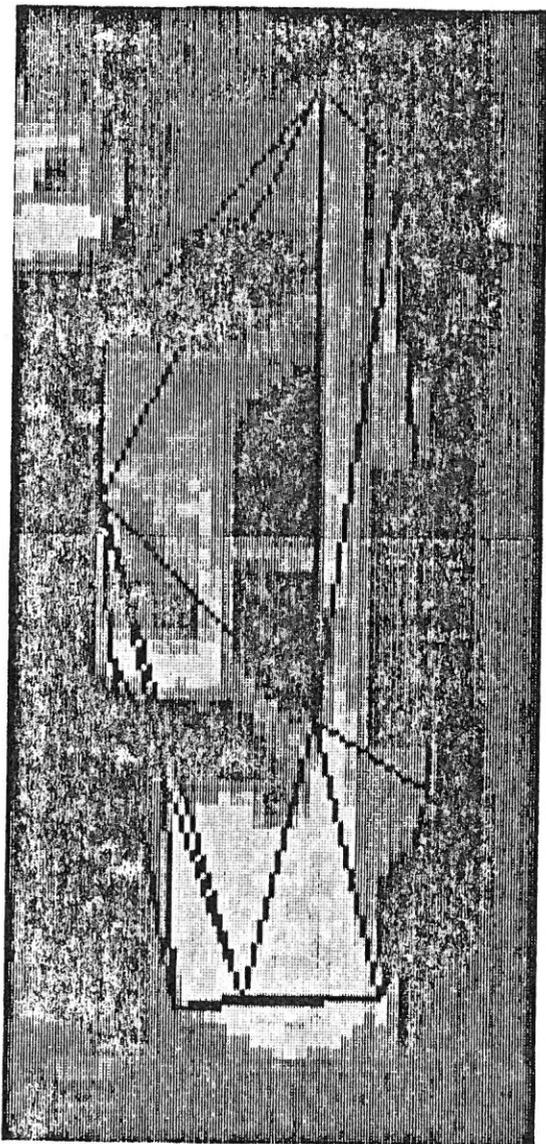
Abb. XI 3D-Rekonstruktion für das Taxi aus Bildfolge 1



d) Genauigkeit der 3D-Rekonstruktion
(siehe Erläuterungen zu Abb. XIa)

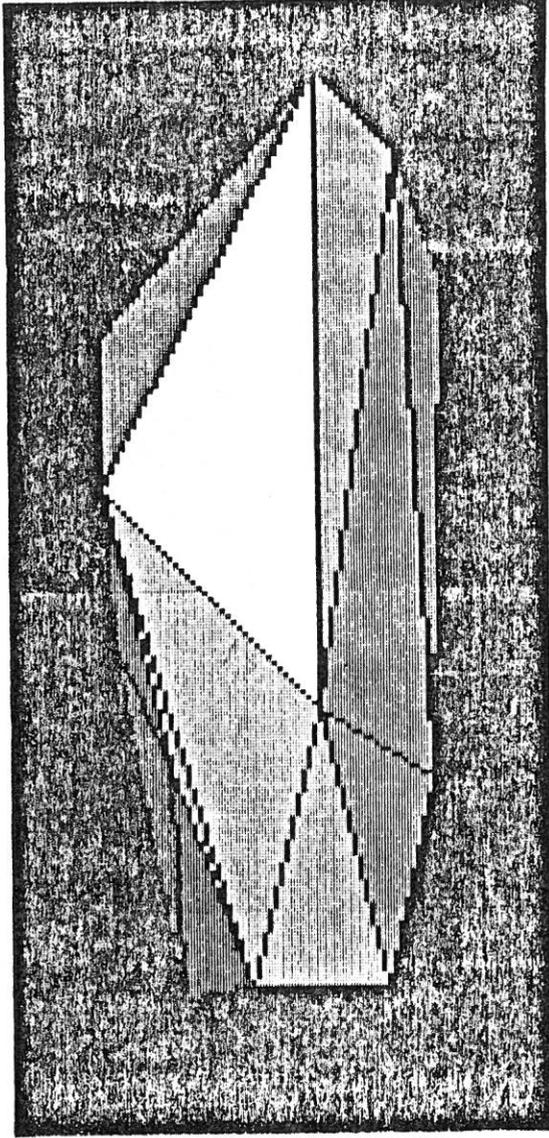


Abb. XII 3D-Rekonstruktion für das Taxi aus Bildfolge 2



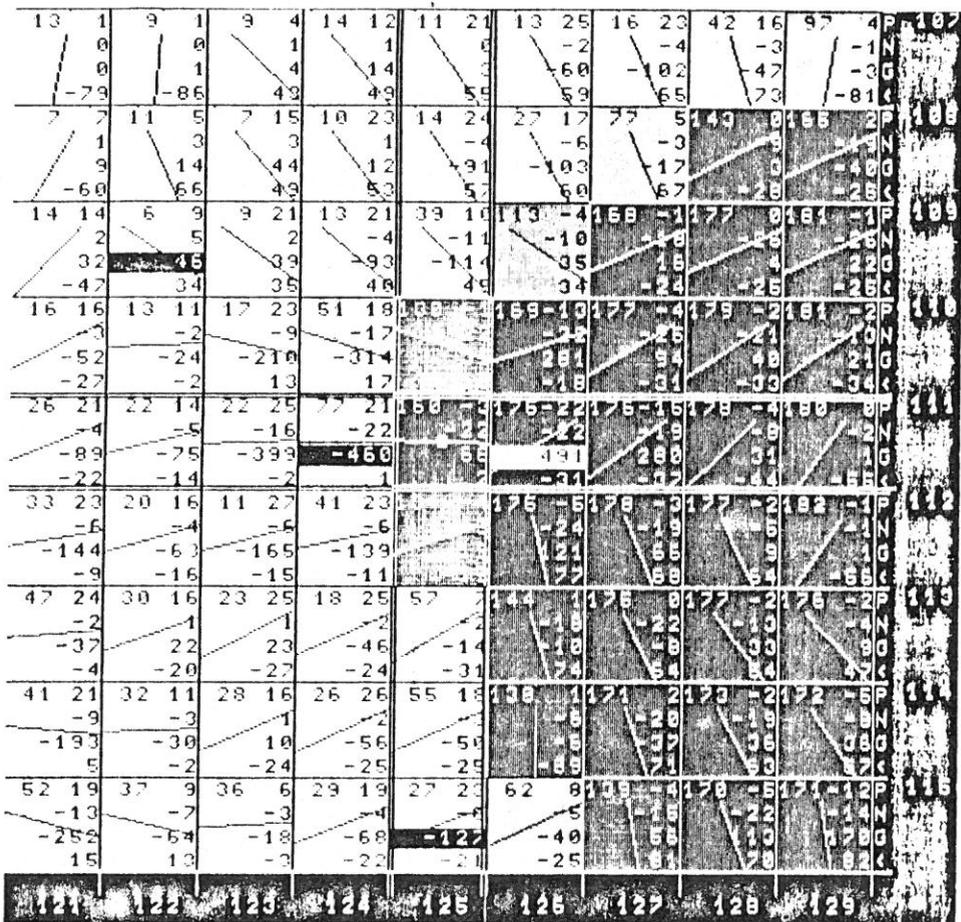
b) sichtbare Linien der konvexen Hülle

Abb. XII 3D-Rekonstruktion für das Taxi aus Bildfolge 2



c) konvexe Hülle schattiert als lambertscher Strichler

Abb. XII 3D-Rekonstruktion für das Taxi aus Bildfolge 2



a) linke obere Ecke des Rückfensters

Abb. XIII Gaußsche Krümmungen in ausgewählten Bildausschnitten

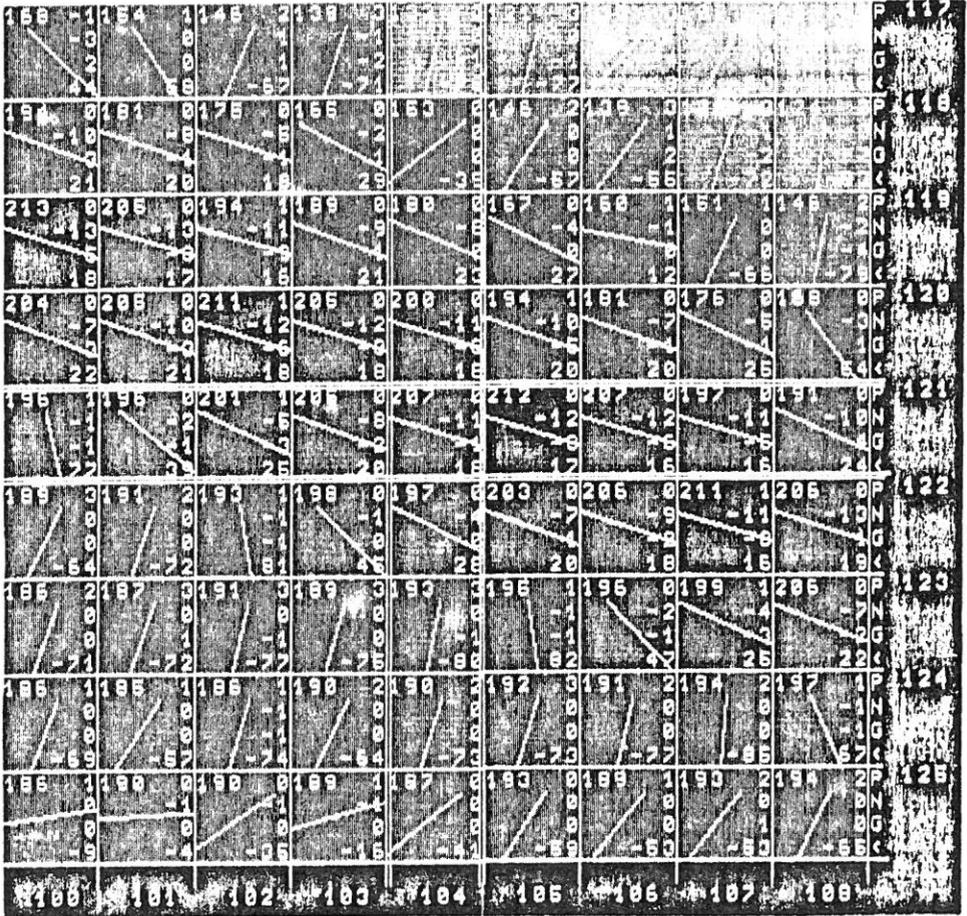
Die Abbildungen stellen 9x9-Ausschnitte aus dem Geobild dar.

Die Richtung der positiven Krümmung ist für jedes Pixel durch ein Geradenstück angegeben, markante Punkte kennzeichnet ein kleines Quadrat.

In jedem Pixel ist links oben der Grauwert angegeben, rechts untereinander positive, negative und Gaußsche Krümmung (bei letztgenannter Extrema in Negativschrift, sofern beide Komponenten die Schwelle überschreiten), sowie der Winkel der Hauptkrümmungsrichtung.

Die Randnumerierung gibt die Zeilen- bzw. Spaltennummern wieder.

Die Beispiele wurden mit einem Operatordurchmesser von 5 gerechnet.



c) Schattengrenze an der linken Wagenunterkante

Abb. XIII Gaußsche Krümmungen in ausgewählten Bildausschnitten

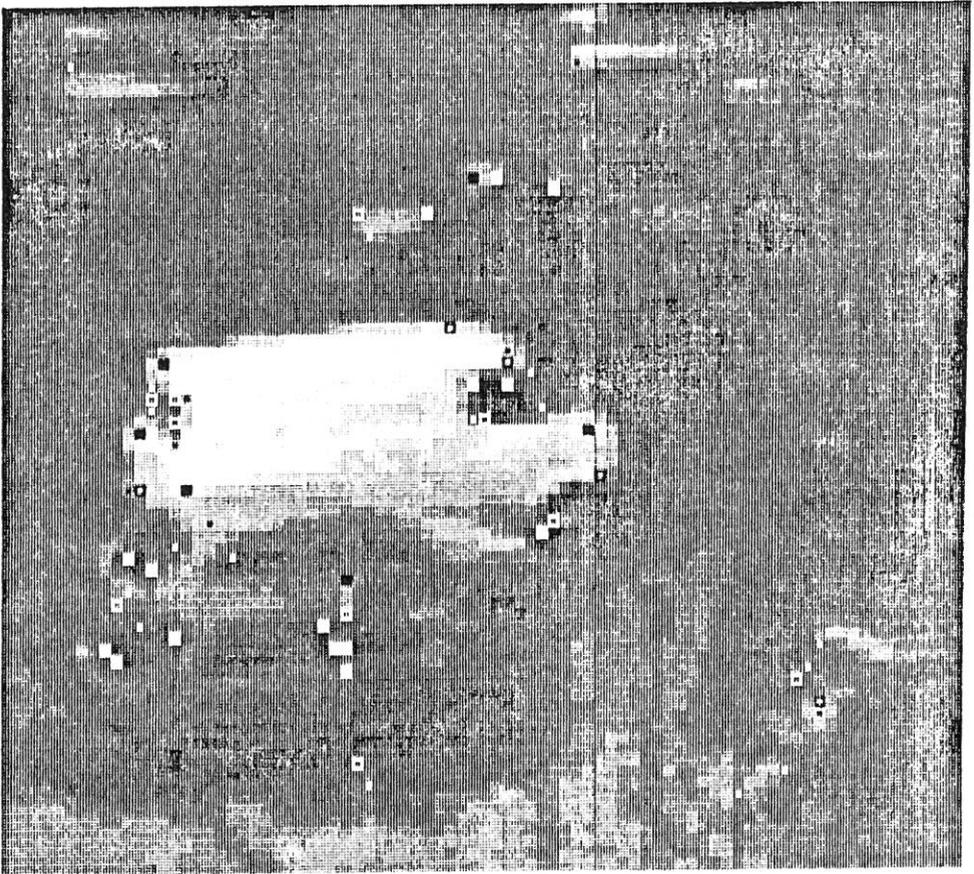


Abb. XIV Messungen zur Reproduzierbarkeit

Punkte markante Punkte aus Bild 1

Quadrate: " " " Bild 2

(zweistufiger Punktelfinder, Bilder Median-gefiltert)

Ich danke allen Mitgliedern, Gästen, Doktoranden und Diplomanden unserer Forschungsgruppe und allen Freunden, die mir nützliche Programme zur Verfügung gestellt oder mit Kritik und Anregungen zur Seite gestanden haben, ganz besonders meinem Betreuer Prof. Dr. H.-H. Nagel, der immer Zeit für klärende Gespräche gefunden hat oder hilfreiche Literaturhinweise geben konnte, wenn unerwartete Probleme auftauchten. Mein aufrichtiger Dank gilt auch den Mitarbeitern des Fachbereichsrechenzentrums, deren Hilfsbereitschaft mir die Arbeit am Fachbereichsrechner sehr erleichtert hat.

Lebenslauf

Geboren am 8. Juni 1952 in Rostock als erstes Kind der Eltern Gerhard Dreschler, Kapitän, und Maria Dreschler, geb. Gorkosch, Schneiderin

Schulen: Von 1959 bis 1963 Besuch der Pestalozzi-Volksschule in Lübeck, danach Besuch des Carl-Jacob-Burckhardt-Gymnasiums in Lübeck, dort

Abitur im Mai 1971.

Zum Wintersemester 71/72 Beginn des Studiums der Informatik an der Universität Hamburg.

Diplomprüfung im Dezember 1977.

Seit Dezember 1977 bin ich als Wissenschaftliche Angestellte am Fachbereich Informatik der Universität Hamburg beschäftigt.

Eidesstattliche Erklärung

Ich versichere hiermit an Eides Statt, diese Arbeit selbständig angefertigt und keine anderen als die angegebenen Hilfsmittel benutzt zu haben.

Hamburg, den 20. März 1981

Leonie Dreschler

