

Methoden und Modelle in der Bildsegmentation
Eine Übersicht

Gunter Hille

IfI-HH-B-80/81

Juli 1981

Zusammenfassung

In dieser Arbeit wird versucht, die bisher auf dem Gebiet der Bildsegmentation vorliegenden Ansätze und Verfahren unter dem Gesichtspunkt der ihnen zugrundeliegenden Modelle zu gliedern. Nicht explizit angegebene Modelle sollen aufgezeigt werden. Die Fortschritte der Bildsegmentation durch Verbesserung von Modellvorstellungen und der benutzten Repräsentationen werden diskutiert. Es wurde die dem Autor zugängliche Literatur bis etwa Ende 1980 berücksichtigt.

Summary

This report tries to survey the articles about the segmentation of pictorial data. The different methods are discussed under common aspects of the involved models. Not explicitly stated models are pointed out. The progress in the segmentation of pictures caused by improved models and used representation of data will be discussed. The survey covers the accessible literature up to the end of 1980.

Inhaltsverzeichnis

1. Einleitung	5
1.1 Definitionen	5
1.2 Segmentation	6
1.3 Modellbildung und Bildfunktion	8
1.4 Die Bedeutung von Modellen in der Bildanalyse	10
1.5 Uniformität als Ausgangspunkt zur Segmentation	12
2. Regionsmodelle	15
2.1 Stetigkeitsannahmen als Hilfsmittel zur Segmentation	15
2.2 Stückweise konstanter Grauwert	18
2.3 Stückweise Normalverteilung des Grauwertes	21
2.4 Stückweise linearer Grauwert	26
2.5 Stückweise Konstanz von Pixeleigenschaften	29
3. Kantenmodelle	35
3.1 Auswertung eindimensionaler Grauwertverläufe zur Kantenbestimmung	38
3.2 Auswertung zweidimensionaler Grauwertverläufe zur Kantenbestimmung	40
3.2.1 Kantenverstärkung	41
3.2.2 Schablonen-Vergleicher	42
3.2.3 Fehlerminimierende Kantenmodelle	44
4. Kooperative Kanten/Regionsverfahren	47
4.1 Exklusive Ansätze	48
4.2 Inklusive Ansätze	50
5. Linienmodelle und Konturmodelle	51
5.1 Linienoperatoren	52
5.2 Stückweise lineare Konturen	54
5.3 Stetige Konturen	58
6. Texturen	61
6.1 Eindimensionale Texturmodelle	62
6.2 Texturmerkmale	64
6.3 Zweidimensionale Texturmodelle	66
6.4 Texturgradient und Texturkanten	68
7. Objektmodelle und Szenenmodelle	70
7.1 Objektmodellierung im Bildbereich	71
7.2 Bildbereichs-Hinweise	71
8. Segmentation in Bildfolgen	79
8.1 Stationäre und nichtstationäre Bildkomponenten	79
9. Zusammenfassung	88
9.1 Von der Bildfunktion zu Bildbereichs-Hinweisen	88
9.2 Von Bildbereichs-Hinweisen zu Szenenbereichs-Hinweisen	89
9.3 Von Szenenbereichs-Hinweisen zum Szenenmodell	90
9.4 Vervollständigung des Deutungszyklusses	91
Autorenverzeichnis	111

Abbildungsverzeichnis

Abbildung 1-1:	Objekte mit gleicher Bildfunktion, Beleuchtung und Betrachtungsrichtung von oben	9
Abbildung 2-1:	Unterschiedliche Segmentationen nach konstantem Grauwert	17
Abbildung 2-2:	Bildfunktion eines Objekts und unimodales Histogramm der Grauwerte	23
Abbildung 3-1:	Ideale Kanten nach (Herskovits and Binford 70)	38
Abbildung 3-2:	Ideale Kanten nach (Shirai 75)	38
Abbildung 3-3:	Kantenschablonen für drei Orientierungen aus : (Holdermann and Kazmierczak 72)	42
Abbildung 3-4:	Kantenmodell mit Übergangsbereich	44
Abbildung 5-1:	Modell der Bildfunktion für ein Linienelement	53
Abbildung 6-1:	Beispiele von Bildmatrix und Cooccurrence-Matrizen für vier Verschiebungsrichtungen (Verschiebungsvektor: 1 Bildpunkt)	63
Abbildung 6-2:	Texturmodell durch Graphen und Grammatik	67
Abbildung 8-1:	Differenzverfahren nach (Jain and Nagel 79)	82
Abbildung 8-2:	Differenzbilder bei verschobenen Kanten	85

Partitioning of an arbitrary intensity image into regions denoting objects is an illusory goal, unless the notion of what constitutes an object is precisely defined. Since objects are often collections of pieces whose association must be learned, general object segmentation seems impossible in principle. It seems more reasonable, on the other hand, to partition an image into regions corresponding to smooth surfaces of uniform reflectance. This is often the implicit goal of programs that attempt to partition an image into regions of uniform intensity. Unfortunately, intensity does not correspond directly to surface characteristics. There is no way of determining whether merging two regions is meaningful, and consequently there is no reliable criterion for terminating the merging process.

(Barrow and Tenenbaum 78)

"Was tun Sie", wurde Herr K. gefragt, "wenn Sie einen Menschen lieben?" "Ich mache einen Entwurf von ihm", sagte Herr K., "und Sorge, daß er ihm ähnlich wird." "Wer? Der Entwurf?" "Nein", sagte Herr K., "Der Mensch."

(B. Brecht, "Geschichten vom Herrn Keuner")

Vorwort

Auf der Suche nach der Entschlüsselung der materiellen Substanzen in unserer Welt haben sich die Chemiker zu einem großen Teil ihrer Forschungstätigkeit mit der Entwicklung von Analysemethoden beschäftigt. Heute lernt ein Chemiestudent schon in den ersten Semestern seines Studiums die Standardmethoden der Analyse. Er ist dann in der Lage, nach wohldefinierten Vorschriften nahezu jede Verbindung in ihre elementaren Bestandteile zu zerlegen.

Während noch vor etwa hundert Jahren die Chemiker bei ihren Analysen durch mehr oder weniger willkürliche Anwendung bekannter oder durch Ausprobieren neuer Methoden ihre Zerlegungen durchführten, werden heute Analysen mit quasi algorithmischen Verfahren ausgeführt; qualitative und quantitative Analyse zählen zu den weitgehend abgeschlossenen Teilgebieten der Chemie. Einen wesentlichen Anteil am Erreichen dieses Erfolges hatten die verbesserten Modellvorstellungen über die Struktur von Verbindungen und genauere Modelle über die Bindungskräfte.

Das Gebiet der Analyse von Bildern mit Hilfe von Rechnern ist mit nur zwanzigjähriger Forschungstätigkeit ein junger Wissenschaftszweig, in dem noch viele offene Probleme auf ihre Lösung warten. Eine grundlegende Aufgabe bei der Analyse von Bildern ist es, im Bild elementare, d.h. in ihren Eigenschaften gleichartige, Gebiete zu finden und deren Eigenschaften für eine Bildinterpretation zu benutzen. Eine Vielzahl von Techniken zur Segmentierung von Bildern ist bisher schon veröffentlicht worden; und ähnlich wie in der Chemie haben sich einige "Standardmethoden" herausgebildet. Trotzdem kann man heute noch nicht einen Satz von Regeln angeben, die für ein beliebiges vorgelegtes Bild eine "zufriedenstellende" Segmentierung liefern

(wobei wir allerdings noch angeben müssen, wann eine Segmentierung als "zufriedenstellend" bezeichnet werden soll).

Auch bei der Segmentierung von Bildern wurden und werden einige Methoden durch "trial and error" gefunden; wenn aber - wie oben behauptet - die größten Erfolge durch Modellbildung erzielt werden, sollten vielleicht auch hier diejenigen Methoden zur Verbesserung von Segmentierungsergebnissen führen, die von guten Modellen der zu interpretierenden Bilder und der abgebildeten Szenen Gebrauch machen.

In der vorliegenden Arbeit wird der Versuch unternommen, die bisher auf dem Gebiet der Bildsegmentierung vorgeschlagenen Methoden daraufhin zu untersuchen, ob sie implizit oder explizit von Modellen ausgehen und wie diese Modelle in den Segmentierungsalgorithmen benutzt werden. In der Übersicht können deshalb nur diejenigen Verfahren berücksichtigt werden, bei denen man aus dem Segmentierungsalgorithmus das ihm zugrundeliegende (oft nur implizite) Modell entnehmen kann. Bei explizit angegebenen Modellvorstellungen soll untersucht werden, wie die Modelle durch Wahl von entsprechenden Algorithmen aufgebaut oder überprüft werden bzw. in gewählten Datenstrukturen repräsentiert werden. Nicht berücksichtigt werden also Verfahren, denen keine Modellvorstellung über die Bildfunktion, das abbildende System oder die Szene zugrunde liegt, d.h. 'ad hoc'-Verfahren oder vom Autor als solche eingeschätzte Verfahren.

Das Ziel dieser Übersicht ist, die bisher verwendeten Segmentierungsverfahren im Hinblick auf die dabei eingesetzten Modelle zu klassifizieren, um so den heutigen Stand der Modellbildung auf dem Gebiet der Segmentierung zu zeigen und damit dem Anwender ein besseres Verständnis der Verfahren zu ermöglichen.

Ein Teil der vorliegenden Übersicht wurde während der Mitarbeit des Autors an einem von der Deutschen Forschungsgemeinschaft geförderten Projekt erarbeitet.

Die Anregung zur Erstellung einer Literaturübersicht unter dem Aspekt der Modellbildung bei der Bildsegmentation gab Prof. H.-H. Nagel. Für seine Unterstützung bei der Vorauswahl der Artikel und für viele fruchtbare Diskussionen möchte ich an dieser Stelle danken. An der Überarbeitung der vorläufigen Fassungen des Manuskriptes haben neben ihm auch B. Radig und H. Westphal hilfreich mitgearbeitet.

1. Einleitung

1.1 Definitionen

Ein Objekt ist ein materieller Gegenstand, der elektromagnetische Strahlung emittieren, absorbieren, streuen oder ablenken kann. Eine Eigenschaft des Objekts ist das Resultat einer Messung einer der Größen Emission, Absorption, Streuung oder Ablenkung für eine definierte Frequenz des elektromagnetischen Spektrums. Zwei Objekte haben gleiche Eigenschaften (sind gleich), wenn sie sich in ihrem Verhalten gegenüber elektromagnetischer Strahlung für alle möglichen Versuchsanordnungen nicht unterscheiden. In diesem Sinne sind Farbung, Transparenz, Absorptionsfaktor, Form, etc. der Objekte verschiedene Objekteigenschaften. Im normalen Sprachgebrauch und auch in den hier betrachteten Anwendungen werden einschränkendere Objekt- und Eigenschaftsbegriffe verwendet. Objekte werden z.B. immer aus Zusammenfassungen von sich in einer 'gewissen' Entfernung befindlichen Anzahl anderer Objekte bestehen und man spricht auch dann von gleichen Eigenschaften, wenn die Meßwerte kleine Differenzen aufweisen oder wenn Meßwerte aus der Messung in einem Frequenzintervall gewonnen wurden.

Unter einer Szene verstehen wir eine beliebige Anordnung von Objekten im dreidimensionalen Raum. Das Bild einer Szene ist das Ergebnis einer Abbildung der Szene auf eine Vektorfunktion $V=V(x,y)$ von zwei Variablen x,y . Die Variablen x und y sind die Bildkoordinaten, das Tupel $(x,y,\text{Vektor der Intensitätsmessungen})$ ist der Bildpunkt oder das Pixel. Die Dimensionalität n des dem Bildpunkt zugeordneten Vektors wird durch das physikalische Sensorsystem bestimmt, die Komponenten des Vektors $V=(v_1,v_2,\dots,v_n)$ sind die Eigenschaften oder Werte des entsprechenden Pixels. Für den Fall $n=1$ wird die Eigenschaft

Grauwert und das skalare Feld Grauwertbild genannt. Wenn der Wertebereich für die Bildkoordinaten und die Eigenschaften nur diskrete Werte aus einem abgeschlossenen Intervall annehmen können, spricht man von einem digitalen Bild oder einer Bildmatrix.

Werden Aussagen über die Szene gemacht, dann befinden wir uns im

- Szenenbereich, wenn z.B. über die räumliche Anordnung der Objekte, deren Reflexionseigenschaften oder den Standort einer Lichtquelle Aussagen gemacht werden.
- Abbildungsbereich, wenn der Sensor beschrieben wird (z.B. sein Standort, Abbildungseigenschaften, physikalische Eigenschaften).
- Bildbereich, wenn auf Bildkoordinaten und Pixelwerte Bezug genommen wird.

Auf diese von Kanade (Kanade 80a) vorgenommene Klassifikation wird in der Übersicht noch mehrfach Bezug genommen werden.

1.2 Segmentation

Sofern nicht anders angegeben, werden im folgenden nur digitale Bilder betrachtet. Auf digitalen Bildern läßt sich eine Nachbarschaftsrelation zwischen Pixeln und damit ein Zusammenhang ("connectedness") zwischen Bildpunkten definieren (Rosenfeld and Kak 76). Jede Teilmenge von zusammenhängenden Bildpunkten des Bildes bezeichnet man als Region. Eine Zerlegung des Bildes in Regionen ist eine Segmentation, wenn sie folgende vier Eigenschaft hat:

1. Die Zerlegung ist vollständig, d.h. jedem Bildpunkt kann eine Region zugeordnet werden.
2. Sie ist eindeutig; der Durchschnitt zweier Regionen ist die leere Menge, ein Bildpunkt kann nicht zu zwei Regionen gehören.
3. Zu jeder Region existiert ein Prädikat, das auf dieser Region erfüllt ist.

4. Die Regionen sind maximal, d.h. bei Hinzunahme weiterer Bildpunkte zu einer Region wird das Prädikat nicht mehr erfüllt.

Mit dieser Definition einer Bildzerlegung (Horowitz and Pavlidis 74, Zucker 76a) lassen sich die zum Teil unterschiedlichen Regionsauffassungen durch verschieden formulierte Prädikate erklären, etwa:

Regionen sind Gebiete im Bild, deren Bildpunkte

- homogenen Grauwert haben (Ernst et al. 76)
- gemeinsame Eigenschaften haben (Zucker 76a)
- Oberflächen in der Szene entsprechen (Kanade 80a).

Die Forderung der Vollständigkeit einer Zerlegung wird von vielen in dieser Übersicht beschriebenen Segmentationsverfahren nicht erfüllt, z.B. von Kantenoperatoren oder Linienoperatoren (vgl. Kap. 3 und Kap. 5). Fischler (Fischler 78) tritt z.B. für eine partielle Zerlegung ein und stützt seine Forderung auf die These, daß eine vollständige Zerlegung auch nicht vom menschlichen bzw. tierischen visuellen System durchgeführt wird:

I claim that efforts towards complete segmentation are misdirected as a technique for general purpose vision, and do not lead to appropriate methods for accomplishing partial scene segmentation...

Auch Mackworth (Mackworth 78) vertritt diese These und schlägt ein "konservatives" Vorgehen (d.h. eine partielle Zerlegung) vor, um Fehler vor der Interpretation der Bilder zu vermeiden. Auf die Problematik der hier vorgestellten Möglichkeit zur Definition einer Segmentation, insbesondere auf die Forderung nach Vollständigkeit und die Problematik der Prädikatsfunktion wird noch einmal im Kapitel 9 eingegangen.

1.3 Modellbildung und Bildfunktion

Durch die Wahl einer bestimmten Prädikatsfunktion oder eines Verfahrens wird eine Zerlegung des Bildes gegenüber anderen möglichen Segmentationen ausgezeichnet. Diese Auswahl wird mit der Absicht durchgeführt, eine für den Betrachter des Bildes bedeutungsvolle Zerlegung des Bildes zu erhalten. Einer solchen zielorientierten Vorgehensweise liegt die Kenntnis einer Klasse von Szenen (etwa: alle möglichen Anordnungen von Polyedern auf einer ebenen Unterlage), Bildern (alle Bilder, deren Grauwertistogramme nur zwei Maxima besitzen) oder Abbildungssystemen zugrunde. Ausgangspunkt ist hierbei immer ein Modell für die Elemente dieser Klasse, die Zugehörigkeit zur Klasse wird durch Vergleich mit dem Modell bestimmt. Die Modellbildung kann stattfinden

- im Szenenbereich als Modell von Objektformen und Objektrelationen
- im Abbildungsbereich durch ein "Kameramodell" oder allgemeines "Abbildungsmodell"
- im Bildbereich durch ein Modell der Bildfunktion, z.B. ein Kantenmodell.

Oft werden Modelle im Bildbereich aus Kenntnissen über Teile des Szenenmodells und Abbildungsmodells gewonnen:

Die ebene Oberfläche eines Körpers mit den Reflexionseigenschaften eines Lambertischen Strahlers erzeugt auf dem Kameratarget an einer durch das Linsensystem definierten Stelle eine der Lichtintensität des abgebildeten Oberflächenelements proportionale Ladungsmenge. Die Bildkoordinaten und Pixelwerte sind die digitalisierten Meßwerte von Position und Ladung auf dem Target. Daher wird ein Oberflächenelement eines Objektes in der Szene auf mehrere benachbarte Bildpunkte gleichen Grauwerts abgebildet.

Da die Abbildung aus dem Szenenbereich in den Bildbereich verschiedene Objekteigenschaften auf einen Intensitätswert abbildet, wird die Umkehrung der Schlußrichtung im allgemeinen

falsch sein (aus zwei Bildkoordinaten und einem Intensitätswert koennen nicht die drei Ortskoordinaten, die Reflektanz und die Neigungen eines Oberflächenelements eindeutig berechnet werden). Benachbarte Bildpunkte mit gleichem Grauwert repräsentieren weder räumlich benachbarte Punkte auf den Objektoberflächen noch Objektflächen mit gleicher Neigung zur Lichtquelle (Horn 77), siehe Abb. 1-1. Die Rückgewinnung der vollständigen Szeneninformation allein aus einem Modell der Bildfunktion ist daher nicht möglich.

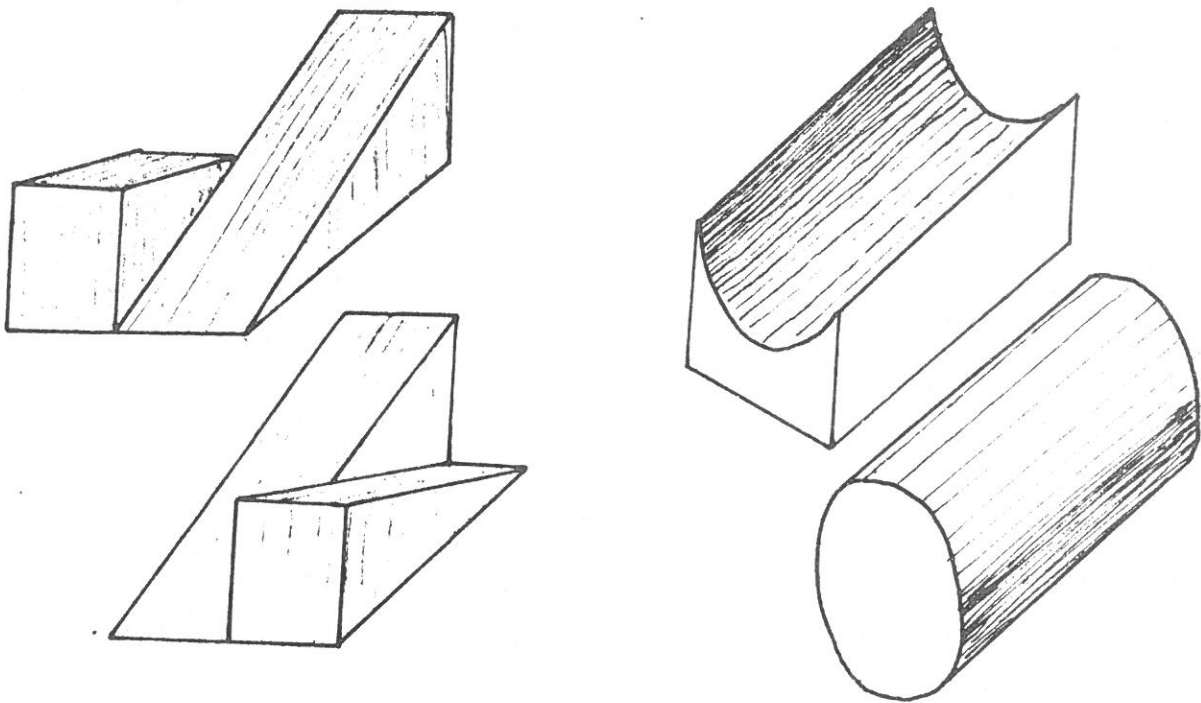


Abbildung 1-1: Objekte mit gleicher Bildfunktion,
Beleuchtung und Betrachtungsrichtung von oben

Die Mehrzahl der Segmentationsverfahren benutzt dennoch nur die Informationen aus dem Bildbereich. Die Bestimmung der Zugehörigkeit von Bildpunkten zu einer Region geschieht in den meisten Fällen durch Angabe von gemeinsamen Eigenschaften der Bildpunkte. Die Konstanz dieser Eigenschaften auf einer Bildpunktmenge entspricht den im Kapitel 1.2 behandelten Prädikaten. Mit einer solchen Betrachtungsweise wird die

Gültigkeit der Umkehr der Schlußrichtung postuliert (aus gleichen Bildpunkteigenschaften folgen gleiche Objekteigenschaften). In den folgenden Abschnitten wird untersucht, ob solche Modelle der Bildfunktion auch Bedeutungen im Szenenbereich haben und welche Kenntnisse (über Objekte, Objekteigenschaften, Relationen zwischen Objekten, etc.) für die Modellbildung benutzt werden.

1.4 Die Bedeutung von Modellen in der Bildanalyse

Bei der Analyse von Bildern sind die Objekte (die reale Welt) im allgemeinen nicht mehr sichtbar. Die Beantwortung von Fragen über das Objekt kann nur durch Analyse des als Bildfunktion vorliegenden Abbildes der realen Welt geschehen oder durch Auswertung des Wissens erfolgen, das der Auswerter über das Objekt hat.

Als Ausgangshypothese verwenden wir den Begriff 'Modell' im Sinne von Minsky (Minsky 77):

Für einen Beobachter B ist ein Gegenstand M ein Modell des Gegenstandes G, insoweit B für die Beantwortung von ihm interessierenden Fragen über G von M Gebrauch machen kann.

Mit einer solchen Betrachtungsweise sind die Bildfunktion und das Wissen des Beobachters über das Objekt verschiedene Modelle für das Objekt, da sie beide zur Beantwortung von Fragen über das Objekt herangezogen werden können. Die Bedeutung von gedanklichen Modellen soll hier jedoch nicht weiter ausgeführt werden, vielmehr wird die Modellierung der realen Welt durch ihr Abbild (die Bildfunktion) näher betrachtet.

In dem von Minsky verwendeten Modellbegriff fehlt ein Hinweis darauf, daß ein Modell in der Regel durch den Beobachter zweckbestimmt geschaffen wird, um die Struktur, die Funktion oder das Verhalten des Originals zu erfassen. Bilder (Bildfunktionen) sind in diesem Sinne aber nicht vom Betrachter hergestellt; sie

sind durch direkte oder indirekte Einwirkung des Objekts auf andere Gegenstände entstanden (etwa bei der Belichtung einer Fotoplatte). Daher ist ein Bild eines Gegenstandes nicht notwendig immer auch ein Modell eines Gegenstandes. Allerdings können Bilder Teile eines Modells sein, wenn der Beobachter diese explizit zur Modellierung einsetzt.

Der Modellbegriff von Minsky muß auch in den möglichen Funktionen des Modells genauer spezifiziert werden. Ein Modell kann eine oder mehrere der aufgeführten Funktionen erfüllen (Klaus and Baur 75):

1. Erkenntnis.

Das Modell dient zur Gewinnung von neuen Informationen über das Original, durch Erkenntnis von Gesetzmäßigkeiten des Modells kann auf Gesetzmäßigkeiten des Originals geschlossen werden.

2. Erklärung.

Aus dem Modell werden Hilfsinformationen gewonnen, die bekannte Eigenschaften des Originals erklären können. Diese Modellfunktion wird hauptsächlich bei der Gewinnung von Bildbereichshinweisen ausgenutzt.

3. Indikation und Demonstration.

Die Modelle werden benutzt, um am Original nicht sichtbare oder meßbare Eigenschaften aufzuzeigen. Ein Modell eines dreidimensionalen Objekts kann etwa verwendet werden, um bei der Bewegung des Objektes auch die Bewegung seines Schwerpunktes darzustellen.

4. Variation und Optimierung.

Es wird versucht, durch Änderung des Modells eine Optimierung des Originals zu erreichen (etwa bei Verfahren der Operationsforschung) oder um durch Vergleich des geänderten Modells mit dem Original ein verbessertes Modell zu erhalten.

5. Verifikation.

Überprüfung einer Hypothese über das Original durch Messungen am Modell.

6. Projektion (Konstruktion).

Gewinnung eines Prototypen des Objektes aus dem Modell. Diese Modellfunktion hat vor allem zusammen mit der Variation und dem Modell-Objekt-Vergleich zur Überprüfung des verwendeten Modells große Bedeutung.

Alle aufgeführten Modellfunktionen werden auch bei Modellierungen in der Bildverarbeitung und Bildanalyse eingesetzt. Die von Kanade vorgeschlagene Methode der zyklischen Bildsegmentation und Bilddeutung entspricht weitgehend dem Vorgehen bei den aufeinander aufbauenden Schritten der "Modellmethode":

1. Auswahl eines Modells entsprechend der gegebenen Aufgabe.
2. Gewinnung von zusätzlichen Informationen über das Modell (aus Schritt 1).
3. Ableitung von Informationen über das Original aus dem Modell (aus Schritt 2).
4. Durchführung der gestellten Aufgabe (Beantwortung von Fragen über das Original) durch Ausnutzung von Informationen aus Schritt 3. Verifizierung der Ergebnisse (durch Projektion) und Gewinnung von Entscheidungen über eine Verbesserung der gewählten Modellvariante durch erneutes Durchlaufen der Schritte 1 bis 4.

Die beschriebenen Modellfunktionen und die vorgestellte Modellmethode haben für alle modellgestützten Methoden Gültigkeit. Man kann deshalb erwarten, daß die richtige Ausnutzung des Wissens über die Modellfunktionen und die Anwendung eines der Modellmethode entsprechenden Verfahrens auch bei der Bildsegmentation zu verbesserten Ergebnissen führen wird.

1.5 Uniformität als Ausgangspunkt zur Segmentation

Wenn der Ausgangspunkt für eine Segmentation die eine Szene repräsentierende Bildfunktion ist, kann man das Auffinden von geeigneten Segmentationsprozeduren als Suche nach der Uniformität von Eigenschaften der Bildfunktion betrachten, die Schlüsse auf gemeinsame Objekteigenschaften im Szenenbereich erlauben. Aus einer lokalen Konstanz des Grauwerts im Bild kann zum Beispiel

auf ebene Objektoberflächen in der Szene geschlossen werden, Ursachen konstanter Grauwertgradienten können gekrümmte Oberflächen oder abnehmende Beleuchtungsintensität sein und Wiederholungen von Grauwerten sind Hinweise auf Flächen gleicher Textur. Obwohl diese Uniformität kein eindeutiger Hinweis auf die Objekteigenschaften ist, wird deren Konstanz dennoch mit Erfolg bei der Bildsegmentation verwendet, da das Zusammentreffen von verschiedenen Objektcharakteristiken, die zu gleichen Eigenschaften der Bildfunktion führen, eine geringere Wahrscheinlichkeit hat als das alleinige Auftreten nur einer Objektcharakteristik, d.h. die in Abb. 1-1 konstruierte Szene tritt in realen Objektkonfigurationen seltener auf als eine Szene mit Objekten unterschiedlicher Färbung oder in anderer Objektanordnung. Abb.1-1 ist ein Beispiel für eine der wenigen möglichen Lagen und Färbungen der Objekte aus der Gesamtheit von Anordnungen und Farben, die zu einem Bild mit gleichem Grauwertverlauf für die zwei Körper führen.

Einige mögliche Interpretationen für die Konstanz von Bildpunkteigenschaften sind in Tabelle 1-1 angegeben.

Die Ausnutzung von Konstanz und Stetigkeit in lokalen Bereichen der Bildfunktion ist eine Anwendung von Wissen über die Konstanz und Stetigkeit in der dreidimensionalen Szenenwelt. Oberflächen sind räumlich stetige Gebiete von Objekten, die Färbung von Flächen ist konstant oder variiert langsam, die Lichtstärke ist eine stetige Funktion des Ortes. Unstetigkeiten sind auf kleine, lokale Bereiche des Raumes beschränkt. Abrupte Helligkeitsänderungen treten an den Rändern von Schlagschatten nur bei nichtdiffuser Beleuchtung auf. Die Belegung des Raumes mit Materie ändert sich in makroskopischen Bereichen nur an den Objekträndern, sie ist konstant an Orten, an denen sich Objekte

I		I		I
I	konstante Größe	I	mögliche Ursache(n)	I
I	der Bildfunktion	I	in der Szene	I
I		I		I
I		I		I
I	Grauwert	I	- ebene Oberfläche	I
I		I	- homogene Reflexions-	I
I		I	eigenschaft	I
I		I	- gleichförmige Beleuchtung	I
I		I		I
I	Grauwertgradient	I	- gekrümmte Oberfläche	I
I	(Betrag)	I	- inhomogene Beleuchtung	I
I		I	einer ebenen Fläche	I
I		I		I
I	Grauwertgradient	I	- linienhafte Objektkanten	I
I	(Richtung)	I	- linienhafte Objektkonturen	I
I		I	- Objektverdeckungen	I
I		I		I
I	Grauwertgruppen	I	- ebene texturierte	I
I		I	Oberflächen	I
I		I	- Gruppen gleicher Objekte	I
I	.	I	.	I
I	etc.	I	etc.	I
I	.	I	.	I

Tabelle 1-1: Uniformität von aus der Bildfunktion abgeleiteten Größen und mögliche Interpretationen in der Szene

befinden. Man vergleiche hierzu den Artikel von Frieden (Frieden 80):

An object is positive. The object to be determined is an energy distribution. Being such, it is necessarily positive at all points, there being no such thing as negative energy.

2. Regionsmodelle

Beim Versuch der Klassifizierung von Segmentationsverfahren wird in den meisten Fällen eine Einteilung in regionsbezogene und kantenbezogene Algorithmen vorgenommen. Die Unterteilung soll auch in dieser Übersicht beibehalten werden, wobei aber innerhalb der beiden Gruppen eine Gliederung nach Umfang, Expliziertheit und Differenziertheit der Modellvorstellungen durchgeführt wird. Auf Ähnlichkeiten in den Modellen von regionalen und kantenorientierten Verfahren wird an den entsprechenden Stellen hingewiesen.

Eine Übersicht von regionalen Segmentierern findet man bei Zucker (Zucker 76a), eine Zusammenstellung von Kantenfindern ist von Davis (Davis 75) vorgenommen worden. Im Kapitel 2.1 werden die Stetigkeitsannahmen diskutiert, die den regionalen Segmentationsverfahren zugrundeliegen. In den Kapiteln 2.2 bis 2.5 werden die Modelle der Bildfunktion und die verwendeten Methoden der regionsorientierten Bilasegmentation beschrieben.

2.1 Stetigkeitsannahmen als Hilfsmittel zur Segmentation

In Binärbildern wird die Intensität eines Bildpunktes mit nur zwei Werten quantisiert. Beispiele solcher Binärbilder sind das Schattentheater, Scherenschnitte und einfarbige Siebarucke. Der auf eine Leinwand geworfene Schatten beim Schattentheater ist die Projektion einer ebenen Objektmaske der aus dünnem Material ausgeschnittenen Figur. Tauschen wir die Figur durch einen undurchsichtigen dreidimensionalen Körper aus, so können wir aus dessen Schattenbild einen Scherenschnitt herstellen. Auf jeden Fall bleiben bei der Betrachtung der Bilder nur Gebiete mit konstanter Helligkeit als Regionen übrig, die uns Rückschlüsse auf ebene oder auch körperhafte Objekte der Szenenwelt erlauben. Die Kenntnis der Abbildung beim Schattenwurf ermöglicht deshalb

ein Modell der Bildfunktion, das Objekten den konstanten Grauwert "schwarz" und dem Hintergrundbereich "weiß" zuordnet.

Diese Überlegungen können - mit gewissen Einschränkungen - auf eine Erhöhung des Wertebereichs bei der Quantisierung verallgemeinert werden, wenn man weiterhin eine Abbildung von in Transmission beleuchteten Objekten betrachtet, wie z.B. bei Bildern lichtmikroskopischer Präparate. Bei gleicher Absorption ist der beobachtete Schwärzungsgrad proportional zur Schichtdicke des Präparats, bei gleicher Schichtdicke proportional der Absorptionsfähigkeit des Präparats. Wenn keine Kenntnis über das vorgelegte Präparat bei der Bildanalyse vorhanden ist, kann nun aber eine Grauwertänderung sowohl von einer Schichtdickenänderung als auch von einer Änderung der Absorptionseigenschaften des Präparats stammen. Wird für zwei benachbarte Pixel gleicher Grauwert gemessen, dann kann dieser sowohl durch die gemeinsame Konstanz beider Objekteigenschaften als auch durch die des Produktes aus Präparatdicke und Absorptionskoeffizient gedeutet werden. Eine Segmentation, die nur einen lokal konstanten Grauwert ausnutzt, zerlegt das Bild dann nur in Regionen mit gleicher Transmissionseigenschaft.

Abb. 2-1 zeigt anhand eines Zahlenbeispiels die Segmentation nach konstantem Grauwert für den angenommenen Zusammenhang zwischen den Objekteigenschaften Präparatdicke D und Färbung F , für die beobachtete Schwärzung gelte $S=D \cdot F$. Im Beispiel a) wird das Bild in alle Regionen mit unterschiedlicher Dicke D richtig zerlegt und auch das Gebiet $F=5$ wird gefunden. Die Interpretation, ob $D=\text{const}$ oder $F=\text{const}$ die charakteristische Eigenschaft für die Einzelregion ist, kann ohne a priori Wissen über die Szene nicht getroffen werden. Im Beispiel b) gibt es eine größere zusammenhängende Region, auf der $D \cdot F=\text{const}$ erfüllt

ist, auf der aber weder Dicke noch Farbung konstant sind. Die Segmentation in Abb. c) erzeugt eine zusätzliche Region, diese trennt die Gebiete mit konstantem Produkt, alle Regionen sind durch die Konstanz einer Eigenschaft beschreibbar. Eine intuitive Auffassung vom Ziel einer Segmentation kann etwa lauten:

Finde diejenigen in der Fläche maximalen Bildregionen, in denen mindestens eine der Eigenschaften "Dicke" oder "Farbung" konstant ist.

Die Beispiele a) und c) würden dieser Definition entsprechen, die Segmentation von b) dagegen nicht.

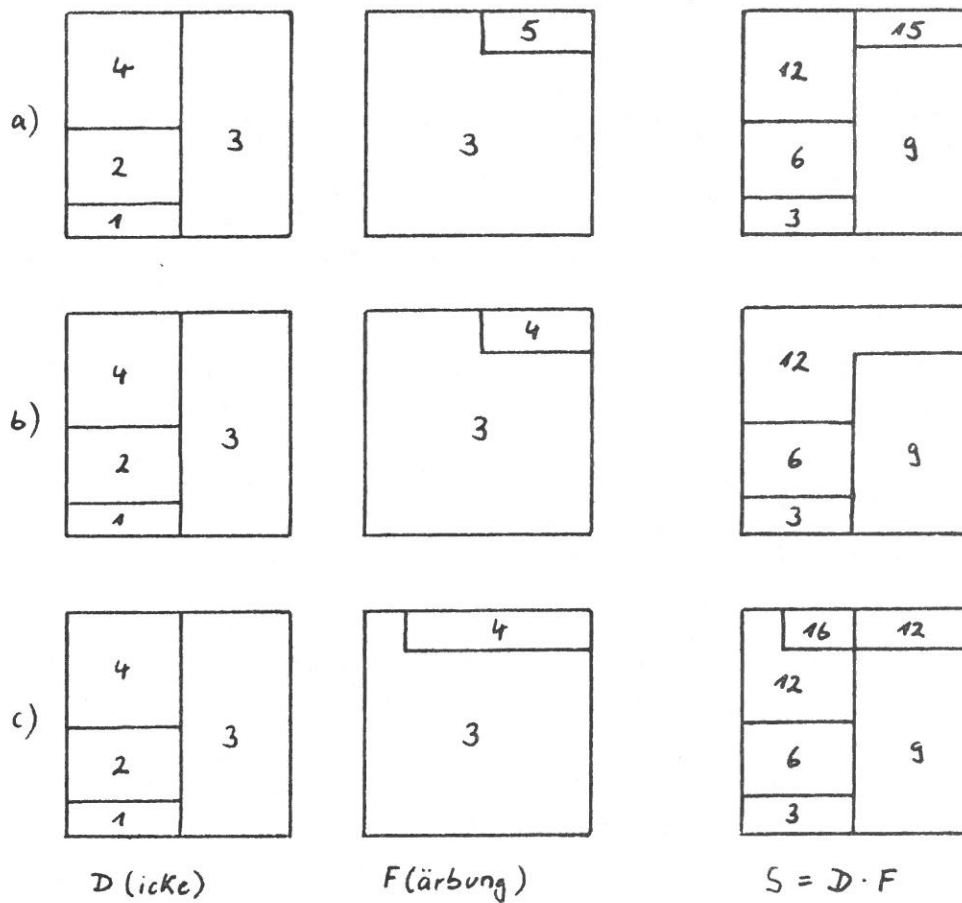


Abbildung 2-1: Unterschiedliche Segmentationen nach konstantem Grauwert

Die bisher gemachten Annahmen über Objekteigenschaften, Art

der abbildenden Funktion und über die Form der Bildfunktion sind leider nur Idealisierungen der in realen Bildern auftretenden Abhängigkeiten. Bei opaken Objekten entspricht jeder Bildpunkt genau einem Oberflächenelement der Szene, aber die in einem Bildpunkt gemessene Intensität ist eine bekannte oder nur teilweise bekannte Funktion verschiedener Größen, wie z.B.

- Entfernung der Objektoberfläche von der Kamera
- Reflektanz der Oberfläche durch unterschiedliche Färbung, Rauigkeit, etc.
- Neigung der Oberfläche relativ zum Sensor
- Ort der Lichtquelle sowie deren Ausdehnung und Strahlungscharakteristik
- Lage anderer Objekte oder Objektflächen relativ zum betrachteten Objekt (Schattierung, Aufhellung)
- Sensoreigenschaften (Linearität, Rauschen, usw.)

Da jeder dieser Faktoren eine zumindest stückweise stetige Funktion des Ortes ist, können Stetigkeitsforderungen für die Intensität der Bildfunktion zur Segmentierung verwendet werden. In den folgenden Kapiteln werden die unterschiedlichen Stetigkeitsannahmen untersucht.

2.2 Stückweise konstanter Grauwert

Das Modell einer stückweise konstanten Bildfunktion geht von der Annahme aus, daß sich die Objekteigenschaft lokal nicht oder nur wenig ändert und daß Eigenschaftsänderungen abrupt erfolgen. In Szenen, in denen als Objekte nur Polyeder auftauchen können ("Blocks-World" Szenen), kann man die Bilder von Polyederflächen und Hintergrund in einer ersten Näherung als Gebiete konstanten Grauwerts auffassen. Änderungen der Bildintensität treten zwischen Objekt und Hintergrund und durch Verdeckung von Objekten oder Verdeckung von Hintergrund durch andere Objekte auf, sie entsprechen räumlichen Diskontinuitäten. Entsprechend wird ein

Sprung im Grauwert durch un stetige Änderung der Oberflächenorientierung an den Polyederkanten hervorgerufen. Zusätzlich können Grauwertänderungen an den Grenzen von Schattengebieten auftreten.

Eines der ersten 'bottom-up'-Verfahren zur Bildsegmentation wurde von Brice und Fennema (Brice and Fennema 70) entwickelt. In einem vorläufigen Segmentations schritt werden alle benachbarten Bildpunkte mit homogenem Grauwert zu Elementargebieten zusammengefaßt. Schon für Bilder einer Polyeder-Szene ergibt sich eine große Anzahl von atomaren Regionen, die nicht den Bildern der gesuchten Oberflächen entsprechen. Wenn die Bedingung des konstanten Grauwertes zu einschränkend ist, ergeben sich zwei mögliche Wege:

1. Lockerung der Forderung nach konstantem Grauwert für Regionen
2. Weiterverschmelzung der Elementarregionen nach anderen Kriterien als der Konstanz des Grauwertes

Brice und Fennema benutzen zwei zusätzliche Verschmelzungskriterien, die von Grauwertdifferenz, Umfang und Form der Regionen abhängen.

Im ersten Fall wird versucht, die physikalischen Eigenschaften des Sensors mitzuerfassen, es wird die realistischere Vorstellung benutzt, daß keine Objekteigenschaft einer Oberfläche der Szenenwelt exakt konstant ist, sondern kleinen lokalen Änderungen unterliegt. Horowitz und Pavlidis (Horowitz and Pavlidis 74) haben bei einer Anwendung des von ihnen entwickelten "Split and Merge"-Verfahrens eine maximale Grauwertdifferenz zwischen Pixeln (und in den weiteren Schritten zwischen zu Regionen zusammengefaßten Bildpunkten) als Entscheidungskriterium für die Segmentation gewählt.

Freuder (Freuder 76) beschreibt einen sequentiellen Verschmelzungsalgorithmus. Dieser arbeitet mit einer "Ähnlichkeits-Funktion", bei der ein Ausdruck

$$\text{Abs}(g(R)-g(R_n))*(A(R)+A(R_n))$$

für die nächste mit der Region R zu verschmelzende Nachbarregion R_n minimal sein soll. Der erste Faktor dieses Ausdrucks begünstigt das Verschmelzen von Regionen annähernd gleichen Grauwerts, der zweite bevorzugt Nachbargebiete kleiner Fläche.

Wenn im Bild Objekte durch unterschiedliche konstante Grauwerte repräsentiert werden, kann die Trennung der Objekte durch eine Schwellwertoperation auf der Bildmatrix erreicht werden. Charakteristisch für solche Verfahren ist die Erstellung einer Häufigkeitsverteilung des Grauwerts. In einem idealisierten Bild, das nur aus einem Objekt und dem kontrastierenden Hintergrund besteht, sind im Histogramm nur die zu Objekt und Hintergrund gehörenden Grauwerte in getrennten Zählern akkumuliert. Jeder Wert zwischen den Extremalwerten kann als Schwellwert zur Trennung des Objekts vom Hintergrund verwendet werden. Eine exakte Berechnung des Histogramms für verschiedene Modelle des Grauwertverlaufs in der Bildmatrix (Kastenfunktion, Keilform, gauß'scher Kegel, usw.) wurde von Wall, Klinger und Castleman durchgeführt (Wall et al. 74), siehe auch Seite 22.

Eine Bildsegmentierung, die durch globale Binarisierung der Bildmatrix erhalten wird, führt für idealisierte Bilder zu den gleichen Ergebnissen wie ein Verschmelzungsalgorithmus, der die Konstanz des Grauwerts in den Regionen fordert. Die Verfahren unterscheiden sich nur in der Vorgehensweise. Für Objekt/Hintergrund-Trennungen sind Realisierungen von Histogrammverfahren mit analogen und digitalen Schaltungen von Ridler und Calvard (Ridler and Calvard 78) erprobt worden. Sie

gehen von dem 'a priori' Wissen aus, daß am Bildrand nur Hintergrund zu finden ist. Eine Schaltfunktion steuert zwei Integratoren, aus deren Signal dann der mittlere Grauwert berechnet und daraus ein Schwellwert abgeleitet wird, der wiederum die Schaltfunktion modifiziert.

2.3 Stückweise Normalverteilung des Grauwertes

Die Annahme, daß Regionen in der Bildfunktion besser durch Gebiete mit einer Streuung der Meßwerte um einen konstanten Erwartungswert (mittleren Grauwert) beschrieben werden, geht von verbesserten Modellen über die physikalischen Eigenschaften der Szene und den nichtidealen Übertragungseigenschaften des abbildenden Systems aus. Die Zugehörigkeit eines Bildpunktes zu einer Region muß aus der Verträglichkeit seines Grauwertes mit einer - im allgemeinen unbekanntem - Grauwertverteilung bestimmt werden. Zur Charakterisierung eines Gebietes können außer dem mittleren Grauwert noch Varianz oder andere statistische Größen herangezogen werden.

Für die Verschmelzungs-Verfahren impliziert eine statistische Vorgehensweise auch eine Vergrößerung der Bildpunktmenge, die zur Parameterberechnung notwendig ist. Muerie und Allen (Muerie and Allen 68) benutzen atomare Bildzellen zwischen 2×2 und 8×8 Bildpunkten zur Berechnung der Wahrscheinlichkeitsverteilung. Zellen werden sequentiell verschmolzen, wobei der Kolmogorov-Smirnov-Test zum Test auf Ähnlichkeit zwischen Zellen dient. Für jedes durch eine Verschmelzungsoperation vergrößerte Gebiet wird die Statistik neu berechnet. Auch Nagao et al. (Nagao et al. 76) klassifizieren landwirtschaftliche Nutzflächen mit einem Test auf Verteilungs-Ähnlichkeit. Der sequentielle Verschmelzungs-Algorithmus von Gupta und Wintz (Gupta and Wintz 75) geht ebenfalls von statistischer Ähnlichkeit zwischen Regionen aus,

die Elementarregionen haben eine Ausdehnung von 2×2 Pixeln.

Für die Auswertung von Luftbildaufnahmen beschreiben Ernst et al. (Ernst et al. 76) einen Algorithmus zur Regionsverschmelzung. Startgebiete werden interaktiv vom Benutzer gewählt, die Verschmelzung erfolgt mit allen benachbarten Pixeln des Gebietes gleichzeitig. Die Autoren machen keine Annahme über die zugrundeliegende Verteilung im Testgebiet, der Algorithmus zerteilt das Bild in Regionen annähernd gleichen Grauwerts.

Eine Erörterung von Problemen, die sich durch die Reihenfolgeabhängigkeit bei sequentiell arbeitenden Verschmelzungsalgorithmen ergeben, ist dem Artikel von Coray et al. (Coray et al. 75) zu entnehmen.

Wall, Klinger und Castleman (Wall et al. 74) haben gezeigt, daß im Falle eines additiven Rauschens das Histogramm des verrauschten Bildes einer Konvolution des ursprünglichen Histogramms mit der Wahrscheinlichkeitsverteilung des Rauschens entspricht. Eine Wahrscheinlichkeitsverteilung von Objekteigenschaften kann ebenfalls die Form der Histogramme beeinflussen. Im Histogramm von Luftbildern werden z.B. Waldgebiete, Wiesen und Acker zu unterschiedlichen Grauerthäufungen mit verschiedener Variation im Grauwert führen (vgl. hierzu die Beschreibung von Texturen durch die Histogrammform im Kap. 6.2).

Histogrammverfahren zur Bildsegmentation gehen von der Annahme aus, daß eine umkehrbar eindeutige Beziehung zwischen Objekteigenschaften und Häufungen im Grauerthistogramm besteht. Die Segmentation des Bildes reduziert sich auf das Auffinden eines (oder mehrerer) Schwellwerte, welche die Häufungen voneinander trennen.

Eine Bestimmung der Parameter der einzelnen Verteilungen in einem Histogramms gelingt aber nur unter den folgenden Voraussetzungen:

1. es werden Annahmen über das Aussehen der Verteilung gemacht;
2. die Mittelwerte sind verschieden;
3. die Streuung der Grauwerte um den mittleren Grauwert ist klein gegen den Abstand der Häufungszentren;
4. die Zahl der zum Merkmal beitragenden Pixel ist genügend groß.

Man beachte, daß hierbei noch immer die stückweise Konstanz des Grauwertes im Bildbereich vorausgesetzt wird. Eine Glockenform der Bildfunktion wie in Abb.2-2a ergibt die in Abb.2-2b gezeigte Histogrammform ohne Häufungspunkte.

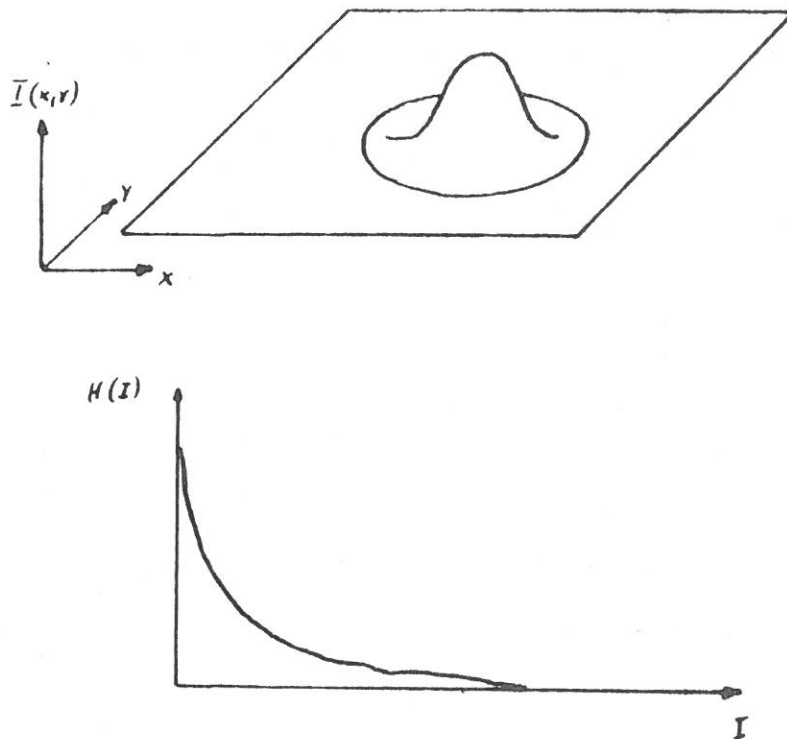


Abbildung 2-2: Bildfunktion eines Objekts und unimodales Histogramm der Grauwerte

In einfachen, nur Objekt und Hintergrund enthaltenden Bildern

gelingt eine Bildsegmentation durch die Wahl einer Schwelle im Minimum zwischen den Objekt und Hintergrund zugeordneten Extrema des Histogramms. Die bekannteste Arbeit mit globaler Schwellwertbildung ist Ohlanders (Ohlander 75, Ohlander et al. 78) Konzept der rekursiven Bildzerlegung. In seinem Verfahren werden die Täler und Gipfel eindimensionaler Histogramme in Bezug auf ihre Form, Höhe und Isoliertheit mit heuristischen Methoden bewertet und der "beste" Häufungspunkt für die Schwellwertbildung gewählt. In den derart segmentierten Bildbereichen werden dann erneut Histogramme erstellt und mit dem gleichen Algorithmus werden die Bereiche rekursiv weiter zerteilt. Da die Intensität als alleiniges Merkmal für die von Ohlander untersuchten Bilder nicht ausreicht, gewinnt er seine Schwellen aus den jeweils besten Häufungen eindimensionaler Histogramme von Merkmalen wie den Intensitäten in drei Farbauszügen, sowie Farbdichte, Kontrast und einer Texturbewertung. Die in dem Verfahren verwendeten Annahmen über die Abbildung von Oberflächen der Szene auf zusammenhängende Bildpunkte sowie die zusätzlichen Annahmen über die Verteilung von Merkmalen in den zu segmentierenden Bildbereichen finden sich in einem Artikel von Shafer (Shafer 80).

Rosenfeld (Rosenfeld and Kak 76, Rosenfeld and Davis 79) zeigt, daß mit einem Modell der Grauwertverteilung im Bild der zu einem minimalen Klassifikationsfehler führende Schwellwert bestimmt werden kann. Mittelwert, Varianz, Objektgröße und Zahl der Objekte sind dabei die das Modell bestimmenden Parameter.

Zur Berechnung der Parameter einer Normalverteilung gehen Weszka und Rosenfeld (Weszka and Rosenfeld 78) von aus dem Histogramm ermittelten Schätzwerten für Mittelwert und Varianz aus und bestimmen einen fehlerminimierenden Schwellwert durch

Anpassung von zwei Normalverteilungskurven an das Histogramm und Wahl des Schwellwertes zwischen den Mittelwerten der Normalverteilungen.

Chow und Kaneko suchen in den von ihnen analysierten Röntgenbildern Teilgebiete mit hoher Varianz des Grauwertes. In den gefundenen Teilgebieten wird ebenfalls versucht, zwei Normalverteilungen an die zugehörigen Grauerthistogramme anzupassen. Nach einer Schwellwertoperation werden dann die Konturen der linken Herzkammer ermittelt (Chow and Kaneko 72).

Um einen dynamischen, über das Bild variierenden Schwellwert zu erhalten, zerlegen Nakagawa und Rosenfeld (Nakagawa and Rosenfeld 79) das Bild in 64 nichtüberlappende Teilbilder, aus deren Histogrammen wiederum Mittelwert und Varianz geschätzt werden. Durch Anpassung von zwei Gaußkurven an die Teilhistogramme wird bei Bimodalität der Schwellwert für dieses Feld errechnet. Die resultierenden 8×8 Histogrammparameter werden geglättet und als dynamische Schwelle für das Bild benutzt. Zur Gewinnung von zwei Schwellwerten werden entsprechend drei Kurven angepaßt, mit Bildern von Objekten auf einem Fließband gelingt auch eine Segmentation der Schattenbereiche. Bei dem letztgenannten Verfahren wird - falls keine eindeutige Bimodalität festzustellen ist - ein Schwellwert auf der Flanke des höchsten Gipfels im Histogramm gewählt. Eine derartige Schwellwertwahl ist nur dann gerechtfertigt, wenn sich tatsächlich zwei Histogramme überlagern, die zu Objekten unterschiedlicher Fläche aber nahezu gleichem Grauwert gehören.

2.4 Stückweise linearer Grauwert

Wenn wir annehmen, daß die Bildfunktion eine stetige Funktion $l(x,y)$ der Bildkoordinaten ist, können wir sie durch eine Taylorreihe in jedem Punkt (x,y) approximieren:

$$l(x+Dx,y+Dy) = l(x,y) + \left(\frac{dl}{dx}\right)*Dx + \left(\frac{dl}{dy}\right)*Dy + \dots$$

Ist die Bildfunktion eine lineare Funktion der Koordinaten, dann kann die Reihenentwicklung nach den ersten partiellen Ableitungen abgebrochen werden, es treten keine höheren Terme auf. Ist dieser lineare Zusammenhang nur in Teilbereichen des Bildes gültig, dann kann das Bild in Gebiete R_n zerlegt werden, in denen in der Umgebung von Punkten (x_i, y_i) gilt

$$l_n(x,y) = a_n(x-x_i) + b_n(y-y_i) + c_n \quad (2.1)$$

mit $\frac{dl_n}{dx} = a_n \quad \frac{dl_n}{dy} = b_n \quad l_n(x_i, y_i) = c_n$

und durch Vergleich mit Gleichung (2.1)

$$l_n(x_i+Dx, y_i+Dy) = a_n Dx + b_n Dy + c_n$$

Ein solches Ebenenmodell der Bildfunktion kann zur Segmentation herangezogen werden, indem in Teilgebieten die Parameter a, b, c berechnet werden und maximale Gebiete gesucht werden, die mit einer ebenen Fläche in der Bildfunktion beschrieben werden können. Die Interpretation von Regionen als geneigte Ebenen in der Bildfunktion geht von einem Szenenmodell aus, bei dem nicht alle physikalischen Eigenschaften auf Objektflächen als konstant angenommen werden. Vielmehr wird angenommen, daß sich jeweils eine physikalische Größe ändert. Dies soll am Beispiel der Lichtintensität näher erläutert werden.

Die von einem beleuchteten Oberflächenelement am Ort r auf

eine Bildkoordinate (x,y) projizierte Lichtintensität ist bei einer beliebigen Beleuchtung der Szene eine Funktion des Ortes r , der Reflektanz R , der Beleuchtungsstärke L und der Neigung N des Oberflächenelements am Ort r :

$$I(x,y) = I(r, R(r), L(r), N(r)) \quad (2.2)$$

Die Intensität in der Umgebung der Bildkoordinaten (x,y) kann wiederum durch eine Taylorreihenentwicklung angenähert werden, wobei nur lineare Terme in einer Bildkoordinate berücksichtigt werden sollen, etwa

$$I(x+Dx,y) = I(x,y) + \left(\frac{dI}{dr} + \frac{dI}{dR} \frac{dR}{dr} + \frac{dI}{dL} \frac{dL}{dr} + \frac{dI}{dN} \frac{dN}{dr} \right) \frac{dr}{dx} Dx \quad (2.3)$$

Aus der Analyse dieses Ausdrucks ist wiederum kein eindeutiger Schluß auf die Objekteigenschaften möglich. Die Faktoren innerhalb des geklammerten Ausdrucks sind ungleich Null, falls sich die Reflektanz, Lichtstärke oder Oberflächenneigung mit der Ortskoordinate ändern. Jeder einzelne der Faktoren und auch deren Linearkombination kann zu einer linearen Variation der Bildfunktion führen. Die Ursache einer linearen Variation in einem Gebiet der Bildfunktion kann eine gekrümmte Objektoberfläche sein. Es kann sich aber genauso gut um eine stetige Änderung der Lichtintensität auf der Oberfläche oder eine gleichförmige Änderung der Reflexionseigenschaft des Objekts handeln. Jede der Objekteigenschaften Oberflächenorientierung, Beleuchtungszustand oder Albedoänderung kann zu einer linearen Variation der Intensität in der Bildfunktion führen.

Erste Versuche zur Modellierung der Bildfunktion durch ein Ebenenmodell unternahm Pavlidis (Pavlidis 72, Pavlidis 73) sowie Holdermann und Kazmierczak (Holdermann and Kazmierczak 72). Pavlidis zerteilte das Bild in eindimensionale Streifen, die

Grauwerte jedes Streifens werden als Funktion von einer Variablen aufgefaßt. Für jeden der Streifen wird die beste Zerteilung in Segmente mit konstanter Ableitung gesucht. Die Koeffizienten der Steigung einer approximierenden Funktion dienen zur Überführung in einen Graphen, an dessen Kanten die Koeffizientendifferenzen notiert sind. Nachfolgend wird eine Verschmelzungsoperation nach minimalen Differenzen im Graphen durchgeführt. Obwohl das Verfahren von Pavlidis die Funktionalapproximation nur in einer Koordinate durchführt, wird durch das nachfolgende Verschmelzen indirekt ein Ebenenmodell der Bildfunktion angewendet.

Holdermann und Kazmierczak bestimmten die Parameter von approximierenden Ebenen in lokalen Bereichen der Bildfunktion mit der Methode der kleinsten Fehlerquadrate. Diese über die von Pavlidis durchgeführte eindimensionale Modellierung hinausgehende 2-D-Modellierung der Bildfunktion wurde zur Kantensuche und Konturverfolgung eingesetzt (Holdermann and Kazmierczak 72).

Zur Bildbeschreibung verwendet Paton (Paton 75) eine lokale Entwicklung in orthonormale Legendre-Polynome und interpretiert die Koeffizienten der Entwicklung als Hinweise auf Täler, Hügel, Ebenen, Sattel, etc. im Grauwertgebirge. Somerville und Mundy (Somerville and Mundy 76) berechnen aus drei benachbarten Bildpunkten die Orientierung des ebenen Dreiecks im $(x,y,\text{Grauwert})$ -Raum, das von diesen drei Pixeln definiert wird. Ebenenelemente mit gleicher Neigung werden zu größeren Ebenenstücken verschmolzen.

Bilder von Polyeder-Szenen, bestehend aus Zylindern und Würfeln werden von Shirai (Shirai 73) segmentiert, in überlappenden Gebieten wird eine Ebene nach der Methode der kleinsten Fehlerquadrate angepaßt und die Varianz der Meßwerte als Maß der "Ebenengüte" benutzt. Das "sloped-facet" Modell von

Haralick (Haralick 80) erlaubt die Berechnung der Ebenenparameter in quadratischen Bildfenstern, in die Beschreibung der Bildfunktion ist ein zusätzlicher Rauschterm aufgenommen worden. Die Hypothese auf Gleichheit der Ebenenparameter in benachbarten Bildfenstern wird mit einem F-Statistik-Test überprüft.

2.5 Stückweise Konstanz von Pixeleigenschaften

Die Hinzunahme weiterer Meßwerte zur Beschreibung eines einzelnen Bildpunktes ist eine Erweiterung des Modells der Bildfunktion in dem Sinne, daß die Bildpunkte mit zusätzlichen Eigenschaften attribuiert werden. Dabei sollen die verwendeten Bildpunkteigenschaften - etwa die Intensitäten der Farbauszüge oder Farbsättigung und Farbart - möglichst genau neuen Objekteigenschaften entsprechen, die erst durch eine Auswertung der spektralen Information bestimmbar werden.

Bei der Ausnutzung der spektralen Information zur Bildsegmentation ergeben sich neue, zum Teil noch nicht vollständig beantwortete Fragestellungen:

- welche spektralen Objekteigenschaften können bei den zur Zeit verwendeten Modellen die Segmentationsergebnisse verbessern;
- wie müssen die Modelle geändert werden, um bei Einbeziehung eines mehrdimensionalen Intensitätsvektors verbesserte Segmentationen zu erhalten;
- welche Transformationen sind auf die Meßwerte anzuwenden, damit das Transformationsergebnis einer Objekteigenschaft entspricht.

Objekttrennung in Schattengebieten und die Bestimmung von gekrümmten, einfarbigen Oberflächen entsprechenden Gebieten der Bildfunktion sind bei einfarbigen Bildvorlagen schwierig durchzuführen. Die Notwendigkeit der Einbeziehung von Farbinformation bei der Segmentation von Bildern wird vielfach mit diesen Schwierigkeiten begründet. Bei konstanter Intensität

im Schattengebiet gelingt keine Objektrennung, da die Farbintensität konstant bleibt und sich vielleicht nur die Farbart ändert. Nutzt man hingegen die zusätzliche Farbinformation aus, so kann ein Wechsel der Farbart Hinweise auf unterschiedliche Objekte im Schattengebiet liefern. Dies kann durch alleinige Analyse der Intensität nicht erreicht werden. Entsprechend werden bei der Segmentation eines Bildes die gekrümmten Oberflächen entsprechenden Teile der Bildfunktion in mehrere Einzelregionen zerteilt, wenn ein Modell der Bildfunktion mit konstantem Grauwert oder auch ein Ebenenmodell der Bildfunktion verwendet wird. Bei Verwendung der Farbart anstelle der Intensität wird die Segmentation nach (annähernd) konstanter Farbart sowohl ebenen als auch gekrümmten Oberflächen entsprechende Bildteile als Elementarregionen ergeben.

Die Komponenten des einem Bildpunkt zugeordneten Vektors werden als Merkmale in einem mehrdimensionalen Merkmalsraum aufgefaßt. Typische Merkmale eines Pixels sind etwa:

- die Intensitäten der drei Farbkanäle einer Fernseh-Kamera
- die Intensitäten der spektralen Kanäle im sichtbaren und unsichtbaren Bereich der elektromagnetischen Strahlung (z.B. bis zu 12 Wellenlängenbereiche in multispektralen Satellitendaten)
- aus Laserlicht-Reflexion ermittelte Intensitäten und Abstände vom Sensor

Der einfachste Ansatz zur Bildsegmentation bei vorliegenden mehrdimensionalen Daten ist, die Konstanz oder zumindest die Ähnlichkeit in allen Merkmalen zu fordern. Obwohl dies eine starke Einschränkung ist, kann sie zur Gewinnung von Elementargebieten zur Bildvorverarbeitung eingesetzt werden.

Nagao et al. (Nagao et al. 79) verwendeten vier Schwellwerte in den Merkmalsdifferenzen zur Verschmelzung von kleinen

Regionen. Das Datenmaterial bestand aus Luftbildaufnahmen im roten, grünen, blauen und infraroten Spektralbereich. Levine und Leemet (Levine and Leemet 76) verwendeten die drei Farbauszüge als Pixeleigenschaften. Startpunkte der Segmentation waren Bildpunkte mit kleiner Grauwertdifferenz zu benachbarten Punkten und verschmolzen wurde jeweils mit den Nachbarbildpunkten mit kleinster Grauwertdifferenz zum betrachteten Aufpunkt. Ähnlichkeitskriterium war der euklidische Abstand im von den Farbkomponenten aufgespannten Merkmalsraum.

Bei der Übertragung von Histogrammverfahren auf mehrdimensionale Bildpunktvektoren wird das Modell der Bildfunktion entsprechend modifiziert. Bei gleichen Voraussetzungen über die Intensitätsverteilung wie im Kapitel 2.3 (stückweise Normalverteilungen) wird die eindimensionale gaußsche Wahrscheinlichkeitsverteilung durch eine multivariate Normalverteilung ersetzt. Die Meßwerte eines Bildpunktes werden durch Punkte oder Vektoren im Merkmalsraum repräsentiert.

Histogrammverfahren, wie das von Ohlander (Ohlander et al. 78) vorgeschlagene rekursive Schwellwertverfahren, gehen von der Voraussetzung aus, daß die Projektion der Punkte auf mindestens eine Merkmalsachse zu einer Häufung im Histogramm führt; durch den Schwellwert wird eine $(n-1)$ -dimensionale Trennebene in den Raum der n Merkmale gelegt, durch die mindestens eine Punktballung von den verbliebenen abgetrennt wird. Nach Abtrennung eines Teilbereichs werden die Histogramme für alle Teilbereiche neu berechnet. Dann wird wiederum in jedem Teilbereich das Merkmal bestimmt, dessen Histogramm die beste Trennung zwischen zwei Maxima ermöglicht. Ohlanders Verfahren geht von der Annahme aus, daß Oberflächen in der Szene durch zusammenhängende Bildpunkte repräsentiert werden und diese

wiederum benachbarten Punkten im Merkmalsraum entsprechen (vgl. Seite 24). Coleman und Andrews schreiben hierzu (Coleman and Andrews 79) :

This model of the feature space when applied to the image segmentation problem, implicitly assumes that numerical difference is directly proportional to perceptual difference, in the human perceptual system. This is an assumption which is almost certainly untrue, at the current state of knowledge about the human perceptual system and the current state of development of features used in digital image pattern-recognition techniques. Nevertheless, the existence of a (almost certainly) nonlinear transformation can be postulated which would map the feature vectors into a new space where the model described previously would be perceptually valid.

Das Vorgehen bei der Klassifikation von Daten wird vom 'a priori' Wissen über die Verteilung geleitet. Die Bestimmung von Trennflächen im Merkmalsraum, z.B. mit einem Bayes- oder Maximum-Likelihood-Klassifikator soll die Klassifikationsfehler einer a priori bekannten Verteilung minimieren. Die Klassifizierung von Datenpunkten durch unüberwachtes Lernen ist ohne Annahmen über die Daten nicht in voller Allgemeinheit lösbar. Um eine Ballungsanalyse durchführen zu können, müssen zumindest einige der auftretenden Probleme durch zusätzliche Annahmen behoben werden:

1. die Zahl der Ballungen im Merkmalsraum ist unbekannt (wenn die Zahl der Objekte der Szene bekannt ist, kann eine Annahme über die maximale Zahl von Ballungen gemacht werden);
2. zwischen den gemessenen Merkmalen bestehen unbekannte Korrelationen, dies kann zu anisotropen Ballungen im Merkmalsraum führen;
3. es gibt kein eindeutiges "bestes" Abstandsmaß im Merkmalsraum, welches einem "wahrnehmbaren" Abstand entspricht;
4. nicht jedes Merkmal liefert den gleichen Beitrag für eine gute Klassifikation.

Die Anwendung von überwachtem Lernen durch Gewinnung von

Trainingsgebieten mit bekannter Klassifikation kann für das Gebiet der Luftbilddauswertung hier nicht erschöpfend behandelt werden, es wird auf die umfangreiche Literatur zur Klassifikation von Luftbildern verwiesen.

Noguchi et al. (Noguchi et al. 78) segmentierten Zellbilder, das Datenmaterial bestand aus Bildern verschiedener spektraler Kanäle. Im Modus des unüberwachten Lernens werden in drei Trainingsgebieten charakteristische Merkmale gewonnen, als weitere Annahmen benutzen sie eine bekannte Normalverteilung im Merkmalsraum und die bekannte Zahl von drei Klassen für Hintergrund, Zytoplasma und Nukleus sowie eine für alle Klassen gleiche a priori Wahrscheinlichkeit. Letztere Annahme läßt es aber fraglich erscheinen, ob nach der Lernphase auch für zu einem späteren Zeitpunkt zu analysierende Zellbilder mit beliebigen Verteilungen der Einzelzellen die Segmentationen zufriedenstellend ausfallen.

Der von Burr und Chien (Burr and Chien 76) benutzte minimale spannende Baum ("Minimal spanning tree") entspricht einer Ballungsanalyse mit der Nächster-Nachbar-Regel ("Nearest-Neighbor-Rule") (siehe (Duda and Hart 75), S.233). Lai und Ehrlich (Lai and Ehrlich 79) segmentieren Bilder mit unvollständig spezifizierten Regionen. Einzelne isolierte Bildpunkte werden mit Marken versehen, die Autoren bestimmen eine stückweise lineare Entscheidungsgrenze, die Punktballungen unterschiedlicher Markierung trennt. Statt der "Nearest-Neighbor"-Methode wird ein iteratives Verschmelzungsverfahren angewendet.

Durch streifenförmige Beleuchtung gewinnen Oshima und Shirai (Oshima and Shirai 79) aus ihren Bilddaten die Ebenenparameter von Oberflächenelementen. In einem zweidimensionalen Merkmalsraum und mit stückweise linearen Entscheidungslinien kann

die Szene grob in ebene, gekrümmte und in der Oberflächenbeschaffenheit unbestimmte Regionen segmentiert werden. Die Analyse einer mit einem Lichtgitter beleuchteten Szene wurde schon vorher von Popplestone und Mitarbeitern (Popplestone et al. 75) sowie von Röcker und Kießling (Roecker and Kiessling 75) beschrieben.

Der Artikel von Coleman und Andrews (Coleman and Andrews 79) geht ausführlich auf die oben genannten Probleme der Segmentation durch Ballungsanalyse ein. Zur Dekorrelation der Merkmale wird eine Karhunen-Loeve-Transformation durchgeführt, ein probabilistisches Abstandsmaß steuert die anschließende Reduktion der Dimensionalität des Merkmalsraums. Zur Bestimmung der Zahl der Ballungen im Merkmalsraum wird zusätzlich die Güte der erhaltenen Punktgruppen bewertet und eine optimale Zahl der Ballungen ermittelt.

Weitere auf der Ballungsanalyse basierende Verfahren werden nicht an dieser Stelle, sondern in den nachfolgenden Kapiteln beschrieben, da die von ihnen benutzten Merkmale von erweiterten Modellvorstellungen ausgehen.

3. Kantenmodelle

Im Kapitel 1.1 sind Regionen definiert worden als eine Menge zusammenhängender Pixel, für die ein gemeinsames Prädikat gültig ist. Dieses Prädikat ist in den meisten Fällen mit der Stetigkeit der Bildfunktion auf den Regionen verbunden. Umgekehrt kann man auch nach den Unstetigkeiten in der Bildfunktion suchen und die gefundenen Unstetigkeitsstellen zur Bildsegmentation verwenden. Wenn die Pixel den Regionen zugeordnet werden, befinden sich die Kantenkoordinaten zwischen den Pixelpositionen der verschiedenen Regionen, jede Region wird durch eine Folge benachbarter Kantenelemente umschlossen.

Entsprechend den bei der Betrachtung der Stetigkeit der Bildfunktion gemachten Aussagen kann eine Kante von verschiedenen Änderungen in der Szene herrühren, wir können also zwischen verschiedenen "Kantentypen" differenzieren:

- **Konturkanten** oder **Verdeckungskanten** treten im Bild an Stellen auf, wo die Oberflächennormalen der Objekte nicht mehr zum Betrachter zeigen - die Objektkante ist eine Konturkante - oder Oberflächen eines Körpers durch andere Flächen (die eigene Oberfläche oder die anderer Objekte) in der Abbildung verdeckt werden (vgl. Barrow und Tenenbaum (Barrow and Tenenbaum '8)).
- **Schattenkanten** findet man an Orten der Szene, wo sich der Beleuchtungszustand von "beleuchtet" in "unbeleuchtet" ändert, also an Konturkanten vom Standpunkt der Lichtquelle aus gesehen.
- **Farbkanten** können solche Gebiete der Szene sein, in denen sich nur eine Farbkomponente deutlich ändert, z.B. die Farbart, während die andere Komponente konstant oder stetig ist.
- **Orientierungskanten** sind Unstetigkeiten in der Oberflächenneigung, sie entsprechen am ehesten unserem umgangssprachlichen Kantenbegriff.
- an **Texturkanten** ändert sich die räumliche Anordnung von gleichartigen Objektgruppen oder Objekteigenschaften.

Durch Analyse des eindimensionalen Verlaufs des Grauwerts in

Polyeder-Szenen konnten Herskovits und Binford (Herskovits and Binford 70) Rückschlüsse auf Orientierung und Verdeckung gewinnen. Shirai (Shirai 73) benutzte ebenfalls drei eindimensionale Grauwertverläufe, die ihm zusätzliche Schlüsse auf die dreidimensionale Form seiner Objekte ermöglichten.

Eine konsequente Ausnutzung der Information über verschiedene Kantentypen haben Barrow und Tenenbaum (Barrow and Tenenbaum 78) mit der Einführung ihrer "intrinsic images" vorgenommen. Aus dem ursprünglichen Bild werden weitere Bildmatrizen abgeleitet, in denen Beleuchtung, Reflektanz, Orientierung, Abstand, usw. notiert werden. Kantenoperatoren auf diesen Teilbildern liefern dann entsprechende Hinweise auf Kontur-, Verdeckungs-, Schatten- oder Orientierungskanten.

Bisher ist noch nicht erwähnt worden, wie man im Bild den Ort einer Kante bestimmt und mit welchen Größen sie beschrieben werden kann. Zur vereinfachten Diskussion benutzen wir das Modell von Regionen als Gebiete konstanten Grauwerts. Nur an den Rändern der Regionen, also an den gesuchten Kantenkoordinaten, ist die Grauwertdifferenz benachbarter Bildpunkte von Null verschieden. Wenn wir den Betrag der Differenz als Maß für die Kantenstärke und die Richtung senkrecht auf die Verbindungslinie der betrachteten Pixel als Kantenrichtung auffassen, haben wir ein einfaches Kantenmodell benutzt. Für ein ideales, unserem Regionsmodell entsprechendes Bild wären die Kantenelemente alle Einträge in der Kantenmatrix mit von Null verschiedener Kantenstärke. Sie können durch eine Schwellwertoperation in eine binäre Kantenmaske überführt werden. Bei Bildern realer Szenen ergeben sich aber folgende Schwierigkeiten:

1. in verrauschten Bildern finden wir Kantenelemente, die nicht zu Objektkonturen gehören;
2. Schwellwertoperationen auf der Kantenmatrix können

schwache Kanten unterdrücken, ausgeprägte Kanten ergeben viele benachbarte Kantenelemente;

5. verschiedene Kantentypen liefern Beiträge in die Matrix der Kantenelemente.

Die Segmentation mit Hilfe von Kanteninformation wird aus diesen Gründen nur durch eine Folge von Operationen durchgeführt, die ausgehend von lokalen Eigenschaften zunehmend globalere Informationen im Bild ausnutzen. Die einzelnen Schritte zur vollständigen Segmentation sind typischerweise:

1. Anwendung eines lokalen Kantenoperators auf der gesamten Bildmatrix.
2. Schwellwertoperation auf der Kantenmatrix (dynamisch oder global).
3. Verdünnung von Kantenelementen.
4. Verbindung von Kantenelementen zu Kantenstücken und Schließen von Lücken zwischen Kantenstücken.
5. Zusammenfassung von Kantenstücken zu geschlossenen Konturen.
6. Bestimmung von Regionseigenschaften innerhalb geschlossener Konturen.

Den Schritten 1. und 4. werden wir besondere Beachtung widmen, da diese stark mit dem Kantenmodell verbunden sind. Während der gewählte Kantenoperator eine Widerspiegelung des Kantenmodells ist, zeigt sich beim Verbinden von Kantenelementen, auf welche Weise Kontextinformation in die Verarbeitung einfließt.

Eine Übersicht über Techniken zur Kantenfindung ist im Artikel von Davis (Davis 75) zu finden, Vergleiche einiger der in den folgenden Abschnitten beschriebenen Verfahren werden in den Artikeln von Abdou und Pratt (Abdou and Pratt 79), Brooks (Brooks 78), Deutsch und Fram (Fram and Deutsch 75, Deutsch and Fram 78) sowie Shaw (Shaw 79) diskutiert.

3.1 Auswertung eindimensionaler Grauwertverläufe zur Kantenbestimmung

Messungen von Sensorsignalen zeigten beim Vorhandensein von Kanten senkrecht zur Abtastrichtung im Bildfenster typische, die Kante charakterisierende Grauwertverläufe. Die ersten Untersuchungen dieser Art unternahmen Herskovits und Binford (Herskovits and Binford 70) und Shirai (Shirai 73), die charakteristischen Intensitätsverläufe senkrecht zur Kantenrichtung veranschaulichen die Abbildungen 3-1 und 3-2.

Diese aus experimentellen Messungen gewonnenen Intensitätsprofile von Kanten wurden von Horn (Horn 77) mit Hilfe von Beleuchtungs- und Objektmodellen auch in ihrer physikalischen Bedeutung erklärt und aus den physikalischen Gesetzen abgeleitet. Ein stufenförmiges Intensitätsprofil ("step edge") ist ein Hinweis auf eine Verdeckungskante, das dachförmige Profil ("roof edge") läßt auf eine konkave Objektkante schließen, während die Hügelform mit großer Wahrscheinlichkeit auf eine korrespondierende konvexe Kante deutet.

Erwartungsgemäß werden Kantenfinder mit diesem einfachen, ungestörten Kantenmodell auch nur bei einfachen Bildvorlagen (Polyeder-Szenen) zum Erfolg führen. Da mit einer Messung des Intensitätsprofils in einer fest vorgegebenen Richtung keine parallel zur Abtastrichtung verlaufenden Kanten detektiert werden können, muß die Berechnung des Profils in mindestens zwei (orthogonalen) Richtungen durchgeführt werden. Pingle und Tenenbaum (Pingle and Tenenbaum 71) beschreiben einen derartigen sequentiellen Kantenfinder, der das Profil der Grauwertdifferenzen auswertet. Der Kantenfinder wurde von den Autoren aufgegeben und durch einen Operator nach Hueckel (siehe Seite 45) ersetzt, da dieser bessere Ergebnisbilder lieferte.

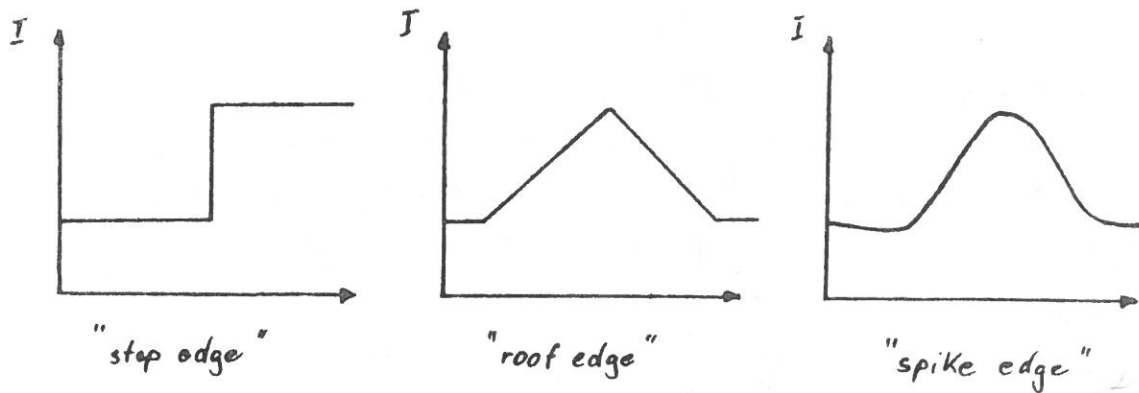


Abbildung 3-1: Ideale Kanten nach (Herskovits and Binford 70)

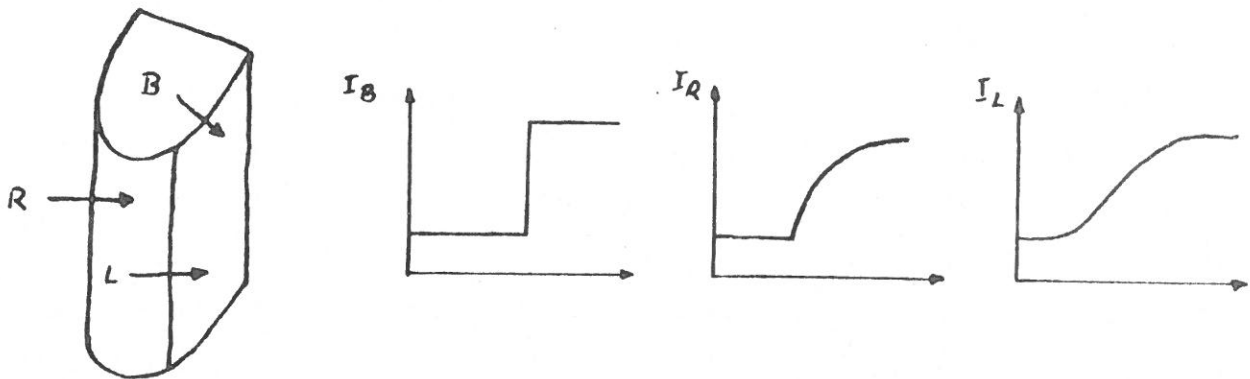


Abbildung 3-2: Ideale Kanten nach (Shirai 75)

Griffith (Griffith 71) konstruiert aus a priori Wissen über die Kantenform sogar sieben verschiedene Intensitätsprofile mit einer linearen Ausdehnung von 15 Pixeln, um für jeden Profiltyp eine optimale Kantenaussage zu treffen. Er notiert aber nur die Werte seines Operators, die durch Konvolution eines Ausschnitts der Bildzeile mit den Masken des Intensitätsprofils gewonnen werden. Eine Unterscheidung nach dem Kantentyp wird allerdings nicht durchgeführt.

Der Berechnung und Interpretation von Intensitätsprofilen wurde nach den Arbeiten von Herskovits und Binford und denen von Shirai eine Zeitlang nicht die angemessene Beachtung geschenkt. In der Folgezeit wurden eine Vielzahl verschiedenster Maskenoperatoren "erfunden" und getestet. Erst als Horn die den Beobachtungen zugrundeliegenden physikalischen Gesetze aufdeckte,

naben sich weitere Autoren mit der Analyse von Intensitätsprofilen beschäftigt. Durch Klassifizierung der Kanten nach dem Intensitätsverlauf auf beiden Seiten einer Kante (konstant oder variierend senkrecht zur Kante, konstant parallel zur Kante) erhalten Barrow und Tenenbaum (Barrow and Tenenbaum 78) Hinweise auf den Kantentyp (Schattenkante, Objekt A verdeckt B, Objekt B verdeckt A) und den Beleuchtungszustand bei Objektverdeckungen (vgl. Seite 36).

Zur Bestimmung der dreidimensionalen Konfiguration seiner "Origami"-Objekte verwendet Kanade eine Markierung von Kreuzungen oder Verbindungen von Linien (Kanade 80b). Er klassifiziert die Verbindungstypen der Linien und verwendet zusätzliche Annahmen über die Oberflächenorientierung zur Konsistenzprüfung der Interpretationen. Eine Einbeziehung des Kantenprofils zur weiteren Einschränkung der Interpretationen soll untersucht werden.

3.2 Auswertung zweidimensionaler Grauwertverläufe zur Kantenbestimmung

Bei der Auswertung des Intensitätsprofils der Bildfunktion in einer Abtastrichtung können der Ort der Kante und der Kantentyp bestimmt werden. Ohne Einbeziehung weiterer Abtastrichtungen können weder die Richtung des Kantenverlaufs noch die Kantenstärke berechnet werden. Zur Berechnung dieser zusätzlichen Parameter muß die Intensität in der Umgebung eines betrachteten Bildpunktes berücksichtigt werden. Bei allen im folgenden aufgeführten Kantendetektoren werden zur Bestimmung von weiteren für die Kante charakteristischen Parametern Bildpunkte aus Fenstern unterschiedlicher Form und Größe herangezogen. Die Vielzahl der bisher veröffentlichten Methoden zur Kantendetektion und des Vergleichs von Kantenoperatoren findet aber keine

Entsprechung in der Zahl der verwendeten Kantenmodelle. Es ist trotzdem überraschend, daß kein Operator existiert, der die im vorigen Kapitel erläuterten Kantentypen zu diskriminieren vermag. Mit wenigen Ausnahmen - diese werden im Kapitel 3.2.3 behandelt - wird das Stufenmodell einer Kante verwendet. Die Beurteilung der Güte eines Kantendetektors wird deshalb in den vergleichenden Artikeln auch meistens nur aufgrund von nicht modellspezifischen Kriterien vorgenommen. Als Vergleichs- und Gütekriterien gelten

- Empfindlichkeit des Operators gegenüber Rauschen
- Richtungsempfindlichkeit
- Vergleich von Operatoren nur anhand spezieller Bildvorlagen
- subjektiver Vergleich.

3.2.1 Kantenverstärkung

Verstärkungsoperatoren bestimmen ein lokales Maß der Kantenstärke durch Konvolution der Bildmatrix mit einem Satz von Operatormasken. Zugrunde liegt die Idee, daß die Differentiation der Bildfunktion genau dort große Werte liefert, wo sich der Grauwert lokal stark ändert. Für jedes Pixel wird durch Konvolution mit entsprechenden Masken ein Satz von Gradientenwerten erhalten, der mit einer nachfolgenden linearen oder nichtlinearen Punktoperation in eine Kantenverstärkungsmatrix überführt wird, aus der Kantenelemente durch Schwellwertbildung extrahiert werden. Diskrete Differentialoperatoren sind z.B. die 2×2 Masken von Roberts (Roberts 65) sowie die 3×3 Masken von Prewitt (Prewitt 70) und Sobel (Duda and Hart 75).

Dem diskreten Differentialoperator liegt das Modell der stückweise konstanten Bildfunktion von Kap. (2.1) zugrunde, mit dem einzigen Unterschied, daß hier der Verlauf der Bildfunktion

an der Kante durch stückweise lineare Grauwertvariationen approximiert wird. Die Operatormasken von Roberts und Prewitt sind identisch mit den Konvolutionsmasken zur Bestimmung der Ebenenparameter nach der Methode der kleinsten Fehlerquadrate. Rosenfeld zeigte, daß der Roberts-Operator als eine Vereinfachung des Operators nach Hueckel aufgefaßt werden kann (Rosenfeld 81). Eine genaue Beschreibung der fehlerminimierenden Methoden findet man bei Brooks (Brooks 78) und Haralick (Haralick 80). Haralick weist auf die Übereinstimmung seiner Kantenmodelle mit den Regionsmodellen hin. Lineare Operatormasken für Taylorentwicklungen bis zur vierten Ordnung und für Maskengrößen bis 8×8 Punkte hat Beaudet (Beaudet 78) errechnet und explizit angegeben.

3.2.2 Schablonen-Vergleicher

Schablonen-Verfahren geben explizit das zugrundeliegende Kantenmodell als $n \times n$ Operatormasken an, wobei für jede zu bestimmende Kantenrichtung auch eine entsprechend orientierte Maske existiert. Für eine Maskengröße von 3×3 Pixeln ergeben sich acht verschiedene Masken ("compass-gradient masks"). Zur Gruppe der Schablonen-Operatoren zählen auch die bei Brooks aufgeführten γ -wertigen und δ -wertigen Schablonen, die Operatormasken von Kirsch (Kirsch 71) und die durch Drehungen entstandenen Prewitt- und Sobel-Operatoren. Robinson (Robinson 77) hat die Operatoren von Prewitt, Sobel, Kirsch sowie die angeführten Schablonen verglichen und beurteilt.

In Abb. 3-5 sind die Bilder der Kantenmodelle für die γ -wertige Schablone in drei möglichen Orientierungen dargestellt.

In der Kantenverstärkungsmatrix wird bei den Schablonenverfahren der maximale Konvolutionswert notiert. Ein Vergleich der γ -wertigen Maske mit dem Prewitt-Operator und der



Abbildung 5-3: Kantenschablonen für drei Orientierungen
aus : (Holdermann and Kazmierczak 72)

5-wertigen Maske mit dem Sobel-Operator zeigt, daß diese identisch sind. Der Unterschied liegt in der durchgeführten Punktoperation, die im einen Fall den mittleren Betrag des Gradienten in orthogonalen Richtungen, beim Schablonenvergleich aber die Richtung und den Betrag des maximalen Gradienten bestimmt. Das Verfahren von Holdermann und Kazmierczak (Holdermann and Kazmierczak 72) überführt das Bild erst in eine Gradientenbetrags-Matrix und konvolviert einen Ausschnitt der Bildmatrix mit acht Masken. Dies entspricht den hier vorgestellten Schablonenverfahren (siehe Kap. 5). Persoon (Persoon 76) beschreibt ein Kantenmodell, das eine lineare Variation des Grauwertes in horizontaler Richtung zuläßt, durch Rotation der erhaltenen 5×5 Maske gewinnt er Schablonen für weitere Richtungen. Die Übereinstimmung zwischen Bild und Kantenmaske wird mit der Methode der kleinsten Fehlerquadrate ermittelt.

Psychophysikalische Forschungsergebnisse waren der Ausgangspunkt für einen von Marr und Hildreth (Marr and Hildreth 80) vorgeschlagene Kantenfinder. Sie gehen davon aus, daß im visuellen System des Menschen Rezeptoren vorhanden sind, die auf Kanten und Linien unterschiedlicher Breite und Orientierung reagieren. Sie simulieren diese Rezeptoren durch Filterung des

Bildes mit verschiedenen Bandpaßfiltern. Auf dem resultierenden geglätteten Bild wird ein zweidimensionaler Laplace-Operator angewendet. In der Matrix der Laplace-transformierten Bildfunktion werden die Nulldurchgänge bestimmt und als Orte hoher Kantenwahrscheinlichkeit interpretiert. Das Verfahren von Marr und Hildreth hat den Vorteil, daß keine weitere Schwellwertoperation notwendig ist. Interessant ist die Hypothese, daß Übereinstimmungen der Nullstellen bei verschiedenen Maskengrößen Indikatoren für eine Intensitätsänderung sind, die durch ein einzelnes physikalisches Phänomen bewirkt wird (Änderung von Reflektanz, Beleuchtung, Abstand oder Orientierung).

3.2.3 Fehlerminimierende Kantenmodelle

Regionsmodelle schließen Kantenbereiche als nicht zum Modell gehörend aus, bisher beschriebene Kantenmodelle sind nur lokal in Kantennähe gültig. Naheliegend ist daher eine Modellierung des gesamten Übergangsbereiches zwischen zwei Regionen, z.B. mit einem Kantenmodell wie in Abb. 3-4.

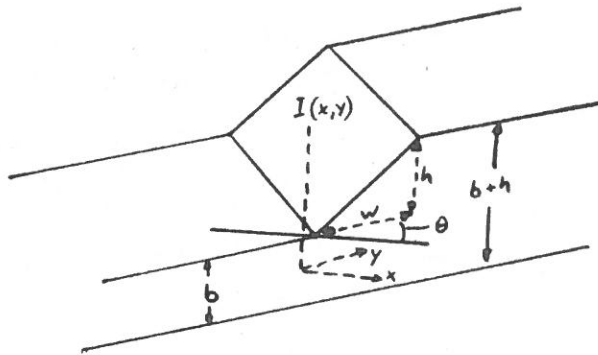


Abbildung 3-4: Kantenmodell mit Übergangsbereich

Die Parameter eines solchen Modells müssen nun aber nicht notwendigerweise die Koeffizienten einer Polynomapproximation sein, sondern können durch Größen wie mittlerer Grauwert, Orientierung, Ausdehnung, Varianz, u.s.w. beschrieben werden. Ein gegebenes Modell läßt sich im allgemeinen durch verschiedene

Parametersätze charakterisieren, wobei die Wahl des Parametersatzes durch die verwendete Methode bestimmt wird.

Hueckels erster Ansatz (Hueckel 71) ging von der Existenz einer stufenförmigen Kante in einer kreisförmigen Umgebung eines gewählten Aufpunktes aus. Die Kante wurde durch fünf Parameter beschrieben, zwei Parameter legten die Intensitäten zu beiden Seiten der Kante fest, drei weitere den Abstand der Kante vom Aufpunkt sowie die Kantenrichtung. Mit Hilfe einer orthonormalen Funktionenbasis und der Forderung nach minimalem Hilbertabstand gibt er eine Näherungslösung mit acht Basisfunktionen an. Als Ergebnis erhält er die beste Schätzung für eine mögliche Kante und ein Maß für die "Güte" der Kante. Mero und Vassy (Mero and Vassy 75) vereinfachten das Verfahren und bestimmten nur zwei Parameter mit quadratischen Operatormasken. Sie verzichteten damit auf die Bestimmung der genauen Lage der Kante innerhalb des Bildfensters.

Nevatia versuchte ebenfalls mit der von Hueckel vorgeschlagenen Methode einen Operator mit weniger Parametern zur Kanten- bzw. Linienextraktion zu entwickeln (Nevatia 77). Er stellte aber fest, daß ein Fortfallen von Termen in der Fourierentwicklung zu schlechteren Operatorantworten führte. Dieser Effekt trat besonders in Bildern mit hohem Rauschanteil auf (solche Bilder wurden von Mero/Vassy nicht getestet). Nevatia vermutet, daß durch eine Erhöhung der Zahl der verwendeten Basisfunktionen auch bei Verwendung anderer Funktionenbasen die Operatoren verbessert werden können.

Methoden zum Auffinden der besten Parameter für komplexe Kantenmodelle werden von Hummel (Hummel 79) angegeben. Mit einer Karhunen-Loeve Transformation des Bildes kann ein Satz orthonormaler Basisfunktionen bestimmt werden, mit denen

Realisationen eines Kantenmodells optimal identifiziert werden können. Hummel benutzte als Fehlerkriterium ebenso wie Hueckel die minimale Norm im Hilbertraum. Der von Hummel vorgeschlagene Lösungsweg kann auch a priori Wissen über die Verteilungsdichte der Kanten - z.B. Bevorzugung horizontaler Kanten - bei der Errechnung optimaler Basisfunktionen berücksichtigen.

Da Konvolutionsoperatoren zur Kantenbestimmung von einer durch die Mitte des Fensters verlaufenden Kante als Voraussetzung ausgehen und aufgrund dieser im allgemeinen nicht zutreffenden Annahme sowohl die Position wie auch die Richtung der Kante fehlerbehaftet sind, beschreiben Iannino und Shapiro (Iannino and Shapiro 79) eine Modifikation des Sobel-Operators. Mit einem iterativen Verfahren werden die Fehler reduziert. Die von Mero und Vassy vernachlässigten Parameter zur Beschreibung der Kante werden hier mit Hilfsverfahren rekonstruiert.

Approximationsverfahren unterscheiden sich von den Schablonenverfahren und Kantenverstärkungsverfahren vor allem darin, daß als Schablone nicht ein explizites Kantenmodell, sondern die Schablonen der Basisfunktionen benutzt werden und die Güte des Vergleichs mit einer wohldefinierten Fehlernorm angegeben werden kann.

4. Kooperative Kanten/Regionsverfahren

Bei vielen der bisher vorgestellten regionalen Verfahren sind implizite Kantenmodelle vorhanden. Es wird aber kein Versuch unternommen, mit dem regionalen Ansatz den Verlauf der Bildfunktion in der Umgebung der Kante zu beschreiben. Entsprechendes gilt für die auf Kantenmodellen basierenden Ansätze, bei denen regionale Aspekte für die Segmentation unberücksichtigt bleiben. Alle Verfahren, bei denen sowohl die Eigenschaften von Kanten als auch die von Regionen zur Bildsegmentation oder zur Gewinnung von Kantenmatrizen herangezogen werden, sollen als kooperative Verfahren bezeichnet werden. In diese Kategorie fallen außerdem solche Ansätze, die andere, zusätzliche Eigenschaften wie Textur oder Oberflächenorientierung in die Segmentation mit einbeziehen.

Die Art des Zusammenwirkens der beteiligten Modelle legt eine weitere Unterteilung der kooperativen Verfahren nahe. Es werde zum Beispiel ein Kantenmodell benutzt, um einen Bildpunkt als Kantenpunkt zu markieren und um ihn dann bei der Bestimmung der Regionen als nicht zur Region gehörig auszuschließen. Entsprechend kann das Modell einer Region verwendet werden, um in einer Kantenmatrix gewisse Punkte als falsch bestimmte Kantenpunkte zu identifizieren. Dieses Vorgehen bezeichnen wir als exklusiven Ansatz.

Wenn eine vollständige Zerlegung des Bildes angestrebt wird, ergeben sich bei dem Modell einer Kante mit einer Breite von mehreren Bildpunkten an den Regionsgrenzen einige nicht den Regionen zuzuordnende Gebiete. Bei der regionalen Bildsegmentation (etwa einem Verschmelzungsalgorithmus) wird bei Ausschließung der Kantenpunkte daher ein neuer Regionstypus ("Kantenregion") eingeführt. Das Auftreten solcher Kantenregionen

konnte an einem von uns implementierten "Split-and-Merge"-Algorithmus verifiziert werden. Ähnliche Beobachtungen werden ausführlicher in einem Artikel von Chen und Pavlidis diskutiert (Chen and Pavlidis 80). Sie bezeichnen nicht sicher bestimmbare Regionen im Kantenbereich als "ambiguous boundary regions".

Inklusive Ansätze lassen sich durch die gleichzeitige Ausnutzung von zwei oder mehreren Modellen zur Erzielung einer Bildsegmentation charakterisieren. Die Hypothese über eine Eigenschaft von Bildpunkten (z.B. Kantenhypothese, Regionshypothese) wird aus einem der beteiligten Modelle postuliert. Dann wird ein zusätzliches Modell zur Unterstützung oder Verwerfung der Hypothese herangezogen.

4.1 Exklusive Ansätze

Obwohl bei Histogrammverfahren nach homogenen Regionen gesucht wird, werden im Histogramm auch die Grauwerte der Kantenelemente mitgezählt. Wenn mit Hilfe von Kantenmodellen die im Übergangsbereich von Regionen liegenden Pixel von der Histogrammbildung ausgeschlossen werden können, erwartet man eine verbesserte Schwellwertwahl und möglicherweise auch exaktere Segmentationsergebnisse.

Durch Erstellen eines Streudiagramms vom mittleren Grauwert im 3×3 Fenster gegen den Grauwert an der betreffenden Pixelposition verbesserten Kirby und Rosenfeld (Kirby and Rosenfeld 79) ihre Ergebnisse durch Histogrammbildung von Pixeln, die ausschließlich in der Nähe der Diagonalen des Streudiagramms liegen. Bildpunkte aus homogenen Gebieten der Bildfunktion liefern Beiträge zu den Diagonalelementen, während Kantenelemente die Zähler in nichtdiagonalen Positionen erhöhen. Allerdings ist dieses Verfahren nicht auf alle Kantentypen anwendbar, es versagt insbesondere bei einem linearen Grauwertverlauf der Kante.

Bildpunkte, die zu Eintragungen in der Diagonalen von Cooccurrence-Matrizen führen, wurden mit dem gleichen Verfahren als Kandidaten homogener Bildregionen ausgewählt. Es ergaben sich aber keine verbesserten Ergebnisse. Die Autoren führen dies darauf zurück, daß Kantenelemente mit gleichem Grauwert in Kantenrichtung ebenfalls als Diagonalelemente der Cooccurrence-Matrix auftraten.

Weitere Untersuchungen bezogen sich auf Streudiagramme von Grauwert und Gradientenwert (Panda and Rosenfeld 78) oder Grauwert und Laplace-Operatorwert (Schachter et al. 79, Weszka and Rosenfeld 79). Die zunehmende Beschäftigung mit expliziten Modellen der Bildfunktion in den letzten Jahren erkennt man deutlich in den Vorbemerkungen von Weszka und Rosenfeld (Weszka and Rosenfeld 79) zu ihrem Verfahren, das zugrundeliegende Modell wird nicht mehr verschwiegen:

We shall assume that the given images consist of objects on a background, where the objects and background each have a unimodal gray-level population. We further assume that the gray-levels of adjacent points interior to the object, or to the background, are highly correlated, while across the edges at which objects and background meet, adjacent points differ significantly in gray level.

Die hier vorgestellten exklusiven Ansätze benutzen alle nur implizit formulierte, nicht durch Parameter beschreibbare Regionsmodelle zum Ausschluß von Bildpunkten aus dem Kantenbild. In der gesichteten Literatur von Kantenoperatoren wurde kein Hinweis darauf gefunden, daß mit Hilfe eines expliziten Regionsmodells Bildpunkte als Kandidaten von Kantenpunkten ausgeschlossen wurden. Vielleicht ließe sich gerade mit dieser Methode das aus einer Bildvorlage mit hohem Rauschanteil gewonnene Kantenbild verbessern. Große Operatorwerte innerhalb eines von einem Regionsmodell identifizierten Bereichs könnten als Operatorantwort auf Rauschanteile der Bildfunktion erkannt

und eliminiert werden.

4.2 Inklusive Ansätze

Einen ersten Versuch zur Vereinigung von Kanten- und Regionsmodellen unternahmen Feldman und Yakimovsky (Feldman and Yakimovsky 74). Ohne hier ihre Forderungen nach semantischer Interpretation zu diskutieren (siehe hierzu Kap. 7.2), soll an dieser Stelle nur auf das kombinierte Modell von Kanten- und Regionsinformation eingegangen werden. Mit Methoden der Entscheidungstheorie wird von den Autoren die Wahrscheinlichkeit einer global besten Interpretation maximiert, indem die Wahrscheinlichkeiten für das richtige Setzen von Kanten, sowie auch die Wahrscheinlichkeiten für das richtige Verschmelzen von Regionen berücksichtigt werden.

Zur Gruppe der kooperativen Verfahren gehören ebenfalls die Versuche, aus der Übereinstimmung der Ergebnisse lokaler Kantenoperatoren und globaleren Verschmelzungs- oder Schwellwertoperationen die Kantensuche zu verbessern ("edge/border coincidence"). Nakagawa und Rosenfeld (Nakagawa and Rosenfeld 78) gewinnen Kantenbilder durch Vergleich von Punkten mit hohem Kantenwert und Regionsgrenzen, die mit variablem Schwellwert bestimmt wurden. Milgram (Milgram 78) benutzt die Übereinstimmung zwischen Kantenstärke und den durch mehrfache Schwellwertoperationen gewonnenen Gebietsrändern. Die Methoden wurden durch Annahmen über die Geschlossenheit von Regionsgrenzen noch verfeinert (Milgram 79, Milgram and Kahl 79).

5. Linienmodelle und Konturmodelle

In vielen Bildern findet man Objekte, die durch eine liniennahe Ausprägung charakterisiert sind, wie z.B. Teilchenspuren in Blaskammerbildern, Straßen und Flüsse in Luftbildern und Adern oder Nerven in Bildern medizinischer Präparate. Dies sind Spezialfälle von Regionen mit großer Ausdehnung in einer Richtung und geringer Ausdehnung senkrecht hierzu. Obwohl die Grenzen solcher linienhafter Objekte mit den beschriebenen Verfahren bestimmt werden können, sind spezielle Linienoperatoren zu ihrer Extraktion aus dem Bild entwickelt worden. Der Grund hierfür ist darin zu sehen, daß man nur Information über den Verlauf der Linien benötigt, aber nicht notwendigerweise deren Breite. Sowohl die Annahmen über den Verlauf der Linien als auch die benutzten Methoden zum Finden linienhafter Objekte unterscheiden sich kaum von Konturfindungs- oder Konturverfolgungsverfahren (bis auf die Forderung nach Geschlossenheit der Konturen), sodaß sie in diesem Abschnitt gemeinsam behandelt werden.

Von Linien im Bildbereich wird man immer dann sprechen, wenn die korrespondierenden Objekte im Szenenbereich als quasi eindimensionale Objekte aufgefaßt werden sollen. Zu ihrer Charakterisierung ist die Kenntnis der zwei- oder dreidimensionalen Ortskoordinaten der Trajektorie eines dem Verlauf des Objektes nachgeführten Raumpunktes ausreichend. Volumetrische Beschreibungen des Objektes sind von untergeordneter Bedeutung (wie breit ist die Straße, wie hoch ist die Böschung?).

Konturen im Bildbereich hingegen entsprechen Rändern von zweidimensionalen (Seen, ebene Oberflächen) oder dreidimensionalen Objekten im Szenenbereich. Daher wird eine

Geschlossenheit für den Verlauf einer Kontur gefordert. Bei linienhaften Objekten ist mit der Bestimmung der Ortskoordinaten das Objekt vollständig beschrieben; Konturen sind nur ein Teil der zur vollständigen Objektmodellierung benötigten Informationen.

Bei Kantenmodellen wird dieses aus der Szene abgeleitete zusätzliche Wissen nicht verwertet. Kantenoperatoren markieren im allgemeinen Punkte, die sich sowohl auf Linien als auch auf Konturen befinden können. Derart erweiterte Annahmen über die Szenenwelt rechtfertigen die hier durchgeführte gesonderte Behandlung der Linien- und Konturmodelle.

5.1 Linienoperatoren

Während maskenorientierte Kantenoperatoren mit Maskengrößen gleicher oder kleinerer Ausdehnung als die Breite der gesuchten Linie zu beiden Seiten der Linie Kantenelemente markieren, sollen Linienoperatoren nur auf der Achse linienhafter Regionen Linienelemente feststellen. Dies wird z.B. durch die entsprechende Form der Masken beim Schablonen-Verfahren von Holdermann und Kazmierczak (Holdermann and Kazmierczak 72) erreicht. Ehrlich (Ehrlich 77) beschreibt eine Maske mit einer Breite von einem und einer typischen Länge von 128 Bildpunkten zum Auffinden stückweise linearer Strukturen in LANDSAT-Bildern. Als Datenmaterial standen Johnston und Rosenfeld (Johnston and Rosenfeld 75) Matrizen von Höhenangaben zur Verfügung. Sie ermittelten hieraus Bergrücken (Schluchten) durch das Aufsuchen aller Punkte, bei denen in ausgewählten Richtungen keine Nachbarpunkte größere (kleinere) Höhenwerte aufwiesen. Hueckel (Hueckel 75) entwickelte seinen Operator auch zum Auffinden von Linienelementen weiter, Frei und Chen (Frei and Chen 77) erweiterten ihre Masken auf ein Linienmodell. Masken

verschiedener Orientierung für Fenstergrößen von 5×5 bis 9×9 Bildpunkten benutzte Paton (Paton 78, Paton 79) zur Detektierung dünner Liniensegmente. Allen vorgestellten Verfahren liegt das Linienmodell einer aus zwei idealen Kanten bestehenden Doppellkante von Abb. 5-1 zugrunde.

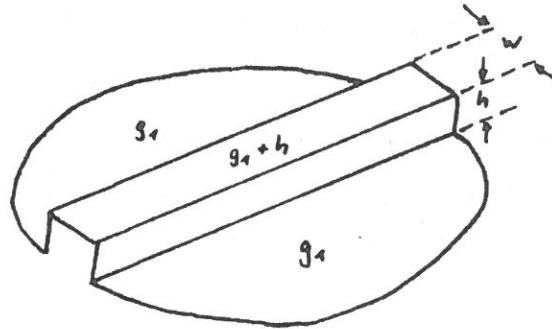


Abbildung 5-1: Modell der Bildfunktion für ein Linienelement

Die Breite der zu bestimmenden Linie wird bei der Wahl der Maske festgelegt. Es treten daher Probleme bei der Detektion von Linien variabler Breite auf. Sind in einem Bild Objekte verschiedener Linienbreite vorhanden, so können diese nur durch eine Maske mittlerer Breite oder durch verschiedene Masken bestimmt werden. Nevatia und Babu (Nevatia and Babu '79, Nevatia and Babu 80) umgehen diesen Nachteil, indem sie eine 5×5 Kantenmaske in sechs Orientierungen über das Bild führen. Die Bestimmung der Kantenrichtung wird (abweichend von den meisten Kantenoperatoren) auf den gesamten Winkelbereich von $0^\circ - 360^\circ$ ausgeweitet. Sie ermitteln antiparallele, benachbarte Kantenpunkte als Kandidaten für Endpunkte der gesuchten Linie. Die Bestimmung von Linien unterschiedlicher Breite ist dann nur von einer geeigneten Wahl der Nachbarschaftsdefinition abhängig. Danker und Rosenfeld ermitteln Kantenbilder mit Hilfe von zwei Kantenoperatoren (Grauwertdifferenz in einem 4×2 -Fenster und ein nichtlinearer Operator in einem 2×7 -Fenster) (Danker and Rosenfeld 80). Mit einem Relaxationsverfahren wird eine

Verstärkung der Operatorwerte von kollinearen wie auch von antiparallelen Kantenpunkten angestrebt. Versuchsergebnisse liegen bisher nur für das Auffinden vertikaler Doppelkanten vor.

5.2 Stückweise lineare Konturen

Kanten- oder Linienoperatoren transformieren das Ausgangsbild in eine Matrix, in der Bildpunkte mit Eigenschaften wie Kantenstärke, Kantenrichtung oder Kantenwahrscheinlichkeiten versehen werden. Für die Weiterverarbeitung werden die Kantenmatrizen in binäre Matrizen überführt, z.B. durch Schwellwertoperationen, Entfernung aller Operatorwerte in der Umgebung von Punkten mit maximaler Kantenstärke (Verdünnungsoperatoren) oder mit "relaxation-labeling" Methoden (Zucker 76b). Eine derart "gewaltsame" Datenreduktion wird mit zwei impliziten Annahmen begründet:

1. der Operatorwert wird als Maß für die Kantenwahrscheinlichkeit interpretiert, die Position der Kante ist demnach der Ort maximaler Wahrscheinlichkeit
2. der maximale Operatorwert wird immer an der Objektkontur angenommen, Operatorwerte senkrecht zur Richtung des Objektrandes können daher vernachlässigt werden.

Die Betrachtung von typischen Kantenbildern nach einer Schwellwertoperation oder anderen Nachbearbeitungen zeigt, daß beide Annahmen im allgemeinen falsch sind. Kantenoperatoren sind empfindlich gegen Rauschen und sprechen auch auf Texturelemente mit Ausdehnungen vergleichbar mit den Maskengrößen an. Auch kann nicht erwartet werden, daß sich jeweils der gleiche Kantenwert entlang einer Objektkontur einstellt.

Zur Rekonstruktion von Objektkonturen aus den Kantenmatrizen muß daher zusätzliche Information verwendet werden. Diese Schlüsse werden durch a priori Wissen über die Fortsetzbarkeit

von Kanten oder Linien gewonnen. In Bildern von Polyeder-Szenen werden z.B. geradlinige Kanten auf kollineare Punkte in der Bildmatrix abgebildet, die Silhouette kreisförmiger Objekte auf Ellipsen.

Die Suche nach Kantenelementen, die sich auf einer gemeinsamen Geraden befinden, ist somit ein erster Schritt zur Modellierung einer Kontur von Objekten.

Globale Verfahren zur Bestimmung von stückweise linearen Konturen werden bei Anwendungen benutzt, in denen a priori Wissen über die Szene vorhanden ist, etwa bei Polyeder-Bildern oder bei der Extraktion von Straßen aus Luftbilddaufnahmen.

Das bekannteste globale Verfahren ist die Hough-Transformation (Duda and Hart 72, Duda and Hart 73). Die Transformationsmethode wurde von Shapiro aufgegriffen und auch auf die Bestimmung nicht geradliniger Konturen erweitert (Shapiro 75, Shapiro 78). Die Erweiterungen werden auf Seite 60 beschrieben.

Bei der Hough-Transformation wird eine Gerade durch ihren Abstand (ρ) von einem gewählten Ursprung und den Winkel (θ) zwischen ihrer Normalen und einer Koordinatenachse parametrisiert. Alle Punkte einer bestimmten Geraden werden dann durch ein Paar (ρ, θ) charakterisiert. In einem zweidimensionalen Zählerfeld von quantisierten Werten (ρ, θ) werden für jeden Kantenpunkt alle möglichen Zähler der durch den Punkt führenden Geraden erhöht. Die auf einer gemeinsamen Geraden liegenden Kantenelemente ergeben dann große Zählerwerte für das die Gerade beschreibende (ρ, θ) -Paar. Da der Wert der Histogramm-Einträge proportional zur Anzahl der auf einer Geraden liegenden Punkte ist, bestimmt man die Geraden durch Maximumsuche oder Schwellwertbildung im Histogramm.

O'Gorman und Clowes (O'Gorman and Clowes 75, O'Gorman and Clowes 76) gewannen ihre Kantenelemente durch einfache Gradientenoperatoren, zur Extraktion von geraden Linien aus dem Hough-transformierten Kantenbild wurden nur Maxima in den Projektionen des Histogramms auf die rho- und theta-Achse bestimmt.

Ebenso wie O'Gorman und Clowes beschränkten sich Perkins und Binford (Perkins and Binford 73) auf die Analyse von Polyeder-Szenen. Kantenrichtung und Kantenposition wurden von einem Hueckeloperator mit einer Maskengröße von 26 Bildpunkten errechnet. Im Gegensatz zu O'Gorman und Clowes wurde aber die Richtungsinformation des Operators nicht verworfen, sondern sie diente zur Berechnung der Normalenparametrisierung der Geraden. Pro Kantenelement wurde daher nur eine Eintragung in das (rho,theta)-Histogramm gemacht. Mit einer Suche nach Gruppen von Kantenpunkten als Funktion des Winkels und für jeden Winkel als Funktion des Abstandsparameters (rho) werden Punktballungen bestimmt und die zugehörigen Geradenparameter für die anschließende Suche nach Eckpunkten im Bild extrahiert.

Anwendungen der Hough-Transformation auf weniger spezielle Bildvorlagen beschreiben Bullock (Bullock 76) sowie Dudani und Luk (Dudani and Luk 77). Bullock transformiert die Ergebnisse eines auf das Bild angewendeten Hueckel-Operators in den Parameterraum, die nach Maximumsuche ermittelten Geradenparameter stützen seine Suche nach Kantenelementen entlang der vorhergesagten Trajektorie. Das Verfahren von Dudani und Luk reduziert die mit dem Hueckel-Operator bestimmten Kantenelemente mit zwei Heuristiken: akzeptiert werden nur zentrale Kantenelemente (Unterdrückung nichtmaximaler Kanten) sowie solche Kantenelemente, in deren Nachbarschaft - innerhalb eines Fensters

von 1×1 Pixel in Kantenrichtung - mindestens n Kantenelemente mit Abweichungen von weniger als 15° zur jeweils betrachteten Kantenrichtung liegen. Das hierbei zugrundeliegende Kanten- bzw. Linienmodell nutzt lokale Eigenschaften - die geradlinige Fortsetzbarkeit von Kantenelementen - zur Datenreduktion aus. Der Vorauswahl folgt dann eine Hough-Transformation und - da den Autoren eine zweidimensionale Ballungsanalyse zu aufwendig erscheint - eine Gruppierung nach den Geradenparametern θ und ρ , sowie den Koordinaten der Kantenelemente. Die Auswertung der Histogramme von Kantenrichtung und Abstandparameter liefert ihnen die Parameter der an die Konturen anzupassenden Geraden.

Lokale Algorithmen werden hauptsächlich bei Bildern angewendet, bei denen auf a priori Wissen über Kantenverläufe verzichtet wird (oder nicht vorhanden ist) und der Aspekt der Modellierung der Kanten durch Fortsetzung bzw. Zusammenfassung von Kantenelementen zu Geradenstücken ("edge linking") im Vordergrund steht.

Ramer (Ramer '75) faßt Kantenelemente zu Strichen zusammen, eine lokale Kostenfunktion steuert die Verschmelzung von Kantenelementen. Kantenstärke, Abstand der Kantenelemente und Winkeldifferenz der Kantenrichtungen gehen als Proportionalitätsfaktoren in die Kostenfunktion ein. Aus Bildern von Objekten mit ebenen oder gekrümmten Oberflächen soll mit seinem Algorithmus eine Strichzeichnung der Szene gewonnen werden. Da die von Ramer gewählte Kostenfunktion auch Abstand und Winkeldifferenz berücksichtigt, enthält sein Verfahren schon ein implizites Modell einer stetigen Konturkante.

Nevatia (Nevatia '76) gruppiert die Kantenelemente nach wachsender Kantenrichtung. Die Verbindung von Kantenstücken wird durch mehrere Schwellwerte gesteuert, die neben Kantenrichtung

und Abstand der Kantenelemente auch die Länge der resultierenden Geradenstücke berücksichtigen.

Obwohl naheliegend, wird von wenigen Autoren außer der Kantenstärke und Kantenrichtung auch der Grauwertverlauf benachbarter Pixel in Kantenrichtung zur Modellbildung herangezogen, etwa zur Unterstützung der Hypothese über die Fortsetzbarkeit einer Konturkante. Griffith (Griffith 71) macht zwar explizite Annahmen über den konstanten Verlauf des Grauwerts zu beiden Seiten einer Kante bzw. Linie. Er benutzt als Operator aber nur eine lineare Maske von 1×16 Bildpunkten, mit der er dann nur den zentralen Grauwert berücksichtigen kann. Eine lineare Variation parallel zur Kante erlaubt der von Marr und Hildreth (Marr and Hildreth 80) beschriebenen Operator.

5.3 Stetige Konturen

Die Verbindung von Kantenelementen zu einer stetigen Konturkurve ist eine Erweiterung des bisher behandelten Modells. Im allgemeinen Fall, d.h. bei unbekannter Form der Objekte, kann keine a priori Parametrisierung der Kontur angegeben werden. Es kann nur die stetige Fortsetzbarkeit von Kantenelementen zur Kontursuche als Hypothese benutzt werden.

Das von Martelli als Kantendetektor beschriebene heuristische Suchverfahren genort nach der hier vorgenommenen Einteilung zur Gruppe der Konturdetektoren oder Liniendetektoren (Martelli 72). Die Kantensuche wird als Suchproblem formuliert. Es soll ein Weg durch einen gerichteten Graphen mit minimaler Kostenfunktion gefunden werden. Gegeben sind ein Startknoten und n Zielknoten. Die Kosten sind eine Funktion des gewählten Weges und der Intensitäten links und rechts des Kantenelementes. Ähnliche Ansätze findet man auch bei Cooper und bei Elliot (Cooper 79, Elliot et al. 79). In beiden angegebenen Arbeiten wird eine

Regionskontur als Markov-Prozeß einer Sequenz von Kanten modelliert. Die Kantenkette hat eine variable Länge; die Wahrscheinlichkeit $p(l)$ für das Auftreten einer Kontur der Länge l ist unbekannt. Benutzt werden einfache Modelle für die Bildfunktion: Regionen haben konstanten Grauwert und sind von einem Hintergrund mit ebenfalls konstantem Grauwert umgeben. Das Bild darf Rauschen enthalten.

Ein mehrstufiges Verfahren zur Extraktion von Konturen wird von Prager (Prager 80) vorgestellt. Er unterscheidet vier Typen von Kantenfortsetzungen und ordnet allen Kantenelementen zwei Konfidenzmaße entsprechend dem Fortsetzungstyp ihrer Enden zu. Mit einem Relaxationsverfahren werden die Wahrscheinlichkeiten für die Existenz eines Kantenelements zwischen zwei Bildpunkten verändert, wobei Nachbarkanten die Wahrscheinlichkeiten verstärken oder bei Unvertraglichkeit von Kantenelementen die Wahrscheinlichkeiten verringern können. In einem nachfolgenden Prozeß werden Kantensegmente an solchen Stellen aufgebrochen, an denen mehrere Kantenelemente zusammenstoßen. Für die verbleibenden Kantenstücke werden Merkmale wie Länge, Abstand der Endpunkte sowie Mittelwert und Varianz des Grauwertes zu beiden Seiten der Kante berechnet. Ein Konfidenzmaß wertet die normalisierte Länge und den Kontrast des Kantensegments aus, ein zweites Konfidenzmaß ist eine Erweiterung des statistischen "Maximum-Likelihood"-Verfahrens von Yakimovsky (Yakimovsky '75, Yakimovsky '76).

Zur Extraktion linienhafter Objekte aus Luftbildaufnahmen beschreibt Groch die Linien durch ihr Grauwertprofil (Groch 80). Mit einer Suchstrategie werden Ansatzstellen (deutlich erkennbare Linienstücke) zur Linienextraktion bestimmt. An diesen Teststellen wird ein objektspezifischer Liniendetektor

eingesetzt. Die Ermittlung von Linienfortsetzungen wird von einem hierarchisch aufgebauten Entscheidungsalgorithmus gesteuert.

Die Modellierung der Kontur als bekannte stetige Ortsfunktion bzw. das Auffinden von Konturstücken mit bekannter Form kann mit einer Erweiterung der Hough-Transformation auf Kurven höherer Ordnung erfolgen. Shapiro (Shapiro 75) hat eine Generalisierung dieser Methoden auf die Detektion von Kreisen und beliebigen Kurven höherer Ordnung beschrieben. Bei Shapiro findet man auch Verweise auf weitere Arbeiten, etwa die Detektion von Kreissegmenten bei der Auswertung von Blasenkammeraufnahmen (Bazin and Benoit 65, Bastien and Dunn 71) und auf medizinische Anwendungen. Aus Röntgenbildern lassen sich z.B. Tumore als kreisförmige Konturen extrahieren (Kimme et al. 75). Die Projektionen der Rippen haben im Röntgenbild einen parabolischen Konturverlauf, der ebenfalls mit einer entsprechenden Hough-Transformation in einem geeigneten Koordinatensystem ermittelt werden kann (Wechsler and Sklansky 75). Über eine Anwendung der Hough-Transformation auf beliebige, mit analytischen Funktionen beschreibbare Konturen berichten Sloan und Ballard (Sloan and Ballard 80).

6. Texturen

In Kap. 2.5 wurde gezeigt, daß die Hinzunahme weiterer Meßwerte zur Beschreibung eines Bildpunktes - etwa die Intensitätswerte aus verschiedenen spektralen Kanälen - in vielen Fällen eine verbesserte Segmentation des Bildes ermöglicht. Zum Bestimmen texturierter Regionen wird nun nicht die Anzahl der einem Pixel zugeordneten Meßwerte erhöht, sondern die Kardinalität der Trägermenge, auf der die lokalen Textureigenschaften bestimmt werden, wobei die Grauwerte der Bildpunkte als die eigentlichen Meßwerte der "Texturprimitive" aufgefaßt werden.

Eine exakte Definition einer Textur ist ebenso wie die Definition einer Region schwierig - wenn nicht sogar unmöglich - sodaß wir hier nur zwei Definitionsversuche anführen wollen:

A region in an image has a constant texture if a set of local characteristics or other local properties of the picture function are constant, slowly varying, or approximately periodic. One may view a textured object as a d-dimensional structure constructed by a machine, or an algorithm whose major parameters remain approximately constant or vary slowly throughout the process of construction.

(Sklansky 78)

The image texture we consider is nonfigurative and cellular. We think of this kind of texture as an organized area phenomena. When it is decomposable, it has two basic dimensions on which it may be described. The first dimension is for describing the primitives out of which the image texture is composed, and the second dimension is for the description of the spatial dependence or interaction between primitives of an image texture. The first dimension is concerned with tonal primitives or local properties, and the second dimension is concerned with the spatial organization of the tonal primitives. ...In summary, to characterize texture, we must characterize the tonal primitives as well as the spatial interrelationship between them.

(Haralick 79)

Eine gute Übersicht über die Ansätze zur Messung und

Charakterisierung von Texturen ist im zitierten Aufsatz von Haralick gegeben worden, der Autor diskutiert ebenfalls die Vor- und Nachteile der einzelnen Methoden.

In den Kapiteln 6.1 bis 6.4 wird nur ein Teil dieser Texturmodelle behandelt werden, wobei die Auswahl sich vorwiegend auf solche Verfahren beschränken wird, die in Bildsegmentationsprogrammen zur Anwendung gekommen sind.

Wegen der Allgemeinheit des Texturbegriffs (er umschließt alle bisher behandelten Regionsmodelle als Spezialfälle) werden sich Fortschritte in der Formulierung eines allgemeinen Texturansatzes auch auf die Verbesserung der Resultate von Bildsegmentationen auswirken.

6.1 Eindimensionale Texturmodelle

Die Charakterisierung von Texturen durch tonale Primitive einerseits und ihre räumliche Anordnung andererseits spiegelt sich ebenfalls in der Vorgehensweise bei der Gewinnung von Deskriptoren zur Texturbeschreibung wider. So wird z.B. ein eindimensionales Modell für ein Texturelement angenommen und im Bild in verschiedenen Richtungen nach dem Auftreten dieses Elementes gesucht. Zur Texturidentifikation wird dann sowohl die räumliche - durch die Suchrichtung und Lage bestimmte - als auch die intensitätsabhängige Information ausgewertet.

Ein Beispiel hierfür sind die von Haralick et al. (Haralick et al. 73) vorgeschlagenen "Cooccurrence-Matrizen" (die Untersuchungen von Anuja und Rosenfeld (Anuja and Rosenfeld 78) wurden auf Seite 47 beschrieben). Für einen gegebenen Verschiebungsvektor (L_x, D_y) wird in einer Matrix P_{ij} notiert, wie oft im Bild von zwei um den Verschiebungsvektor dislozierten Bildpunkten der eine den Grauwert i und der andere den Grauwert j annimmt. Abbildung 6-1 illustriert dies für ein 4×4 Bildfenster

mit vier Intensitätsstufen und vier Verschiebungsrichtungen.

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

0°	$P_H = \begin{pmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$	90°	$P_V = \begin{pmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$
135°	$P_{LD} = \begin{pmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$	45°	$P_{RD} = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

Abbildung 6-1: Beispiele von Bildmatrix und Cooccurrence-Matrizen für vier Verschiebungsrichtungen (Verschiebungsvektor: 1 Bildpunkt)

Der hohe Datenaufwand zur Speicherung der Matrizen bei großer Grauwertauflösung und für alle erlaubten Verschiebungsrichtungen erzwingt eine Normierung des Bildes auf wenige Graustufen und die Auswahl einiger Verschiebungsrichtungen. Aus den Cooccurrence-Matrizen leitet Haralick 14 Merkmale ab, die er zur Texturklassifikation vorschlägt (typische Eigenschaften sind z.B. Uniformität, Entropie, Kontrast, Korrelation zwischen Grauwerten).

Deutsch und Belnkamp (Deutsch and Belnkamp '72) haben mit Cooccurrence-Matrizen für nur ein Grauwertpaar Bildsegmentationen an künstlich erzeugten Texturen untersucht. Chen und Pavlidis (Chen and Pavlidis '78) leiteten aus den Cooccurrence-Matrizen ein Kriterium zur Charakterisierung uniformer Texturen ab und berichteten über deutliche Verbesserungen des 'Split and Merge' Verfahrens bei Bildvorlagen mit Textur.

Analog zu den Cooccurrence-Matrizen notiert Galloway (Galloway '75) die Häufigkeit des Auftretens von Sequenzen konstanten

Grauwerts in "Run-Length"-Matrizen. Sie berechnet mit den von Haralick vorgeschlagenen Textur-Merkmalen fünf Deskriptoren für jeweils vier Richtungen. Mit Hilfe dieser Deskriptoren wird eine Klassifikation von Luftbildern in 9 Kategorien durchgeführt.

Ehrich und Foith (Ehrich and Foith 78) repräsentieren den Grauwertverlauf einer Bildzeile in einer relationalen Baumstruktur. Mit einer rekursiven Prozedur wird die Bildzeile in Segmente zerteilt, die Knoten des Baumes enthalten Informationen über relative Extrema, aus denen Texturmerkmale - wie Segmentkontrast oder Segmentlänge - abgeleitet werden können. Ihr Ansatz läßt sich mit gekoppelten Relationalbäumen auch auf zweidimensionale Texturen ausweiten, Bildsegmentationen auf der Grundlage dieser Datenstruktur sind bisher jedoch nicht durchgeführt worden.

Ebenfalls ein eindimensionales Modell für das Texturelement wird von Davis und Mitiche (Davis and Mitiche 80) verwendet. Als Maß für die Kantenstärke wählten sie die symmetrische Differenz von mittlerem Grauwert in eindimensionalen Streifen unterschiedlicher Breite. Mit Annahmen über die Verteilung der Meßwerte und die lineare Ausdehnung des Texturelementes berechnen sie die zu einem minimalen Klassifikationsfehler führende Streifenbreite des Kantenfinders.

6.2 Texturmerkmale

Da Texturen nicht notwendigerweise aus Wiederholungen von Pixelgruppen mit identischen Grauwertrelationen zwischen den einzelnen Pixeln einer Gruppe bestehen müssen - man denke z.B. an Holzmaserungen, Wiesen oder Wälder in Luftbildern - können Modelle mit probabilistischen Merkmalen wichtige Schlüsse auf Texturunterschiede liefern. Verbindend für die im folgenden dargestellten Verfahren ist die gemeinsame Vorstellung von

texturierten Gebieten als Regionen mit großen lokalen Variationen des Grauwerts, die sich in Begriffen wie

- busyness
- texture coarseness
- roughness

wiederfindet. Gesucht wird nach Merkmalen, die sich zur Diskriminierung von Gebieten unterschiedlicher "Texturiertheit" verwenden lassen. Da räumliche Wiederholungen von Grauwertänderungen Maxima im Fourierspektrum entsprechen, gewinnt Bajcsy (Bajcsy 73a, Bajcsy 73b) ihre Deskriptoren durch Bestimmung der Fourierkoeffizienten in verschiedenen Bildfenstern. Das Energiespektrum der fouriertransformierten Bildfenster wird von ihr in eine radiale Funktion und eine azimutale Funktion zerlegt. Maxima der Radialfunktion liefern Hinweise auf kreisförmige Texturprimitive, Maxima in der azimutalen Funktion sind Hinweise auf gleichorientierte Texturelemente.

Mit Hilfe eines eindimensionalen "Fleckfinders" versuchten Hayes et al. (Hayes et al. 74) die Texturrauhigkeit ("coarseness") zur Unterscheidung heranzuziehen, aber auch eine Ausweitung auf einen zweidimensionalen Fleckfinder ergab enttäuschende Resultate. Zucker et al. (Zucker et al. 75) berichteten über Versuche, durch Variation der Fenstergröße eines Fleckfinders und Analyse der Histogramme des Detektorsignals unterschiedliche Texturen zu diskriminieren. Es ergab sich nur dann ein Maximum im Histogramm, wenn die Fenstergröße annähernd gleichgroß mit dem Texturelement war. Dem Fleckfinder lag ein einfaches Texturelementmodell zugrunde, bei dem allein die Differenz von mittleren Grauwerten berücksichtigt wurde.

Mitchell und Carlton (Mitchell and Carlton 78)

charakterisierten die Textur durch die Zahl der lokalen Extrema in horizontaler und vertikaler Richtung und benutzten dies als Maß zur Segmentation von Infrarotbildern mit einer "nearest-neighbor" Ballungsanalyse.

Ando und Doi (Ando and Doi 78) bestimmten in lokalen Bildfenstern Grauerthistogramme, die Diskriminierung verschieden texturierter Gebiete erreichten sie durch Berechnung eines Maßes für die Ähnlichkeit von Histogrammen (ähnlich den Verfahren von Muerle und Allen, Gupta und Wintz, Nagao et al., vgl. Kap. 2.3). In interaktiv gewählten Testgebieten wurden die für die Klassifikation notwendigen Parameter bestimmt.

6.3 Zweidimensionale Texturmodelle

Für die Teilklasse von Texturen, die aus örtlichen Wiederholungen gleicher Texturelemente gebildet wird, können zweidimensionale Texturelementmodelle konstruiert werden. Bis heute sind nur einfache zweidimensionale Texturmodelle untersucht worden, fast ausschließlich zusammenhängende Bildpunkte konstanten Grauwerts (binäre Texturen).

Tsuji und Tomita (Tsuji and Tomita 75) sowie Tomita et al. (Tomita et al. 75, Tomita et al. 78) erzeugen binäre Testbilder durch Schwellwertbildung mit Histogrammverfahren bzw. durch Verschmelzungsalgorithmen. Zugrunde liegt ein Regionsmodell konstanten Grauwerts. Für die Regionen werden dann Eigenschaften wie mittlerer Grauwert, Fläche, Ausrichtung, Rundheit, Krümmung, Momente usw. berechnet. Die Texturklassen werden durch Analyse der Eigenschaftshistogramme (Schwellwertbildung an den Minima der Histogramme) gewonnen.

Die Repräsentation von Texturen durch einen regulären Graphen schlägt Zucker (Zucker 76c) vor, jeder Knoten entspricht einem Texturelement, die gleichartige Zusammensetzung des Bildes durch

Textureprimitive bedeutet, daß jeder Knoten im Graphen in identischer Weise mit den Nachbarknoten verbunden ist (vgl. hierzu die Ansätze von Ehrich und Foith, Kap. 6.1).

Expliziten Gebrauch von der Darstellung von Texturen durch einen Graphen machen Tsai und Fu (Tsai and Fu 78). Für ein durch einen Baum beschriebenes Texturelement wird eine spezielle stochastische Baumgrammatik konstruiert, alle von dieser Grammatik akzeptierten Texturen entsprechen dann einer Texturklasse (siehe Abb. 6-2). In einer Lernphase wird in Trainingsgebieten die repräsentative Grammatik der betreffenden Textur automatisch erzeugt.

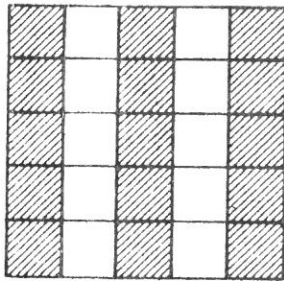


Figure 1

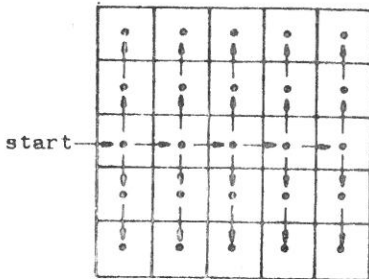


Figure 2

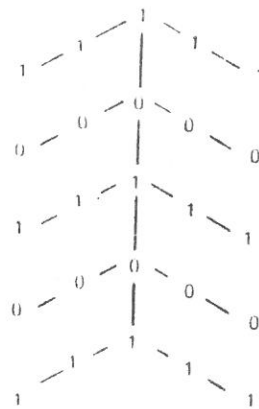


Figure 3

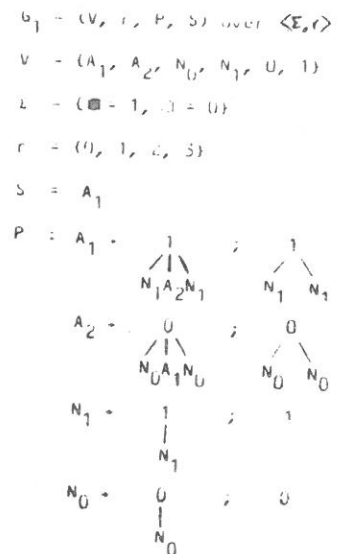


Figure 4

Abbildung 6-2: Texturmodell durch Graphen und Grammatik

6.4 Texturgradient und Texturkanten

Ähnlich wie bei einem Modell der Region als ein Gebiet homogenen Grauwerts die Regionsgrenzen durch Analyse des Grauwertgradientenbildes erhalten werden, kann ein Maß für die lokale Texturänderung zur Bildsegmentation texturierter Objekte benutzt werden ("Texturkanten-Operator"). Wichtiger noch sind aber die Schlüsse, die aus einer örtlichen Änderung der Textur über die Objekte des Szenenbereichs gezogen werden können. Eine Veränderung der Größe von (als konstant auf der Objektoberfläche angenommenen) Texturprimitiven liefert z.B. Hinweise auf Entfernungsänderungen zwischen Objekt und Beobachter. Formänderungen der Primitive oder geänderte Texturdichte können ebenfalls durch Orientierungsänderungen bewirkt sein. Bajcsy und Liebermann berechneten aus Messungen des Texturgradienten Informationen über die Entfernung texturierter Objekte vom Standort des Beobachters (Bajcsy and Lieberman 74, Bajcsy and Lieberman 76). Das Verfahren arbeitet mit einem einfachen Szenenmodell. Es darf nur ein Texturelement in der Szene auftreten; und die Szene darf keine nichttexturierten Objekte enthalten. Aus dem Fourier-transformierten Bild wird das Energiespektrum und das Histogramm der Phasenwinkel ermittelt. Daraus ließ sich die "Wellenlänge" der Textur bestimmen, die zur Gewinnung der Entfernungsinformation herangezogen wurde.

Da die Eigenschaft "Textur" nicht wie die Eigenschaft "Grauwert" exakt definiert werden kann, existiert auch kein wohldefinierter Texturgradient-Operator. Rosenfeld (Rosenfeld 75) schlägt z.B. die Größe und Richtung der maximalen Zunahme des Wertes eines Texturoperators als Meßwert für einen Texturgradienten vor (oder auch nur die Differenz des Operatorwertes an benachbarten Pixelpositionen).

Konturen von texturierten Regionen bestimmten Yachida et al. (Yachida et al. 79) mit rekursiven Schwellwertverfahren, der in drei Testgebieten trainierte Kantendetektor benutzte als Eigenschaften die Charakteristiken des Grauwertistogramms sowie Grauwertgradient pro Fläche und Grauwert-Kantenrichtung. Davis und Mitiche (Davis and Mitiche 80) errechneten als einziges Attribut ihres eindimensionalen Kantendetektors die symmetrische Differenz von Mittelwerten benachbarter Pixel.

7. Objektmodelle und Szenenmodelle

Die bisher besprochenen Verfahren sind in der Mehrzahl ohne jede Vorkenntnis über die zu untersuchenden Bilder ausgekommen. Die dabei durchgeführte Modellbildung für homogene Regionen, Kanten, Linien und Texturen beschränkte sich auf den Bereich der Bildfunktion. Eigenschaften der Objekte in der Szene, die diese Bildbereichs-Hinweise bewirkten, blieben unbestimmbar oder es waren mehrere Deutungen möglich. Die Modellierung bekannter, im Bild erwarteter Objekte muß nun aber nicht notwendig auf solche Bildbereichs- oder Szenenbereichs-Hinweise beschränkt bleiben (vgl. (Nagel 79)), sondern kann sich auf eine vollständigere zweidimensionale oder dreidimensionale Objektmodellierung erstrecken. Die Hinzunahme von "Bildsemantik" führt in folgenden Fällen zu einer Einschränkung der Klasse der analysierbaren (segmentierbaren) Bilder:

- die Objektmodelle sind zu speziell auf die im Bild erwarteten Objekte zugeschnitten (es sind keine "generischen" Beschreibungen)
- die Anzahl der Modelle ist zu klein, um die Vielfalt tatsächlicher Objekte zu beschreiben.

Die semantischen Segmentationsverfahren werden in Zukunft eine zunehmende Bedeutung gewinnen, wenn diese Einschränkungen überwunden werden können. Ansätze hierfür sind schon erkennbar, etwa durch

- Wahl geeigneter rechnerinterner Datenstrukturen zur Modellierung von Objekten, Objekteigenschaften und Relationen zwischen Objekten;
- Entwicklung effizienter Methoden zum Vergleich von Bildinhalten mit den Modellen (Bildinterpretation);
- Herausnahme der Modelle aus den Programmen, Speicherung der Modelle in einer einfach modifizierbaren Datenbasis;
- Erhöhung der Zahl der verfügbaren Modelle (wobei Konflikte bei der Interpretation möglichst

ausgeschlossen werden).

7.1 Objektmodellierung im Bildbereich

Die einfachste Form der Objektmodellierung ist die rechnerinterne Speicherung eines oder mehrerer Objekte als Grauwertbild, zum Beispiel durch eine Kameraaufnahme eines Prototypen gewonnen. Mit Korrelationsverfahren wird durch Vergleich verschiedener Bildausschnitte mit der Schablone des Objektbildes der Ort maximaler Übereinstimmung bestimmt. Eine Anwendung dieser Methode wird aber nur bei speziellen Aufgabenstellungen zum Erfolg führen, z.B. in industriellen Anwendungen von wiederkehrend identischen Objekten gleicher Positionierung vor dem Bildaufnahme-Sensor. Schablonenverfahren versagen bei gedrehten Objekten, bei Objekten gleicher Form aber unterschiedlicher Größe und bei Formunterschieden zwischen Objekt und Maske. Die Bildsegmentation mit Schablonenverfahren beschränkt sich hierbei auf die Trennung des Objekts vom Hintergrund und die Bestimmung der Abweichung des Objekts von der Schablone.

7.2 Bildbereichs-Hinweise

Die Invarianz des Objektmodells gegenüber Translationen, Rotationen und Größenskalierungen von Objekten der Szene ist eine grundlegende Forderung, die bei der Wahl einer Modellbeschreibung berücksichtigt werden muß. Der Vergleich zwischen Modellbeschreibung und einer aus dem Bild gewonnenen Objektbeschreibung muß auch bei geänderter Beobachterstandpunkt oder geänderter Lage des Objekts in der Szene möglich sein.

Ein erster Schritt hierzu ist die Modellierung der zweidimensionalen Form der Objekte im Bildbereich und das Beschreiben von Relationen zwischen Objekten bzw. Objektflächen.

Zur Erkennung von Rippen in Röntgenaufnahmen des Brustkorbes

bettet Ballard (Ballard 78) sein zweidimensionales Modell des Brustkorbes in sein Analyseprogramm ein, in dem die Lage der einzelnen Rippen zueinander in einem relationalen Graphen aufgenommen wird. Im prozeduralen Teil des Programms wird dann gezielt nach einer besten Ausprägung der generischen Beschreibung im vorliegenden Röntgenbild gesucht und aus den berechneten Ergebnissen neue Information in den relationalen Graphen aufgenommen.

Nevatia und Price (Nevatia and Price 78) repräsentieren die a priori Information über die von ihnen untersuchten Luftbilder als Graphenstruktur, an den Knoten werden Eigenschaften von Objekten (Position, Größe, Orientierung, Form, Farbe und Textur) notiert, während Relationen die Lage der Objekte zueinander (über, unter, links-von, rechts-von) beschreiben. Eine Übereinstimmung zwischen dem zweidimensionalen Modell und dem zu analysierenden Bild wird mit einer Ähnlichkeits-Funktion zwischen den Eigenschaften der durch Segmentation gewonnenen Elementarregionen und der symbolischen Bildbeschreibung errechnet.

You und Fou (You and Fou 78) beschreiben Silhouetten von verschiedenen Flugzeugtypen durch ein Tupel von vier Eigenschaften der Konturstücke (Vektor zwischen Endpunkten, Länge, Winkeländerung, Symmetriemaß) und einer zugehörigen attributierten Grammatik; die Erkennung der Flugzeugtypen wird von einem endlichen Automaten ausgeführt.

Zur Identifizierung von flächigen industriellen Werkstücken beschreibt Perkins (Perkins 76, Perkins 78) ein Analysesystem, mit dem auf einem Förderband liegende Bauteile beliebiger Orientierung erkannt werden können. Bei Beschränkung auf nur eine Auflageart der Werkstücke ist ein zweidimensionales Modell ausreichend. Als zusätzliche Vereinfachung wird auf die

Modellierung der Objektoberfläche verzichtet. Im Bild werden Kantensegmente mit dem Hueckel-Operator bestimmt und durch eine Nachverarbeitung zu Silhouettenstücke der Objektberandungen verbunden. Die Objektmodelle werden ebenfalls als Silhouettenstücke in einer Datenbasis beschrieben. Ein heuristisches Konfidenzmaß wird zum Vergleich von Objektmodell und einer aus dem Bild gewonnenen "Objektskizze" herangezogen.

Mit zunehmender Ausnutzung von semantischem Wissen über die zu analysierende Szene verändert sich die Aufgabe einer reinen Bildsegmentation in Richtung auf eine Bestimmung der semantisch richtigen Bildinterpretation, wobei eine Wechselwirkung zwischen Segmentation und Interpretation nicht ausgeschlossen ist.

Preparata und Ray (Preparata and Ray 72) zerlegten Farbbilder natürlicher Szenen in Einheitszellen und verschmolzen Zellen gleicher Farbe (entsprechende Verfahren wurden in Kap. 2.3 vorgestellt). Diese Elementarregionen und ihre Attribute (Größe, Form, Farbe, etc.) wurden in einen sogenannten "Szenengraphen" überführt, wobei Kanten wiederum mit Relationen zwischen den Regionen attribuiert werden und an den Knoten die Regionsbeschreibungen notiert werden. Der bekannte Szeneninhalte wird ebenfalls mit einem - als semantische Karte bezeichneten - Graphen beschrieben. Das Interpretationsproblem besteht dann im Auffinden einer isomorphen Abbildung zwischen Teilgraphen aus dem "Szenengraphen" und der "semantischen Karte". Das Auffinden isomorpher Wege in beiden Graphen geschieht durch Hypothesenbildung über die Regionen unter Zuhilfenahme der entsprechenden Attribute, die Verifikation der Hypothesen wird von Gewichtungsfaktoren an den Kanten der "semantischen Karte" gestützt. Das Verfahren von Preparata und Ray arbeitet bis zur Erstellung des Szenengraphen ohne spezielles Szenenwissen.

Kontextinformation und die semantische Information werden erst bei der Interpretation ausgenutzt.

Die Formulierung der Bildinterpretation als Suche nach einem Graphen- oder Teilgraphenisomorphismus zwischen einem "atomaren Graphen" und einem "semantischen Graphen" wird auch von Pavlidis (Pavlidis 79) beschrieben, wobei der atomare Graph dem Szenengraph bei Preparata und Ray entspricht.

Ebenfalls mit graphorientierten Methoden arbeitet das Analyseprogramm von Radig (Radig 78). Es faßt die Parameter einer Ebenenapproximation der Bildfunktion mit den Koordinaten, dem Grauwert und der Streuung zu einem mehrkomponentigen Vektor zusammen. Die aus einem Bild gewonnenen Vektoren bilden die Knoten eines Graphen, dessen Kanten durch die Nachbarschaftsrelation initialisiert sind. Ein Bereichsmodell wird durch einen Parametersatz von Zwangsbedingungen zwischen benachbarten Vektoren formuliert. Bei Anwendung der im Modell geforderten Konsistenzen zerfällt der Graph des Bildes in Subgraphen. Die Subgraphen werden als Bereiche aufgefaßt, aus den Werten der in ihnen enthaltenen Vektoren werden Bereichsbeschreibungen ermittelt. Mit diesen Beschreibungen konnten die gefundenen Bereiche auch in einer Bildsequenz verfolgt werden.

Von Feldman und Yakimovsky (Feldman and Yakimovsky 74) wurde die semantische Information zur Steuerung ihres Segmentationsprogramms ausgenutzt. In einem kontextfreien Verarbeitungsschritt wurde das Bild in elementare Regionen mit Regionsattributen zerlegt und zusätzlich die Regionsgrenzen (Kanten) mit ihren Kantenattributen bestimmt. Den Regionen und Kanten werden semantische Marken ("Interpretationen") zugeordnet, die Wahrscheinlichkeit für eine Marke wird aus den lokal

beobachteten Attributen berechnet. Ziel der Segmentation ist eine Maximierung der Wahrscheinlichkeit für eine global richtige Interpretation auf der Basis der gemessenen Attribute. Die Wahrscheinlichkeit der Gesamtinterpretation wird aus den Produkten der Wahrscheinlichkeiten für eine lokal richtige Regionsinterpretation (bestimmt aus den Regionsattributen) und den lokal richtigen Kanteninterpretationen (errechnet aus Kanten- und Regionsattributen) gebildet.

Der interpretationsgestützte Segmentationsalgorithmus von Tenenbaum und Barrow (Tenenbaum and Barrow 76) ist eine Erweiterung des semantischen Verfahrens von Felman und Yakimovsky. Zur Analyse von Kantentypen in Strichzeichnungen einer Polyeder-Szene entwickelte Waltz (Waltz 75) eine Methode zur semantischen Filterung von Kombinationsmöglichkeiten beim Zusammentreffen verschieden markierter Kantentypen. Tenenbaum und Barrow verwendeten seine Methode der "constraints-relaxation" zur Elimination von Interpretationsmarken in vorsegmentierten Bildern.

Die Konsistenz einer Regionsinterpretation wird durch relationale Zwangsbedingungen bestimmt, die unter Auswertung des Szeneninhalts aufgestellt werden können. Ein Beispiel für eine Zwangsbedingung ist etwa, daß eine mit der Marke "Gras" versehene Region keine benachbarte Region "Himmel" haben kann und die Relation "Gras liegt über Himmel" gleichzeitig erfüllt ist. Die Verschmelzung von Nachbarregionen erfolgt iterativ in vier Schritten:

1. Verschmelzung von Regionen gleicher Interpretation mit der "Likelihood-Methode";
2. Bei Verschmelzung von Regionen unterschiedlicher, aber miteinander vertraglicher Interpretation werden Nachbargebiete mit schwachen Kanten zuerst verschmolzen;

3. Nach jeder Verschmelzung erfolgt eine Neuinterpretation der Regionen;
4. Die Konsistenz der Interpretation wird mit dem von Waltz entwickelten Verfahren überprüft.

Für die analysierten Bilder werden in einer interaktiven Phase ausgewählte Regionen mit Interpretationen versehen. Das starke Ineinandergreifen von Segmentation und Interpretation ist bei dem Verfahren von Tenenbaum und Barrow besonders deutlich zu erkennen.

Die Schwierigkeiten, die bei einer Trennung von Segmentation und Interpretationsphase auftreten (Bestimmen der Korrespondenzen zwischen Szenenmodell und der aus dem Bild gewonnenen Datenstruktur) werden mit diesem Ansatz zwar umgangen, die Programme verlieren dadurch aber Modularität und Universalität.

Das Bildanalyseprogramm von Ohta et al. (Ohta et al. 78) weist ebenfalls eine starke Verflechtung zwischen dem Segmentationsprozeß (bottom-up-Verfahren) und der semantischen Bildinterpretation (top-down-Verfahren) auf. Mit der von Ohlander entwickelten rekursiven Schwellwerttechnik wird eine vorläufige Segmentation in Regionen mit homogenen spektralen Eigenschaften vorgenommen. Diese Regionen werden in einer Datenstruktur mit ihren Attributen und topologischen Relationen abgelegt. Ein Szenenmodell - in Form eines semantischen Netzes - unterstützt mit a priori Wissen über den Inhalt der Szene diese "plangenerierende" Phase. Im top-down gesteuerten Interpretationsprozeß wird Szenenwissen in der Form eines Produktionensystems zum Verschmelzen von Regionen eingesetzt. Nach einem Entscheidungsprozeß in der Interpretationsphase kann der "Plan" der Segmentierungsphase geändert werden. Der zyklische Datenfluß im Analyseprogramm zeigt eine mögliche Realisierung des von Kanade und Nagel (Kanade 80a, Nagel 79) beschriebenen Modells

zur algorithmischen Bilddeutung.

Die bisher zitierten Arbeiten dieses Abschnitts haben in ihren Modellen bereits den Projektionsvorgang vom 3-D-Szenenbereich in den 2-D-Bildbereich vollzogen, wenn sie die geometrischen Relationen der Objekte durch Beziehungen wie "oben", "unten", "benachbart" in den Modellen als zweidimensionale Relationen implementieren.

Im ACRONYM-System von Binford und Mitarbeitern (Brooks et al. 79a, Brooks et al. 79b) werden Objekte in einer hierarchischen Struktur als generalisierte Kegel modelliert. In Sichtbarkeitsgraphen wird das Wissen über eine spezielle Ausprägung des Objekts einem Vergleichsalgorithmus übermittelt, der Übereinstimmungen zwischen dem aus dem Bild erzeugten Bildgraphen und dem Sichtbarkeitsgraphen ermittelt.

Ebenfalls ein hierarchisches System wurde von Bertelsmeier und Radig konzipiert (Bertelsmeier and Radig 77a, Bertelsmeier and Radig 77b). Aus Punkten, Konturen und Regionen wird bei der Vorverarbeitung des Bildes eine relationale Struktur erzeugt. Die Relationalstruktur soll diese primitiven Objekte und ihre Beziehungen zueinander aufnehmen. Die derart erhaltene Skizze eines Bildes wird zu Objekten gruppiert, für die beschreibende Prototypen definiert sind. Diese wiederum können zu Strukturen der nächsthöheren Abstraktionsebene zusammengefaßt werden (Szenenmodell).

Kraasch und Zach (Kraasch and Zach 78) verwendeten das relationale Modell zur Beschreibung dreidimensionaler Polyeder. Obwohl die Modellierung im zweidimensionalen Bildbereich erfolgte, zeigten sie, daß mit einem Modell des bilderzeugenden Systems aus einer zeitlichen Folge von Bildern - unter Ausnutzung der gefundenen Zuordnungen markanter Punkte in den Einzelbildern

- die dreidimensionale Objektform rekonstruiert werden kann (Kraasch et al. 79a, Kraasch et al. 79b, Radig et al 80). Die Zuordnung der ermittelten, zu gleichen Objektpunkten gehörigen 2-D-Punktkoordinaten zu 3-D-Szenenkoordinaten benutzte eine erste Implementation des von Bonde und Nagel beschriebenen Verfahrens zur 3-D-Rekonstruktion von Objektpunkten aus 2-D-Beschreibungen der Punkte im Bildkoordinatensystem (Bonde and Nagel 79).

Ein modulares System zur Bildsegmentation und Bildinterpretation, das viele der auf Seite 70 gestellten Anforderungen einer semantischen Bildinterpretation erfüllt, wird von Levine und Shaheen beschrieben (Levine and Shaheen 79). Eine als "Kurzzeitgedächtnis" bezeichnete relationale Datenbasis enthält die Rohdaten der momentanen Interpretation der zu analysierenden Szene. Das Kurzzeitgedächtnis wird während der Analyse von den gewonnenen Erkenntnissen verändert (man vergleiche diese Beschreibung mit den Szenenbereichs-Hinweisen in Kap. 9.2). Eine zweite relationale Datenbasis enthält eine Verkörperung der Modellrepräsentation oder eine Beschreibung der zu analysierenden Klasse von Szenen. Diese zweite Datenbasis verändert sich nicht während der Analyse. Verschiedene Prozesse, von denen bisher aber nur "low-level"-Implementationen existieren, erzeugen die Datenstrukturen im Kurzzeitgedächtnis. Durch andere Prozesse werden mögliche Interpretationen der Regionen initialisiert und verifiziert. Für diese Funktion werden die im Langzeitgedächtnis abgelegten Objekteigenschaften herangezogen.

8. Segmentation in Bildfolgen

Der Analyse von zeitlichen oder räumlichen Bildfolgen ist während der letzten fünf Jahre eine stetig wachsende Bedeutung zugekommen. Für das zu analysierende Bild ergibt sich bei Bildfolgen eine dritte Bildkoordinate, die etwa bei Folgen von Tomogrammen des menschlichen Körpers eine räumliche Koordinate oder bei Fernsehbildsequenzen eine zeitliche Koordinate sein kann, entlang der eine Grauwertvariation des Bildes auftritt. In den Abschnitten dieses Kapitels soll gezeigt werden, welche Methoden der Segmentation von statischen (zweidimensionalen) Bildern auf Bildfolgen übertragen werden können und inwieweit neue Techniken bei der Segmentation von Bildfolgen eingesetzt werden. Hierbei soll insbesondere der Einfluß von verbesserten Modellen der im Bild enthaltenen Szenenkomponenten auf die Bildsegmentation untersucht werden.

8.1 Stationäre und nichtstationäre Bildkomponenten

Wenn sich in einer Szene Objekte in ihrer relativen Lage zueinander oder zu einem stationären Szenenteil bewegen, ist es naheliegend, in einer Folge von Bildern dieser Szene systematische Verschiebungen von Bildpunkten als die das Objekt repräsentierenden Bildpunkte zu interpretieren. Damit eröffnet sich ein Weg, stationäre und nichtstationäre Bildkomponenten durch Analyse von Grauwertänderungen entlang der zeitlichen Bildkoordinate zu trennen und eine Segmentation des Bildes in bewegte Objekte und stationäre Komponenten der Szene durchzuführen. Eine äquivalente Aufgabe ist das Auffinden einer Folge binärer Objektmasken, die jeweils alle dem bewegten Objekt zuzuordnenden Bildpunktkoordinaten enthalten.

Die bei der Analyse der Grauwertänderungen auftretenden Phänomene lassen sich auf folgende Änderungen in der Szene bzw.

des Bildaufnahmesystems zurückführen:

1. konstanter Grauwert an gleichen Bildkoordinaten

- a. es handelt sich um einen stationären Teil des Bildes
- b. die Objektoberfläche eines bewegten Objektes führt im Folgebild zum gleichen Grauwert wie die entsprechende Objektoberfläche des vorherigen Bildes (z.B. bei der Bewegung von Oberflächen homogener Färbung mit einer Verschiebung, die klein gegenüber den Objektmessungen in Translationsrichtung ist).

2. Grauwertänderung an der gleichen Bildkoordinate

- a. Rauscheffekte im Bildaufnahmesystem
- b. Verdeckung des stationären Hintergrunds durch ein bewegtes Objekt
- c. Freigabe von Hintergrund durch ein bewegtes Objekt
- d. Verdeckung von Objektteilen
 - durch stationäre Szenenkomponenten
 - durch andere bewegte Objekte
 - durch Objektrotation oder Veränderung des Beobachtungsstandpunktes
- e. Freigabe von Objektteilen
- f. Formveränderung des Objekts bei nichtstarrten Körpern
- g. Veränderung des Objektbildes durch eine allgemeine dreidimensionale Objektbewegung in der Szene, z.B.
 - Vergrößerung oder Verkleinerung des Objektbildes durch Annäherung an den Betrachter oder Entfernung vom Betrachter
 - Lichtintensitätsänderungen durch andere Orientierung von Objektoberflächen relativ zur Lichtquelle.

Die Segmentation einer Bildfolge in stationäre und nichtstationäre Bildkomponenten ohne a priori Wissen über die

Szene kann wegen der Vielzahl der möglichen Interpretationen einer zeitlichen Grauwertveränderung nur durch zusätzliche, einschränkende Annahmen erfolgen.

Einer der frühesten Versuche, aus Bildfolgen die nichtstationären Bildkomponenten zu extrahieren, ist von Potter (Potter 75a, Potter 75b, Potter 77) durchgeführt worden. In seiner erstgenannten Veröffentlichung ging er von Objekten aus, die durch eine zeitlich konstante Objektmaske beschreibbar sind; es wurden daher nur die unter den Punkten a. bis c. aufgeführten Effekte untersucht. In einem über das Bild gelegten Raster wurden für jeden Bildpunkt in Zeilen- und Spaltenrichtung die Abstände zur nächsten Grauwertdiskontinuität bestimmt. Die Berechnung wiederholte sich im Folgebild und Bildpunkte mit gleichen "Geschwindigkeitsattributen" werden als Koordinaten korrespondierender Objektpunkte interpretiert. Durch Analyse der Geschwindigkeitsattribute wird für jeden Bildpunkt entschieden, ob er zum bewegten Objekt oder zum Hintergrund gehört oder ob es sich um verdeckten oder freigegebenen Hintergrund handelt.

In (Potter 75b) wird eine veränderte Methode beschrieben, bei der eine kreuzförmige Maske im ersten Bild am Objektbild verankert wird und ebenfalls die Abstände vom Zentrum des Kreuzes bis zur Objektkante im Bild gemessen werden. Mit einer heuristischen Methode wird in den Folgebildern die Maske gesucht und durch Zusammenfassen von Bereichen gleicher Verschiebung die Bildsegmentation durchgeführt. Mithilfe verschieden geformter Schablonen versuchte Potter (Potter 77) durch Messung der zeitlichen Veränderung der Armlängen der Masken auch Objektverdeckungen zu erfassen. Da die Verfahren von Potter konstante Objektgrößen erfordern, können sie als Vereinfachung von Kreuzkorrelationsverfahren mit einer aus dem ersten Bild

gewonnenen Maske betrachtet werden.

Jain et al. (Jain et al. 77) befassen sich explizit mit der Extraktion einer binären Maske des bewegten Objekts ohne Ausnutzung von speziellem Szenenwissen. Die Analyse setzt ebenso wie Potters Verfahren voraus, daß die Objekte eine gleichförmige Bewegung ausführen. Die Lageänderung zwischen zwei Bildern ist gering gegenüber den Abmessungen der Objektbilder. Das erste Bild einer Szenenfolge wird als Referenzbild gewählt. Mit einem Likelihood-Kriterium wird zwischen dem aktuellen Bild und dem Referenzbild die Grauwertänderung bestimmt, und bei jedem Überschreiten eines Schwellwertes zwischen aktuellem Bild und Referenzbild für diese Bildkoordinate ein Zähler erhöht. Bei der Bewegung eines Objekts mit homogenem (aber vom Hintergrund verschiedenem) Grauwert werden aber nur an Objekträndern senkrecht zur Bewegungsrichtung Änderungen bemerkt (vgl. Fall b.), sie entsprechen verdecktem Hintergrund an der Vorderkante des Objekts sowie freigegebenem Hintergrund am hinteren Objektrand . Bei stetiger Lageänderung kann aus der Monotonie benachbarter Zähler die Bewegungsrichtung des Objekts ermittelt werden (siehe Abb. 8-1). Die Objektmaske des nichtstationären Bildes kann durch Vergleich der Grauwerte von Referenzbild und aktuellem Bild (an Positionen hoher Zählerwerte) ermittelt werden.

Mit der Einführung einer binären Maske (Jain and Nagel 79), in der Unterschiede zwischen momentanem Bild und Vorgängerbild notiert werden, kann eine Objektmaske für das momentane wie auch für das Referenzbild bestimmt werden, wenn das Objekt den im ersten Bild verdeckten Hintergrund vollständig freigegeben hat.

Dreschler und Nagel (Dreschler and Nagel 78) bestimmen aus aufeinanderfolgenden Bildern zusammenhängende Gebiete mit großer Grauwertdifferenz, diese werden in einer ersten Hypothese als

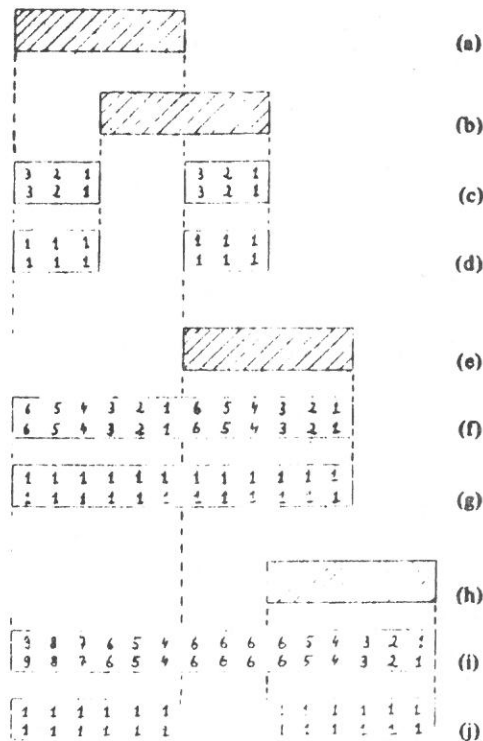


Fig. 1. (a) Object image in the reference frame. (b), (c), (d) Object image, FODP, and SODP after three successive displacements of the object image by a sample area each. (e), (f), (g) Object image, FODP, and SODP, respectively, after the object image has been displaced exactly by its diameter D projected onto the direction of displacement. (h), (i), (j) Object image, FODP, and SODP, respectively, after the object image has been displaced by more than its diameter D projected onto the direction of displacement.

Abbildung 8-1: Differenzverfahren nach (Jain and Nagel 79)

Kandidaten für bewegte Objekte betrachtet. Für die Kandidatenregionen werden acht Eigenschaften (Grauwert, Fläche, Schwerpunkt, Umfang, etc.) berechnet. Nach der Verarbeitung des letzten Bildes wird eine Ballungsanalyse im Merkmalsraum durchgeführt und die erhaltenen Ballungen zur Extraktion der nichtstationären Bildkomponenten verwendet. Mit dem Verfahren konnte in einer Straßenszene ein Fahrzeug auch bei einer Drehung von 90 Grad aus dem Bildmaterial extrahiert werden. Das verwendete Verfahren der Analyse mittels eines minimalen spannenden Baumes toleriert kleine Änderungen der Meßwerte im Merkmalsraum zwischen den Bildern einer Bildfolge. Die Änderung

der Meßwerte wurde aber nicht bei der Analyse ausgewertet.

Mit einem von Moravec (Moravec 77) vorgeschlagenen Auffälligkeits-Operator ("conspicuousness-Operator") bestimmen Barnard und Thompson (Barnard and Thompson 78) in den Einzelbildern einer Bildsequenz die durch ein hohes Auffälligkeitsmaß ausgezeichneten Punkte. In einem Streudiagramm wird ein Punkt aus einem gegebenen Bild gegen alle Punkte des Folgebildes geprüft. Aus dem Streudiagramm von Verschiebungsvektoren können nichtstationäre Bildkomponenten aus Ballungen mit von Null verschiedenen Verschiebungsvektoren extrahiert werden. Der stationäre Hintergrund wird aus Ballungsgebieten verschwindender Disparität bestimmt. Mit dem Verfahren von Barnard und Thompson können bewegte Objekte separiert werden, wenn hinreichend viele Punkte einer Oberfläche mit dem Auffälligkeits-Operator extrahiert werden.

Die Ansätze von Potter, Jain und Nagel, Dreschler und Nagel sowie Barnard und Thompson können mit Verfahren der Segmentation von statischen Szenen verglichen werden, bei denen die Segmentierung des Bildes durch eine Modellierung von Regionen erreicht werden soll (vgl. Kap. 2).

Bei Potter zum Beispiel wird statt einer räumlichen Homogenität des Grauwerts die zeitliche Konstanz der Geschwindigkeitsattribute für einen zum gleichen Objekt gehörenden Bildpunkt gefordert, der Ansatz von Dreschler und Nagel erlaubt die langsame Veränderung (zumindest einiger) komplexer Formattribute. Beim zuletztgenannten Verfahren wird damit nicht mehr die Konstanz des Objektbildes in den Folgebildern vorausgesetzt. Auch mit der schwächeren Forderung nach Ähnlichkeit aufeinanderfolgender Objektbilder läßt sich das bewegte Objekt in einer Bildsequenz verfolgen.

Korrespondenzen zwischen den Segmentationsalgorithmen von Einzelbildern und Bildfolgen ergeben sich auch bei der Betrachtung des zeitlichen Grauwertverlaufs an festen Pixelpositionen. Eine Unstetigkeit im zeitlichen Grauwertverlauf ("zeitliche Grauwertkante") kann als Bewegung der Objektkontur gedeutet werden - dieses Modell liegt z.B. bei Jain et al. vor (Jain et al. 77). Kantenoperatoren, die bei der Analyse statischer Bilder benutzt wurden, können zur Detektion von zeitlichen Grauwertkanten herangezogen werden.

Batali und Ullman (Batali and Ullman 79) berichten von einer Übertragung eines von Marr und Ullman (Marr and Ullman 79) entwickelten Konvolutionsverfahrens auf eine Bildfolge. Die Bilder wurden mit einer Laplace-Maske konvolviert, die Nullstellen der sich ergebenden Bildfunktion sind die Orte größter Grauwertänderung. Aus der zeitlichen Ableitung der Konvolution bestimmten die Autoren lokale Geschwindigkeitsrichtungen. Die Objektkonturen werden aus den Unstetigkeiten der Richtungsinformation ermittelt.

Bei einer Übertragung von räumlichen Kantenmodellen auf Bildfolgen ist aber Vorsicht geboten, z.B. hat der Betrag des Grauwertgradienten weder eine eindeutige Interpretation als "räumliche Kantenstärke" noch erlaubt er Rückschlüsse auf den Betrag der Geschwindigkeit einer Kante. Dies wird in Abb. 8-2 veranschaulicht. Für zwei eindimensionale Kantentypen und zu verschiedenen Verschiebungen der Kante wird die zeitliche Grauwertdifferenz berechnet. Man sieht, daß die Breite der von Null verschiedenen Differenzwerte möglicherweise mehr Aufschluß über die Verschiebung liefert als der Betrag der Differenzwerte.

Die Segmentation einer Bildfolge in stationäre und

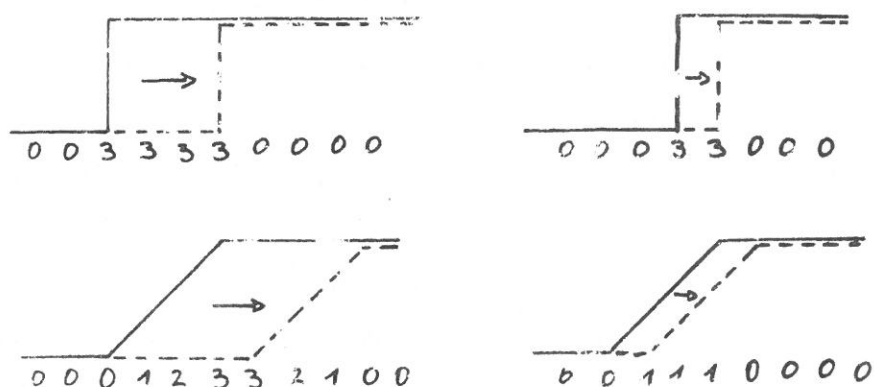


Abbildung 8-2: Differenzbilder bei verschobenen Kanten

nichtstationäre Bildteile kann auch durch eine Teilung der Aufgabe in eine Segmentation der Einzelbilder (statische Segmentation) und nachfolgender Suche nach Veränderungen zwischen den Segmentbeschreibungen aufeinanderfolgender Bilder erfolgen (eine nach den Methoden unterscheidende Gliederung der Verfahren hat Nagel (Nagel 78) vorgenommen, dort findet man ebenfalls eine umfangreiche Bibliographie).

Aggarwal und Duda (Aggarwal and Duda 75) benutzten in einer Vorstudie für die geplante Analyse von Wolkenbewegungen aus Satellitenbildern synthetische Bilder von zweidimensionalen starren Polygonen. Als tolerierte Veränderungen waren zweidimensionale Verschiebungen sowie gegenseitige Objektverdeckungen zugelassen. Es durfte aber höchstens eine Kante verdeckt werden oder neu im Folgebild auftreten. Um die lineare Geschwindigkeit und die Winkelgeschwindigkeit der Objekte zu bestimmen, unterschieden sie zwischen wahren (Objekt-) und falschen (Verdeckungs-) Eckpunkten sowie wahren (zeitlich konstanten) und falschen (zeitlich veränderlichen) Winkelbeziehungen zwischen Kanten. In den Bildern einer Folge wurden korrespondierende Objekte durch einen Vergleich von Polygonen unter der Annahme einer konstanten Topologie

identifiziert. Diese Annahme einer zeitlich konstanten Objektsilhouette wurde in einer späteren Arbeit modifiziert (Roach and Aggarwal 79), bei der die Bewegung konvexer Polyeder verfolgt wurde. Das Objektmodell wurde erweitert um lokale Eigenschaften wie Zentroidposition von Teilflächen, relative Positionen von Teilflächen zu anderen Flächen oder zu anderen Objekten. Unsicherheiten bei der Segmentation eines Einzelbildes wurden in das Folgebild übernommen. Die Verdeckung von Objekten wurde im Programm zu lösen versucht, indem eine Beschränkung auf den Vergleich zwischen sichtbarem Teil des Objekts mit dem vollständigen Objektmodell aus dem vorigen Bild zugelassen war.

Tsuji et al. (Tsuji et al. 79) untersuchten die Analyse von aus Strichzeichnungen bestehenden Zeichentrickfilmen ohne a priori Wissen. Ausgehend von einem aus Konturstücken bestehenden Bild wird nach Korrespondenzen zwischen Konturelementen in der Bildfolge gesucht. Regionen werden durch ihre Zentroidkoordinaten, ihren Umfang sowie einer Liste der Konturelemente beschrieben (vgl. diesen Ansatz mit den kooperativen Regions/Konturmodellen in Kap. 4). Durch Gruppierung von Elementen mit ähnlichen Geschwindigkeitsvektoren kann dieses Verfahren auch bei nichtstarrten Objekten die richtigen Zuordnungen während der Bildsequenz aufrechterhalten.

9. Zusammenfassung

Ziel dieser Übersicht ist es, die in den bisher veröffentlichten Segmentationsprogrammen verwendeten Modelle der Bildfunktion, des Abbildungssystems, der Objekte oder der Szene zu charakterisieren und die Modellannahmen zu verdeutlichen. An dieser Stelle soll noch einmal ihre Stellung innerhalb des Deutungsprozesses von Kanade aufgezeigt werden. Zusätzlich wird untersucht, wie ein Übergang von einem Modell zu Modellen einer anderen Stufe im Zyklus ausgeführt werden kann.

9.1 Von der Bildfunktion zu Bildbereichs-Hinweisen

Unter Bildbereichs-Hinweisen versteht Kanade die 2-D-Merkmale eines Bildes; homogene Bildregionen und Kanten sind Beispiele für Bildbereichs-Hinweise. Wenn Modelle der Bildfunktion zur Gewinnung von Bildbereichs-Hinweisen erstellt werden, müssen diese aus der gewählten Realisation des Modells leicht entnehmbar sein. Die in den Kapiteln 2 und 3 vorgestellten Regionsmodelle und Kantenmodelle der Bildfunktion vermitteln diesen Übergang zu den Bildbereichs-Hinweisen. Regionale Segmentationsverfahren bestimmen homogene Teilbereiche der Bildfunktion durch Zusammenfassung von Bildpunkten gleicher oder ähnlicher Intensität oder aus Messungen von in den Teilgebieten abgeleiteten Eigenschaften. Ein Vergleich von aus gleichen Bildvorlagen gewonnenen Segmentationen zeigt jedoch große Unterschiede in den Resultaten. Ein möglicher Grund für das unbefriedigende Verhalten regionaler Segmentierer kann sein, daß beim Übergang von der Bildfunktion zu den Bildbereichs-Hinweisen die Forderung nach einer **vollständigen** Bildzerlegung noch nicht erfüllt werden kann. Die Betrachtung der Ergebnisse von mit Kantenmodellen arbeitenden Verfahren unterstützt diese Behauptung. Unabhängig vom gewählten Modell des Kantentyps generieren die kantenorientierten Segmentationsverfahren nach

Schwellwert- oder Verdünnungsoperationen nur für eine geringe Zahl von Bildkoordinaten Eintragungen in die resultierenden binären Kantenmatrizen. Zu diesem Zeitpunkt der Verarbeitung werden weder Fortsetzbarkeit der Kantenelemente noch Geschlossenheit von aus benachbarten Kantenelementen bestehenden Ketten gefordert. Ein subjektiver 'Vergleich durch Betrachten' von nicht mit dem Originalbild unterlegten, aber vollständig segmentierten Konturbildern eines regional arbeitenden Verfahrens mit den binären Kantenmasken würde vermutlich letztere als bessere Bildrepräsentationen bewerten.

9.2 Von Bildbereichs-Hinweisen zu Szenenbereichs-Hinweisen

Szenenbereichs-Hinweise sind solche Eigenschaften der Szene, die eine Ursache der Bildbereichs-Hinweise sein können, zum Beispiel Objektkanten, Oberflächenorientierungen, Reflektanz, Beleuchtungsverhältnisse und Abstand vom Beobachter. Zur Gewinnung von Szenenbereichs-Hinweisen sind die Modelle der Bildfunktion allein nicht ausreichend, es werden zusätzliche Modelle benötigt. Eine Gruppe ließe sich als 'Modelle von Bildbereichs-Hinweisen' bezeichnen. Beispiele hierfür sind Verfahren, die mit Hilfe von Regeln über die Fortsetzbarkeit von Kantenstücken Schlüsse auf mögliche Kantenkonfigurationen ziehen (Waltz 75, Kanade 80b). Die zweite Gruppe umfaßt die Modellierung der physikalischen Gesetze bei der Beleuchtung von Objekten und des daraus resultierenden Intensitätsverlaufs im Bild (Abbildungsmodell). Zu dieser Gruppe gehören die auf der 'Shape-from-shading'-Theorie von Horn aufbauenden Verfahren (Horn 75, Horn 77). Bei Kenntnis der Richtungen von Lichtquelle und Beobachter sowie für ein bekanntes Reflexionsgesetz auf den Objektflächen können aus Intensitätsverläufen im Bild die Oberflächenorientierungen berechnet werden. Zur Gewinnung von Szenenbereichs-Hinweisen ist es nicht ausreichend, Modelle über

Lichtquelle, Beobachter und Reflexionsverhalten zu erstellen. Erst die Einbeziehung von Teilausprägungen dieser Modelle erlaubt die Bestimmung von Orientierung, Reflektanz oder Objektkontur.

Das von Barrow und Tenenbaum vorgeschlagene Arbeiten mit 'innewohnenden Bildern' (intrinsic images) zeigt deutliche Parallelen zu den 'Shape-from-Shading'-Verfahren. Ihre Attributierung von Pixeln mit Kanten-, Regions- und Segmenteigenschaften könnte um weitere Attribute wie Orientierung und Albedo erweitert werden (Barrow and Tenenbaum 78). Während Barrow und Tenenbaum nur eine Korrespondenz von Pixeln zu Oberflächenelementen der Szene voraussetzen, setzt Woodham zusätzliche Annahmen über Stetigkeit von Oberflächen und Korrespondenzen zwischen Pixeln und Punkten im Gradientenraum ein (Woodham 77, Woodham 80). Als weiterer Szenenbereichs-Hinweis läßt sich der Abstand vom Beobachter aus der Veränderung der Texturdichte oder der Größe von Texturprimitiven erschließen (Bajcsy 73a, Bajcsy and Lieberman 74, Bajcsy and Lieberman 76) (Kanade and Kender 80).

Auch bei den in diesen Abschnitten beschriebenen Verfahren ist die Zuordnung von Szenenbereichs-Hinweisen zu allen Pixeln nicht das primäre Ziel. Wichtiger ist die vollständige Auswertung der insgesamt zur Verfügung stehenden Information. Auch auf dieser Stufe der Bilddeutung ist die Durchführbarkeit einer vollständigen Segmentation fraglich.

9.3 Von Szenenbereichs-Hinweisen zum Szenenmodell

Ein Szenenmodell besteht aus einer Szenenbeschreibung, die Beschreibungen der Objekte sowie Relationen zwischen Objekten bzw. Objektteilen enthält. A priori Wissen über die in der Szene erwarteten Szenenkomponenten wird im Szenenmodell durch explizite Konstruktion eines Szenenprototyps für die zu beschreibenden

Bilder als 'generische Szenenbeschreibung' in den Deutungszyklus eingebracht (siehe Kap. 7.2). Das Szenenmodell enthält ebenfalls generische Beschreibungen des bilderzeugenden Systems und der Szenenbeleuchtung, deren Ausprägungen schon zur Gewinnung der Szenenbereichs-Hinweise benutzt wurden (Nagel 79).

Obwohl nur wenige Autoren bei der Bildsegmentation und Bildinterpretation bis zu 3-D-Modellen gelangen, ist die wachsende Bedeutung solcher Modelle heute schon erkennbar. Besondere Bedeutung werden Szenenmodelle insbesondere bei der Analyse räumlicher (Shani 80) und zeitlicher Bildfolgen erhalten.

Wenn geeignete Strukturen zur rechner-internen Repräsentation von Szenenmodellen auch außerhalb der eigentlichen Analyseprogramme in Form von Modelldaten zur Verfügung stehen, kann durch deren fortlaufende Erweiterung möglicherweise ein 'Weltausschnittsmodell' aufgebaut werden. Vielleicht gelingt es dann, eine größere Klasse von Szenen mit einem einzigen Programm zu analysieren, sodaß die Vielzahl der heute bestehenden Segmentations- und Interpretationsprogramme auf wenige erprobte und verstandene Programme reduziert werden kann.

9.4 Vervollständigung des Deutungszyklusses

Die noch nicht behandelten Schritte beim Durchlaufen eines vollständigen Deutungszyklusses sind die Gewinnung einer Ausprägung des Szenenmodells (das instantiierte Modell), das Erzeugen einer 2-D-Szenenprojektion (die Szenenskizze) und die Erstellung eines synthetischen Bildes mit Hilfe der Szenenprojektion.

Die Synthese eines Bildes aus einem Objekt- oder Szenenmodell wurde bisher überwiegend dem Gebiet der 'Computer-Graphics' zugeordnet. Im Verlauf der letzten Jahre haben sich jedoch deutliche Parallelen bei der Entwicklung von Modellen der

Bildfunktion (Hsu et al. 78), (Enamoto and Katayama 76), der Szenenbereichs-Hinweise (Fishman and Schachter 80), (Horn and Sjoberg 79) und der Szenenmodelle (Badler and Bajcsy 78) ergeben. Es bleibt abzuwarten, ob sich diese Entwicklung bis zu gleichen oder zumindest ähnlichen Modellen bei Computer-Graphics-Methoden und in der Bildanalyse fortsetzt.

Die Szenenprojektion als Zwischenschritt zwischen Szenenmodell und synthetischem Bild enthält eine abstrakte Beschreibung der Szenenausprägung bei speziellen Bedingungen für Blickrichtung und Beleuchtung. Sie wird durch zusätzliche Beschreibungen der Bedeutung von Kanten und Regionen ergänzt. Die formale Ähnlichkeit der sowohl in der Szenenprojektion wie auch der in den Bildbereichs-Hinweisen enthaltenen Information deutet ebenfalls auf die Benutzung ähnlicher Modelle hin.

Der Vergleich des in einem Durchlauf des Deutungszyklusses gewonnenen synthetischen Bildes mit dem Originalbild liefert Hinweise auf Gebiete der Bildfunktion, in denen unvollständige Modelle oder Ausprägungen von Modellen benutzt wurden. Wenn es gelingt, das diese Inkonsistenz verursachende Modell zu identifizieren und dieses Modell oder die verwendete Ausprägung zu modifizieren, wird jeder weitere Durchlauf des Deutungszyklusses zu einem verbesserten Gesamtmodell führen, sodaß eine vollständige Bildsegmentation und Bildinterpretation möglich wird. Dieser Weg wird in den kommenden Jahren zu beschreiten sein.

LITERATURVERZEICHNIS

Abkürzungen:

IJCAI : International Joint Conference on
Artificial Intelligence

IJCAI-69 : Washington/DC, USA, May 7-9, 1969

IJCAI-71 : London, UK, Sept. 1-3, 1971

IJCAI-73 : Stanford/CA, USA, Aug. 20-23, 1973

IJCAI-75 : Tblisi/Georgia, USSR, Sept. 3-8, 1975

IJCAI-77 : Cambridge/MA, USA, Aug. 22-25, 1977

IJCAI-79 : Tokyo, Japan, Aug. 20-23, 1979

IJCPR : International Joint Conference on
Pattern Recognition

IJCPR-73 : Washington/DC, USA, Oct. 30 - Nov. 1, 1973

IJCPR-74 : Copenhagen, Denmark, Aug. 13-15, 1974

IJCPR-76 : Coronado/CA, USA, Nov. 8-11, 1976

IJCPR-78 : Kyoto, Japan, Nov 7-10, 1978

Abdou and Pratt 79

Quantitative Design and Evaluation of Enhancement/
Thresholding Edge Detectors

I.E. Abdou, W.K. Pratt

Proc. IEEE 67 (1979) 753-763

Aggarwal and Duda 75

Computer Analysis of Moving Polygonal Images

J.K. Aggarwal, R.O. Duda

IEEE Transactions on Computers 23 (1975) 966-976

Ahuja and Rosenfeld 78

A Note on the Use of Second-Order Gray-Level
Statistics for Threshold Selection

N. Ahuja, A. Rosenfeld

IEEE Transactions on Systems, Man, and Cybernetics
8 (1978) 895-898

Ando and Doi 78

Region Segmentation and Classification by Textural
Features

S. Ando, Y. Doi

Proc. 4th IJCPR (1978) 553-555

Badler and Bajcsy 78

Three-dimensional Representations for Computer
Graphics and Computer Vision

N. Badler, R. Bajcsy

Computer Graphics 12 , No. 3 (1978) 153-160

Bajcsy and Lieberman 74

Computer Description of Real Outdoor Scenes

R. Bajcsy, L.I. Lieberman

Proc. 2nd IJCPR (1974) 174-179

Bajcsy and Lieberman 76

Texture Gradient as a Depth Cue

R. Bajcsy, L.I. Lieberman

Computer Graphics and Image Processing 5 (1976)
52-67

- Bajcsy 75a
 Computer Description of Textured Surfaces
 R. Bajcsy
 Proc. 3rd IJCAI (1973) 572-579
- Bajcsy 75b
 Computer Identification of Visual Surfaces
 R. Bajcsy
 Computer Graphics and Image Processing 2 (1973)
 118-130
- Ballard 78
 Model-directed Detection of Ribs in Chest
 Radiographs
 D.H. Ballard
 Proc. 4th IJCPR (1978) 907-910
- Barnard and Thompson 78
 Visual Disparity as a Cue for Depth and Velocity
 in Real World Scenes
 S.T. Barnard, W.B. Thompson
 IEEE Conference on Pattern Recognition and Image
 Processing Chicago/Ill. (May 31 - June 2 1978)
 402-404
- Barrow and Tenenbaum 78
 Recovering Intrinsic Scene Characteristics from
 Images
 H.G. Barrow, J.M. Tenenbaum
 In A.R. Hanson, E.M. Riseman (eds.)
 Computer Vision Systems
 Academic Press, New York (1978) 3-26
- Bastien and Dunn 71
 Global Transformations in Pattern Recognition of
 Bubble Chamber Photographs
 P. Bastien, L. Dunn
 IEEE Transactions on Computers 20 (1971) 995-1001
- Batali and Ullman 79
 Motion Detection and Analysis
 J. Batali, S. Ullman
 In L.S. Baumann (ed.)
 Proc. Image Understanding Workshop
 Science Applications, Inc., Arlington/VA 22209,
 Los Angeles/CA (1979) 69-75
- Bazin and Benoit 65
 Off-line Global Approach to Pattern Recognition
 for Bubble Chamber Pictures
 M. J. Bazin, J. W. Benoit
 IEEE Trans. Nucl. Sci. 12 (1965) 291
- Beaudet 78
 Rotationally Invariant Image Operators
 P.R. Beaudet
 Proc. 4th IJCPR (1978) 579-583
- Bertelsmeier and Radig 77a
 Kontextunterstuetzte Analyse von Szenen mit
 bewegten Objekten
 R. Bertelsmeier, B. Radig

- In H.-H. Nagel (ed.)
GI/NTG Fachtagung Digitale Bildverarbeitung
Springer, Berlin-Heidelberg-New York(1977) 101-128
- Bertelsmeier and Radig 77b
Context-Guided Analysis of Scenes with Moving
Objects
R. Bertelsmeier, B. Radig
Technical Report IfI-HH-B-41/77, University
Hamburg (1977)
- Bonde and Nagel 79
Deriving a 3-D Description of a Moving Rigid
Object from Monocular TV-Frame Sequences
T. Bonde, H.-H. Nagel
IEEE Workshop on Computer Analysis of Time-Varying
Imagery (1979) 44-45
- Brice and Fennema 70
Scene Analysis Using Regions
C.R. Brice, C.L. Fennema
Artificial Intelligence 1 (1970) 205-226
- Brooks et al. 79a
Progress Report on a Model-Based Vision System
R.A. Brooks, R. Greiner, T.O. Binford
In: R. Bajcsy, (ed.)
Proc. NSF Workshop on the Representation of
Three-Dimensional Objects
(May 1-2 1979)
- Brooks et al. 79b
The ACRONYM Model-Based Vision System
R.A. Brooks, R. Greiner, T.O. Binford
IJCAI (1979) 105-113
- Brooks 78
Rationalizing Edge Detectors
M.J. Brooks
Computer Graphics and Image Processing 8 (1978)
277-285
- Bullock 76
Finding Structure in Outdoor Scenes
B.L. Bullock
In C.H. Chen (ed.)
Pattern Recognition and Artificial Intelligence
Academic Press, New York (1976) 61-85
- Burr and Chien 76
The Minimal Spanning Tree in Visual Data
Segmentation
D.J. Burr, R.T. Chien
Proc. 3rd IJCP (1976) 519-523
- Chen and Pavlidis 78
Segmentation by Texture Using a Co-Occurrence
Matrix and a Split-And-Merge Algorithm
C.H. Chen, T. Pavlidis
Proc. 4th IJCP (1978) 565-569
- Chen and Pavlidis 80
Image Segmentation as an Estimation Problem

- P.C. Chen, T. Pavlidis
Computer Graphics and Image Processing 12 (1980)
153-172
- Chow and Kaneko 72
Automatic Boundary Detection of the Left Ventricle
from Cineangiograms
C.K. Chow, T. Kaneko
Computers and Biomedical Research 5 (1972) 388-410
- Coleman and Andrews 79
Image Segmentation by Clustering
G.B. Coleman, H.C. Andrews
Proc. IEEE 67 (1979) 773-785
- Cooper 79
Maximum Likelihood Estimation of Markov-Process
Blob Boundaries in Noisy Images
D. B. Cooper
IEEE Transactions on Pattern Analysis and Machine
Intelligence 1 (1979) 372-384
- Coray et al. 75
Order Independence in Local Clustering Algorithms
G. Coray, A. Noetzel, S.M.Selkow
Computer Graphics and Image Processing 4 (1975)
120-132
- Danker and Rosenfeld 80
Strip Detection Using Relaxation
A. Danker, A. Rosenfeld
Pattern Recognition 12 , No. 2 (1980) 97-100
- Davis and Mitiche 80
Edge Detection in Textures
L.S. Davis, A. Mitiche
Computer Graphics and Image Processing 12 (1980)
25-39
- Davis 75
A Survey of Edge Detection Techniques
L.S. Davis
Computer Graphics and Image Processing 4 (1975)
248-270
- Deutsch and Belnkamp 72
Texture Descriptors Using Neighborhood Information
E.S. Deutsch, N.J. Belnkamp
Computer Graphics and Image Processing 1 (1972)
145-168
- Deutsch and Fram 78
A Quantitative Study of the Orientation Bias of
Some Edge Detector Schemes
E.S. Deutsch, J.R. Fram
IEEE Transactions on Computers 27 (1978) 205-213
- Dreschler and Nagel 78
Using 'Affinity' for Extracting Images of Moving
Objects from TV-Frame Sequences
L. Dreschler, H.-H. Nagel
Technical Report IfI-HH-B-44/78, Fachbereich

- Informatik, Universitaet Hamburg (February 1978)
- Duda and Hart 72
Use of the Hough-Transformation to Detect Lines
and Curves in Pictures
R.O. Duda, P.E. Hart
Communications of the ACM 15 (1972) 11
- Duda and Hart 73
Pattern Classification and Scene Analysis
R.O. Duda, P.E. Hart
John Wiley & Sons , New York (1973)
- Dudani and Luk 77
Locating Straight-Line Edge Segments on Outdoor
Scenes
S.A. Dudani, A.L. Luk
Proc. IEEE Conference on Pattern Recognition and
Image Processing, Troy/NY (June 6-8 1977)
- Ehrich and Foith 78
Topology and Semantics of Intensity Arrays
R.W. Ehrich, J.P. Foith
In A.R. Hanson, E.M. Riseman (eds.)
Computer Vision Systems
Academic Press, New York (1978) 111-127
- Ehrich 77
Detection of Global Edges in Textured Images
R.W. Ehrich
IEEE Transactions on Computers 26 (1977) 589-603
- Elliot et al. 79
Implementation, Interpretation and Analysis of a
Suboptimal Boundary Finding Algorithm
H. Elliot, D.E. Cooper, P. Symosek
IEEE Conf. PRIP, Chicago/Ill. (August 6-8 1979)
122-129
- Enamoto and Katayama 76
Structure Lines of Images
H. Enamoto, T. Katayama
Proc. 3rd IJCPR (1976) 811-815
- Ernst et al. 76
Processing of Remote Sensing Data by a Region
Growing Algorithm
D. Ernst, B. Barger, F. Holdermann
Proc. 3rd IJCPR (1976) 679-683
- Feldman and Yakimovsky 74
Decision Theory and Artificial Intelligence I. A
Semantics-based Region Analyser
J.A. Feldman, Y. Yakimovsky
Artificial Intelligence 5 (1974) 349-371
- Fischler 78
On the Representation of Natural Scenes
M.A. Fischler
In A.R. Hanson, E.M. Riseman (eds.)
Computer Vision Systems
Academic Press, New York (1978) 47-52
- Fishman and Schachter 80

- Computer Display of Height Fields
 B. Fishman, B. Schachter
 Computers & Graphics 5 , No. 2-4 (1980) 53-60
- Fram and Deutsch 75
 On the Quantitative Evaluation of Edge Detection
 Schemes and their Comparison with Human
 Performance
 J.R. Fram, E.S. Deutsch
 IEEE Transactions on Computers 24 (1975) 616-628
- Frei and Chen 77
 Fast Boundary Detection: a Generalization and a
 New Algorithm
 W. Frei, C.C. Chen
 IEEE Transactions on Computers 26 (1977) 988-998
- Freuder 76
 Affinity: A Relative Approach to Region Finding
 E.C. Freuder
 Computer Graphics and Image Processing 5 (1976)
 254-264
- Frieden 80
 Statistical Models for the Image Restoration
 Problem
 B.R. Frieden
 Computer Graphics and Image Processing 12 (1980)
 40-59
- Galloway 75
 Texture Analysis Using Gray Level Run Lengths
 Mary M. Galloway
 Computer Graphics and Image Processing 4 (1975)
 172-179
- Griffith 71
 Mathematical Models For Automatic Line Detection
 A.K. Griffith
 2nd IJCAI (1971) 17-26
- Groch 80
 Automatisierung der Extraktion linienhafter
 Objekte aus Grauwertbildern
 W.D. Groch
 PhD Thesis, University Karlsruhe (1980)
- Gupta and Wintz 75
 A Boundary Finding Algorithm and its Application
 J.N. Gupta, P.A. Wintz
 IEEE Transactions on Circuits and Systems 22
 (1975) 351-362
- Haralick et al. 73
 Textural Features for Image Classification
 R. Haralick, K. Shanmugan, I. Dinstein
 IEEE Transactions on Systems, Man, and Cybernetics
 3 (1973) 610-621
- Haralick 79
 Statistical and Structural Approaches to Texture
 R.M. Haralick
 Proc. IEEE 67 (1979) 786-804

- Haralick 80
 Edge and Region Analysis for Digital Image Data
 R.M. Haralick
 Computer Graphics and Image Processing 12 (1980)
 60-73
- Hayes et al. 74
 Texture Coarseness: Further Experiments
 K.C. Hayes, Jr., A.N. Shah, A. Rosenfeld
 IEEE Transactions on Systems, Man, and Cybernetics
 4 (1974) 467-472
- Herskovits and Binford 70
 On Boundary Detection
 A. Herskovits, T.O. Binford
 MIT AI Memo 183, Cambridge, Mass. (1970)
- Holdermann and Kazmierczak 72
 Preprocessing of Gray-Scale Pictures
 F. Holdermann, H. Kazmierczak
 Computer Graphics and Image Processing 1 (1972)
 66-80
- Horn and Sjoberg 79
 Calculating the Reflectance Map
 B. K. P. Horn, R. W. Sjoberg
 Applied Optics 18 (1979) 1770-1779
- Horn 75
 Obtaining Shape from Shading Information
 B. K. P. Horn
 In P. H. Winston (ed.)
 The Psychology of Computer Vision
 McGraw Hill(1975)
- Horn 77
 Understanding Image Intensities
 B.K.P. Horn
 Artificial Intelligence 8 (1977) 201-231
- Horowitz and Pavlidis 74
 Picture Segmentation by a Directed Split-and-Merge
 Procedure
 S.L. Horowitz, T. Pavlidis
 Proc. 2nd IJ CPR (1974) 424-433
- Hsu et al. 78
 Web Representation of Image Data
 S. Hsu, J.L. Mundy, P.R. Beaudet
 Proc. 4th IJ CPR (1978) 675-680
- Hueckel 71
 An Operator which Locates Edges in Digitized
 Pictures
 M. Hueckel
 Journal of the ACM 18 (1971) 113-125
- Hueckel 75
 A Local Visual Operator which Recognizes Edges and
 Lines
 M. Hueckel
 Journal of the ACM 20 (1975) 634-647
- Hummel 79

- Feature Detection Using Basic Functions
 R.A. Hummel
 Computer Graphics and Image Processing 9 (1979)
 40-55
- Iannino and Shapiro 79
 An Iterative Generalization of the Sobel Edge
 Detection Operator
 A. Iannino, S.D. Shapiro
 IEEE Conf. PRIP, Chicago/Ill. (1979) 130-137
- Jain and Nagel 79
 On the Analysis of Accumulative Difference
 Pictures from Image Sequences of Real World Scenes
 R. Jain, H.-H. Nagel
 IEEE Transactions on Pattern Analysis and Machine
 Intelligence 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, H.-H. Nagel
 IJCAI (1977) 612-618
- Johnston and Rosenfeld 75
 Digital Detection of Pits, Peaks, Ridges, and
 Ravines
 E.G. Johnston, A. Rosenfeld
 IEEE Transactions on Systems, Man, and Cybernetics
 5 (1975) 472-480
- Kanade and Kender 80
 Mapping Image Properties into Shape Constraints:
 Skewed Symmetry, Affine-Transformable Patterns,
 and the Shape-from-Texture Paradigm
 T. Kanade, R. Kender
 Technical Report , Carnegie Mellon University
 (1980)
- Kanade 80a
 Region Segmentation: Signal vs. Semantics
 T. Kanade
 Computer Graphics and Image Processing 13 , No. 4
 (1980) 279-297
- Kanade 80b
 A Theory of Origami World
 T. Kanade
 Artificial Intelligence 13 (1980) 279-311
- Kimme et al. 75
 Finding Circles by an Array of Accumulators
 C. Kimme, D. Ballard, J. Sklansky
 Communications of the ACM 18 (1975) 120
- Kirby and Rosenfeld 79
 A Note on the Use of (Gray Level, Local Average
 Gray Level) Space as an Aid in Threshold Selection
 R.L. Kirby, A. Rosenfeld
 IEEE Transactions on Systems, Man, and Cybernetics
 9 (1979) 860-864
- Kirsch 71

- Computer Determination of the Constituent
Structure of Biological Images
R.A. Kirsch
Computers and Biomedical Research 4 (1971) 315-328
- Klaus and Buhr 73
Woerterbuch der Philosophie
G. Klaus, M. Buhr
Rowohlt Taschenbuch Verlag (1973)
- Kraasch and Zach 78
SERF - Eine Untersuchung zur Segmentation und
symbolischer Beschreibung von Fernseh Bildern
R. Kraasch, W. Zach
Technical Report IfI-HH-M-59/78, Fachbereich
Informatik, Univ. Hamburg (1978)
- Kraasch et al. 79a
Automatic Three-Dimensional Description of Simple
Moving Objects
R. Kraasch, B. Radig, W. Zach
IEEE Workshop on Computer Analysis of Time-Varying
Imagery (1979)
- Kraasch et al. 79b
Automatische dreidimensionale Beschreibung
bewegter Gegenstaende
R. Kraasch, B. Radig, W. Zach
In J. P. Foith (ed.)
Angewandte Szenenanalyse, Informatik Fachberichte
20
Springer, Berlin-Heidelberg-New York(1979) 208-215
- Lai and Ehrich 79
Segmentation of Images with Incompletely Specified
Regions
P.F. Lai, R.W. Ehrich
IEEE Transactions on Systems, Man, and Cybernetics
9 (1979) 864-868
- Levine and Leemet 76
A Method for Non-Purposive Picture Segmentation
M.D. Levine, J. Leemet
Proc. 3rd IJCP (1976) 494-498
- Levine and Shaheen 79
A Modular Computer Vision System for Picture
Segmentation and Interpretation
M.D. Levine, S.I. Shaheen
IEEE Conf. PRIP (1979) 523-539
- Mackworth 78
Vision Research Strategy: Black Magic, Metaphors,
Mechanisms, Miniworlds and Maps
A. K. Mackworth
In A. R. Hanson, E.M. Riseman (eds.)
Computer Vision Systems
Academic Press, New York(1978) 53-59
- Marr and Hildreth 80
Theory of Edge Detection

- D. Marr, E. Hildreth
Proc. Royal Society of London B 207 (1980) 187-217
- Marr and Ullman 79
Directional Selectivity and its Use in Early
Visual Processing
D. Marr, S. Ullman
Technical Report AI Memo 524, MIT, Cambridge MA
(June 1979)
- Martelli 72
Edge Detection Using Heuristic Search Methods
A. Martelli
Computer Graphics and Image Processing 1 (1972)
169-182
- Mero and Vassy 75
A Simplified and Fast Version of the Hueckel
Operator for Finding Optimal Edges in Pictures
L. Mero, Z. Vassy
4th IJCAI (1975) 650-655
- Milgram and Kahl 79
Recursive Region Extraction
D.L. Milgram, D.J.Kahl
Computer Graphics and Image Processing 9 (1979)
82-88
- Milgram 78
Results in FLIR Target Detection and
Classification
D.L. Milgram
in: Proc. Image Understanding Workshop, Cambridge
(May 1978) 118-124
- Milgram 79
Region Extraction Using Convergent Evidence
D.L. Milgram
Computer Graphics and Image Processing 11 (1979)
1-12
- Minsky 77
Materie, Geist, Modell
M. L. Minsky
In P.Eisenberg (ed.)
Semantik und kuenstliche Intelligenz
de Gruyter(1977) 11-20
- Mitchell and Carlton 78
Image Segmentation Using a Local Extrema Texture
Measure
O.R. Mitchell, S.G. Carlton
Pattern Recognition 10 (1978) 205-210
- Moravec 77
Towards Automatic Visual Obstacle Avoidance
H.P. Moravec
IJCAI (1977) 584
- Muerle and Allen 68
Experimental Evaluation of Techniques for
Automatic Segmentation of Objects in a Complex

- Scene
 J.L. Muerle, D.C. Allen
 In Cheng et al. (eds.)
 Pictorial Pattern Recognition
 Thompson, Washington, DC (1968) 3-13
- Nagao et al. 76
 Agricultural Land Use Classification of Aerial
 Photographs by Histogram Similarity Methods
 M. Nagao, H. Tanabe, K. Ito
 Proc. 3rd IJCPR (1976) 669-672
- Nagao et al. 79
 Region Extraction and Shape Analysis in Aerial
 Photographs
 M. Nagao, T. Matsuyama, Y. Ikeda
 Computer Graphics and Image Processing 10 (1979)
 195-223
- Nagel 78
 Analysis Techniques for Image Sequences
 H.-H. Nagel
 Proc. 4th IJCPR (1978) 186-211
- Nagel 79
 Ueber die Repraesentation von Wissen zur
 Auswertung von Bildern
 H.-H. Nagel
 In J.P. Foith (ed.)
 Angewandte Szenenanalyse
 Springer Verlag, Berlin-Heidelberg-New York(1979)
 3-21
- Nakagawa and Rosenfeld 78
 Edge/Border Coincidence as an Aid in Edge
 Extraction
 Y. Nakagawa, A. Rosenfeld
 IEEE Transactions on Systems, Man, and Cybernetics
 8 (1978) 899-901
- Nakagawa and Rosenfeld 79
 Some Experiments on Variable Thresholding
 Y. Nakagawa, A. Rosenfeld
 Pattern Recognition 11 (1979) 191-204
- Nevatia and Babu 79
 Linear Feature Extraction and Description
 R. Nevatia, K.R. Babu
 IJCAI (1979) 639-641
- Nevatia and Babu 80
 Linear Feature Extraction and Description
 R. Nevatia, K.R. Babu
 Computer Graphics and Image Processing 13 , No. 3
 (1980) 257-269
- Nevatia and Price 78
 Locating Structures in Aerial Images
 R. Nevatia, K. Price
 Proc. 4th IJCPR (1978) 686-690
- Nevatia 76
 Locating Object Boundaries in Textured

- Environments
R. Nevatia
IEEE Transactions on Computers 25 (1976) 1170-1175
- Nevatia 77
Evaluation of a Simplified Hueckel Edge-Line
Detector
R. Nevatia
Computer Graphics and Image Processing 6 (1977)
582-588
- Noguchi et al. 78
A Method for Segmenting a Clump of Cells into
Cellular Characteristic Parts Using Multispectral
Informations
Y. Noguchi, Y. Tenjin, T. Sugishita
Proc. 4th IJCPR (1978) 872-874
- O'Gorman and Clowes 73
Finding Picture Edges Through Collinearity of
Feature Points
F. O'Gorman, M.B. Clowes
3rd IJCAI (1973) 543-555
- O'Gorman and Clowes 76
Finding Picture Edges Through Collinearity of
Feature Points
F. O'Gorman, M.B. Clowes
IEEE Transactions on Computers 25 (1976) 449-456
- Ohlander et al. 78
Picture Segmentation Using a Recursive Region
Splitting Method
R. Ohlander, K. Price, D.R. Reddy
Computer Graphics and Image Processing 8 (1978)
313-333
- Ohlander 75
Analysis of Natural Scenes
R.B. Ohlander
PhD Thesis, Comp.Sc. Dept., Carnegie-Mellon Univ.,
Pittsburgh/Pa. (April 1975)
- Ohta et al. 78
An Analysis System for Scenes Containing Objects
with Substructures
Y. Ohta, T. Kanade, T. Sakai
Proc. 4th IJCPR (1978) 752-754
- Oshima and Shirai 79
A Scene Description Method Using Three-Dimensional
Information
M. Oshima, Y. Shirai
Pattern Recognition 11 (1979) 9-17
- Panda and Rosenfeld 78
Image Segmentation by Pixel Classification in
(Gray Level, Edge Value) Space
D.P. Panda, A. Rosenfeld
IEEE Transactions on Computers 27 (1978) 875-879
- Paton 75

- Picture Description Using Legendre Polynomials
K. Paton
Computer Graphics and Image Processing 4 (1975)
40-54
- Paton 78
Detection of Thin Line Segments
K. Paton
Proc. 4th IJ CPR (1978) 704-706
- Paton 79
Line Detection by Local Methods
K. Paton
Computer Graphics and Image Processing 9 (1979)
316-332
- Pavlidis 72
Segmentation of Pictures and Maps Through
Functional Approximation
T. Pavlidis
Computer Graphics and Image Processing 1 (1972)
360-372
- Pavlidis 73
Waveform Segmentation Through Functional
Approximation
T. Pavlidis
IEEE Transactions on Computers 22 (1973) 689-697
- Pavlidis 79
Hierarchies in Structural Pattern Recognition
T. Pavlidis
Proc. IEEE 67, No. 5 (1979) 737-744
- Perkins and Binford 75
A Corner Finder for Visual Feedback
W.A. Perkins, T.O. Binford
Computer Graphics and Image Processing 2 (1973)
355-374
- Perkins 76
A Multilevel Vision Recognition System
W.A. Perkins
Proc. 3rd IJ CPR (1976) 739-744
- Perkins 78
A Model-Based Vision System for Industrial Parts
W.A. Perkins
IEEE Transactions on Computers 27 (1978) 126-143
- Persoon 76
A New Edge Detection Algorithm and Its Application
in Picture Processing
E. Persoon
Computer Graphics and Image Processing 5 (1976)
425-446
- Pingle and Tenenbaum 71
An Accommodating Edge Follower
K.K. Pingle, J.M. Tenenbaum
Proc. 2nd IJCAI (1971) 1-7
- Popplestone et al. 75
Forming Models of Plane-and-Cylinder Faceted

- Bodies from Light Stripes
R. Popplestone, C. Brown, A. Ambler, G. Crawford
Proc. 4th IJCAI (1975) 664-668
- Potter 75a
Velocity as a Cue to Segmentation
J. Potter
IEEE Transactions on Systems, Man, and Cybernetics
5 (1975) 390-394
- Potter 75b
Scene Segmentation by Velocity Measurements
Obtained with a Cross-Shaped Template
J. Potter
Proc. 4th IJCAI (1975) 803-810
- Potter 77
Scene Segmentation Using Motion Information
J.L. Potter
Computer Graphics and Image Processing 6 (1977)
558-581
- Prager 80
Extracting and Labeling Boundary Segments in
Natural Scenes
J.M. Prager
IEEE Transactions on Pattern Analysis and Machine
Intelligence 2 (1980) 16-27
- Preparata and Ray 72
An Approach to Artificial Nonsymbolic Cognition
F.R. Preparata, S.R. Ray
Information Sciences 4 (1972) 65-86
- Prewitt 70
Object Enhancement and Extraction
J.M. Prewitt
In B.S. Lipkin, A. Rosenfeld (eds.)
Picture Processing and Psychopictorics
Academic Press, New York (1970) 75-149
- Radig et al 80
Matching Symbolic Descriptions for 3-D
Reconstruction of Simple Moving Objects
B. Radig, R. Kraasch, W. Zach
Proc. 5th ICPR (1980)
- Radig 78
Auswertung von digitisierten Fernsehbildern zur
Beschreibung bewegter Objekte
B. Radig
PhD Thesis, Universitaet Hamburg (1978)
and Report IFI-HH-B-59/79
- Ramer 75
Extraction of Line Structures from Photographs of
Curved Objects
U. Ramer
Computer Graphics and Image Processing 4 (1975)
81-103
- Ridler and Calvard 78
Picture Thresholding Using an Iterative Selection

- Method
 T.W. Ridler, S. Calvard
 IEEE Transactions on Systems, Man, and Cybernetics
 8 (1978) 630-632
- Roach and Aggarwal 79
 Computer Tracking of Objects Moving in Space
 J. Roach, J.K. Aggarwal
 IEEE Transactions on Pattern Analysis and Machine
 Intelligence 1 (1979) 127-135
- Roberts 65
 Machine Perception of Three-Dimensional Solids
 L. G. Roberts
 in Tippet et al. (eds.)
 Optical and Electro-Optical Information Processing
 MIT Press, Cambridge/Mass.(1965) 159-197
- Robinson 77
 Edge Detection by Compass Gradient Masks
 G.S. Robinson
 Computer Graphics and Image Processing 6 (1977)
 492-501
- Roecker and Kiessling 75
 Methods for Analyzing Three Dimensional Scenes
 F. Roecker, A. Kiessling
 Proc. IJCAI (1975) 669-673
- Rosenfeld and Davis 79
 Image Segmentation and Image Models
 A. Rosenfeld, L.S. Davis
 Proc. IEEE 67 (1979) 764-772
- Rosenfeld and Kak 76
 Digital Picture Processing
 A. Rosenfeld, A.C. Kak
 Academic Press, New York (1976)
- Rosenfeld 75
 A Note on Automatic Detection of Texture Gradients
 A. Rosenfeld
 IEEE Transactions on Computers 24 (1975) 988-991
- Rosenfeld 81
 The Max Roberts Operator is a Hueckel-Type Edge
 Detector
 A. Rosenfeld
 IEEE Transactions on Pattern Analysis and Machine
 Intelligence 3 (1981) 101-103
- Schachter et al. 79
 Some Experiments in Image Segmentation by
 Clustering of Local Feature Values
 B.J. Schachter, L.S. Davis, A. Rosenfeld
 Pattern Recognition 11 (1979) 19-28
- Shafer 80
 MOOSE: User's Manual, Implementation Guide,
 Evaluation
 S. Shafer
 Technical Report Ifl-HH-B-70/80, University
 Hamburg (1980)

- Shani 80
 A 3-D Model-Driven System for the Recognition of
 Abdominal Anatomy from CT Scans
 U. Shani
 Proc. 5th IJCP (1980)
- Shapiro 75
 Transformations for the Computer Detection of
 Curves in Noisy Pictures
 S. Shapiro
 Computer Graphics and Image Processing 4 (1975)
 328-338
- Shapiro 78
 Feature Space Transforms for Curve Detection
 S.D. Shapiro
 Pattern Recognition 10 (1978) 129-143
- Shaw 79
 Local and Regional Edge Detectors: Some
 Comparisons
 G.B. Shaw
 Computer Graphics and Image Processing 9 (1979)
 135-149
- Shirai 73
 A Step Toward Context-Sensitive Recognition of
 Irregular Objects
 Y. Shirai
 Computer Graphics and Image Processing 2 (1973)
 298-307
- Sklansky 78
 Image Segmentation and Feature Extraction
 J. Sklansky
 IEEE Transactions on Systems, Man, and Cybernetics
 8 (1978) 237-247
- Sloan and Ballard 80
 Experience With the Generalized Hough Transform
 K.R. Sloan, D.H. Ballard
 Proc. 5th ICPR (1980) 174-179
- Somerville and Mundy 76
 One Pass Contouring of Images through Planar
 Approximation
 C. Somerville, J.L. Mundy
 Proc. 3rd IJCP (1976) 745-748
- Tenenbaum and Barrow 76
 IGS: A Paradigm for Integrating Image Segmentation
 and Interpretation
 J.M. Tenenbaum, H.G. Barrow
 In C.h. Chen (ed.)
 Pattern Recognition and Artificial Intelligence
 Academic Press, New York (1976) 472-507
- Tomita et al. 73
 Detection of Homogeneous Regions by Structural
 Analysis
 F. Tomita, M. Yachida, S. Tsuji
 Proc. 3rd IJCAI (1973) 564-571

- Tomita et al. 78
 Classification of Textures by a Structural Analysis
 F. Tomita, Y. Shirai, S. Tsuji
 Proc. 4th IJCPR (1978) 556-558
- Tsai and Fu 78
 Image Segmentation and Recognition by Texture Discrimination: A Syntactic Approach
 W.H. Tsai, K.S. Fu
 Proc. 4th IJCPR (1978) 560-564
- Tsuji and Tomita 73
 A Structural Analyzer for a Class of Textures
 S. Tsuji, F. Tomita
 Computer Graphics and Image Processing 2 (1973) 216-231
- Tsuji et al. 79
 Three Dimensional Movement Analysis of Dynamic Line Images
 S. Tsuji, M. Osada, M. Yachida
 Proc. 6th IJCAI (1979) 896-901
- Wall et al. 74
 Analysis of Image Histograms
 R.J. Wall, A. Klinger, K.R. Castleman
 Proc. 2nd IJCPR (1974) 341-344
- Waltz 75
 Understanding Line Drawings of Scenes with Shadows
 D. Waltz
 In P.H. Winston (ed.)
 The Psychology of Computer Vision
 McGraw-Hill, New York(1975)
- Wechsler and Sklansky 75
 Automatic Detection of Ribs in Chest Radiographs
 H. Wechsler, J. Sklansky
 Technical Report TP-75-1, Univ. of Calif. at Irvine (February 1975)
- Weszka and Rosenfeld 78
 Threshold Evaluation Techniques
 J.S. Weszka, A. Rosenfeld
 IEEE Transactions on Systems, Man, and Cybernetics 8 (1978) 622-629
- Weszka and Rosenfeld 79
 Histogram Modification for Threshold Selection
 J.S. Weszka, A. Rosenfeld
 IEEE Transactions on Systems, Man, and Cybernetics 9 (1979) 38-52
- Woodham 77
 A Cooperative Algorithm for Determining Surface Orientation from a Single View
 R. J. Woodham
 Proc. 5th IJCAI (1977) 635-641
- Woodham 80
 Photometric Method for Determining Surface Orientation from Multiple Images

- R.J. Woodham
Optical Engineering 19 (1980) 139-144
- Yachida et al. 79
Boundary Detection of Textured Regions
M. Yachida, M. Ikeda, S. Tsuji
Proc. 6th IJCAI (1979) 992-994
- Yakimovsky 75
Boundary and Object Detection in Real World Images
Y. Yakimovsky
Proc. 4th IJCAI (1975) 695-704
- Yakimovsky 76
Boundary and Object Detection in Real World Images
Y. Yakimovsky
Journal of the ACM 23 (1976) 599-618
- You and Fou 78
A Syntactic Approach to Shape Recognition
K.C. You, K.S. Fou
Proc. Image Understanding Workshop, Cambridge (May
1978) 105-112
- Zucker et al. 75
Picture Segmentation by Texture Discrimination
S. Zucker, A. Rosenfeld, L.S. Davis
IEEE Transactions on Computers 24 (1975) 1228-1233
- Zucker 76a
Region Growing: Childhood and Adolescence
S.W. Zucker
Computer Graphics and Image Processing 5 (1976)
382-399
- Zucker 76b
Relaxation Labelling and the Reduction of Local
Ambiguities
S.W. Zucker
Technical Report 451, Comp.Sc. Center, Univ. of
Maryland (1976)
- Zucker 76c
Toward A Model of Texture
S.W. Zucker
Computer Graphics and Image Processing 5 (1976)
190-202

Autorenverzeichnis

Abdou 37
 Aggarwal 86, 87
 Ahuja 49, 62
 Allen 21
 Ando 66
 Andrews 32

 Babu 53
 Badler 92
 Bajcsy 65, 68, 90, 92
 Ballard 60, 72
 Barnard 84
 Barrow 1, 35, 36, 40, 75, 90
 Bastien 60
 Batali 85
 Bazin 60
 Beaudet 42
 Belnkamp 63
 Benoit 60
 Bertelsmeier 77
 Binford 36, 38, 56
 Bonde 78
 Brice 19
 Brooks 42, 77
 Bullock 56
 Burr 33

 Calvard 20
 Carlton 65
 Castleman 20
 Chen 48, 52, 63
 Chien 33
 Chow 25
 Clowes 56
 Coleman 32
 Cooper 58
 Coray 22

 Danker 53
 Davis 15, 24, 37, 64, 69
 Deutsch 37, 63
 Doi 66
 Dreschler 82
 Duda 33, 55, 86
 Dudani 56
 Dunn 60

 Ehrich 33, 52, 64
 Elliot 58
 Enomoto 91
 Ernst 7, 22

 Feldman 50
 Feldmann 74
 Fennema 19
 Fischler 7
 Fishman 92
 Foith 64
 Fou 72

Fram 57
 Frei 52
 Freuder 20
 Frieden 14
 Fu 67

 Galloway 63
 Griffith 39, 58
 Groch 59
 Gupta 21

 Haralick 28, 42, 61
 Hart 33, 55
 Hayes 65
 Herskovits 36, 38
 Hildreth 43, 58
 Holdermann 27, 43, 52
 Horn 9, 38, 89, 92
 Horowitz 7, 19
 Hsu 91
 Hueckel 45, 52
 Hummel 45

 Iannino 46

 Jain 82, 85
 Johnston 52

 Kahl 50
 Kanade 6, 7, 76, 89, 90
 Kaneko 25
 Katayama 91
 Kazmierczak 27, 43, 52
 Kender 90
 Kießling 34
 Kimme 60
 Kirby 48
 Kirsch 42
 Klinger 20
 Kraasch 77

 Lai 33
 Leemet 31
 Levine 31, 78
 Lieberman 90
 Liebermann 68
 Luk 56

 Mackworth 7
 Marr 43, 58, 85
 Martelli 58
 Mero 45
 Milgram 50
 Minsky 10
 Mitchell 65
 Mitiche 64, 69
 Moravec 84
 Muerle 21
 Mundy 28

 Nagao 21, 30
 Nagel 70, 76, 78, 82, 86, 91

Nakagawa 25, 50
 Nevatia 45, 53, 57, 72
 Noguchi 33

 O'Gorman 56
 Ohlander 24, 31
 Ohta 76
 Oshima 33

 Panda 49
 Paton 28, 53
 Pavlidis 7, 19, 27, 48, 63, 74
 Perkins 56, 72
 Persoon 43
 Pingle 38
 Poppiestone 34
 Potter 81
 Prager 59
 Pratt 37
 Preparata 73
 Prewitt 41
 Price 72

 Radig 74, 77
 Ramer 57
 Ray 73
 Ridler 20
 Roach 87
 Roberts 41
 Robinson 42
 Rosenfeld 6, 24, 25, 42, 48, 49, 50, 52, 53, 62, 68
 Röcker 34

 Schachter 49, 92
 Shafer 24
 Shaheen 78
 Shapiro 46, 55, 60
 Shirai 28, 33, 36, 38
 Sjoberg 92
 Sklansky 60, 61
 Sloan 60
 Sobel 41
 Somerville 28

 Tenenbaum 1, 35, 36, 38, 40, 75, 90
 Thompson 84
 Tomita 66
 Tsai 67
 Tsuji 66, 87

 Ullman 85

 Vassy 45

 Wall 20
 Waltz 75, 89
 Wechsler 60
 Weska 49
 Weszka 24
 Wintz 21
 Woodnam 90

Yachida 69
Yakimovsky 50, 59, 74
You 72

Zach 77
Zucker 7, 15, 54, 65, 66