

Ein Ansatz zur Segmentation
von Farbbildern
durch hierarchische Ballungsanalyse

Carsten Schröder

FBI-HH-B-132/87

20. Oktober 1987

Universität Hamburg
Fachbereich Informatik
Bodenstedtstraße 16
D-2000 Hamburg 50

Zusammenfassung

Die Schwierigkeiten der meisten bisher verwendeten Bereichsegmentationsverfahren liegen in der ausschließlichen Verwendung entweder nur lokaler Informationen des Bildes, wie z.B. 'Ähnlichkeit' der einzelnen Pixel bezüglich ihrer Farbe, oder aber nur globaler Informationen, wie z.B. der Farbverteilung des gesamten Bildes. Hier wird ein Segmentationsverfahren für Farbbilder durch hierarchische Ballungsanalyse mittels des "minimal spannenden Baumes" (MST) vorgeschlagen, das in einem 5D-Merkmalraum neben Farbinformationen auch Informationen über den räumlichen Zusammenhang der einzelnen Pixel im Bild, und damit sowohl globale als auch lokale Informationen verwendet. Zur Zerlegung der Baumstruktur des aufgebauten MST in Teilbäume, die den gesuchten Bereichen entsprechen, erscheinen statistische Testverfahren geeignet, die auf dem Prinzip des "Likelihood-Verhältnistests" beruhen. Diese von anderen Autoren bereits erfolgreich zur Kanten- und Änderungsdetektion verwendeten Testverfahren werden für die Anwendung auf Farbbildern auf multivariate Normalverteilungen verallgemeinert. Es wird untersucht, inwieweit sich die Ergebnisse des Testverfahrens verbessern, wenn die Bildfunktion zuvor stückweise durch bivariate Polynome höherer Ordnung approximiert wird. Die Ergebnisse des vorgeschlagenen Segmentationsverfahrens auf Farbbildern von Realweltszenen werden gezeigt und diskutiert.

Abstract

The difficulties with most image segmentation techniques are caused by only using either local information, e.g., the similarity of the pixels, or global information, e.g., the overall color-distribution. We suggest a segmentation technique for color images by hierarchical clustering based on the minimal spanning tree (MST) in a 5-D feature space that combines position information with color features, and therefore, uses local as well as global information.

Statistical test procedures based on the likelihood ratio test principle seem to be well suited for partitioning the MST into subtrees corresponding to the desired regions. In order to apply these test procedures to color images they have been generalized for use with multivariate normal distributions. The effects of a piecewise approximation of the image by higher order bivariate polynomials on the test procedures are investigated. Results of the suggested segmentation algorithm applied to color images of real-world scenes are presented and discussed.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Bereiche: Versuch einer Definition	3
1.2	Grundsätzliche Vorgehensweise in SISSY	6
1.3	Überblick	7
2	Stand der Kunst	8
2.1	Lokale Verfahren	8
2.2	Globale Verfahren	10
2.3	Lokale/Globale Verfahren	13
2.4	Vorschläge zur Verbesserung	16
2.5	Probleme mit linienartigen Strukturen	20
3	Ballungsanalyse im Farben-Orts-Merkmalraum	22
3.1	Idee des Verfahrens	22
3.2	Eine anschauliche Darstellung des Verfahrens	25
3.3	Vorteile des minimal spannenden Baumes	25
3.4	Berechnung des minimal spannenden Baumes	27
4	Einfache Zerlegungsverfahren	30
4.1	Eine globale Schwelle	30
4.2	Eine lokale Schwelle	32
4.2.1	Verschmelzen zu kleiner Bereiche	34
4.2.2	Eine Schwelle für zu kurze Kanten	35
4.2.3	Segmentation eines synthetischen Bildes mit feinen Bildstrukturen	35
4.2.4	Probleme durch Rauschen	37

5	Statistische Zerlegungsverfahren	41
5.1	Das Yakimovsky-Kriterium	41
5.1.1	Herleitung des Yakimovsky-Tests für Farbe	43
5.1.2	Anwendung des hergeleiteten Tests	45
5.2	Hotellings T^2 -Test	48
5.2.1	Herleitung des T^2 -Tests von Hotelling	49
5.2.2	Anwendung des T^2 -Tests	54
5.3	Das Behrens-Fisher Problem	56
5.4	Approximation der Bildfunktion durch Polynome höherer Ordnung	59
5.4.1	Approximation durch Polynome 0. Ordnung	62
5.4.2	Approximation durch Polynome 1. Ordnung	63
5.4.3	Approximation durch Polynome 2. Ordnung	65
5.5	Zusammenfassung	66
6	Versuch einer Bewertung	69
6.1	Vergleich der einzelnen Zerlegungsverfahren	69
6.2	Robustheit	73
6.3	Reproduzierbarkeit	74
6.4	Vergleich mit Ergebnissen des Kantenfinders	77
6.5	Nachteile des minimal spannenden Baumes	79
7	Zusammenfassung und Ausblick	81
A	Bilder	85
B	Schätzung der Verteilungsparameter	101
B.1	Schätzung des Mittelwertvektors μ	101
B.2	Schätzung der Kovarianzmatrix Σ	102
C	Schätzung der Parametermatrix B	104
D	Exponentenvereinfachung	107
	Literaturverzeichnis	109

Ein Ansatz zur Segmentation
von Farbbildern
durch hierarchische Ballungsanalyse

Carsten Schröder

FBI-HH-B-132/87

20. Oktober 1987

Universität Hamburg
Fachbereich Informatik
Bodenstedtstraße 16
D-2000 Hamburg 50

Zusammenfassung

Die Schwierigkeiten der meisten bisher verwendeten Bereichssegmentationsverfahren liegen in der ausschließlichen Verwendung entweder nur lokaler Informationen des Bildes, wie z.B. 'Ähnlichkeit' der einzelnen Pixel bezüglich ihrer Farbe, oder aber nur globaler Informationen, wie z.B. der Farbverteilung des gesamten Bildes. Hier wird ein Segmentationsverfahren für Farbbilder durch hierarchische Ballungsanalyse mittels des "minimal spannenden Baumes" (MST) vorgeschlagen, das in einem 5D-Merkmalraum neben Farbinformationen auch Informationen über den räumlichen Zusammenhang der einzelnen Pixel im Bild, und damit sowohl globale als auch lokale Informationen verwendet. Zur Zerlegung der Baumstruktur des aufgebauten MST in Teilbäume, die den gesuchten Bereichen entsprechen, erscheinen statistische Testverfahren geeignet, die auf dem Prinzip des "Likelihood-Verhältnistests" beruhen. Diese von anderen Autoren bereits erfolgreich zur Kanten- und Änderungsdetektion verwendeten Testverfahren werden für die Anwendung auf Farbbildern auf multivariate Normalverteilungen verallgemeinert. Es wird untersucht, inwieweit sich die Ergebnisse des Testverfahrens verbessern, wenn die Bildfunktion zuvor stückweise durch bivariate Polynome höherer Ordnung approximiert wird. Die Ergebnisse des vorgeschlagenen Segmentationsverfahrens auf Farbbildern von Realweltszenen werden gezeigt und diskutiert.

Abstract

The difficulties with most image segmentation techniques are caused by only using either local information, e.g., the similarity of the pixels, or global information, e.g., the overall color-distribution. We suggest a segmentation technique for color images by hierarchical clustering based on the minimal spanning tree (MST) in a 5-D feature space that combines position information with color features, and therefore, uses local as well as global information.

Statistical test procedures based on the likelihood ratio test principle seem to be well suited for partitioning the MST into subtrees corresponding to the desired regions. In order to apply these test procedures to color images they have been generalized for use with multivariate normal distributions. The effects of a piecewise approximation of the image by higher order bivariate polynomials on the test procedures are investigated. Results of the suggested segmentation algorithm applied to color images of real-world scenes are presented and discussed.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Bereiche: Versuch einer Definition	3
1.2	Grundsätzliche Vorgehensweise in SISSY	6
1.3	Überblick	7
2	Stand der Kunst	8
2.1	Lokale Verfahren	8
2.2	Globale Verfahren	10
2.3	Lokale/Globale Verfahren	13
2.4	Vorschläge zur Verbesserung	16
2.5	Probleme mit linienartigen Strukturen	20
3	Ballungsanalyse im Farben-Orts-Merkmalraum	22
3.1	Idee des Verfahrens	22
3.2	Eine anschauliche Darstellung des Verfahrens	25
3.3	Vorteile des minimal spannenden Baumes	25
3.4	Berechnung des minimal spannenden Baumes	27
4	Einfache Zerlegungsverfahren	30
4.1	Eine globale Schwelle	30
4.2	Eine lokale Schwelle	32
4.2.1	Verschmelzen zu kleiner Bereiche	34
4.2.2	Eine Schwelle für zu kurze Kanten	35
4.2.3	Segmentation eines synthetischen Bildes mit feinen Bildstrukturen	35
4.2.4	Probleme durch Rauschen	37

5	Statistische Zerlegungsverfahren	41
5.1	Das Yakimovsky-Kriterium	41
5.1.1	Herleitung des Yakimovsky-Tests für Farbe	43
5.1.2	Anwendung des hergeleiteten Tests	45
5.2	Hotellings T^2 -Test	48
5.2.1	Herleitung des T^2 -Tests von Hotelling	49
5.2.2	Anwendung des T^2 -Tests	54
5.3	Das Behrens-Fisher Problem	56
5.4	Approximation der Bildfunktion durch Polynome höherer Ordnung	59
5.4.1	Approximation durch Polynome 0. Ordnung	62
5.4.2	Approximation durch Polynome 1. Ordnung	63
5.4.3	Approximation durch Polynome 2. Ordnung	65
5.5	Zusammenfassung	66
6	Versuch einer Bewertung	69
6.1	Vergleich der einzelnen Zerlegungsverfahren	69
6.2	Robustheit	73
6.3	Reproduzierbarkeit	74
6.4	Vergleich mit Ergebnissen des Kantenfinders	77
6.5	Nachteile des minimal spannenden Baumes	79
7	Zusammenfassung und Ausblick	81
A	Bilder	85
B	Schätzung der Verteilungsparameter	101
B.1	Schätzung des Mittelwertvektors μ	101
B.2	Schätzung der Kovarianzmatrix Σ	102
C	Schätzung der Parametermatrix B	104
D	Exponentenvereinfachung	107
	Literaturverzeichnis	109

Kapitel 1

Einleitung

Eine Aufgabe bei der automatischen Auswertung von TV-Bildfolgen, die in den letzten Jahren für den Einsatz in autonomen mobilen Robotern an praktischer Bedeutung gewonnen hat, ist die *geometrische Szenenrekonstruktion* [Dreschler-Fischer 87].

Hierbei geht es darum, die dreidimensionale Struktur der Objekte einer beobachteten Szene, sowie die relative Anordnung der Objekte zueinander, aus Bildvorlagen zu ermitteln. Außerdem ist man daran interessiert, aus einer Folge von Bildern die Bewegungsparameter der Objekte zu bestimmen.

Bei der Bildentstehung geht durch die Projektion der dreidimensionalen Szene auf die zweidimensionale Bildebene die Information über die dritte Dimension, die sogenannte *Tiefeninformation*, jedoch verloren.

Wie kann nun aus zweidimensionalen Bildern die Tiefeninformation zurückgewonnen werden? Eine Möglichkeit hierfür ist durch das *Structure from Motion Theorem* gegeben [Ullman 79], nach dem unter der Annahme orthographischer Projektion aus 3 Ansichten von 4 nicht koplanaren Punkten eines starren Körpers sowohl die dreidimensionale Struktur als auch die Bewegungsparameter eindeutig bestimmt werden können (bis auf eine Spiegelung an der Bildebene).

Aus einer monokularen Bildfolge, aufgenommen von einer stationären Kamera, können danach Struktur und Bewegung für *bewegte starre Körper* ermittelt werden. Dieser Ansatz ist in den Arbeiten an dem System MORIO verfolgt worden [Dreschler + Nagel 82]. Über nicht bewegte Objekte kann hier jedoch keine Aussage gemacht werden, da bei einer stationären Kamera die notwendigen verschiedenen Ansichten nicht vorhanden sind. Durch die Forderung nach starren Körpern kann über im Laufe der Zeit sich verformende Objekte mit Hilfe dieses Ansatzes ebenfalls keine Aussage gemacht

werden.

Eine zweite Möglichkeit zur Rekonstruktion der Tiefeninformation ist die Verwendung von Stereobildpaaren, die auch dem Menschen die Wahrnehmung von Tiefe und Entfernung ermöglichen. Hierbei wird von zwei in einem bestimmten Abstand voneinander aufgestellten Kameras gleichzeitig jeweils ein Bild der Szene aufgenommen. Sind die Kameraparameter, dies sind die relative Lage der beiden Kameras zueinander sowie die Brennweiten bei der Aufnahme, bekannt, so können aus den beiden Abbildungen eines Objektpunktes der Szene in den Bildern durch sogenannte *Triangulation* dessen 3D-Koordinaten eindeutig bestimmt werden [Duda + Hart 73]. Über Bewegung kann mit diesem Ansatz natürlich keine Aussagen gemacht werden, da das verwendete Stereobildpaar ja zu einem festen Zeitpunkt aufgenommen wird. Auch die Gruppierung von Objektpunkten zu ganzen Objekten ist hier nur schwer möglich, da Informationen über die Bewegung von Objektpunkten dazu eine wesentliche Hilfe darstellen.

Im Projekt SISSY,¹ an dem im Arbeitsbereich 'Kognitive Systeme' seit einiger Zeit gearbeitet wird, sollen Folgen von Stereobildern verwendet werden, so daß die beiden genannten Ansätze miteinander kombiniert werden können [Dreschler-Fischer 86]: Aus den einzelnen Stereobildern werden die 3D-Koordinaten der Objektpunkte rekonstruiert; die Folge der Stereobilder kann dazu verwendet werden, die Bewegungsparameter von Objektpunkten zu ermitteln. Hiermit können dann Objektpunkte zu ganzen Objekten gruppiert und möglicherweise sogar die Verformungen von Objekten bestimmen werden.

Ein wesentliches Problem, das gelöst werden muß, bevor aus Stereobildern Tiefeninformation rekonstruiert und aus Bildfolgen Bewegung ermittelt werden kann, ist das sogenannte *Korrespondenzproblem*: Punkte, oder allgemeiner, Merkmale, die Abbildungen ein und desselben Objektpunktes bzw. Objektteils der Szene sind, müssen in jeweils zwei Bildern identifiziert werden.

Wir wollen hier kurz das Vorgehen zur Lösung des Korrespondenzproblems im Projekt SISSY schildern, das sich grob in folgende zwei Schritte gliedert:

1. Zunächst werden aus den einzelnen Bildern geeignete Merkmale *extrahiert*. Als Merkmale sollen hier markante Punkte, Kanten und homogene Bereiche verwendet werden [Bartsch et al. 86, Bartsch 87, Drewniok]. Die Verfahren, die diese Merkmale aus den Bildern extra-

¹Stereo Image Sequences SYstem

hieren, sollen dabei unabhängig voneinander arbeiten, um durch die Redundanz der Ergebnisse die Lösung des Korrespondenzproblems zu erleichtern. Als Bilder sollen hierbei Farbbilder verwendet werden, da die zusätzliche Information sowohl bei der Extraktion der Merkmale als auch bei der Korrespondenzanalyse eine wertvolle Hilfe darstellen kann.

Die extrahierten Merkmale werden dann samt der sie beschreibenden Attribute in eine Datenstruktur eingetragen, die auch die räumlichen Beziehungen der einzelnen Merkmale zueinander im Bild in Form von Relationen bereitstellt [Sprengel 86]. Hierdurch entsteht eine reichhaltige *symbolische Beschreibung* der einzelnen Bilder.

2. Mit Hilfe der symbolischen Bildbeschreibungen wird dann im zweiten Schritt versucht, diejenigen Merkmale in verschiedenen Bildern einander *zuzuordnen*, die Abbildungen ein und desselben Objektteils der Szene sind.

Das Problem der Zuordnung von Merkmalen zwischen den einzelnen Bildern ist der eigentliche Kern des Korrespondenzproblems und Forschungsschwerpunkt im Projekt Sissy.

Diese Arbeit beschäftigt sich nun mit der *Segmentation* von Farbbildern, also dem Zerlegen der Bilder in *Bereiche*, wie sie im Projekt Sissy durchgeführt wird, um Merkmale für die Korrespondenzanalyse zu erhalten.

Segmentation gehört zu dem ersten Schritt der Analyse, der auf jedem Bild einer Stereobildfolge separat ausgeführt wird, und so wird im folgenden nur noch von Einzelbildern die Rede sein.

Um das Problem der Bereichsegmentation zu analysieren, wollen wir zunächst eine Definition des Begriffs *Bereich* versuchen.

1.1 Bereiche: Versuch einer Definition

Ein Blick in die Literatur läßt sogleich die Schwierigkeiten deutlich werden: eine einheitliche und zufriedenstellende Definition für diesen Begriff gibt es nicht. Meist liest man sehr Allgemeines, wie z.B.:

A region is an area in a picture whose points have a common property [Zucker 76, S. 382].

Regions of an image segmentation should be uniform and homogeneous with respect to some characteristics [Haralick + Shapiro 85, S. 100].

Diese Definitionen lassen sich zusammenfassen und formal schreiben [Horowitz + Pavlides 74]:

Sei \mathcal{X} die Menge der Bildpunkte (Pixel) eines Bildes und P ein Prädikat definiert auf einer Menge zusammenhängender Pixel. Eine Segmentation in Bereiche ist dann definiert als eine Partition der Menge \mathcal{X} in disjunkte Teilmengen $\mathcal{S}_1, \dots, \mathcal{S}_n$, so daß gilt:

1. $\mathcal{X} = \bigcup_{i=1}^n \mathcal{S}_i$,
2. $\mathcal{S}_i \cap \mathcal{S}_j = \{\}$ für alle $i \neq j$,
3. $P(\mathcal{S}_i) = \text{wahr}$ für alle i ,
4. $P(\mathcal{S}_i \cup \mathcal{S}_j) = \text{falsch}$ für alle $i \neq j$, wenn \mathcal{S}_i und \mathcal{S}_j benachbart.

Diese Definition ist stark ausgerichtet an der direkten Verwendung in einem Segmentationsverfahren: das Prädikat P ist dabei eine lokale Eigenschaft einer Menge von Pixeln, wie z.B. gleicher oder "ähnlicher" Grauwert. Dies wird in einem Verfahren eingesetzt, um Pixelmengen zu vereinigen oder zu zerteilen, um so zu einer Segmentation zu gelangen (siehe hierzu Abschnitt 2.1).

Diese Definition sagt jedoch nicht direkt etwas über das globale Ziel eines Segmentationsverfahrens aus; vielmehr wird bei einem konkreten Verfahren ein Prädikat verwendet, mit dem man ein bestimmtes Ziel zu erreichen versucht. Bei vielen Verfahren läßt sich ein solches Prädikat jedoch gar nicht angeben; erst eine durchgeführte Segmentation definiert dann durch ihr Ergebnis den Begriff des Bereichs.

Was ist nun das Ziel eines Segmentationsverfahrens? Orientiert man sich an den Zielen der Bildanalyse, z.B. einer Interpretation der Bilder, der Objekterkennung o.ä., so ist man daran interessiert, eine Segmentation zu erreichen, bei der den Bereichen eine Bedeutung zugewiesen werden kann. Kanade 80 schreibt (S. 279 f):

Regions in the image correspond to surfaces in the scene.

Often the ultimate goal of image analysis is to obtain a segmentation which separates out semantically meaningful objects or parts of objects.

Man könnte also definieren:

Ein Bereich im Bild soll einer Objektoberfläche oder einem Teil einer Objektoberfläche in der beobachteten Szene entsprechen.

Dies ist auch die intuitive Vorstellung, die wir von der Aufgabe eines Segmentationsverfahrens haben. Eine solche Segmentation ist jedoch ohne den Einsatz szenenspezifischen Wissens im allgemeinen nicht möglich. Der größte Teil der existierenden Verfahren setzt solches Wissen nicht ein. Sie verwenden aber als Prädikat meist die lokale Konstanz und Stetigkeit von bestimmten Eigenschaften des Bildes, wie z.B. konstanten Grauwert oder konstanten Grauwertgradienten, und machen damit — bewußt oder unbewußt — Gebrauch von Wissen über die Konstanz und Stetigkeit von Objekteigenschaften der beobachteten Szene. So sind z.B. Oberflächen räumlich stetige Gebiete von Objekten, und die Reflektanzeigenschaften dieser Oberflächen sind oft konstant. Angenommen wird stets auch, daß gleiche Bildeigenschaften durch gleiche Objekteigenschaften zustande kommen. Dies ist sogar der eigentliche Grundgedanke bei der Segmentation: man faßt jeweils einander 'ähnliche' Pixel des Bildes zu Bereichen zusammen, weil man annimmt, daß diese 'Ähnlichkeit' durch gleiche Objekteigenschaften verursacht wird.

Daß all diese Annahmen in Realweltszenen oft nicht gerechtfertigt sind, zeigen die — gemessen an dem hohen Ziel der Segmentation in Objektoberflächen — zum Teil enttäuschenden Ergebnisse der einzelnen Verfahren. Einen guten Überblick über die den einzelnen Segmentationsverfahren zugrunde liegenden Annahmen über Eigenschaften von Objekten der dreidimensionalen Welt gibt *Hille 80*.

Die Probleme, die hier bei der Definition des Begriffs Bereich bestehen, haben ihre Auswirkungen auf die Beurteilung der Ergebnisse eines Segmentationsverfahrens: ohne eine zufriedenstellende formale Definition des Problems läßt sich auch ein Gütekriterium nicht angeben. Beurteilt werden die Ergebnisse meist, indem eine ikonische Darstellung der Segmentation in Form eines synthetischen Bildes erstellt wird, die dann von einem Menschen mit dem ursprünglichen Bild verglichen wird. Hierbei beurteilt er die Segmentation dann nach Kriterien, die ihm von seinem Sehsystem vorgegeben werden: er vergleicht die einzelnen Bereiche mit den Objekten, die er im ursprünglichen Bild erkennt. Falls es im menschlichen Sehsystem überhaupt eine Trennung der Verarbeitung in Segmentation und eine darauf aufbauende Objekterkennung gibt, so geschieht dies auf jeden Fall unbewußt und kann zur Beurteilung nicht herangezogen werden [*Mohnhaupt 87*].

1.2 Grundsätzliche Vorgehensweise in Sissy

Wir wollen an dieser Stelle kurz etwas über die grundsätzliche Philosophie sagen, die auf der Ebene der Merkmalsdetektion im Projekt Sissy verfolgt wird, um zu sehen, welche Anforderungen sich hieraus für ein Segmentationsverfahren ergeben.

Ein wichtiger Punkt ist, daß kein szenenspezifisches Wissen, also spezielles Vorwissen über die beobachtete Szene, bei der Analyse der Stereobildfolge verwendet werden soll, um die möglichen Anwendungen des zu entwickelnden Systems nicht einzuschränken. Für ein Verfahren zur Bereichssegmentation folgt hieraus, daß dann die im vorigen Abschnitt schon genannten Annahmen gemacht werden müssen, um überhaupt zu Ergebnissen zu kommen, auch wenn diese Annahmen nicht immer gerechtfertigt sind.

Weitere Grundsätze ergeben sich aus dem eigentlichen Projektziel: es soll ermittelt werden, welches Wissen für die Korrespondenzanalyse sinnvoll eingesetzt werden kann [Dreschler-Fischer 86]. Dazu gehört es auch, festzustellen, welche Bildmerkmale hierfür geeignet sind. Auf der Ebene der Merkmalsdetektion werden daher zunächst so viele Informationen aus den Bildern extrahiert wie möglich.

Aus diesem Grunde soll neben einem Kantendetektor auch ein Bereichssegmentierer verwendet werden, obwohl die Bereichssegmentation das duale Problem zur Kantendetektion ist. Perfekte Verfahren für beide Probleme vorausgesetzt, beinhalten die Ergebnisse jeweils die gleiche Information: Bereichsgrenzen entsprechen den Kanten und die von Kanten umschlossenen Gebiete entsprechen den Bereichen [Ballard + Brown 82]. Verfahren zur Kantendetektion und zur Bereichssegmentation liefern durch ihre unterschiedliche Vorgehensweise jedoch durchaus nicht die gleichen Ergebnisse. Wir wollen am Ende dieser Arbeit die Ergebnisse unseres Segmentationsverfahrens mit denen des im Projekt Sissy verwendeten Kantendetektors [Drewniok] vergleichen, um zu zeigen, wie sich diese gegenseitig ergänzen.

Einige Verfahren führen eine Segmentation durch, indem sie zunächst Kanten detektieren, diese dann zu geschlossenen Linienzügen vervollständigen, um durch die hiervon umschlossenen Bildgebiete zu Bereichen zu kommen. Solche Verfahren bringen im Vergleich zu einem Kantendetektor jedoch keine neuen Informationen und sollen aus diesem Grunde in dieser Arbeit nicht behandelt werden.

Aus dem oben genannten Projektziel folgt auch, daß auf der Ebene der Merkmalsdetektion aufgrund von unsicheren Ergebnissen noch keine voreiligen Entscheidungen getroffen werden sollen. Dies spiegelt sich insbesondere

dadurch wieder, daß die Ergebnisse der voneinander unabhängig arbeitenden Verfahren im Nachhinein nicht zu einer konsistenten Beschreibung verschmolzen werden sollen, da hierdurch eventuell Informationen verloren gehen könnten. Alle Ergebnisse werden also in die schon erwähnte relationale Datenstruktur eingetragen, auch wenn diese sich zum Teil widersprechen. Die entgültige Entscheidung soll erst auf einer höheren Ebene durch die Verfahren zur Korrespondenzanalyse getroffen werden.

Für diese Arbeit ergibt sich aus dem Projektziel jedoch noch folgende Schwierigkeit: zum gegenwärtigen Zeitpunkt ist noch nicht klar, wie die Ergebnisse eines Segmentationsverfahrens bei der Korrespondenzanalyse genau eingesetzt werden sollen, so daß sich noch keine genauen Anforderungen an ein solches Verfahren angeben lassen.

1.3 Überblick

Im folgenden Kapitel sollen einige existierende Segmentationsverfahren vorgestellt und verglichen werden. Es wird jedoch keine vollständige Literaturübersicht gegeben; vielmehr wird versucht, die existierenden Verfahren in drei verschiedene Klassen einzuteilen.

Aus den dabei deutlich werdenden Vor- und Nachteilen der den einzelnen Klassen zugrundeliegenden Vorgehensweisen ergibt sich dann in Kapitel 3 der Ansatz für ein Segmentationsverfahren durch hierarchische Ballungsanalyse mittels des "minimal spannenden Baumes". Dabei verwenden wir einen Merkmalsraum, der sowohl Farbinformationen als auch Informationen über den räumlichen Zusammenhang der einzelnen Pixel im Bild enthält.

Der Schwerpunkt der Arbeit liegt dann auf der Entwicklung von Verfahren zur Zerlegung der aufgebauten Baumstruktur, um dadurch zu einer Zerlegung des Merkmalsraumes in Ballungen zu gelangen, die dann gleichzeitig die Segmentation des Bildes in Bereiche ergibt.

In Kapitel 4 verwenden wir für diese Zerlegung einfache Adhoc-Verfahren, im darauffolgenden Kapitel entwickeln wir dann statistische Testverfahren für diese Aufgabe.

Kapitel 6 versucht eine Bewertung unseres Verfahrens. Die Arbeit endet mit einem kurzen Ausblick.

Kapitel 2

Stand der Kunst

Die mittlerweile recht große Zahl existierender Verfahren läßt sich grob in drei Klassen einteilen [Kanade 80]:

- Verfahren, die Bereiche aufgrund lokaler räumlicher Informationen im Bild verschmelzen,
- Verfahren, die das Bild aufgrund globaler Informationen des gesamten Bildes in Bereiche zerteilen,
- Verfahren, die sowohl lokale als auch globale Informationen nutzen.

Im folgenden sollen für jede dieser Klassen einige Verfahren exemplarisch dargestellt werden.

2.1 Lokale Verfahren

Das wohl einfachste Verfahren besteht darin, benachbarte Pixel mit gleichem Grauwert oder einer Grauwertdifferenz unterhalb eines Schwellwertes zu Bereichen zu verschmelzen. Für Bereiche eines Bildes, die Objektoberflächen in komplexen Realweltszenen entsprechen, ist ein so einfaches Prädikat in der Regel nicht erfüllt, und es müssen dementsprechend komplexere Prädikate zum Bereichverschmelzen gewählt werden.

In einem der frühesten Verfahren [Muerle + Allen 68] wird das Bild zu Beginn in Elementarbereiche von 2×2 , 4×4 oder 8×8 Pixeln aufgeteilt, für die jeweils ihre Grauwertverteilung berechnet wird. Beginnend in der linken oberen Ecke des Bildes, wird der erste Elementarbereich mit Hilfe eines statistischen Tests auf Ähnlichkeit mit einem Nachbarn geprüft:

Bei Ähnlichkeit werden beide zu einem neuen Bereich verschmolzen und dessen Grauwertverteilung wird neu berechnet. Das Verschmelzen mit benachbarten Elementarbereichen wird solange fortgesetzt, bis dies nicht mehr möglich ist. Mit dem nächsten Elementarbereich wird dieser Prozeß des Verschmelzens erneut durchgeführt usw., bis alle Elementarbereiche des Bildes behandelt worden sind.

Das Ergebnis dieses Verfahrens ist ganz offensichtlich reihenfolgeabhängig: das Segmentationsergebnis ist abhängig von der Wahl des Startpunktes und der Reihenfolge, in der die Elementarbereiche abgearbeitet werden. Da aber im allgemeinen kein Startpunkt vor einem anderen, und keine Reihenfolge vor einer anderen ausgezeichnet ist, so ist auch ein Verfahren mit einer bestimmten Wahl von Startpunkt und Reihenfolge nicht vor einem anderen ausgezeichnet, und kann somit per definitionem kein optimales Ergebnis liefern. Verfahren, deren Ergebnisse reihenfolgeabhängig sind, sollten daher vermieden werden.

Ein in der Literatur viel beachtetes Verfahren stammt von *Brice + Fenema 70*. Sie erweitern die Datenstruktur eines Bildes so, daß zwischen den einzelnen Pixeln Platz zum Einfügen von Elementarkanten zur Verfügung steht. In einem ersten Schritt wird dann zwischen jeweils zwei Pixeln, die nicht den gleichen Grauwert haben, eine Elementarkante eingefügt. Einer solchen Elementarkante wird eine Wichtung proportional zur Grauwertdifferenz der beiden Pixel zugewiesen. Das Bild wird hierdurch in viele kleine Bereiche gleichen Grauwertes aufgeteilt, die jeweils durch eine geschlossene Kette von Elementarkanten umrandet werden. In einem zweiten Schritt werden diese Bereiche durch die Anwendung zweier Heuristiken zu größeren Bereichen verschmolzen.

Phagocyte Heuristic: Zunächst sei die Länge W_{ij} des 'schwächsten' Teiles der gemeinsamen Kante I_{ij} zwischen zwei Bereichen \mathcal{R}_i und \mathcal{R}_j definiert als die Anzahl der Elementarkanten, deren Wichtung einen Schwellwert θ_0 unterschreitet. P_i sei der Umfang, also die Anzahl der umrandenden Elementarkanten des Bereiches \mathcal{R}_i . Nach der "Phagocyte Heuristic" werden zwei benachbarte Bereiche dann verschmolzen, wenn gilt:

$$\frac{W_{ij}}{\min(P_i, P_j)} > \theta_1.$$

Weakness Heuristic: Nach Anwendung der "Phagocyte Heuristic" werden jetzt zwei benachbarte Bereiche verschmolzen, wenn gilt:

$$\frac{W_{ij}}{I_{ij}} > \theta_2.$$

Durch beide Heuristiken werden Bereiche verschmolzen, die eine schwache gemeinsame Kante haben. Die "Phagocyte Heuristic" berücksichtigt aber zusätzlich noch die Form der zu verschmelzenden Bereiche: Bereiche mit einer einfacheren glatteren Form werden durch sie bevorzugt.

Die einfache Art und Weise, wie hier Formkriterien für Bereiche in den Algorithmus eingebaut wurden, hat sicherlich zu dem Ansehen dieses Verfahrens viel beigetragen, auch wenn das verwendete Formkriterium wohl eher für Bilder mit einfachen geometrischen Figuren ("Blockwelten") als für Bilder von Realweltszenen geeignet ist. Die Ergebnisse dieses Verfahrens sind ebenfalls reihenfolgeabhängig.

Die Aufzählung ließe sich hier noch fortführen, die gemeinsame Eigenschaft der Verfahren dieser Klasse wird aber so schon deutlich:

Die Pixel werden nur aufgrund lokaler Informationen des Bildes zu Bereichen verschmolzen.

Die ausschließliche Verwendung lokaler Informationen ist aber zugleich auch die Schwäche dieser Verfahren. Sie müssen die Entscheidung zum Verschmelzen oder Zerlegen von Bereichen treffen, bevor ihnen globalere Informationen des Bildes zur Verfügung stehen. Beispielsweise könnte ein lokales Ähnlichkeitskriterium besser gewählt werden, wenn dazu Informationen über die globale Grauwert- oder Farbverteilung des gesamten Bildes zur Verfügung stünden.

Alle frühen Ansätze zur Bereichssegmentation fallen in diese Klasse, und der Übersichtsartikel von Zucker 76 bildet einen gewissen Abschluß in der Untersuchung lokaler Verfahren.

2.2 Globale Verfahren

Zucker 76 schlägt am Ende seines Übersichtsartikels für weitere Untersuchungen die Verwendung globaler Informationen durch Histogramme be-

stimmter Bildmerkmale¹ vor. Dieser Ansatz wurde dann auch von *Ohlander et al.* 78 verfolgt:

Durch die Verwendung von Farbbildern sind als Bildmerkmale die drei Komponenten der RGB-Farbdarstellung vorhanden. Hieraus werden zusätzlich die Farbdarstellung YIQ der amerikanischen N.T.S.C. Farbfernsehnorm und die Farbdarstellung IHS (Intensity, Hue, Saturation) berechnet [*Bal-lard + Brown* 82, *Kender* 76], so daß insgesamt neun redundante Merkmale zur Verfügung stehen. Beginnend mit dem gesamten Bild als einem Bereich, zerlegt dieses Verfahren Bereiche nach einem bestimmten Kriterium in Unterbereiche. Hierfür werden für den gerade zu zerlegenden Bereich neun eindimensionale Histogramme der oben genannten Merkmale berechnet. Danach wird dann das Merkmal ausgewählt, dessen Histogramm nach bestimmten Kriterien das 'beste' Maximum enthält. Alle diejenigen Pixel des zu zerlegenden Bereiches, deren Wert dieses Merkmals innerhalb eines durch zwei Schwellwerte bestimmten Intervalls um dieses Maximum liegt, werden markiert. Die markierten zusammenhängenden Unterbereiche werden in eine Liste noch zu bearbeitender Bereiche eingefügt. Auf diese, sowie auf die verbleibende Fläche des gerade bearbeiteten Bereiches, wird das beschriebene Verfahren rekursiv wieder angewendet. Kann kein 'bestes' Maximum gefunden werden oder ist der betrachtete Bereich zu klein, gilt der Bereich als endgültig segmentiert.

Dieses Verfahren ist inzwischen häufig mit verschiedenen Abwandlungen und zum Teil unterschiedlichen Anwendungen implementiert worden [*Ohta et al.* 80, *Shafer* 80, *Shafer + Kanade* 82], und hat sich zu einem Standardverfahren der Bereichsegmentation entwickelt.

Anstatt in eindimensionalen Histogrammen nach Maxima zu suchen, versuchen *Schachter et al.* 75, in einem dreidimensionalen 'Farbraum' Ballungen zu finden. Die einzelnen Pixel des Bildes werden dann entsprechend der Ballung, zu deren Entstehen sie beitragen, markiert, und zusammenhängende Komponenten des Bildes mit gleicher Markierung ergeben die gewünschten Bereiche. Die Verwendung des dreidimensionalen Farbraumes hat den Vorteil, daß Ballungen gefunden werden können, deren Projektionen auf die drei Achsen des Farbraumes sich überlappen und sich dort nicht voneinander trennen lassen (die Projektionen des Farbraumes auf seine drei Achsen entsprechen den eindimensionalen Histogrammen).

Das Problem des 'Überlappens' haben auch *Ohlander et al.* 78 erkannt. Aus diesem Grunde führen sie die erwähnte Umrechnung von der RGB-

¹Den Begriff *Merkmal* benutzen wir hier nun im Sinne der Theorie der Mustererkennung. Gemeint sind nicht die Merkmale Punkte, Kanten oder Bereiche.

Farbdarstellung in andere Farbdarstellungen durch. Hierdurch werden zwar keine grundsätzlich neuen Informationen gewonnen; sämtliche Informationen sind in dem dreidimensionalen RGB-Farbraum schon enthalten, und bei der Umrechnung in eine andere Farbdarstellung wird lediglich auf ein anderes Koordinatensystem übergegangen. Hierdurch erhalten sie aber andere Projektionen des Farbraumes, und es besteht die Hoffnung, daß sich in mindestens einem der dann neun Projektionen die Ballungen nicht überlappen.

Das Markieren der Pixel je nach Zugehörigkeit zu einer Ballung ist eine *Klassifikation* von Daten im Sinne der klassischen Theorie der Mustererkennung, und die geschilderten Verfahren des Suchens von Maxima in einem Histogramm oder von Ballungen in einem Farbraum entsprechen dem Paradigma der Klassifikation durch *unüberwachtes Lernen* [Duda + Hart 73]. Coleman + Andrews 79 definieren die Bereichssegmentation denn auch als unüberwachte Klassifikation der Pixel durch Ballungsanalyse in einem n -dimensionalen Merkmalsraum. Sie verwenden in ihrem sehr rechenintensiven Verfahren bis zu 15 Merkmale, die zum Teil erst aus den Bildern errechnet werden müssen. Zu Beginn werden die Merkmale mit Hilfe einer Karhunen-Loève Transformation [Rosenfeld + Kak 82] dekorreliert. Danach stellen sie die der Gesamtverteilung der Merkmale inhärente Anzahl von Ballungen fest, und entscheiden, welche Merkmale zu einer guten Separierung der Ballungen beitragen. In dem Merkmalsraum, der durch die ausgewählten Merkmale aufgespannt wird, führen sie dann mit dem k -Mittelwert Verfahren [Tou + Gonzalez 74] eine Ballungsanalyse durch.

Alle Verfahren dieser Klasse gehen von der Annahme aus, daß eine umkehrbareindeutige Beziehung zwischen Objekteigenschaften und Ballungen im Merkmalsraum besteht [Hille 80]. Diese Annahme läßt sich in drei Teile gliedern:

1. Ein zusammenhängender Bereich im Bild mit ähnlichen Eigenschaften entspricht einer Objektfläche der beobachteten Szene.
2. Ein zusammenhängender Bereich im Bild mit ähnlichen Eigenschaften bildet im Merkmalsraum eine Ballung.
3. Eine Ballung im Merkmalsraum wird durch einen zusammenhängenden Bereich im Bild mit ähnlichen Eigenschaften gebildet.

Die erste Annahme liegt allen Verfahren zur Bereichssegmentation zugrunde. Auf deren Gültigkeit wurde schon in Abschnitt 1.1 eingegangen. Die zweite

Annahme ist zweifelsohne richtig. Durch die in diesem Abschnitt betrachteten Verfahren wird mit der dritten Annahme aber die Umkehrbarkeit der zweiten Annahme postuliert. Daß diese Annahme nicht immer gültig ist, ist offensichtlich: Über das ganze Bild verteilte viele kleine Flecken gleicher Farbe führen im Merkmalsraum zu einer Ballung, die somit aber nicht durch einen zusammenhängenden Bereich zustande kommt.

Die hier beschriebenen Verfahren haben aus diesem Grunde Schwierigkeiten mit verrauschten Bildern und texturierten Bereichen in Bildern. Ihnen fehlt die lokale Information, ob die einer bestimmten Ballung entsprechenden Pixel im Bild tatsächlich räumlich zusammenhängen.

2.3 Lokale/Globale Verfahren

Schachter et al. 79 schlagen vor, die Ergebnisse der Ballungsanalyse durch Einbeziehung lokaler Informationen mit einem der folgenden Verfahren zu verbessern:

Rauschbeseitigung: Unterscheidet sich die Klassifikation eines Pixels von derjenigen eines Großteils seiner Nachbarn, so wird seine Klassifikation derjenigen seiner Nachbarn angepaßt.

Diskrete Relaxation: Anstatt ein Pixel nur einer Ballung zuzuordnen, kann man es all den Ballungen zuordnen, zu denen es möglicherweise gehört. Danach kann festgestellt werden, ob die Nachbarn eines Pixels im Bild die Zuordnung zu einer bestimmten Ballung unterstützen oder nicht, um so die 'beste' auszuwählen.

Probabilistische Relaxation: Den Pixeln wird ein Wahrscheinlichkeitsvektor zugeordnet, der die Wahrscheinlichkeit der Zugehörigkeit zu allen gefundenen Ballungen angibt. Diese Wahrscheinlichkeitsvektoren können dann iterativ an die Wahrscheinlichkeitsvektoren der Nachbarn angepaßt werden. Zum Schluß wird die Ballung mit der höchsten Wahrscheinlichkeit ausgewählt.

Das Verfahren der Rauschbeseitigung wurde schon von *Ohlander et al. 78* angewendet. *Rosenfeld + Kak 82* geben an, daß mit probabilistischer Relaxation wesentlich bessere Resultate erzielt werden als durch Rauschbeseitigung. Auch *Hanson + Riseman 78* verwenden probabilistische Relaxation, um ihre durch Ballungsanalyse gewonnene Segmentation zu verbessern. Wir wollen deren Verfahren und insbesondere das Relaxationsverfahren hier deshalb vorstellen.

Sie führen zu Beginn eine Ballungsanalyse in einem zweidimensionalen Merkmalsraum (zwei Komponenten einer bestimmten Farbdarstellung) durch. Hierfür geben sie kein genaues Verfahren an, sondern verweisen lediglich auf die Literatur zur Mustererkennung; für die gezeigten Ergebnisse wurde sie sogar von Hand vorgenommen. In dem nachfolgenden Relaxationsprozeß verwenden sie nur die Mittelpunkte der im Merkmalsraum gefundenen Ballungen. Bei der Ballungsanalyse brauchen deshalb nicht alle Parameter für die Verteilung einer Ballung ermittelt zu werden. Der Relaxationsprozeß soll nun etwas formaler dargestellt werden.

Sei $\mathcal{X} = \{x_1, \dots, x_n\}$ die Menge der Pixel des Bildes und die Menge der Ballungen im Merkmalsraum sei gegeben durch $\mathcal{C} = \{C_1, \dots, C_m\}$. Die Ballungen C_a definieren uns die Klassen, denen wir die Pixel x_i zuordnen wollen. Wir bezeichnen diese Klassen im folgenden ebenfalls mit C_a . Zu Beginn wird jedem Pixel x_i ein m -dimensionaler Wahrscheinlichkeitsvektor (p_{i1}, \dots, p_{im}) zugeordnet, wobei eine Komponente p_{ia} die Wahrscheinlichkeit angibt, daß das Pixel x_i zu der Klasse C_a gehört. Hanson und Riseman wählen hierfür:

$$p_{ia} = \frac{\frac{1}{d_{ia}}}{\sum_{a=1}^m \frac{1}{d_{ia}}} \quad (2.1)$$

wobei d_{ia} der Euklidische Abstand von dem Punkt im Merkmalsraum, der dem Pixel x_i entspricht, zum Ballungsmittelpunkt von C_a ist. Hiermit gilt:

$$0 \leq p_{ia} \leq 1 \quad \text{und} \quad \sum_{a=1}^m p_{ia} = 1.$$

Außerdem werden für den Relaxationsprozeß sogenannte Kompatibilitätskoeffizienten $c_{ij}(a, b)$ benötigt, die für jedes Paar im Bild benachbarter Pixel x_i und x_j angeben, ob die Zuordnung des Pixel x_j zur Klasse C_b die Zuordnung des Pixels x_i zur Klasse C_a unterstützt oder nicht. Die Kompatibilitätskoeffizienten sollen folgende Eigenschaften haben:

- $c_{ij}(a, b) \in [-1, +1]$,
- $c_{ij}(a, b) > 0$, wenn die Zuordnung $x_i \in C_a$ von der Zuordnung $x_j \in C_b$ unterstützt wird,
- $c_{ij}(a, b) < 0$, wenn die Zuordnung $x_i \in C_a$ von der Zuordnung $x_j \in C_b$ nicht unterstützt wird,

- $c_{ij}(a, b) = 0$, wenn die Zuordnung $x_i \in C_a$ von der Zuordnung $x_j \in C_b$ unabhängig ist.

Hanson + Riseman 78 machen zwei Vorschläge für die Wahl der Kompatibilitätskoeffizienten:

1. $c_{ij}(a, b) = 1$ für $a = b$ und $c_{ij}(a, b) = -1$ für $a \neq b$,
2. $c_{ij}(a, b) = 1$ für $a = b$ und $c_{ij}(a, b) = -\frac{d_{ab}}{d_{max}}$ für $a \neq b$

d_{ab} = Euklidischer Abstand der Ballungsmittelpunkte von C_a und C_b ,
 $d_{max} = \max(d_{ab})$ für alle a und b .

Im zweiten Vorschlag werden mit den Abständen der Ballungsmittelpunkte auch Informationen über die Beziehungen zwischen den einzelnen Ballungen im Merkmalsraum zur Verfügung gestellt. Seine Wahl verspricht daher die besseren Ergebnisse. Die obige Definition der Kompatibilitätskoeffizienten legt es auch nahe, hierfür bedingte Wahrscheinlichkeiten einzusetzen. Diesen Ansatz verfolgen Nagin et al. 82. Aus den oben beschriebenen Wahrscheinlichkeitsvektoren läßt sich die bedingte Wahrscheinlichkeit, daß ein Pixel x_i der Klasse C_a zugeordnet wird, wenn sein rechter Nachbar x_j der Klasse C_b zugeordnet wird, nach folgender Formel berechnen:

$$P(x_i \in C_a | x_j \in C_b) = \frac{P(x_i \in C_a, x_j \in C_b)}{P(x_j \in C_b)} = \frac{\sum_{i=1}^n p_{ia} p_{j(i)b}}{\sum_{j=1}^n p_{jb}} \quad (2.2)$$

Wir berechnen nun

$$\log\left(\frac{P(x_i \in C_a | x_j \in C_b)}{P(x_i \in C_a)}\right) \quad (2.3)$$

Nach einer Skalierung auf das Intervall $[-1, +1]$ hat dies dann die gewünschte Eigenschaft der Kompatibilitätskoeffizienten.

Mit dieser Wahl der Kompatibilitätskoeffizienten gehen weitere statistische Eigenschaften des Bildes in das Verfahren ein, und zwar die relativen Häufigkeiten von bestimmten Merkmalskonfigurationen. Liegen im Bild z.B. auf der rechten Seite von 'roten' Pixeln häufig 'grüne' Pixel, so ist die bedingte Wahrscheinlichkeit aus Gleichung 2.2 hierfür hoch. Die Wahrscheinlichkeit dafür, daß ein Pixel als 'rot' klassifiziert wird, kann dann durch

die Kompatibilitätskoeffizienten gestützt werden, wenn es auf seiner rechten Seite ein Pixel mit hoher Wahrscheinlichkeit für die Klasse 'Grün' hat.

Eine Schwierigkeit besteht jedoch bei der Berechnung der Kompatibilitätskoeffizienten. Hierfür werden ja die Wahrscheinlichkeitsvektoren der einzelnen Pixel benötigt. Um korrekte Kompatibilitätskoeffizienten zu erhalten, müßten eigentlich die Wahrscheinlichkeitsvektoren verwendet werden, die eine perfekte Segmentation beschreiben. Zur Verfügung stehen zu Beginn jedoch nur diejenigen, die mit Hilfe der Ballungsanalyse ermittelt wurden.

Sind die Anfangswahrscheinlichkeiten bestimmt und die Kompatibilitätskoeffizienten festgelegt, so werden die Komponenten der Wahrscheinlichkeitsvektoren durch folgende Iterationsformel an die ihrer Nachbarpixel angepaßt:

$$p_{ia}(t+1) = \frac{p_{ia}(t)(1+q_{ia}(t))}{\sum_{a=1}^m p_{ia}(t)(1+q_{ia}(t))}, \quad (2.4)$$

wobei

$$q_{ia}(t) = \frac{1}{\text{card}(\mathcal{N})} \sum_{j \in \mathcal{N}} \left(\sum_{b=1}^m c_{ij}(a,b) p_{jb}(t) \right) \quad (2.5)$$

\mathcal{N} = Menge der Pixel in der Nachbarschaft von A_i .

Eine Anzahl von weniger als zehn Iterationen reicht im allgemeinen aus, um eine wesentliche Verbesserung des Segmentationsergebnis zu erreichen. Das hier vorgestellte Relaxationsverfahren lehnt sich an die Darstellungen in [Nagin et al. 82] und [Rosenfeld + Kak 82] an.

2.4 Vorschläge zur Verbesserung

Ballungsanalyse in einem mehrdimensionalen Merkmalsraum mit anschließender Anwendung von probabilistischer Relaxation ist derzeit das wohl beste Bereichssegmentationsverfahren. Hier sollen die Probleme dieses Verfahrens diskutiert sowie einige Überlegungen zur Verbesserung angestellt werden.

Das Problem der Ballungsanalyse ist in der Theorie der Mustererkennung bisher nicht zufriedenstellend gelöst. Es tauchen hierbei folgende Probleme auf [Hille 80]:

- Die Verteilung der Daten im Merkmalsraum ist nicht bekannt.
- Die verwendeten Merkmale sind oft miteinander korreliert.
- Die Anzahl der Ballungen im Merkmalsraum ist im allgemeinen nicht bekannt.

Alle existierenden Verfahren machen bestimmte Annahmen über die Daten, um Schwierigkeiten durch die oben genannten Probleme zu entgehen.

Unbekannte Verteilung

In den meisten Fällen wird angenommen, daß sich die Verteilung der Daten im Merkmalsraum aus mehreren, der Anzahl der Ballungen entsprechenden, multivariaten Normalverteilungen zusammensetzt. Ist die Anzahl der Ballungen bekannt, so können alle Parameter der gemischten Verteilung, dies sind die Mittelwertvektoren sowie die Kovarianzmatrizen, z.B. mit Hilfe eines Maximum-Likelihood-Schätzers 'gelernt' werden [Duda + Hart 73]. Dies Verfahren ist jedoch sehr aufwendig. Durch die Annahme, daß die Merkmale nicht miteinander korreliert sind, läßt sich die Parameterschätzung erheblich vereinfachen, da dann von vornherein bekannt ist, daß die Nicht-diagonalelemente der Kovarianzmatrizen den Wert Null haben. Ist diese Annahme nicht gerechtfertigt, so könnte man zu einer anderen Farbdarstellung übergehen, von der bekannt ist, daß deren Komponenten in der Regel sehr wenig miteinander korreliert sind [Ohta et al. 80].

Unbekannte Anzahl

Das größte Problem ist jedoch die unbekannt Anzahl der Ballungen. Zum einen muß für die Anwendung der meisten Verfahren die Anzahl bekannt sein, und zum anderen ist sie entscheidend wichtig für ein gutes Segmentationsergebnis. Wird z.B. eine Ballung nicht gefunden, die zwischen zwei anderen liegt, so werden die Pixel, die zu ihr gehören, teils der einen, teils der anderen Nachbarballung zugeordnet, und der Bereich, der diese Ballung gebildet hat, wird dadurch in viele kleine Stücke zerteilt. Der gleiche Fehler tritt auf, wenn anstatt einer zwei Ballungen gefunden werden. Daß diese Segmentationsfehler auch mit einem Relaxationsverfahren kaum rückgängig zu machen sind, zeigen Hanson + Riseman 78 und Shafer 80.

Ein Verfahren, bei dem die Anzahl der Ballungen nicht bekannt zu sein braucht, ist die hierarchische Ballungsanalyse mittels des "minimal spannenden Baumes" (MST) [Zahn 71, Duda + Hart 73]. Hierbei betrachtet man

die besetzten Punkte des Merkmalsraumes als Knoten eines vollständigen gewichteten Graphen, d.h. eines Graphen, in dem jeder Knoten mit jedem anderen Knoten über eine Kante einer bestimmten Länge verbunden ist. Die Länge der Kanten entspricht der Entfernung der Knoten im Merkmalsraum nach einem zu definierenden Abstandsmaß. Für diesen Graphen wird dann ein minimal spannender Baum berechnet. Dies ist ein Baum, der alle Knoten des Graphen enthält, und dessen Summe seiner Kantenlängen minimal ist. Die Partition des Merkmalsraumes in Ballungen erfolgt dann durch Zerlegen des MST in Teilbäume. Dieses Verfahren macht, im Gegensatz zu den Verfahren, die eine bestimmte Verteilung der Daten im Merkmalsraum zugrunde legen, keine Annahmen über die Form der Ballungen, und man sagt, daß es deshalb den Daten im Merkmalsraum keine Struktur 'aufzwingt', sondern tatsächlich Struktur 'findet'. Ein Problem dieses Verfahrens ist, daß es besonders anfällig gegen Rauschen ist, da hierdurch die Ränder der Ballungen 'verschmiert' werden und somit zwei Ballungen ineinander übergehen können. Diese Ballungen können dann nicht mehr voneinander getrennt werden, da die Kante des MST, die die beiden miteinander verbindet, sehr kurz ist.

Um festzustellen, welche Formen Ballungen in einem Farbmerkmalsraum überhaupt haben, und wie dicht diese Ballungen beieinander liegen, haben wir für die von uns zu verarbeitenden Bilder einmal die drei Projektionen des 3D-Farbmerkmalsraumes auf die Rot-Grün-, die Rot-Blau- und die Grün-Blau-Ebene berechnet (siehe Abbildung 2.1). Es zeigt sich dabei, daß die einzelnen Ballung nicht sehr scharfe Ränder haben, daß sie sehr dicht beieinander liegen und oft ineinander übergehen. Weiterhin fällt auf, daß die meisten Ballungen entlang der Raumdiagonalen des 3D-Farbmerkmalsraumes, also der Intensitätsachse liegen. Die drei Farbmerkmale sind also recht stark miteinander korreliert. Eine Möglichkeit, die Ballungen schärfer voneinander zu trennen, wäre, diejenigen Pixel des Bildes, die an Bereichsgrenzen liegen, bei der Berechnung der Ballungen nicht mit zu berücksichtigen. Diesem Vorgehen liegt die Annahme zugrunde, daß Pixel, die im Bild zwischen Bereichen liegen, im Merkmalsraum auch zwischen Ballungen liegen. Bestimmen ließen sich diese Pixel durch Anwendung eines einfachen Kantentfinders, z.B. des Sobel-Operators. Berücksichtigt man diese Pixel nicht, so müßten die einzelnen Ballungen schärfer gegeneinander abgegrenzt sein. Dies ist bis zu einem gewissen Grade auch der Fall (siehe Abbildung 2.2), für die Anwendung des MST-Verfahrens sind die einzelnen Ballungen aber wohl immer noch nicht scharf genug voneinander getrennt.

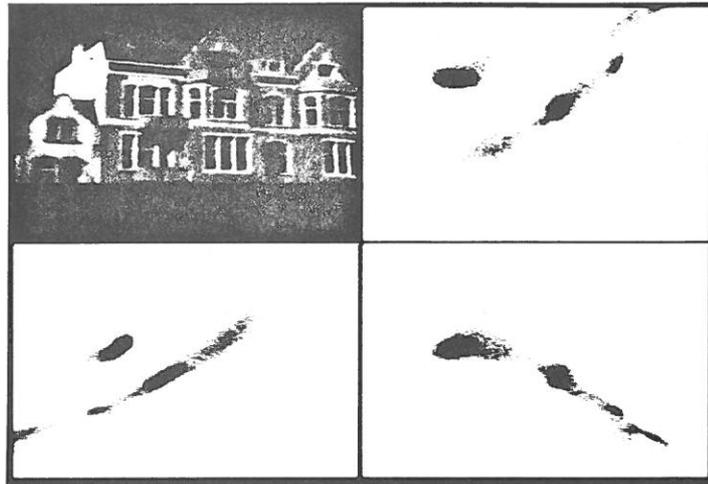


Abbildung 2.1: Projektionen des 3D-Farbmerkmalsraumes von Bild 12 auf die Rot-Grün-, Rot-Blau- und Grün-Blau-Ebene. Die Rot-Achse zeigt nach rechts, die Grün-Achse nach unten und die Blau-Achse nach oben und nach links. Die einzelnen Häufigkeiten wurden logarithmisch auf 256 Graustufen skaliert.

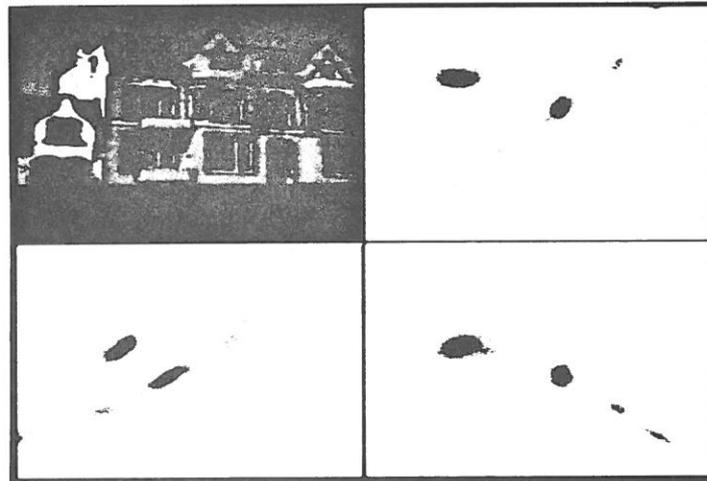


Abbildung 2.2: Wie Abbildung 2.1, jedoch ohne Berücksichtigung der Pixel, die durch einen Sobel-Operator als Kantenelemente detektiert wurden.

Abstandsmaß im Merkmalsraum

Ein Vorschlag zur Verbesserung an anderer Stelle wäre die Verwendung eines anderen Abstandsmaßes im Merkmalsraum für die Berechnung der Wahrscheinlichkeitsvektoren in Gleichung 2.1. Hanson und Riseman verwenden den Euklidischen Abstand der Punkte von den Ballungsmittelpunkten. Hierdurch wird aber die Größe und die Form der einzelnen Ballungen nicht mit berücksichtigt. Dies kann geschehen durch Verwendung der sogenannten *Mahalanobisdistanz* r , wobei gilt:

$$r^2 = (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})$$

- \mathbf{x} : Punkt im Merkmalsraum
- $\boldsymbol{\mu}$: Mittelwertvektor einer Ballung
- $\boldsymbol{\Sigma}$: Kovarianzmatrix der Ballung

Ein Nachteil ist aber, daß hierfür nicht nur die Mittelwertvektoren der einzelnen Ballungen berechnet werden müssen, sondern auch noch die jeweiligen Kovarianzmatrizen.

2.5 Probleme mit linienartigen Strukturen

Es soll nun noch ein Problem angesprochen werden, das die Grenzen von Relaxationsverfahren aufzeigt. *Nagin et al. 82* haben dieses Problem anhand eines synthetischen Bildes untersucht, in dem neben großen homogenen Gebieten auch sehr feine Linien vorhanden sind [*Nagin et al. 82, S.267, Fig. 3d*] (siehe Bild 1 in Anhang A und die Angaben hierzu in Abschnitt 4.2.3). Die Ballungsanalyse führen sie von Hand durch, so daß hierdurch alle Pixel, die zu diesen Linien gehören, richtig klassifiziert werden. Danach bestimmen sie die Kompatibilitätskoeffizienten für das Relaxationsverfahren durch die Verwendung von bedingten Wahrscheinlichkeiten, wie es weiter oben erläutert wurde. Nach Anwendung des Relaxationsverfahrens sind die feinen Linien jedoch so gut wie verschwunden. Der Grund hierfür liegt darin, daß die Struktur der feinen Linien für das Bild nicht typisch genug ist, da sie nur aus relativ wenigen Pixeln besteht. Sie geht in der Statistik der Kompatibilitätskoeffizienten völlig unter und kann sich deshalb während des Relaxationsverfahrens gegenüber den großen homogenen Bereichen nicht behaupten. *Nagin et al. 82* versuchen dieses Problem dadurch zu lösen, daß sie das Bild in mehrere rechteckige Sektoren aufteilen und das Segmentationsverfahren dann auf jeden dieser Sektoren anwenden. Da die Berechnung der

Kompatibilitätskoeffizienten hierbei nun auf kleine lokale Gebiete des Bildes beschränkt ist, werden feine Strukturen durch sie besser repräsentiert. Dadurch bleiben diese Strukturen während des Relaxationsprozesses erhalten. Der Nachteil dieses Verfahrens liegt darin, daß die in den einzelnen Sektoren gefundenen Bereiche über die künstlich eingefügten Sektorengrenzen teilweise wieder miteinander verschmolzen werden müssen.

Kapitel 3

Ballungsanalyse im Farben-Orts-Merkmalraum

Das Ziel des in Abschnitt 2.3 besprochenen Relaxationsverfahrens liegt darin, falsche Klassifikationen von Pixeln, die durch die Ballungsanalyse zustande kommen, nachträglich zu korrigieren. Das am Ende des letzten Kapitels diskutierte Problem macht jedoch deutlich, daß durch ein Relaxationsverfahren immer noch nicht genügend lokale Information zur Verfügung gestellt wird, um befriedigende Segmentationsergebnisse zu erzielen. Es wäre daher wünschenswert, die Ballungsanalyse so zu modifizieren, daß die Fehler, die einen Relaxationsprozeß überhaupt erst notwendig machen, von vornherein vermieden werden.

Im folgenden schlagen wir ein Verfahren zur Bereichsegmentation vor, bei dem die Information über den räumlichen Zusammenhang von Pixeln im Bild direkt im Merkmalsraum bei der Ballungsanalyse zur Verfügung steht.

3.1 Idee des Verfahrens

Hinzunahme der Ortskoordinaten

In unserem Verfahren verwenden wir an Stelle des dreidimensionalen Farbmerkmalsraumes einen fünfdimensionalen Farben-Orts-Merkmalraum. Jedes Pixel des Bildes wird durch einen Merkmalsvektor mit fünf Komponenten repräsentiert, der sich aus den drei Farbkomponenten sowie den beiden Bildkoordinaten zusammensetzt. Durch die Hinzunahme der Bildkoordinaten wird erreicht, daß im Bild räumlich getrennte Bereiche im Merkmalsraum separate Ballungen bilden, oder anders ausgedrückt, die Pixel müssen

sich sowohl in ihrer Farbe als auch in ihren Bildkoordinaten 'ähnlich' sein, um eine Ballung im Merkmalsraum zu bilden. Die Annahme, daß eine Ballung im Merkmalsraum durch einen zusammenhängenden Bereich im Bild erzeugt wird, ist hierdurch gerechtfertigt; ihre Gültigkeit muß nicht, wie bei den globalen Verfahren in Abschnitt 2.2, postuliert werden. Man kann hierbei dann allerdings nicht mehr von Ballungen im üblichen Sinne sprechen; die Punkte im Merkmalsraum sind höchstens einfach besetzt, da sich alle Pixel durch ihre Ortskoordinaten voneinander unterscheiden. Jedem besetzten Punkt im Merkmalsraum entspricht genau ein Pixel im Bild. Die 'Ballungen' bilden hierdurch im Merkmalsraum Hyperflächen.

Verwendung des minimal spannenden Baumes

Um dem Problem der unbekanntenen Anzahl von Ballungen zu entgehen und um keine Annahmen bezüglich der Form der Ballungen machen zu müssen — was sinnvoll auch nicht möglich wäre, da sich die Form der Hyperflächen durch eine der bekannten Verteilungsfunktionen nicht beschreiben läßt —, verwenden wir die in Abschnitt 2.4 schon angesprochene hierarchische Ballungsanalyse mittels des minimal spannenden Baumes (MST) [Zahn 71, Duda + Hart 73].

Wir wollen dieses Verfahren hier noch einmal kurz skizzieren: Die besetzten Punkte des Merkmalsraumes werden als Knoten eines vollständigen gewichteten Graphen betrachtet. Jeder Knoten eines solchen Graphen ist mit jedem anderen Knoten über eine Kante einer bestimmten Länge verbunden. Die Länge der Kanten entspricht der Entfernung der Knoten im Merkmalsraum nach einem zu definierenden Abstandsmaß. Für diesen Graphen wird dann ein minimal spannender Baum berechnet, also ein Baum, dessen Summe seiner Kantenlängen minimal ist. Die Partition des Merkmalsraumes in einzelne Ballungen erfolgt dann durch Zerlegen des MST in Teilbäume, indem Kanten des Baumes, die ein bestimmtes Kriterium erfüllen, aus dem MST entfernt werden. Jeder so entstandene Teilbaum bildet damit eine Ballung im Merkmalsraum.

Die Abstandsfunktion

Wir müssen nun eine Abstandsfunktion in unserem Farben-Orts-Merkmalraum definieren, die die Länge der Verbindungskanten zwischen den einzelnen Knoten des Graphen bestimmt.

Naheliegender ist hier die Verwendung des Euklidischen Abstandes der Merkmalsvektoren. Wir werden aber eine für unser Verfahren wesentliche

Einschränkung vornehmen: Die Abstandsfunktion soll nur für diejenigen Punkte im Merkmalsraum den Euklidischen Abstand der Merkmalsvektoren liefern, die im Bild benachbarten Pixeln entsprechen (4-Zusammenhang). Für alle anderen Paarungen, also von Punkten, die nicht im Bild benachbarten Pixeln entsprechen, soll die Abstandsfunktion den Wert 'Unendlich' liefern. Durch die Verwendung des 4-Zusammenhanges läßt sich der Euklidische Abstand der Merkmalsvektoren noch auf den Euklidischen Abstand der Farbvektoren reduzieren, da für benachbarte Pixel die Differenz der einen Ortskoodinate immer Eins und die der anderen immer Null ist. Bei der Berechnung des MST spielen nur die relativen Längen der Graphkanten zueinander eine Rolle. Bleiben diese erhalten, so ändert sich der berechnete MST nicht. Wir können daher die Berechnung einer Quadratwurzel sparen und verwenden nun schließlich das Quadrat des Euklidischen Abstandes der Farbvektoren. Wir wollen dies noch einmal formal schreiben:

Sei A_i ein Pixel mit
 den Bildkoordinaten (r_i, c_i) ,
 dem Farbvektor $\mathbf{c}_i = (R_i, G_i, B_i)$
 sowie dem Merkmalsvektor $\mathbf{x}_i = (R_i, G_i, B_i, r_i, c_i)$.
 Unsere Abstandsfunktion ist dann gegeben durch:

$$d(\mathbf{x}_1, \mathbf{x}_2) = \begin{cases} \|\mathbf{c}_1 - \mathbf{c}_2\|^2 & : |r_1 - r_2| + |c_1 - c_2| = 1 \\ \infty & : \text{sonst} \end{cases}$$

Diese Abstandsfunktion ist allerdings keine Metrik mehr, da sie die Dreiecksungleichung nicht erfüllt. Wir verwenden in unserem Verfahren aber lediglich die Symmetrieeigenschaft, so daß dies nicht weiter stört.

Zerlegen des minimal spannenden Baumes

Die Partition des Merkmalsraumes in verschiedene Ballungen erfolgt nun durch Zerlegen des MST in Teilbäume. Dies geschieht, indem Kanten des MST, die ein noch zu definierendes Kriterium erfüllen, entfernt werden. Beispielsweise könnten Kanten entfernt werden, deren Länge eine bestimmte Schwelle überschreitet.

Durch die von uns verwendete Abstandsfunktion entspricht jeder Teilbaum des zerlegten MST nun einem zusammenhängenden Bildbereich, da die Kanten des MST im Merkmalsraum ja nur Punkte miteinander verbinden, die im Bild im 4-Zusammenhang stehenden Pixeln entsprechen. Damit stellt die Partition des Merkmalsraumes in Teilbäume schon das Segmentierungsergebnis dar!

3.2 Eine anschauliche Darstellung des Verfahrens

Um deutlich zu machen, was wir durch die Hinzunahme der Bildkoordinaten in Verbindung mit unserer Abstandsfunktion erreichen, wollen wir unser Verfahren hier noch einmal in einer anderen, anschaulichen Weise erläutern.

Unsere Abstandsfunktion liefert für zwei Punkte des Merkmalsraumes, die im Bild Pixeln entsprechen, die nicht benachbart sind (4-Zusammenhang), den Abstand 'Unendlich'. Eine Kante unendlicher Länge wird jedoch niemals Kante des MST sein, wir brauchen diese Kanten also gar nicht erst zu betrachten. Dies entspricht dann der Vorstellung, daß der MST auf einem zwar zusammenhängenden, aber unvollständigen Graphen berechnet wird. Der Graph enthält nur noch die Kanten mit endlicher Länge, also Kanten zwischen Punkten, die im Bild benachbarten Pixeln entsprechen. Hierdurch wird dieser Graph *planar*, er ist kreuzungsfrei, und wir können ihn somit auf unser Bild legen.

Jedes Pixel entspricht dann einem Knoten des Graphen und ist mit jedem seiner 4 Nachbarn über eine Kante des Graphen verbunden. Die Länge einer solchen Kante ist jeweils die Differenz der Farbvektoren der einzelnen Pixel (genauer: das Quadrat der Differenz).

Für diesen Graphen berechnen wir nun einen minimal spannenden Baum. Abbildung 4.2 zeigt einen Ausdruck eines solchen MST (Seite 38). Der MST überdeckt dabei die gesamte Bildfläche. Zerlegen wir den Baum in Teilbäume, so entsprechen diese Teilbäume zusammenhängenden Bildbereichen.

Ein MST enthält per definitionem die jeweils kürzesten Kanten des Graphen: lokal gesehen sind immer Pixel über Kanten des MST miteinander verbunden, die sich bezüglich des definierten Abstandsmaßes sehr 'ähnlich' sind. In einem homogenen Bereich des Bildes hängen also alle Pixel über kurze Kanten in einem Teilbaum zusammen. Diese Teilbäume, die homogene Bereiche des Bildes überdecken, sind dann jeweils über längere Kanten zum Gesamtbaum verbunden. Das Entfernen von Kanten über ein Längenkriterium wird uns dann die Partition in homogene Bereiche liefern.

3.3 Vorteile des minimal spannenden Baumes

Die Verwendung des minimal spannenden Baumes zur Bereichssegmentation wurde bereits von *Burr + Chien 76* vorgeschlagen. Da sie dabei aber auf Grauwertbildern arbeiten, stehen ihnen mit dem einen Grauwert pro Pixel nicht genügend Merkmale zur Verfügung, um sicher klassifizieren zu

können. Sie fassen daher mehrere Pixel zu quadratischen Elementarbereichen zusammen, für die dann weitere Merkmale, wie z.B. Textur, berechnet werden. Diese Elementarbereiche werden dann als Knoten des Graphen aufgefaßt, für den ein MST berechnet wird.

Im Gegensatz hierzu stehen uns bereits auf der Ebene der Pixel mit den drei Farbkomponenten genügend Merkmale für die Klassifikation zur Verfügung; wir können daher mit der Berechnung des MST bereits auf der Pixelebene beginnen. Hierdurch bleibt die volle räumliche Auflösung erhalten und es wird möglich, Bereiche zu erkennen, die feinen Strukturen im Bild entsprechen. Wir werden dies im folgenden Kapitel bei der Segmentation des synthetischen Bildes bestätigt finden, das wir am Ende des letzten Kapitels schon angesprochen haben.

Auch *Freuder 76* verwendet in einem ähnlichen Verfahren eine Baumstruktur zur Segmentation von Bildern. Dabei faßt er zu Beginn ebenfalls 'ähnliche' Pixel zu Elementarbereichen zusammen, und berechnet hierfür verschiedene Attribute. Er verwendet allerdings keinen MST, sondern erzeugt auf andere Weise eine Baumstruktur: Bereiche werden schrittweise mit ihren jeweils ähnlichsten Nachbarbereichen verschmolzen, für die die Attribute dann jeweils neu berechnet werden. Ein so durch Verschmelzen erhaltener Bereich kann in einem Baum nun als Vaterknoten der beiden verschmolzenen Bereiche betrachtet werden. Dies geschieht solange, bis das gesamte Bild zu einem Bereich verschmolzen wurde. Das komplette Bild entspricht der Wurzel des Baumes. Die durch dieses iterative Verschmelzen erzeugte Baumstruktur wird schließlich dazu verwendet, das Bild in Bereiche zu zerlegen.

Sowohl *Freuder 76* wie auch *Burr + Chien 76* geben als Vorteil der verwendeten Bäume an, daß sie eine globalere Sicht auf die Daten — sprich die Bilder — ermöglichen. Im Falle des minimal spannenden Baumes spiegelt sich dies in der Struktur des Baumes sowie in den Längen der einzelnen Kanten wieder: Die Pixel des Bildes werden so organisiert und durch Teilbäume des MST zusammengefaßt, daß ein globales Kriterium optimiert wird.

Weiterhin weisen sie auf die Möglichkeit hin, daß die Zerlegung eines solchen Baumes, und damit die Segmentation des Bildes, erst durch Prozesse auf einer höheren Ebene der Verarbeitung vorgenommen werden könnte; die Baumstruktur enthält dann eine Vororganisation der Daten, trifft jedoch noch keine entgültige Entscheidung.

Vorstellen könnte man sich auch, daß auf der unteren Ebene der Verarbeitung zwar eine Segmentation vorgenommen wird, diese jedoch *zusammen* mit der berechneten Baumstruktur und einer Angabe über die Sicherheit

der gefällten Entscheidung als Ergebnis abgeliefert wird. Auf einer höheren Ebene kann dieser ‘Segmentierungsvorschlag’ aufgrund weiterer Informationen mit Hilfe des MST dann eventuell revidiert werden.

3.4 Berechnung des minimal spannenden Baumes

Die Verfahren zur Berechnung eines minimal spannenden Baumes beruhen auf folgender MST-Eigenschaft [Sedgewick 83]:

Gegeben sei ein Graph $G = (\mathcal{N}, \mathcal{E})$ mit einer Menge von Knoten \mathcal{N} , einer Menge von Kanten \mathcal{E} , und es gelte $\text{card}(\mathcal{N}) = n$ sowie $\text{card}(\mathcal{E}) = e$.

Sei die Menge der Knoten \mathcal{N} des Graphen, für den der MST berechnet werden soll, in zwei Teilmengen \mathcal{U} und \mathcal{V} zerlegt, dann ist die kürzeste Kante $(u, v) \in \mathcal{E}$, die einen Knoten u der Teilmenge \mathcal{U} mit einem Knoten v der Teilmenge \mathcal{V} verbindet, eine Kante im MST.

Es gibt zwei bekannte Algorithmen [Kruskal 56, Prim 57], die im folgenden kurz erläutert werden sollen.

Der Algorithmus von Prim

Der Algorithmus von Prim ergibt sich in einfacher Weise aus der MST-Eigenschaft. Die Menge der Knoten \mathcal{N} wird in zwei Teilmengen \mathcal{U} und \mathcal{V} zerlegt, wobei \mathcal{U} zu Beginn nur einen Knoten und \mathcal{V} alle restlichen Knoten enthält. Der MST wird dann Kante für Kante aufgebaut, indem in jedem Schritt die kürzeste Kante (u, v) genommen wird, die einen Knoten u der Teilmenge \mathcal{U} mit einem Knoten v der Teilmenge \mathcal{V} verbindet. Der Knoten v wird dann der Menge \mathcal{V} entnommen und der Menge \mathcal{U} zugeordnet. Dies wiederholt sich solange, bis $\mathcal{U} = \mathcal{N}$ und $\mathcal{V} = \emptyset$.

Um die kürzeste Kante zwischen \mathcal{U} und \mathcal{V} zu ermitteln, werden zwei Felder **CLOSEST** und **LOWCOST** benötigt. **CLOSEST**(i) enthält denjenigen Knoten in \mathcal{U} , der am dichtesten an Knoten i in \mathcal{V} liegt. **LOWCOST**(i) enthält die Länge der Kante (**CLOSEST**(i), i). In jedem Schritt wird nun einfach **LOWCOST** durchsucht, um den Knoten zu finden, der am dichtesten an \mathcal{U} ist. Da der gefundene Knoten nun zur Menge \mathcal{U} gehört, müssen die beiden Felder entsprechend angepaßt werden. (Diese Darstellung ist entnommen aus [Aho et al. 83].)

Der Prim-Algorithmus benötigt $n-1$ Schritte, da dies genau die Zahl der Kanten des MST ist. In jedem Schritt muß zuerst das Feld **LOWCOST** komplett durchsucht werden, danach müssen beide Felder ganz durchgegangen werden, um sie an die neue Situation anzupassen. Beide Felder haben die Länge n , die Komplexität des Prim-Algorithmus ist somit von der Ordnung $\mathcal{O}(n^2)$.

Bei diesem Algorithmus wird in jedem Schritt für alle Knoten, die noch nicht in der Menge \mathcal{U} enthalten sind, jeweils nur die kürzeste Kante zu dem bereits aufgebauten Teilbaum benötigt. Es ist aus diesem Grunde nicht erforderlich, alle n Kanten abzuspeichern; für Graphen mit sehr vielen Knoten ein großer Vorteil.

Der Algorithmus von Kruskal

In dem Algorithmus von Kruskal wird der MST aufgebaut, indem in jedem Schritt die jeweils kürzeste Kante verwendet wird, die keinen Zyklus erzeugt, wenn sie zu den bereits ausgewählten hinzugefügt wird.

Der Test auf Zyklensfreiheit läßt sich auf Mengenoperationen zurückführen: Die Überprüfung, ob eine Kante einen Zyklus erzeugen würde, entspricht einem Test, ob zwei Elemente in derselben Menge sind. Fällt dieser Test negativ aus, so soll die überprüfte Kante in den Baum aufgenommen werden, dies entspricht dann der Vereinigung von zwei Mengen. Hierfür gibt es sehr effiziente Algorithmen (sogenannte *Union-Find* Algorithmen), deren Komplexität, für eine Anzahl x von Tests, von der Ordnung $\mathcal{O}(x\alpha(x))$ ist [Sedgewick 83]. Die Funktion $\alpha(x)$ ist dabei eine Art Umkehrfunktion der Ackermann-Funktion $A(x, y)$, denn es gilt $\alpha(A(x, x)) \simeq x$ (für den genauen Zusammenhang siehe [Tarjan 75]). Damit steigt die Funktion extrem langsam gegen unendlich und kann für praktische Probleme als konstant angenommen werden. Die Komplexität des Kruskal-Algorithmus wird damit also bestimmt durch das Suchen nach der jeweils kürzesten Kante im Graphen. Im schlimmsten Falle muß die Liste der Kanten vollständig sortiert werden, die Komplexität ist deshalb von der Ordnung $\mathcal{O}(e \log_2 e)$.

Da in jedem Schritt des Algorithmus jeweils nur die kürzeste Kante benötigt wird und nicht alle Kanten des Graphen für den MST gebraucht werden, ist in der Regel weniger Aufwand erforderlich, wenn man die Kanten in einer Datenstruktur, wie sie für den Sortieralgorithmus **HEAPSORT** [Wirth 79] benötigt wird, partiell geordnet hält. Die kürzeste Kante steht jeweils an der Spitze des sogenannten 'Heap' zur Verfügung, und nach Entnehmen dieser Kante kann die partielle Ordnung mit maximal $\log_2 e$ Operationen

wieder hergestellt werden. Im schlimmsten Fall entartet aber auch dies zu vollständigem Sortieren.

Der Kruskal-Algorithmus hat den Nachteil, daß alle n Kanten des Graphen für die Sortierung abgespeichert werden müssen.

Der Vergleich der Komplexitäten beider Algorithmen zeigt, daß der Prim-Algorithmus eher für vollständige Graphen geeignet ist, wohingegen der Kruskal-Algorithmus bei unvollständigen Graphen verwendet werden sollte.

Wir haben bei unserem Problem einen unvollständigen Graphen vorliegen. Die Anzahl der Knoten n ist gleich der Anzahl der Pixel des Bildes, also $n = r \cdot c$ (mit r und c als Anzahl der Zeilen und Spalten des Bildes). Die Anzahl der Kanten unseres Graphen ist damit $e = 2 \cdot r \cdot c - (r + c) \simeq 2n$. Nach obiger Diskussion ist also der Algorithmus von Kruskal für unser Problem geeignet.

Nach der Berechnung des MST müssen wir nun ein Kriterium angeben, mit dessen Hilfe wir den MST so zerlegen können, daß die dabei entstehenden Teilbäume den gewünschten Bereichen entsprechen. Es wird sich zeigen, daß die Güte unseres Segmentationsverfahrens entscheidend von der Wahl dieses Zerlegungskriteriums abhängt. Wir wollen ihr deshalb die folgenden zwei Kapitel widmen.

Kapitel 4

Einfache Zerlegungsverfahren

Wir wollen nun ein Verfahren zum Zerlegen des MST in Teilbäume entwickeln. Dabei werden wir so vorgehen, daß wir zunächst ein Kriterium angeben und es auf unseren Bildern erproben. Die Analyse der Ergebnisse zeigt dann gewisse Schwächen des Verfahrens auf, die im darauffolgenden Schritt durch die Formulierung eines neuen Kriteriums beseitigt werden sollen.

4.1 Eine globale Schwelle

Eine naheliegende Idee ist die Verwendung der Kantenlängen des MST zur Formulierung eines Zerlegungskriteriums [Zahn 71].

Homogene Bereiche eines Bildes sollten jeweils durch kurze Kanten des MST zu einem Teilbaum verbunden sein. Der Übergang von einem Bereich zu einem anderen wird sich dann durch eine längere Kante ausdrücken. Das Entfernen dieser langen Kanten müßte uns daher die gewünschte Segmentation liefern. Hiermit formulieren wir also das folgende

Zerlegungskriterium 1

Eine Kante wird entfernt, wenn sie länger ist als eine globale Schwelle. Diese Schwelle kann z.B. aus der Verteilung der Kantenlängen im gesamten MST ermittelt werden. Sei $\bar{\ell}$ die mittlere Kantenlänge im MST und σ deren Standardabweichung, so wird eine Kante der Länge ℓ entfernt, wenn gilt:

$$\ell > \bar{\ell} + \theta \cdot \sigma,$$

wobei θ ein vorher angegebener Schwellwertfaktor ist.

Bild 10 zeigt nun die Segmentation von Bild 9 mit Hilfe dieses Zerlegungskriteriums. Der Schwellwert wurde hierbei jedoch noch nicht automatisch aus der Verteilung der Kantenlängen ermittelt, sondern explizit angegeben. Eine Kante der Länge ℓ wurde entfernt, wenn $\ell > \theta_1$. Die verwendete Schwelle war $\theta_1 = 350$ (die mittlere Kantenlänge im MST betrug dabei $\bar{\ell} = 215$). Es wurden insgesamt 6562 Bereiche erzeugt. Das Ergebnis zeigt deutlich die Schwächen dieses Zerlegungskriteriums:

In Gebieten des Bildes mit hohem Kontrast — dem Vordergrund mit den parkenden Autos — wurden sehr viele kleine Bereiche erzeugt, die meist nur ein Pixel groß sind; das Ergebnisbild läßt sich daher in diesem Gebiet von dem Originalbild kaum unterscheiden.

In Gebieten des Bildes mit geringem Kontrast — dem Hintergrund mit den Büschen — wurden nur wenige, sehr große Bereiche erzeugt. Selbst das Postauto, das mit seiner gelben Farbe deutlich ins Auge springt, wurde zusammen mit den Büschen zu einem großen Bereich verschmolzen. Der Grund hierfür liegt darin, daß es am unteren Rand des gelben Gebietes einen sehr kontrastarmen Übergang in den dunklen Hintergrund gibt.

Wie läßt sich dieses Ergebnis erklären? Dazu vergegenwärtige man sich, welche Auswirkungen hoher Kontrast im Bild auf den MST hat.

- Hoher Kontrast bedeutet lange Kanten im MST.
- Geringer Kontrast bedeutet kurze Kanten im MST.

In Gebieten des Bildes mit geringem Kontrast werden die Kanten, die zwei Bereiche miteinander verbinden und deshalb entfernt werden sollten, sehr kurz sein, eventuell so kurz wie die Kanten zwischen Pixeln in Gebieten des Bildes mit hohem Kontrast, die nicht entfernt werden sollen. Eine zufriedenstellende Segmentation, sowohl in kontrastreichen wie auch in kontrastarmen Gebieten des Bildes, ist mit einer einzigen, für das ganze Bild gültigen Schwelle aus diesem Grunde nicht zu erreichen.

Bei genauerer Betrachtung zeigt sich auch, daß sich unser Segmentationsverfahren bei Verwendung des Zerlegungskriteriums 1 in keiner Weise von dem primitiven Verschmelzungsverfahren unterscheidet, bei dem zwei Pixel miteinander verschmolzen werden, wenn ihre Farbdifferenz eine vorgegebene Schwelle unterschreitet. Es muß hierfür nur folgendes gezeigt werden:

Zwei benachbarte Pixel u und v , die im MST nicht über eine Kante miteinander verbunden sind, deren Differenz aber die Schwelle unterschreitet, sind über einen Pfad miteinander verbunden, der aus Kanten besteht, die kürzer als oder gleich lang wie die Schwelle sind (das bedeutet, daß sich die beiden Pixel nach dem Zerlegen des MST im gleichen Teilbaum befinden, und somit zum gleichen Bereich gehören). Dies kann leicht mit Hilfe der in Abschnitt 3.4 genannten MST-Eigenschaft geschehen:

Nehmen wir an, es gibt keinen solchen Pfad, oder anders gesagt, es gibt nur einen Pfad, in dem mindestens eine Kante existiert, die länger als die Schwelle ist. Durch diese Kante werden zwei Teilmengen \mathcal{U} und \mathcal{V} der Knotenmenge des Graphen miteinander verbunden, sie ist jedoch länger als die Kante (u, v) , die ebenfalls die Teilmengen \mathcal{U} und \mathcal{V} miteinander verbinden würde. Der MST müßte also an Stelle der längeren Kante die Kante (u, v) enthalten, da er sonst nicht minimal wäre. Dies widerspricht jedoch der obigen Annahme, daß u und v nicht über eine Kante miteinander verbunden sind. Es muß also einen Pfad von u nach v geben, dessen Kanten kürzer oder gleich der Schwelle sind. \square

4.2 Eine lokale Schwelle

Eine Anpassung an lokale Kontraständerungen wird möglich, wenn für jede Kante des MST lokal ein eigener Schwellwert berechnet wird. Eine Kante kann entfernt werden, wenn sie im Vergleich zu den Kanten ihrer 'Umgebung' im MST zu lang ist [Zahn 71]. Die Umgebung einer Kante ist dabei in einfacher Weise durch den MST definiert als alle diejenigen Kanten und Knoten, die von der betrachteten Kante aus im MST erreichbar und nicht weiter als eine vorgegebene Tiefe von ihr entfernt sind. Aus der Verteilung der Kantenlängen der 'Umgebung' kann dann die lokale Schwelle berechnet werden. Hiermit erhalten wir dann

Zerlegungskriterium 2

Eine Kante des MST teilt diesen in zwei Teilbäume. Beide Teilbäume werden bis zu einer vorgegebenen Tiefe τ verfolgt. Dabei werden mit den mittleren Kantenlängen $\bar{\ell}_1$ bzw. $\bar{\ell}_2$ und deren Standardabweichungen σ_1 bzw. σ_2 die Verteilungen der Kantenlängen in beiden Teilbäumen berechnet.

Enthält einer der beiden Teilbäume nicht eine vorgegebene Mindestanzahl von Kanten (hierfür wurde ebenfalls der Wert τ verwendet), so soll die betrachtete Kante in diesem Falle nicht entfernt werden, da hierdurch ein Bereich

entstehen würde, der sehr klein ist und somit für die auf der Segmentation aufbauenden Verfahren keine Aussagekraft mehr hätte.

Anderenfalls wird die Kante markiert, wenn gilt:

$$(\ell > \bar{\ell}_1 + \theta \cdot \sigma_1) \wedge (\ell > \bar{\ell}_2 + \theta \cdot \sigma_2),$$

wobei ℓ die Länge der Kante und θ ein vorgegebener Schwellwertfaktor ist.

In einem ersten Durchlauf durch den MST geschieht dies für jede seiner Kanten. Damit das Verfahren nicht reihenfolgeabhängig ist, werden diejenigen Kanten, die den für sie berechneten Schwellwert überschreiten, zunächst jedoch nur markiert und erst in einem zweiten Durchlauf tatsächlich entfernt.

Bild 11 zeigt die Segmentation bei Verwendung dieses Kriteriums. Die Teilbäume wurden hierbei bis zu einer Tiefe $\tau = 7$ verfolgt; als Schwellwertfaktor wurde $\theta = 1.0$ gewählt. Hierfür wurden 637 Bereiche erzeugt.

Wie erwartet, paßt sich das Verfahren mit diesem Zerlegungskriterium unterschiedlichen Kontrasten im Bild besser an. Gebiete hohen Kontrastes werden nicht mehr in viele sehr kleine Bereiche zerlegt. Beispielsweise konnten die Heck- und Windschutzscheiben der parkenden Autos sowie deren Nummernschilder als jeweils ein größerer Bereich gefunden werden. Der Hintergrund wurde nicht mehr zu einem großen Bereich verschmolzen. Das Postauto konnte von dem dunkleren Gebiet der Büsche getrennt werden.

Bild 13 zeigt das Segmentationsergebnis bei Verwendung des gleichen Zerlegungskriteriums für ein anderes Bild (Bild 12). Die Teilbäume wurden auch hier bis zu einer Tiefe $\tau = 7$ verfolgt, als Schwellwertfaktor wurde $\theta = 1.5$ gewählt. Dafür wurden 650 Bereiche erzeugt. Auch hier zeigen sich gute Ergebnisse in Bildteilen mit schwachem Kontrast, wie z.B. den Balkonen. Deren Mauerflächen unterscheiden sich nur gering voneinander, sind aber trotzdem in verschiedene Bereiche zerlegt worden. Bemerkenswert sind auch die weißen Giebelränder des Hauses (siehe Bild 12). Dort betragen die Unterschiede von Pixel zu Pixel in jedem der drei Farbkanäle ca. 20 Intensitätsstufen (bei einer Auflösung von 256 Intensitätsstufen pro Farbkanal). Trotz dieser großen Differenzen werden die Pixel zu je einem Bereich zusammengefaßt und nicht in mehrere Bereiche zerlegt. Dies liegt daran, daß für die Entscheidung, ob eine Kante aus dem MST entfernt werden soll, ihre Länge mit den Kantenlängen ihrer 'Umgebung' verglichen wird. Da diese in dem betrachteten Gebiet der Giebelränder aber alle vergleichbar lang sind, werden hierdurch lange Kanten eben nicht entfernt. Diesen Vorteil des Vergleichs von Kantenlängen hat schon *Freuder 76* in einem ähnlichen Verfahren erwähnt (vergleiche Abschnitt 3.3).

Wir sehen allerdings auch zwei grobe Segmentationsfehler: Die Frontseite des Hauses auf der linken Seite ist links oben teilweise mit dem Dach und rechts mit der Frontseite des Nachbarhauses verschmolzen. Hier hat das Rauschen die Kanten des MST, die die jeweiligen Bereiche miteinander verbindet, extrem verkürzt. Wir werden auf dieses Problem in Abschnitt 4.2.4 noch zurückkommen.

Bild 19 zeigt das Segmentationsergebnis für ein Bild mit wesentlich einfacheren Strukturen (Bild 18). Das Ergebnis läßt sich daher einfacher bewerten. Die Teilbäume wurden wieder bis zu einer Tiefe von $\tau = 7$ verfolgt; als Schwellwertfaktor wurde $\theta = 1.5$ gewählt. Hierfür wurden 171 Bereiche erzeugt. Es zeigt sich, daß sich alle wesentlichen Strukturen des Bildes in der Segmentation wiederfinden: Die Oberflächen der Objekte sowie deren Schatten konnten als Bereiche segmentiert werden. Die Oberfläche des Tisches ist jedoch mit dem Hintergrund auf der rechten Seite verschmolzen, da beide — wieder bedingt durch Rauschen — durch eine sehr kurze Kante des MST verbunden sind. Wie oben schon erwähnt, werden wir auf dieses Problem in Abschnitt 4.2.4 noch zurückkommen.

4.2.1 Verschmelzen zu kleiner Bereiche

Dieses Segmentationsergebnis zeigt aber auch, daß durch das Verfahren Bereiche erzeugt werden, die so klein sind, daß sie für diejenigen Prozesse, die auf den Segmentationsergebnissen aufbauen, keine brauchbaren Informationen enthalten. Solche Bereiche entstehen bevorzugt an den Übergängen größerer Bereiche (siehe hierzu z.B. den unteren Rand des Pyramidenstumpfes in Bild 19). Das nachträgliche Verschmelzen eines kleinen Bereiches mit einem Nachbarbereich ist mit Hilfe des MST nun auf sehr einfache Weise möglich:

Entstanden ist der Bereich durch Entfernen von Kanten, die den Teilbaum des MST, dem unser Bereich entspricht, mit anderen Teilbäumen verbinden. Zum Verschmelzen wird nun einfach die kürzeste dieser Kanten wieder eingefügt. Dadurch wird der Bereich mit demjenigen Nachbarbereich verschmolzen, dem er am ähnlichsten ist, denn kurzer Abstand im Merkmalsraum bedeutet Ähnlichkeit bezüglich der Merkmale.

Wir verschmelzen Bereiche nun solange mit ihren jeweils ähnlichsten Nachbarn, bis ihre Größe eine vorgegebene Mindestgröße θ_y nicht mehr unterschreitet.

Bild 20 zeigt die Segmentation für $\theta_g = 30$. Hiermit wurden jetzt noch 49 Bereiche erzeugt. Bild 14 zeigt die Segmentation des Hausbildes mit $\theta_g = 20$. Es wurden nur noch 256 Bereiche erzeugt.

Die Wahl des Schwellwertes θ_g kann durch Vorwissen über die zu segmentierenden Bilder beeinflusst werden. Die interessanten Strukturen des Blockbildes sind z.B. alle recht groß, so daß die Schwelle hier entsprechend groß gewählt werden kann. Denkbar wäre auch, daß die auf der Segmentation aufbauenden Verfahren der Korrespondenzanalyse eine Vorgabe für die von ihnen erwartete Mindestgröße machen.

4.2.2 Eine Schwelle für zu kurze Kanten

In Bild 20 fällt auf, daß der relativ homogene Hintergrund in mehrere Bereiche zerlegt wurde. Die Analyse der bei der Segmentation entfernten Kanten zeigte, daß diese Bereiche jeweils durch extrem kurze Kanten miteinander verbunden waren. Die Länge dieser Kanten beträgt $\ell \leq 10$, d.h., die Pixel unterscheiden sich in jedem der drei Farbkanäle nur um ein bis zwei Intensitätswerte voneinander. Das Entfernen solcher Kanten soll vermieden werden, und wir erhalten

Zerlegungskriterium 3

Wir modifizieren das Zerlegungskriterium 2, indem wir für eine zu entfernende Kante eine minimale Länge θ_{min} fordern. Eine Kante wird dann entfernt, wenn gilt:

$$(\ell > \bar{\ell}_1 + \theta \cdot \sigma_1) \wedge (\ell > \bar{\ell}_2 + \theta \cdot \sigma_2) \wedge (\ell > \theta_{min}).$$

Dieses Zerlegungskriterium hat außerdem den Vorteil, daß wir hierdurch Rechenzeit einsparen können: Für eine Kante, deren Länge den Schwellwert θ_{min} unterschreitet, brauchen wir nicht in die beiden Teilbäume abzusteigen, um die Verteilungsparameter zu berechnen.

Mit $\theta_{min} = 10$ erhalten wir für unser Blockbild dann 41 Bereiche (siehe Bild 21). Die Segmentation des Hausbildes ändert sich für $\theta_{min} = 10$ nicht.

4.2.3 Segmentation eines synthetischen Bildes mit feinen Bildstrukturen

In Abschnitt 2.4 hatten wir gesehen, daß Verfahren zur Bereichsegmentation durch Ballungsanalyse mit anschließender Relaxation besondere Schwierigkeiten mit feinen Bildstrukturen haben [Nagin et al. 82]. Um hier die Möglichkeiten unseres Verfahrens zu untersuchen, wollen wir es auf dem synthetischen Bild erproben, das Nagin et al. 82 in ihrem Artikel auf Seite 267

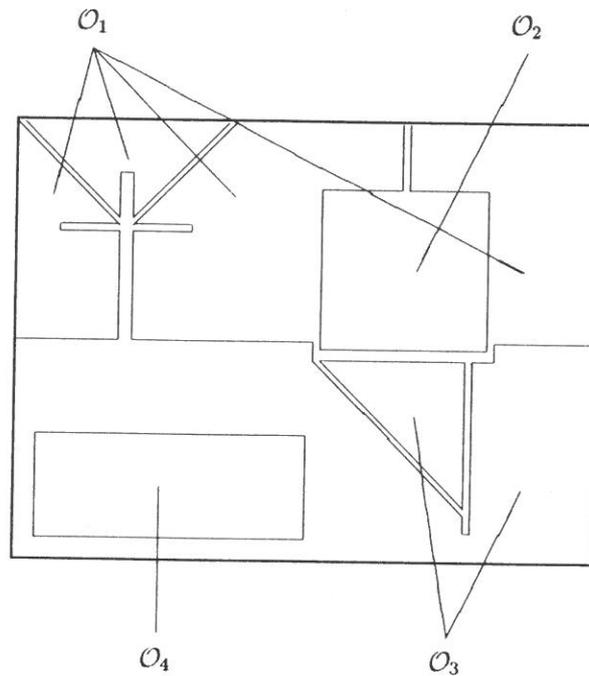


Abbildung 4.1: Skizze des synthetischen Bildes

als Fig. 3d zeigen. Dazu mußte dieses Bild nach den Angaben aus dem Artikel jedoch erst erzeugt werden.

Nagin et al. 82 verwenden für ihr Bild offenbar eine Auflösung von 6 bit pro Pixel. Wir haben daher die von ihnen angegebenen Grauwerte auf eine 8-bit-Auflösung umgerechnet. Es ergaben sich dabei folgende Werte (siehe hierzu Abbildung 4.1):

- Fläche O_1 : $\mu_1 = 120$,
- Fläche O_2 : $\mu_2 = 160$,
- Fläche O_3 : $\mu_3 = 192$,
- Fläche O_4 : $\mu_4 = 46 \dots 69$,

Standardabweichung des Rauschens: $\sigma = 12$.

Bei der Fläche O_4 handelt es sich um eine 'Rampe', deren Intensität, beginnend von oben, alle zwei Zeilen um jeweils eins steigt. Mit diesen Wer-

ten wurde dann ein in allen drei Farbkanälen identisches Farbbild erzeugt, das anschließend in jedem Farbkanal mit einem normalverteilten Rauschen überlagert wurde (Bild 1).

Unser Verfahren sollte mit den feinen Linienstrukturen keinerlei Schwierigkeiten haben, da der MST die Linien 'entlanglaufen' kann, um deren Pixel jeweils zu Teilbäumen zusammenzufassen. Dies können wir ganz deutlich in Abbildung 4.2 sehen, die den für unser synthetisches Bild berechneten MST darstellt.

Bild 2 zeigt die Segmentation dieses Bildes mit unserem Verfahren. Es wurden 111 Bereiche erzeugt mit den folgenden Parametern: $\tau = 7$, $\theta = 1.5$, $\theta_g = 0$ und $\theta_{min} = 0$. Unsere Erwartung zeigt sich bestätigt: Die feinen Linien des Bildes bleiben bei der Segmentation erhalten.

Viel auffallender als dieses positive Ergebnis ist jedoch, daß die beiden in ihrer Farbe visuell deutlich unterschiedlichen Flächen \mathcal{O}_1 und \mathcal{O}_2 in der Segmentation nicht voneinander getrennt werden konnten. Dieses Problem soll im nachfolgenden Abschnitt näher untersucht werden.

4.2.4 Probleme durch Rauschen

Wie in Abschnitt 2.4 schon erwähnt, ist eine Ballungsanalyse mittels eines MST recht anfällig gegen Rauschen. Ein verstärktes Rauschen ist gleichbedeutend mit einer höheren Varianz der einzelnen Merkmale. Ballungen im Merkmalsraum werden daher unschärfer und gehen eventuell sogar ineinander über. In unserem Falle führt Rauschen dazu, daß die Ränder der Bereiche unschärfer werden, d.h., die Farbdifferenzen der Pixel werden teilweise geringer.

Dies ist genau der Grund für die Probleme bei der Segmentation des synthetischen Bildes, die wir am Ende des vorigen Abschnitts gesehen haben. Eine Untersuchung des bei der Segmentation des synthetischen Bildes berechneten MST zeigte, daß die Flächen \mathcal{O}_1 und \mathcal{O}_2 durch eine sehr kurze Kante verbunden sind, da die Farbwerte eines in der Fläche \mathcal{O}_1 liegenden Pixels durch das Rauschen extrem erhöht, die Farbwerte eines benachbarten Pixels aus der Fläche \mathcal{O}_2 dagegen extrem erniedrigt wurden. Die Kante, die diese beiden Pixel und damit auch die Flächen \mathcal{O}_1 und \mathcal{O}_2 verbindet, ist sogar kürzer als die mittleren Kantenlängen $\bar{\ell}_1$ und $\bar{\ell}_2$ ihrer 'Umgebung'; die Kante könnte also nur bei einem negativen (!) Schwellwertfaktor θ entfernt werden. Was dies für den Rest der Segmentation bedeuten würde, sei der Vorstellungskraft des Lesers überlassen.

Eine naheliegende Idee, dieses Problem zu lösen, besteht darin, die Origi-

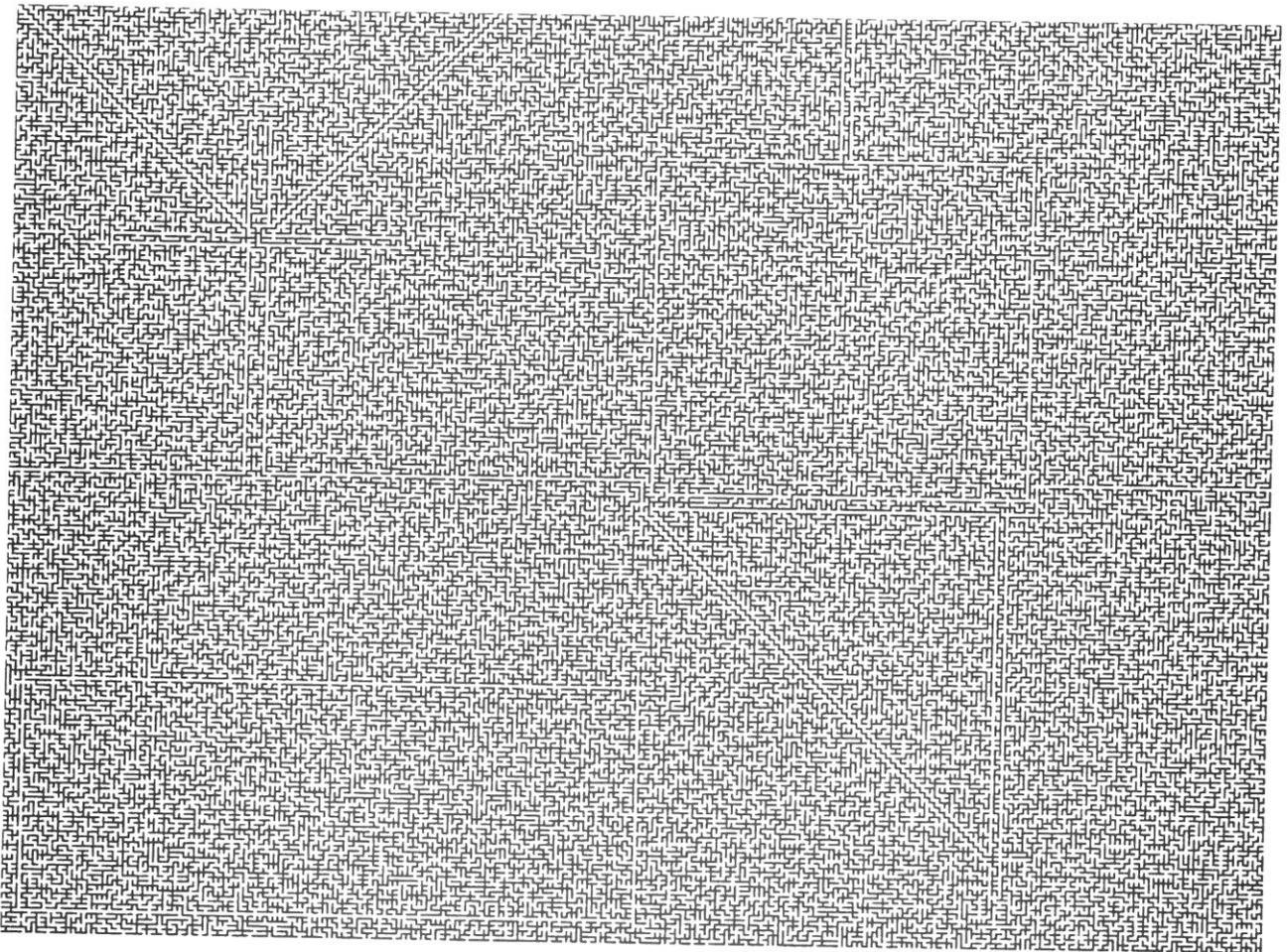


Abbildung 4.2: Berechneter MST für unser synthetisches Bild (vergleiche Abbildung 4.1 sowie Bild 1). Die Längen der einzelnen Kanten sind dieser Darstellung nicht zu entnehmen.

nalbilder vor der Segmentation zu filtern, um das Rauschen zu unterdrücken. Hierbei bedarf es jedoch besonderer Vorsicht: Mittelwert- bzw. Tiefpaßfilterung dürfen nicht verwendet werden, da hierdurch scharfe Diskontinuitäten verloren gehen würden, sprunghafte Farbübergänge würden geglättet werden. Eine sprunghafte Änderung in der Farbe signalisiert aber gerade eine Grenze zwischen zwei Bereichen, und so ist es natürlich wünschenswert, daß diese nicht geglättet wird, da hierdurch eine erfolgreiche Segmentation verhindert werden könnte. Es muß also eine Filterung durchgeführt werden, die das Rauschen beseitigt, scharfe Diskontinuitäten jedoch beibehält. Eine Medianfilterung kommt dieser Forderung näher, da sie nicht über die einzelnen Intensitäten mittelt, sondern die im Originalbild vorhandenen Intensitäten bei der Filterung verwendet. Wie führt man aber auf Farbbildern eine Medianfilterung durch? Im dreidimensionalen Farbraum ist der Median nicht definiert. Die Filterung darf aber auch nicht auf den einzelnen Farbkanälen separat durchgeführt werden, da hierdurch Farbvektoren entstehen könnten, die vorher im Bild nicht vorhanden waren. Wir haben hierzu einige Versuche durchgeführt, jedoch alle mit negativem Ausgang: die Segmentationsergebnisse waren nach einer Filterung schlechter als vorher. Der Einfluß einer Filterung auf die Segmentation war zudem sehr schwer zu beurteilen, da nicht nachvollzogen werden konnte, wie eine Filterung den MST und dadurch auch die Segmentation verändert. Dies scheint allgemein für die Medianfilterung zu gelten. Bovik et al. 87 schreiben (Seite 193):

It is very difficult to make any strong statements regarding the efficacy of the technique for any of the many other interesting edge detectors without massive experimentation, as the analysis of median filters remains intractable for all but the most simple assumptions.

Ein grundsätzliches Problem der von uns verwendeten Ballungsanalyse mittels eines MST besteht darin, daß zwei Ballungen — hier Bereiche — jeweils nur durch eine einzige Kante miteinander verbunden sind (in der englischsprachigen Literatur wird dieses Verfahren auch als *single linkage clustering* bezeichnet). Bei unserem Zerlegungskriterium für den MST wird für jede Kante getestet, ob diese in die Verteilung der Kantenlängen ihrer 'Umgebung' paßt. Für die Entscheidung zum Entfernen einer Kante steht damit also lediglich eine einzige Größe zur Verfügung: die Länge der Kante. Wird diese durch Rauschen gestört, so ist eine korrekte Entscheidung kaum mehr möglich.

Wir haben die Möglichkeit, zu erproben, wie sich unser Verfahren bei geringerem Rauschen verhält, indem wir das gleiche synthetische Bild noch

einmal erzeugen, diesmal mit einer Standardabweichung für das Rauschen von $\sigma = 6$ (Bild 7). Mit den Parametern von $\tau = 7$, $\theta = 3.0$, $\theta_g = 0$ sowie $\theta_{min} = 0$ erhalten wir dann eine Segmentation in 10 Bereiche (Bild 8). Diese Segmentation ist nun fast perfekt. Es wurden nur noch 3 zusätzliche Bereiche erzeugt und Flächen \mathcal{O}_1 und \mathcal{O}_2 sind nicht mehr miteinander verschmolzen.

Bisher haben wir mehr oder weniger mit Adhoc-Verfahren versucht, unseren MST in Teilbäume und damit das Bild in Bereiche zu zerlegen. Dabei haben wir die Längen der MST-Kanten für das jeweilige Kriterium benutzt. Es hat sich gezeigt, daß wir keine korrekte Entscheidung mehr treffen können, wenn diese Kantenlängen durch Rauschen verfälscht sind. Im folgenden Kapitel wollen wir nun versuchen, dem Problem des Rauschens Herr zu werden, indem wir einen etwas fundierteren Ansatz durch die Verwendung von statistischen Testverfahren untersuchen. Wir werden dabei dann nicht mehr die Kantenlängen benutzen, sondern auf die Farbvektoren der einzelnen Pixel zurückgreifen.

Kapitel 5

Statistische Zerlegungsverfahren

5.1 Das Yakimovsky-Kriterium

Im vorigen Kapitel haben wir gesehen, daß *eine* Größe für die Entscheidung zum Entfernen einer Kante des MST nicht ausreicht, wenn diese durch Rauschen verfälscht ist. Es besteht also Anlaß, über ein anderes Zerlegungskriterium nachzudenken. Eine Möglichkeit für ein solches wäre ein von Yakimovsky 76 zur Detektion von Kanten¹ vorgeschlagenes Verfahren:

Hierbei werden zwei disjunkte, in ihrer Form vorgegebene, räumlich benachbarte Flächen \mathcal{A}_1 und \mathcal{A}_2 des Bildes als 'Umgebungen' einer Kante definiert. Dann wird ein Likelihood-Verhältnistest angegeben, um zwischen den beiden folgenden Hypothesen zu entscheiden:

H_0 : Die Grauwerte der Pixel aus den Flächen \mathcal{A}_1 und \mathcal{A}_2 stammen aus *einer* Normalverteilung $N(\mu_0, \sigma_0)$.

H_1 : Die Grauwerte der Pixel aus den Flächen \mathcal{A}_1 und \mathcal{A}_2 stammen aus *zwei* Normalverteilungen $N(\mu_1, \sigma_1)$ und $N(\mu_2, \sigma_2)$.

Eine hier nicht wiederholte Herleitung ergibt das folgende Verhältnis der Wahrscheinlichkeiten für die beiden Hypothesen:

¹Gemeint sind hier Kanten im Bild, nicht Kanten im MST

$$\frac{P(H_1)}{P(H_0)} = \frac{(\sigma_0)^{m+n}}{(\sigma_1)^m(\sigma_2)^n} \quad (5.1)$$

m und n sind hierbei jeweils die Zahl der Pixel der Flächen \mathcal{A}_1 bzw. \mathcal{A}_2 , die zur Schätzung der Verteilung $N(\mu_1, \sigma_1)$ bzw. $N(\mu_2, \sigma_2)$ verwendet wurden.

Überschreitet dieses Verhältnis eine Schwelle θ , so wird die Nullhypothese H_0 zurückgewiesen, und es wird ein Kantenelement zwischen den beiden Flächen \mathcal{A}_1 und \mathcal{A}_2 eingefügt.

Wird dieses Verfahren für die Verwendung auf Farbbildern angepaßt — die Herleitung einer Gleichung 5.1 entsprechenden Verhältnisses muß dann für eine dreidimensionale Normalverteilung erfolgen — so läßt es sich als Zerlegungskriterium für unseren Baum verwenden; bei Überschreiten der Schwelle θ , also der Ablehnung der Nullhypothese, wird eine Kante unseres MST entfernt werden.

Die Form der beiden Flächen \mathcal{A}_1 und \mathcal{A}_2 muß dabei nicht wie bei Yakimovsky explizit vorgegeben werden, sondern entspricht wieder in einfacher Weise der 'Umgebung' einer MST-Kante, wie dies auch schon bei den bisher verwendeten Zerlegungskriterien der Fall war. Hierdurch findet eine bessere Anpassung an die lokalen Eigenschaften des Bildes statt. Ein Unterschied besteht lediglich darin, daß jetzt die Pixel selbst und nicht mehr die MST-Kanten betrachtet werden.

Yakimovsky verwendet die durch den von ihm vorgeschlagenen Test detektierten Kantenelemente auch zur Segmentierung des Bildes in Bereiche. Dabei müssen die einzelnen detektierten Kantenelemente jedoch zu geschlossenen Kantenzügen vervollständigt werden. Nagel 76 zeigt, daß sich hierbei Probleme ergeben.

In Kapitel 1 haben wir erwähnt, daß wir kein Verfahren verwenden wollen, das Bilder aufgrund der von Kantendetektoren gelieferten Informationen in Bereiche segmentiert. Wir wollen den von Yakimovsky vorgeschlagenen Test hier jedoch auf eine ganz andere Weise verwenden: Mit dem Test soll die Entscheidung getroffen werden, ob eine MST-Kante unseres Baumes entfernt werden soll oder nicht. Die Bereiche selbst werden dabei aber durch Teilbäume unseres MST definiert.

Der Test von Yakimovsky unterscheidet sich auch für sich genommen von anderen Methoden, die zur Kantendetektion eingesetzt werden: Es wird nicht die Stärke der Diskontinuität an einer lokalen Stelle des Bildes gemessen, wie dies z.B. bei Gradientenoperatoren geschieht, sondern es wird

geprüft, ob die Gesamtfläche $\mathcal{A}_1 \cup \mathcal{A}_2$ ein Homogenitätskriterium erfüllt. Diesen Test in einem Verfahren zu Bereichssegmentation zu verwenden, ist deshalb keineswegs abwegig sondern liegt vielmehr nahe.

5.1.1 Herleitung des Yakimovsky-Tests für Farbe

In diesem Abschnitt wollen wir nun den Test von Yakimovsky für Farbe, oder allgemeiner, für multivariate Normalverteilungen herleiten. Dazu müssen wir etwas Statistik betreiben und werden aus diesem Grunde etwas formaler. Zuvor jedoch noch einige Bemerkungen zur verwendeten Notation:

Wir verwenden im folgenden Matrixschreibweise. Dabei werden Skalare klein, Vektoren klein und fett sowie Matrizen groß und fett geschrieben. Große kalligraphische Zeichen stehen für Mengen, nicht für Matrizen!

Wir haben zwei Proben $\mathcal{X}_1 = \{\mathbf{x}_{11}, \dots, \mathbf{x}_{1m}\}$ und $\mathcal{X}_2 = \{\mathbf{x}_{21}, \dots, \mathbf{x}_{2n}\}$, deren Beobachtungen \mathbf{x}_{1i} und \mathbf{x}_{2j} jeweils unabhängig aus den multivariaten Normalverteilungen $N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ bzw. $N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ gezogen worden sind, und wollen die folgende Hypothese testen:

$$H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2.$$

$$H_1: \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Etwas anders formuliert bedeutet dies:

H_0 : Die einzelnen Beobachtungen der Proben \mathcal{X}_1 und \mathcal{X}_2 kommen aus der gleichen Normalverteilung $N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$.

H_1 : Die einzelnen Beobachtungen der Proben \mathcal{X}_1 und \mathcal{X}_2 kommen aus unterschiedlichen Normalverteilungen $N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ und $N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$.

Wir erhalten einen Test für unsere Hypothese durch das folgende Likelihood-Verhältniskriterium [Bickel + Doksum 77]:

$$\lambda = \frac{\sup L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2 | H_1)}{\sup L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | H_0)} \quad (5.2)$$

Die Wahrscheinlichkeitsdichte einer multivariaten Normalverteilung ist gegeben durch

$$p(\mathbf{x}_i) = \frac{1}{\sqrt{(2\pi)^d \det(\boldsymbol{\Sigma})}} \exp \left[-\frac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right],$$

und die Likelihood-Funktionen aus Gleichung 5.2 sind dann gegeben durch das Produkt der Wahrscheinlichkeitsdichten:

$$L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | H_0) = \prod_{i=1}^m p(\mathbf{x}_{1i}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \prod_{j=1}^n p(\mathbf{x}_{2j}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

$$L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2 | H_1) = \prod_{i=1}^m p(\mathbf{x}_{1i}, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \prod_{j=1}^n p(\mathbf{x}_{2j}, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$$

Hiermit erhalten wir dann:

$$L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | H_0) = \left(\frac{1}{(2\pi)^d \det(\boldsymbol{\Sigma}_0)} \right)^{\frac{m+n}{2}}$$

$$\cdot \exp \left[-\frac{1}{2} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{x}_{1i} - \boldsymbol{\mu}_0) + \sum_{j=1}^n (\mathbf{x}_{2j} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{x}_{2j} - \boldsymbol{\mu}_0) \right) \right]$$

$$L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2 | H_1) = \left(\frac{1}{(2\pi)^d} \right)^{\frac{m+n}{2}} \left(\frac{1}{\det(\boldsymbol{\Sigma}_1)} \right)^{\frac{m}{2}} \left(\frac{1}{\det(\boldsymbol{\Sigma}_2)} \right)^{\frac{n}{2}}$$

$$\cdot \exp \left[-\frac{1}{2} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_1^{-1} (\mathbf{x}_{1i} - \boldsymbol{\mu}_1) + \sum_{j=1}^n (\mathbf{x}_{2j} - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}_2^{-1} (\mathbf{x}_{2j} - \boldsymbol{\mu}_2) \right) \right]$$

Das Supremum der Likelihood-Funktionen erhalten wir durch die Maximumlikelihood-Schätzungen der Verteilungsparameter (siehe hierzu Anhang B, Gleichungen B.1 und B.2):

$$\hat{\boldsymbol{\mu}}_0 = \frac{1}{m+n} \left(\sum_{i=1}^m \mathbf{x}_{1i} + \sum_{j=1}^n \mathbf{x}_{2j} \right) \quad (5.3)$$

$$= \frac{1}{m+n} (m\hat{\boldsymbol{\mu}}_1 + n\hat{\boldsymbol{\mu}}_2)$$

$$\hat{\boldsymbol{\mu}}_1 = \frac{1}{m} \sum_{i=1}^m \mathbf{x}_{1i} \quad (5.4)$$

$$\hat{\boldsymbol{\mu}}_2 = \frac{1}{n} \sum_{j=1}^n \mathbf{x}_{2j} \quad (5.5)$$

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{1}{m+n} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_0)(\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_0)^T + \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_0)(\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_0)^T \right) \quad (5.6)$$

$$\hat{\Sigma}_1 = \frac{1}{m} \sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)(\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)^T \quad (5.7)$$

$$\hat{\Sigma}_2 = \frac{1}{n} \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)(\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)^T \quad (5.8)$$

Nach Einsetzen dieser Schätzungen in die Likelihood-Funktionen und Vereinfachung des Exponenten (Anhang D) ergibt sich für das Likelihood-Verhältnis aus Gleichung 5.2:

$$\lambda = \left(\frac{\det(\hat{\Sigma}_0)^{m+n}}{\det(\hat{\Sigma}_1)^m \det(\hat{\Sigma}_2)^n} \right)^{\frac{1}{2}} \quad (5.9)$$

Hiermit erhalten wir nun folgenden Test für unsere Nullhypothese:

Die Nullhypothese wird zurückgewiesen, wenn gilt

$$\lambda > \theta.$$

Dies ist die Verallgemeinerung des Tests von Yakimovsky für multivariate Normalverteilungen. Wollen wir ihn für unsere drei Farben benutzen, so müssen wir eine dreidimensionale Normalverteilung einsetzen. Die Kovarianzmatrizen aus Gleichung 5.9 sind dann 3×3 -Matrizen. Die beiden Proben \mathcal{X}_1 und \mathcal{X}_2 sind dabei zwei Mengen von Beobachtungen aus jeweils einer Fläche des Bildes. Die Beobachtungen \mathbf{x}_{1i} bzw. \mathbf{x}_{2j} entsprechen den Farbvektoren der einzelnen Pixel.

Interessant ist auch die folgende Analogie zur Ballungsanalyse: Der hier durchzuführende Test entspricht einem Test, ob die Farbvektoren der beiden Proben in einem dreidimensionalen Farbmerkmalsraum eine oder zwei Ballungen bilden.

Steht wie bei Schwarz-Weiß-Bildern nur ein Intensitätswert zur Verfügung, so muß eine eindimensionale Normalverteilung eingesetzt werden, und unsere Herleitung entspricht dann genau der von Yakimovsky 76. Unser Ergebnis aus Gleichung 5.9 geht dann in Gleichung 5.1 über.

5.1.2 Anwendung des hergeleiteten Tests

Den hier hergeleiteten Test wollen wir nun verwenden für das folgende

Zerlegungskriterium 4

Die 'Umgebung' einer MST-Kante und damit unsere beiden Proben \mathcal{X}_1 und \mathcal{X}_2 sei wieder durch zwei Teilbäume definiert.

Um die Berechnung des Verhältnisses aus Gleichung 5.9 zu vereinfachen, wollen wir die beiden Teilbäume jedoch nicht bis zu einer vorgegebenen Tiefe τ durchlaufen, sondern bis in beiden Teilbäumen die gleiche vorgegebene Anzahl von Pixeln erreicht worden ist. Dies bedeutet, daß unsere beiden Proben \mathcal{X}_1 und \mathcal{X}_2 die gleiche Anzahl von $m = n$ Beobachtungen haben. Dabei wird ein Teilbaum ausgehend von der zu untersuchenden Kante des MST im Sinne einer *breadth-first* Strategie durchlaufen; zuerst alle Pixel der Tiefe 1, dann diejenigen der Tiefe 2 usw. Ist die vorgegebene Anzahl n von Pixeln erreicht, bevor alle Pixel einer bestimmten Tiefe betrachtet wurden, so bleibt der Rest unberücksichtigt.

Eine Kante des MST wird nun entfernt, wenn gilt:

$$\lambda^{\frac{2}{n}} = \frac{\det(\hat{\Sigma}_0)^2}{\det(\hat{\Sigma}_1) \det(\hat{\Sigma}_2)} > \theta. \quad (5.10)$$

Wir fordern weiterhin eine Mindestgröße von θ_g für die Bereiche.

Wenn wir mit diesem Zerlegungskriterium unser synthetisches Bild segmentieren, so stellen wir fest, daß an den Rändern der eigentlichen Flächen zusätzlich viele sehr kleine Bereiche entstehen. Ein Vergleich mit dem für dieses Bild berechneten MST zeigt außerdem, daß diese in der unmittelbaren Umgebung derjenigen Kanten des Baumes entstehen, die zwei Flächen miteinander verbinden und entfernt worden sind. Der Grund hierfür läßt sich jedoch leicht angeben: Das Verhältnis 5.10 hat für eine Kante, die beispielsweise die beiden Flächen \mathcal{O}_1 und \mathcal{O}_2 miteinander verbindet, einen hohen Betrag, so daß die Kante entfernt wird. Haben wir nun eine sehr dicht benachbarte Kante, die z.B. zwei Pixel der Fläche \mathcal{O}_1 miteinander verbindet, so wird der Teilbaum, der an dieser Kante hängt, teilweise über die soeben betrachtete Kante in die Fläche \mathcal{O}_2 'hineinlaufen', und damit die Werte der Kovarianzmatrix für diesen Teilbaum vergrößern. Das Likelihood-Verhältnis hat damit auch für diese Kante einen hohen Betrag, so daß sie daraufhin entfernt wird. Das 'Hineinlaufen' in den anderen Bereich können wir zwar nicht verhindern, aber es gibt eine andere Möglichkeit, die Segmentationsfehler, die hierdurch zustande kommen, zu verhindern. Die Kanten, die von dem eigentlichen Übergang zwischen zwei Bereichen etwas entfernt liegen, werden ein geringeres Likelihood-Verhältnis haben als diejenige Kante, die die beiden Bereiche miteinander verbindet. Dies führt uns zu

Zerlegungskriterium 5

Eine Kante des MST wird entfernt, wenn sie das Zerlegungskriterium 4 erfüllt

und das Verhältnis 5.10 für diese Kante zusätzlich ein *lokales Maximum* innerhalb der Umgebung von maximal sechs möglichen Nachbarkanten ist (siehe Abbildung 5.1).

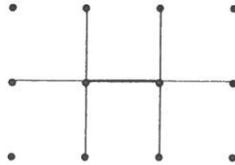


Abbildung 5.1: Die sechs möglichen Nachbarkanten einer MST-Kante.

Mit diesem Zerlegungskriterium und den Parametern $n = 15$, $\theta = 120$ und $\theta_g = 0$ erhalten wir nun eine fast perfekte Segmentation des synthetischen Bildes in 11 Bereiche (Bild 3). Ein kleiner zusätzlicher Bereich findet sich immer noch direkt am Übergang zwischen den Flächen \mathcal{O}_3 und \mathcal{O}_4 am linken unteren Rand des Bildes. Weiterhin sehen wir, daß ein Teil der zur Fläche \mathcal{O}_2 gehörenden dünnen Linie am oberen Rand des Bild als eigener Bereich abgetrennt wurde, und daß die Verbindungslinien zwischen dem mittleren und rechten Teil der Fläche \mathcal{O}_1 einen eigenen Bereich darstellen.

Diese Segmentationsfehler können wir erklären, wenn wir den für dieses Bild berechneten MST untersuchen (Abbildung 4.2, Seite 38). Hierbei stellen wir fest, daß der MST den mittleren und rechten Teil der Fläche \mathcal{O}_1 gar nicht direkt über die dünnen Linien verbindet, und daß der obere Teil der dünnen Linie von Fläche \mathcal{O}_2 nicht direkt mit dem unteren Teil sondern nach rechts mit der Fläche \mathcal{O}_1 verbunden ist. Das Rauschen dieses Bildes ist also so hoch, daß einige Kanten zwischen zwei unterschiedlich gefärbten Flächen kürzer sind als Kanten innerhalb einer Fläche. Gegen solch starkes Rauschen, das uns sogar einen 'falschen' MST verursacht, sind wir mit unserem Verfahren dann machtlos.

Wir wollen nun noch das Block- sowie unser Hausbild mit Hilfe dieses Zerlegungskriteriums segmentieren. Das Blockbild wurde mit den Parametern $n = 15$, $\theta = 130$ sowie $\theta_g = 30$ in 46 Bereiche segmentiert. Bild 22 zeigt das Ergebnis. Der Tisch konnte wie erwartet vom Hintergrund auf der rechten Seite getrennt werden. Dafür ist jetzt das weiße Papier in der Mitte mit dem Papier im Hintergrund verschmolzen, da der mittlere Farbwert dieser beiden Flächen fast gleich ist. Diese Segmentation unterscheidet sich in ihrer Qualität jedoch nicht grundlegend von der Segmentation mit Hilfe des Zerlegungskriteriums 3 in Bild 21.

Bild 15 zeigt die Segmentation des Hausbildes in 329 Bereiche. Die verwendeten Parameter waren $n = 15$, $\theta = 160$ und $\theta_g = 20$. Hier konnten die gravierenden Segmentationsfehler aufgrund von Rauschen vermieden werden: die Frontseite des linken Hauses konnte sowohl vom Dach als auch von der Front des rechten Hauses getrennt werden. Ein großer qualitativer Unterschied zu der Segmentation mit Hilfe des lokalen Zerlegungskriteriums in Bild 14 ist hier jedoch ebenfalls nicht festzustellen. Wir werden bei der abschließenden Bewertung in Kapitel 6 darauf noch einmal zurückkommen.

5.2 Hotellings T^2 -Test

Hsu et al. 82 verwenden den von Yakimovsky vorgeschlagenen Test zur Änderungsdetektion zwischen zwei zu unterschiedlichen Zeitpunkten aufgenommen Bildern, um dadurch Hinweise auf Bewegung zu bekommen. Dabei verwenden sie in jedem Bild eine quadratische Maske, die in beiden Bildern jeweils an die gleiche Stelle gelegt wird. Sie testen damit dann nicht die Homogenität von räumlich, sondern die von *zeitlich* benachbarten Bereichen.

In ihrer Arbeit zeigen sie auch, daß das Likelihood-Verhältnis aus Gleichung 5.1, das wir hier für multivariate Normalverteilungen hergeleitet haben, unter bestimmten Annahmen auf eine bekannte und tabellierte Verteilung, nämlich die Studentsche t -Verteilung, zurückgeführt werden kann. Dies hat den Vorteil, daß man den Schwellwert θ dann nicht mehr interaktiv ermitteln (ausprobieren) muß; man gibt vielmehr ein Konfidenzniveau vor und entnimmt den dazugehörigen Schwellwert aus einer Tabelle der Verteilungsquantile. Dies erscheint uns als ein großer Vorteil, so daß wir diesen Ansatz hier weiter verfolgen wollen.

Um das Likelihood-Verhältnis auf eine bekannte Verteilung zurückführen zu können, muß jedoch die Annahme gemacht werden, daß die Varianzen (in unserem Falle Kovarianzmatrizen) der zu untersuchenden Bildflächen gleich sind. Wir wollen hierzu kurz die Argumentation von [*Hsu et al. 82*] wiedergeben:

Die systematischen Grauwertstrukturen eines Bildes lassen sich stückweise als Flecken konstanten Grauwertes beschreiben (genau dies wird angenommen, wenn man die Pixel des Bildes als Proben aus einer Normalverteilung betrachtet und ihren Mittelwert schätzt).

Die Grauwertvarianz eines solchen Flecks enthält dann keine Anteile inadäquat modellierter Grauwertstrukturen; sie wird allein durch das Rauschen verursacht.

Das Rauschen sei ein durch die Aufnahmeapparatur verursachtes *additives* Rauschen. Aus diesem Grunde erscheint es gerechtfertigt, anzunehmen, daß das Rauschen zweier benachbarter Bereiche der gleichen Verteilung unterliegt.

Da die Grauwertvarianz aber nur durch das Rauschen verursacht wird, folgt hieraus, daß die Varianzen zweier benachbarter Flecken gleich sind.

Im folgenden Abschnitt wollen wir nun die Herleitung für eine multivariate Normalverteilung versuchen.

5.2.1 Herleitung des T^2 -Tests von Hotelling

Wir haben zwei Proben $\mathcal{X}_1 = \{\mathbf{x}_{11}, \dots, \mathbf{x}_{1m}\}$ und $\mathcal{X}_2 = \{\mathbf{x}_{21}, \dots, \mathbf{x}_{2n}\}$, die jeweils unabhängig aus den multivariaten Normalverteilungen $N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ bzw. $N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$ gezogen worden sind, und wollen die folgende Hypothese testen:

H_0 : Die einzelnen Beobachtungen der Proben \mathcal{X}_1 und \mathcal{X}_2 kommen aus der gleichen Normalverteilung $N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$.

H_1 : Die einzelnen Beobachtungen der Proben \mathcal{X}_1 und \mathcal{X}_2 kommen aus unterschiedlichen Normalverteilungen $N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ und $N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$.

Es wird dabei vorausgesetzt, daß die Kovarianzmatrizen unbekannt aber gleich sind ($\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$)!

Einen Test für unsere Hypothese konstruieren wir wieder durch das folgende Likelihood-Verhältnis:

$$\lambda = \frac{\sup L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma} | H_1)}{\sup L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | H_0)} \quad (5.11)$$

Die Likelihood-Funktionen sind dabei wieder gegeben durch das Produkt der Wahrscheinlichkeitsdichten (beachte, daß sich die Likelihood-Funktion

für die Hypothese H_1 gegenüber Abschnitt 5.1.1 geändert hat):

$$L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0 | H_0) = \left(\frac{1}{(2\pi)^d \det(\boldsymbol{\Sigma}_0)} \right)^{\frac{m+n}{2}} \cdot \exp \left[-\frac{1}{2} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{x}_{1i} - \boldsymbol{\mu}_0) + \sum_{j=1}^n (\mathbf{x}_{2j} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{x}_{2j} - \boldsymbol{\mu}_0) \right) \right]$$

$$L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma} | H_1) = \left(\frac{1}{(2\pi)^d \det(\boldsymbol{\Sigma})} \right)^{\frac{m+n}{2}} \cdot \exp \left[-\frac{1}{2} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_{1i} - \boldsymbol{\mu}_1) + \sum_{j=1}^n (\mathbf{x}_{2j} - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_{2j} - \boldsymbol{\mu}_2) \right) \right]$$

Hieraus erhalten wir die folgenden Maximumlikelihood-Schätzungen für die Verteilungsparameter (Anhang B, Gleichungen B.1 und B.2):

$$\begin{aligned} \hat{\boldsymbol{\mu}}_0 &= \frac{1}{m+n} \left(\sum_{i=1}^m \mathbf{x}_{1i} + \sum_{j=1}^n \mathbf{x}_{2j} \right) \\ &= \frac{1}{m+n} (m\hat{\boldsymbol{\mu}}_1 + n\hat{\boldsymbol{\mu}}_2) \\ \hat{\boldsymbol{\mu}}_1 &= \frac{1}{m} \sum_{i=1}^m \mathbf{x}_{1i} \\ \hat{\boldsymbol{\mu}}_2 &= \frac{1}{n} \sum_{j=1}^n \mathbf{x}_{2j} \\ \hat{\boldsymbol{\Sigma}}_0 &= \frac{1}{m+n} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_0)(\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_0)^T + \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_0)(\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_0)^T \right) \\ \hat{\boldsymbol{\Sigma}} &= \frac{1}{m+n} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)(\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)^T + \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)(\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)^T \right) \end{aligned} \quad (5.12)$$

Nach Einsetzen dieser Schätzungen und Vereinfachung des Exponenten (Anhang D) erhalten wir für das Likelihood-Verhältnis:²

$$\lambda = \left(\frac{\det(\hat{\boldsymbol{\Sigma}}_0)}{\det(\hat{\boldsymbol{\Sigma}})} \right)^{\frac{m+n}{2}}$$

²Im folgenden verwenden wir $|\mathbf{M}|$ für $\det(\mathbf{M})$.

$$= \left(\frac{\left| \sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_0)(\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_0)^T + \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_0)(\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_0)^T \right|}{\left| \sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)(\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)^T + \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)(\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)^T \right|} \right)^{\frac{m+n}{2}}$$

Wir können nun $\hat{\boldsymbol{\Sigma}}_0$ in Abhängigkeit von $\hat{\boldsymbol{\mu}}_1$ und $\hat{\boldsymbol{\mu}}_2$ ausdrücken. Eine kleine Zwischenrechnung zeigt, daß gilt:

$$\begin{aligned} \hat{\boldsymbol{\Sigma}}_0 &= \frac{1}{m+n} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_0)(\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_0)^T + \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_0)(\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_0)^T \right) \\ &= \frac{1}{m+n} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)(\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)^T + \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)(\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)^T \right. \\ &\quad \left. + \frac{mn}{m+n} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \right) \end{aligned}$$

Verwenden wir noch folgende Abkürzung

$$\begin{aligned} \mathbf{A} &= (m+n)\hat{\boldsymbol{\Sigma}} \\ &= \sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)(\mathbf{x}_{1i} - \hat{\boldsymbol{\mu}}_1)^T + \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)(\mathbf{x}_{2j} - \hat{\boldsymbol{\mu}}_2)^T, \end{aligned}$$

so erhalten wir für das Likelihood-Verhältnis

$$\lambda = \left(\frac{\left| \mathbf{A} + \frac{mn}{m+n} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \right|}{|\mathbf{A}|} \right)^{\frac{m+n}{2}}$$

Wir verwenden nun folgende Rechenregeln für partitionierte Matrizen [Morison 67, S. 65f]:

Sei

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix}$$

mit \mathbf{M}_{11} , \mathbf{M}_{12} , \mathbf{M}_{21} und \mathbf{M}_{22} als Untermatrizen, dann gilt

$$|\mathbf{M}| = |\mathbf{M}_{11}| \cdot |\mathbf{M}_{22} - \mathbf{M}_{21}\mathbf{M}_{11}^{-1}\mathbf{M}_{12}| \quad \text{wenn } \mathbf{M}_{11} \text{ regulär (5.13)}$$

$$\text{und } |\mathbf{M}| = |\mathbf{M}_{22}| \cdot |\mathbf{M}_{11} - \mathbf{M}_{12}\mathbf{M}_{22}^{-1}\mathbf{M}_{21}| \quad \text{wenn } \mathbf{M}_{22} \text{ regulär. (5.14)}$$

Hiermit können wir nun folgende Umformungen durchführen:

$$\begin{aligned}
 \lambda &= \left(\frac{|1| \cdot \left| \mathbf{A} + \sqrt{\frac{mn}{m+n}}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2) \sqrt{\frac{mn}{m+n}}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \right|}{|\mathbf{A}|} \right)^{\frac{m+n}{2}} \\
 &= \left(\frac{\begin{vmatrix} 1 & \sqrt{\frac{mn}{m+n}}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \\ -\sqrt{\frac{mn}{m+n}}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2) & \mathbf{A} \end{vmatrix}}{|\mathbf{A}|} \right)^{\frac{m+n}{2}} \\
 &= \left(\frac{|\mathbf{A}| \cdot \left| 1 + \frac{mn}{m+n}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \mathbf{A}^{-1}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2) \right|}{|\mathbf{A}|} \right)^{\frac{m+n}{2}} \\
 &= \left(1 + \frac{mn}{m+n}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \mathbf{A}^{-1}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2) \right)^{\frac{m+n}{2}}
 \end{aligned}$$

Wir verwenden noch folgende Abkürzungen

$$\begin{aligned}
 \mathbf{S} &= \frac{1}{m+n-2} \mathbf{A} \\
 &= \frac{m+n}{m+n-2} \hat{\boldsymbol{\Sigma}}
 \end{aligned}$$

und

$$T^2 = \frac{mn}{m+n}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \mathbf{S}^{-1}(\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2), \quad (5.15)$$

und erhalten damit dann das Likelihood-Verhältnis

$$\lambda = \left(1 + \frac{T^2}{m+n-2} \right)^{\frac{m+n}{2}}$$

T^2 ist die sogenannte *Hotellings T^2 -Statistik* mit $m+n-2$ Freiheitsgraden, eine Verallgemeinerung der Studentischen t -Statistik.

Ist die Nullhypothese $H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ wahr, dann hat die Größe

$$\hat{F} = \frac{m+n-d-1}{(m+n-2)d} T^2$$

eine F -Verteilung mit d und $m+n-d-1$ Freiheitsgraden [Anderson 58], wobei d wieder die Dimension der Verteilung unserer Proben ist.

Da λ eine monotone Funktion von T^2 ist, erhalten wir den folgenden Test zum Konfidenzniveau α [Anderson 58]:

Die Nullhypothese wird zurückgewiesen, wenn gilt

$$T^2 > \frac{(m+n-2)d}{m+n-d-1} \mathcal{F}_{d, m+n-d-1}(\alpha) \quad (5.16)$$

wobei \mathcal{F} die tabellierten Quantile der F -Verteilung sind.

Der Parameter α gibt uns hierbei an, wie groß die Wahrscheinlichkeit ist, einen Fehler 1. Art zu machen, daß heißt, die Nullhypothese abzulehnen, obwohl sie wahr ist.

Dieser Test aus Gleichung 5.16 ist eine Verallgemeinerung des doppel- t -Tests für univariate Normalverteilungen, denn für $d = 1$ gilt

$$\frac{(m+n-2)d}{m+n-d-1} \mathcal{F}_{d, m+n-d-1}(\alpha) = (t_{m+n-2}(\alpha))^2.$$

Setzen wir eine univariate Normalverteilung ein, so geht unsere Herleitung genau in die Herleitung von [Hsu et al. 82, Kapitel 5, Gleichungen 6–16] über.

Bei der Berechnung von T^2 nach Gleichung 5.15 können wir die Invertierung der Matrix \mathbf{S} vermeiden, wenn wir mit Hilfe der Gleichungen 5.13 und 5.14 noch folgende Umformung vornehmen:

$$\begin{aligned} T^2 &= \frac{mn}{m+n} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \mathbf{S}^{-1} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2) \\ &= \frac{|\mathbf{S}| \cdot \left| 1 + \sqrt{\frac{mn}{m+n}} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \mathbf{S}^{-1} \sqrt{\frac{mn}{m+n}} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2) \right|}{|\mathbf{S}|} - 1 \\ &= \frac{\left| \begin{array}{cc} 1 & -\sqrt{\frac{mn}{m+n}} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \\ \sqrt{\frac{mn}{m+n}} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2) & \mathbf{S} \end{array} \right|}{|\mathbf{S}|} - 1 \\ &= \frac{\left| \mathbf{S} + \frac{mn}{m+n} (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2) (\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2)^T \right|}{|\mathbf{S}|} - 1 \end{aligned} \quad (5.17)$$

5.2.2 Anwendung des T^2 -Tests

Wir erhalten hiermit dann

Zerlegungskriterium 6

Aus den beiden Teilbäumen, die durch die zu betrachtende MST-Kante definiert werden, werden nach der *breadth-first* Strategie wieder jeweils n Pixel als Beobachtungen entnommen. Eine Kante wird dann entfernt, wenn durch den in Formel 5.16 angegebenen Test die Nullhypothese zurückgewiesen wird. Das Konfidenzniveau α kann dabei frei gewählt werden. Außerdem muß der in Formel 5.16 berechnete Wert ein lokales Maximum sein, und es wird weiterhin eine Mindestgröße θ_g von den Bereichen gefordert.

Bild 4 zeigt die Segmentation des synthetischen Bildes in 232 Bereiche mit folgenden Parametern: Für die Anzahl der Proben je Teilbaum haben wir wieder $n = 15$ gewählt, die Mindestgröße war $\theta_g = 0$, und als Konfidenzniveau wurde $\alpha = 0.001$ gewählt und. Der Tabellenwert hierfür ist damit $\mathcal{F}_{d,m+n-d-1}(\alpha) = \mathcal{F}_{d,2n-4}(\alpha) = \mathcal{F}_{3,26}(0.001) = 7.36$ [Abramowitz + Stegun 68, S. 989]. Mit den gleichen Parametern für n und α sowie $\theta_g = 30$ haben wir auch unser Blockbild segmentiert (Bild 23). Hierfür erhielten wir 604 Bereiche. Wie wir sehen, sind beide Bilder in extrem viele Bereiche segmentiert worden.

Wie kommt dieses schlechte Ergebnis zustande? Um das Likelihood-Verhältnis auf eine bekannte Verteilung zurückführen zu können, mußten wir die Annahme machen, daß die Kovarianzmatrizen der beiden betrachteten Bildflächen gleich sind ($\Sigma_1 = \Sigma_2 = \Sigma$). Sollte diese Annahme nicht gerechtfertigt sein, so können wir den oben hergeleiteten Test natürlich nicht anwenden. Wir wollen deshalb einmal die Schätzungen $\hat{\Sigma}_1$, $\hat{\Sigma}_2$ sowie $\hat{\Sigma}$ für eine Kante des synthetischen Bildes aus der Fläche \mathcal{O}_3 untersuchen, die durch den Test entfernt worden ist. Wir berechnen diese Schätzungen nach den Gleichungen 5.7, 5.8 sowie 5.12, und erhalten die folgenden Matrizen:

$$\hat{\Sigma}_1 = \begin{pmatrix} 156.8 & 32.0 & -6.6 \\ 32.0 & 142.8 & 23.2 \\ -6.6 & 23.2 & 80.5 \end{pmatrix}$$

$$\hat{\Sigma}_2 = \begin{pmatrix} 77.6 & -16.0 & 39.6 \\ -16.0 & 33.0 & -2.6 \\ 39.6 & -2.6 & 88.9 \end{pmatrix}$$

$$\hat{\Sigma} = \begin{pmatrix} 117.2 & 8.0 & 16.5 \\ 8.0 & 87.9 & 10.3 \\ 16.5 & 10.3 & 84.7 \end{pmatrix}$$

Wir sehen hier, daß die drei Kovarianzmatrizen alles andere als gleich sind. Bei dem Bild einer natürlichen Szene mit komplexeren Bildstrukturen ließe sich dies dadurch erklären, daß die Annahme, das Bild lasse sich durch Flecken stückweise konstanter Farbe beschreiben, schon falsch ist. Die Varianzen würden dann teilweise auch durch inadäquat modellierte Bildstrukturen verursacht; die Varianzen könnten also demnach nicht gleich sein. Das hier vorliegende synthetischen Bild entspricht jedoch genau unseren Modellvorstellungen: die Flächen wurden mit konstanter Farbe eingefärbt und anschließend mit einem normalverteilten Rauschen überlagert. Wie läßt sich also erklären, daß die Kovarianzmatrizen nicht gleich sind?

Betrachten wir einmal die Diagonalelemente der Kovarianzmatrizen, so fällt auf, daß ein Rauschen mit einer geringeren Standardabweichung als $\sigma = 12$ pro Farbkanal geschätzt wurde. Dies läßt sich durch unseren MST erklären: Der Baum verbindet über eine Kante immer Pixel miteinander, die sich sehr ähnlich sind. Man kann daher annehmen, daß in einem kleinen Teil des Baumes immer ähnliche Pixel zusammenhängen. Der Baum *gruppiert* also ähnliche Pixel zusammen. Dadurch verändert sich für eine solche Menge von Pixeln jedoch die Verteilung. Die Standardabweichung des Rauschens ist lokal für diese Gruppe von Pixeln geringer, als für eine Menge von Pixeln, die nicht durch den MST gruppiert, sondern wahllos betrachtet werden. Genau genommen bedeutet dies sogar, daß man Zweifel an der statistischen *Unabhängigkeit* der einzelnen Beobachtungen haben muß. In diesem Falle dürfte ein statistisches Testverfahren dann gar nicht angewendet werden. Wir haben mit der Verallgemeinerung des Yakimovsky-Tests, dem die Annahme der statistischen Unabhängigkeit ebenfalls zugrunde liegt, jedoch schon gute Resultate erzielt, so daß die Unabhängigkeit der Beobachtungen wohl doch angenommen werden darf.

Sollte ein kleiner Fleck des Bildes nun zudem noch zufällig eine geringere Varianz haben, so wird gerade dieser durch einen Teilbaum unseres MST überdeckt werden, und so ließen sich dann auch die ungleichen Kovarianzmatrizen erklären: Der MST reagiert sehr sensibel auf lokale Änderungen der Varianz, indem er gerade solche Flächen mit möglichst kleiner Varianz zusammengruppiert. Aus einer relativ kleinen Stichprobe, wie wir sie hier verwenden, wird man die Varianz einer größeren Fläche also nicht korrekt schätzen können. Dies führt dann zu den ungleichen Kovarianzmatrizen.

Für nicht synthetisch erzeugte, sondern tatsächlich mit einer Videokamera aufgenommene Bilder können wir auch noch einen Grund angeben, warum die Varianzen tatsächlich nicht gleich sind: Die visuelle Begutachtung unseres Bildmaterials legt den Verdacht nahe, daß das Rauschen signalabhängig ist. In dunklen Bildbereichen ist das Rauschen stärker als in hellen. Da das Signal (Bildhelligkeit) ortsabhängig ist, ist somit auch das Rauschen ortsabhängig und wir können aus diesem Grunde nicht die Annahme gleicher Kovarianzmatrizen machen (siehe hierzu auch [Mottschall 85]).

5.3 Das Behrens-Fisher Problem

Die Ergebnisse des vorigen Abschnitts haben gezeigt, daß wir die Annahme, die Kovarianzmatrizen seien gleich, nicht machen können. Wir wollen nun versuchen, einen Test herzuleiten, dem diese Annahme nicht zugrunde liegt. Hierdurch kommen wir zurück auf die Hypothese, die wir schon in Abschnitt 5.1.1 getestet haben:

H_0 : Die einzelnen Beobachtungen der Proben \mathcal{X}_1 und \mathcal{X}_2 kommen aus der gleichen Normalverteilung $N(\mu_0, \Sigma_0)$.

H_1 : Die einzelnen Beobachtungen der Proben \mathcal{X}_1 und \mathcal{X}_2 kommen aus unterschiedlichen Normalverteilungen $N(\mu_1, \Sigma_1)$ und $N(\mu_2, \Sigma_2)$.

Das Likelihood-Verhältnis aus Gleichung 5.2, aus dem wir einen Test für diese Hypothese konstruiert hatten, läßt sich unglücklicherweise jedoch nicht auf eine bekannte und tabellierte Verteilung zurückführen, wie dies für das Likelihood-Verhältnis aus Gleichung 5.11 der Fall war.

Dieses Problem, das sich durch die voneinander verschiedenen Kovarianzmatrizen Σ_1 und Σ_2 ergibt, ist in der Statistik bekannt als das sogenannte *Behrens-Fisher Problem*, für das es im eindimensionalen Fall mehrere Lösungen gibt [Scheffé 70].

Für unseren Fall einer multivariaten Normalverteilung gibt es eine Verallgemeinerung einer Lösung für den eindimensionalen Fall [Anderson 58], die wir im folgenden kurz skizzieren wollen.

Zur Vereinfachung nehmen wir zunächst an, daß die Größe der beiden Proben \mathcal{X}_1 und \mathcal{X}_2 gleich ist, d.h.: $m = n$. Nun transformieren wir das Problem auf folgende Weise: Aus den Proben $\mathcal{X}_1 = \{\mathbf{x}_{11}, \dots, \mathbf{x}_{1n}\}$ und $\mathcal{X}_2 = \{\mathbf{x}_{21}, \dots, \mathbf{x}_{2n}\}$ bilden wir eine neue Probe $\mathcal{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_n\}$, wobei gilt: $\mathbf{y}_i = \mathbf{x}_{1i} - \mathbf{x}_{2i}$. Da \mathcal{X}_1 und \mathcal{X}_2 aus den Verteilungen $N(\mu_1, \Sigma_1)$ bzw.

$N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ stammen, bedeutet dies, daß \mathcal{Y} aus der Verteilung $N(\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)$ stammt. Vorausgesetzt wird hierbei, daß die Numerierung der einzelnen Beobachtungen der Proben \mathcal{X}_1 und \mathcal{X}_2 unabhängig von den Beobachtungen selbst ist!

Nach dieser Transformation können wir nun die folgende Hypothese testen:

$$H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0}.$$

$$H_1: \boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 \neq \mathbf{0}.$$

Ein Test hierfür wird wie gehabt mit Hilfe des Likelihood-Verhältniskriteriums konstruiert. Die Herleitung hierfür entspricht der Herleitung in Abschnitt 5.2.1. Sie ist etwas einfacher, und wir wollen sie hier nicht wiederholen.

Analog zu Gleichung 5.15 ist das Likelihood-Verhältnis dann eine monotone Funktion von

$$T^2 = n\hat{\boldsymbol{\mu}}^T \mathbf{S}^{-1} \hat{\boldsymbol{\mu}} \quad (5.18)$$

wobei

$$\begin{aligned} \hat{\boldsymbol{\mu}} &= \frac{1}{n} \sum_{i=1}^n \mathbf{y}_i \\ &= \hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2 \\ \mathbf{S} &= \frac{1}{n-1} \sum_{i=1}^n (\mathbf{y}_i - \hat{\boldsymbol{\mu}})(\mathbf{y}_i - \hat{\boldsymbol{\mu}})^T \\ &= \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_{1i} - \mathbf{x}_{2i} - \hat{\boldsymbol{\mu}}_1 + \hat{\boldsymbol{\mu}}_2)(\mathbf{x}_{1i} - \mathbf{x}_{2i} - \hat{\boldsymbol{\mu}}_1 + \hat{\boldsymbol{\mu}}_2)^T \end{aligned}$$

T^2 hat wieder die T^2 -Verteilung von Hotelling mit $n-1$ Freiheitsgraden [Anderson 58]. Somit erhalten wir dann den folgenden Test zum Konfidenzniveau α :

Die Nullhypothese wird zurückgewiesen, wenn gilt

$$T^2 > \frac{(n-1)d}{n-d} \mathcal{F}_{d, n-d}(\alpha). \quad (5.19)$$

Bei der Berechnung von T^2 können wir wieder die Invertierung der Matrix \mathbf{S} vermeiden, wenn wir Gleichung 5.18 mit Hilfe der Rechenregeln für partitionierte Matrizen aus den Gleichungen 5.13 und 5.14 umformen (siehe

Abschnitt 5.2.1). Wir erhalten dann:

$$T^2 = \frac{|\mathbf{S} + n\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^T|}{|\mathbf{S}|} - 1.$$

Hiermit ergibt sich

Zerlegungskriterium 7

Wir übernehmen Zerlegungskriterium 6, verwenden aber Formel 5.19 für den Test.

Wir wählen nun folgende Parameter für die Segmentation des synthetischen Bildes: $n = 15$, $\theta_g = 0$ und das Konfidenzniveau $\alpha = 0.001$. Der Tabellenwert hierfür ist dann $\mathcal{F}_{d,n-d}(\alpha) = \mathcal{F}_{3,12}(0.001) = 10.80$ [Abramowitz + Stegun 68, S. 989]. Bild 5 zeigt die Segmentation in 189 Bereiche. Das Ergebnis ist jedoch nicht besser als vorher: Das Bild wurde immer noch in viele Bereiche zerlegt.

Die Lösung des Behrens-Fisher Problems, dessen Verallgemeinerung für den mehrdimensionalen Fall wir hier benutzt haben, stammt von Henry Scheffé. Dieser schreibt hierüber in einem Übersichtsartikel [Scheffé 70, S. 1503], 26 Jahre nachdem er diese Lösung veröffentlicht hat:³

These articles were written before I had much consulting experience, and since then I have never recommended the solution in practice. The reason is that the estimate s_d requires putting in random order the elements of the larger sample, and the value of s_d . . . depend very much on the result of this randomization of the data. The effect of this in practice would be deplorable.⁴

Der Grund für die schlechten Ergebnisse liegt also in der Annahme, die Durchnummerierung unserer Beobachtungen sei unabhängig von den Beobachtungen selbst.

Es sollte möglich sein, eine der anderen Lösungen aus Scheffés Übersichtsartikel, die diese Annahme nicht machen, auf multivariate Normalverteilungen zu verallgemeinern. Der mathematische Aufwand hierfür dürfte jedoch recht hoch sein und so wollen wir diesen Ansatz hier nicht weiter verfolgen.

³Leider fanden wir dies Bemerkung erst, nachdem wir uns die Mühe gemacht hatten, diesen Ansatz zu versuchen

⁴ s_d ist eine Schätzung für die Standardabweichung und entspricht unserem $\hat{\Sigma}$

5.4 Approximation der Bildfunktion durch Polynome höherer Ordnung

Bisher haben wir angenommen, das Bild ließe sich durch Flecken stückweise konstanter Farbe beschreiben. Ist diese Annahme nicht gerechtfertigt, so beschreiben die Kovarianzmatrizen nicht nur das Rauschen, sondern sie müssen auch diejenigen Bildstrukturen repräsentieren, die sich nur durch Terme höherer Ordnung beschreiben lassen. Es liegt also nahe, das Bild stückweise durch bivariate Polynome höherer Ordnung zu approximieren, so daß die Kovarianzmatrizen wieder nur das Rauschen beschreiben. Wir folgen mit dieser Argumentation wieder *Hsu et al. 82*, die für das eindimensionale Problem von Grauwertbildern die Bildfunktion durch bivariate Polynome 1. und 2. Ordnung approximieren und dann einen Test herleiten, den sie wiederum auf eine bekannte Verteilung zurückführen können.

Zunächst wollen wir dies allgemein formulieren und dabei von der Ordnung des Polynoms abstrahieren. Da wir für jedes Pixel unseres Bildes drei Intensitätswerte haben, müssen wir hier eine Vektorfunktion approximieren. Wir wollen zunächst auch von der Dimension dieser Funktion abstrahieren.

Nehmen wir an, eine Beobachtung \mathbf{x}_i aus einer Probe \mathcal{X} lasse sich durch das folgende sogenannte *Allgemeine Lineare Modell*⁵ beschreiben:

$$\mathbf{x}_i = \beta_1 z_{1i} + \beta_2 z_{2i} + \dots + \beta_q z_{qi} + \varepsilon_i \quad (5.20)$$

mit

$$\begin{aligned} \mathbf{x}_i &= (x_{1i}, \dots, x_{di})^T \\ \beta_g &= (\beta_{1g}, \dots, \beta_{dg})^T \\ \varepsilon_i &= (\varepsilon_{1i}, \dots, \varepsilon_{di})^T \end{aligned}$$

und d als der Dimension.

Hierbei sind die β_g die unbekanntenen Polynomparameter, die z_{gi} bekannte Variable und ε_i die voneinander unabhängigen Fehlerterme mit der Verteilung $N(\mathbf{0}, \Sigma)$, die das Rauschen ausdrücken.

In Matrixschreibweise läßt sich dies einfacher wie folgt formulieren:

$$\mathbf{x}_i = \mathbf{B}\mathbf{z}_i + \varepsilon_i$$

⁵linear bedeutet hier: linear in den Parametern β_g !

mit

$$\begin{aligned}\mathbf{B} &= (\beta_1, \dots, \beta_q) \\ \mathbf{z}_i &= (z_{1i}, \dots, z_{qi})^T\end{aligned}$$

oder für die gesamte Probe \mathcal{X} , wobei wir die Probe hier einfach als Matrix \mathbf{X} betrachten:

$$\mathbf{X} = \mathbf{B}\mathbf{Z} + \mathbf{E}$$

mit

$$\begin{aligned}\mathbf{X} &= (\mathbf{x}_1, \dots, \mathbf{x}_n) \\ \mathbf{Z} &= (\mathbf{z}_1, \dots, \mathbf{z}_n) \\ \mathbf{E} &= (\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n)\end{aligned}$$

\mathbf{Z} wird hierbei auch *Designmatrix* genannt. Wie wir später noch genauer sehen werden, enthält \mathbf{Z} die Bildkoordinaten, aus denen die einzelnen Beobachtungen \mathbf{x}_i der Probe \mathbf{X} entnommen werden. \mathbf{Z} bestimmt somit die Fläche des Bildes, die unsere Probe \mathbf{X} bildet. Durch geeignete Wahl von \mathbf{Z} und die entsprechende Dimensionierung der Parametermatrix \mathbf{B} wird außerdem der Grad des approximierten Polynoms bestimmt.

Unsere Situation stellt sich somit wie folgt dar: Zwei Flächen unseres Bildes, die durch die Designmatrizen \mathbf{Z}_1 und \mathbf{Z}_2 bestimmt werden, bilden die beiden Proben $\mathbf{X}_1 = (\mathbf{x}_{11}, \dots, \mathbf{x}_{1m})$ sowie $\mathbf{X}_2 = (\mathbf{x}_{21}, \dots, \mathbf{x}_{2m})$. Beide Flächen zusammengenommen werden durch die Designmatrix $\mathbf{Z}_0 = (\mathbf{Z}_1 \mathbf{Z}_2)$ bestimmt und es gelte $\mathbf{X}_0 = (\mathbf{X}_1 \mathbf{X}_2) = (\mathbf{x}_{11}, \dots, \mathbf{x}_{1m}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2n})$. Wir wollen nun testen, ob sich die beiden Proben \mathbf{X}_1 und \mathbf{X}_2 besser durch *zwei* Polynome approximieren lassen als die Gesamtprobe \mathbf{X}_0 durch *ein* Polynom. Dies läßt sich durch die folgende Hypothese ausdrücken.

H_0 : Eine Beobachtung \mathbf{x}_{0l} ($l = 1 \dots m + n$) der Probe \mathbf{X}_0 kommt aus der Verteilung $N(\mathbf{B}_0 \mathbf{z}_{0l}, \boldsymbol{\Sigma}_0)$.

H_1 : Die Beobachtungen \mathbf{x}_{1i} ($i = 1 \dots m$) der Probe \mathbf{X}_1 und \mathbf{x}_{2j} ($j = 1 \dots n$) der Probe \mathbf{X}_2 kommen aus den Verteilungen $N(\mathbf{B}_1 \mathbf{z}_{1i}, \boldsymbol{\Sigma}_1)$ bzw. $N(\mathbf{B}_2 \mathbf{z}_{2j}, \boldsymbol{\Sigma}_2)$.

Für diese Hypothese konstruieren wir wie gehabt einen Likelihood-Verhältnistest:

$$\lambda = \frac{\sup L(\mathbf{B}_1 \mathbf{Z}_1, \mathbf{B}_2 \mathbf{Z}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2 | H_1)}{\sup L(\mathbf{B}_0 \mathbf{Z}_0, \boldsymbol{\Sigma}_0 | H_0)}.$$

Die Likelihood-Funktionen sind dann

$$\begin{aligned}
 L(\mathbf{B}_0 \mathbf{Z}_0, \boldsymbol{\Sigma}_0 | H_0) &= \left(\frac{1}{(2\pi)^d \det(\boldsymbol{\Sigma}_0)} \right)^{\frac{m+n}{2}} \\
 &\cdot \exp \left[-\frac{1}{2} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \mathbf{B}_0 \mathbf{z}_{1i})^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{x}_{1i} - \mathbf{B}_0 \mathbf{z}_{1i}) \right. \right. \\
 &\quad \left. \left. + \sum_{j=1}^n (\mathbf{x}_{2j} - \mathbf{B}_0 \mathbf{z}_{2j})^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{x}_{2j} - \mathbf{B}_0 \mathbf{z}_{2j}) \right) \right] \\
 L(\mathbf{B}_1 \mathbf{Z}_1, \mathbf{B}_2 \mathbf{Z}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2 | H_1) &= \left(\frac{1}{(2\pi)^d} \right)^{\frac{m+n}{2}} \left(\frac{1}{\det(\boldsymbol{\Sigma}_1)} \right)^{\frac{m}{2}} \left(\frac{1}{\det(\boldsymbol{\Sigma}_2)} \right)^{\frac{n}{2}} \\
 &\cdot \exp \left[-\frac{1}{2} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \mathbf{B}_1 \mathbf{z}_{1i})^T \boldsymbol{\Sigma}_1^{-1} (\mathbf{x}_{1i} - \mathbf{B}_1 \mathbf{z}_{1i}) \right. \right. \\
 &\quad \left. \left. + \sum_{j=1}^n (\mathbf{x}_{2j} - \mathbf{B}_2 \mathbf{z}_{2j})^T \boldsymbol{\Sigma}_2^{-1} (\mathbf{x}_{2j} - \mathbf{B}_2 \mathbf{z}_{2j}) \right) \right]
 \end{aligned}$$

Hieraus ergeben sich für die Kovarianzmatrizen folgenden Maximumlikelihood-Schätzungen (vergleiche Anhang B, Gleichung B.2):

$$\begin{aligned}
 \hat{\boldsymbol{\Sigma}}_0 &= \frac{1}{m+n} \left(\sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\mathbf{B}}_0 \mathbf{z}_{1i})(\mathbf{x}_{1i} - \hat{\mathbf{B}}_0 \mathbf{z}_{1i})^T \right. \\
 &\quad \left. + \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\mathbf{B}}_0 \mathbf{z}_{2j})(\mathbf{x}_{2j} - \hat{\mathbf{B}}_0 \mathbf{z}_{2j})^T \right) \quad (5.21)
 \end{aligned}$$

$$\hat{\boldsymbol{\Sigma}}_1 = \frac{1}{m} \sum_{i=1}^m (\mathbf{x}_{1i} - \hat{\mathbf{B}}_1 \mathbf{z}_{1i})(\mathbf{x}_{1i} - \hat{\mathbf{B}}_1 \mathbf{z}_{1i})^T \quad (5.22)$$

$$\hat{\boldsymbol{\Sigma}}_2 = \frac{1}{n} \sum_{j=1}^n (\mathbf{x}_{2j} - \hat{\mathbf{B}}_2 \mathbf{z}_{2j})(\mathbf{x}_{2j} - \hat{\mathbf{B}}_2 \mathbf{z}_{2j})^T \quad (5.23)$$

Für die Parametermatrizen erhalten wir nach einer etwas längeren Herleitung die folgenden Maximumlikelihood-Schätzungen (Anhang C, Gleichung C.1):

$$\hat{\mathbf{B}}_0 = (\mathbf{X}_1 \mathbf{Z}_1^T + \mathbf{X}_2 \mathbf{Z}_2^T) (\mathbf{Z}_1 \mathbf{Z}_1^T + \mathbf{Z}_2 \mathbf{Z}_2^T)^{-1} \quad (5.24)$$

$$\hat{\mathbf{B}}_1 = \mathbf{X}_1 \mathbf{Z}_1^T (\mathbf{Z}_1 \mathbf{Z}_1^T)^{-1} \quad (5.25)$$

$$\hat{\mathbf{B}}_2 = \mathbf{X}_2 \mathbf{Z}_2^T (\mathbf{Z}_2 \mathbf{Z}_2^T)^{-1} \quad (5.26)$$

Nach Einsetzen dieser Schätzungen in die Likelihood-Funktionen können wir wie gehabt deren Exponenten vereinfachen und erhalten schließlich folgendes Likelihood-Verhältnis:

$$\lambda = \left(\frac{\det(\hat{\Sigma}_0)^{m+n}}{\det(\hat{\Sigma}_1)^m \det(\hat{\Sigma}_2)^n} \right)^{\frac{1}{2}} \quad (5.27)$$

Wegen des Gruppierungseffektes des MST können wir auch hier nicht die Annahme machen, daß die Kovarianzmatrizen Σ_1 und Σ_2 gleich sind. Dies führt dazu, daß wir $\hat{\mathbf{B}}_0$, und damit auch $\hat{\Sigma}_0$ nicht in Abhängigkeit von $\hat{\mathbf{B}}_1$ und $\hat{\mathbf{B}}_2$ ausdrücken können. Weiterhin haben die Bildflächen, aus denen wir die Proben \mathbf{X}_1 und \mathbf{X}_2 nehmen, beliebige Form, da sie wie gehabt durch unseren MST bestimmt werden sollen. Wir können daher nicht wie *Hsu et al.* 82 vereinfachende Annahmen über die Designmatrizen machen.

All dies führt dazu, daß sich das Likelihood-Verhältnis aus Gleichung 5.27 nicht weiter vereinfachen und auf eine bekannte und tabellierte Verteilung zurückführen läßt. Wir können somit keinen Test zu einem vorgegebenen Konfidenzniveau angeben und benutzen daher das Verhältnis aus Gleichung 5.27 direkt für den Test:

Die Nullhypothese wird zurückgewiesen, wenn gilt

$$\lambda > \theta.$$

Bei der Herleitung des Tests in diesem Abschnitt haben wir vollkommen von der Ordnung des zu approximierenden Polynoms, von der Dimension der Verteilung und auch von der speziellen Anwendung des Test abstrahiert. Wir wollen den Test nun für Polynome 1. und 2. Ordnung verwenden. Dabei werden wir dann sehen, wie die Parametermatrizen \mathbf{B} und insbesondere die Designmatrizen \mathbf{Z} für unsere Anwendung aussehen. Zuvor wollen wir jedoch zeigen, daß dieser Test die Verallgemeinerung des Tests von Yakimovsky als Spezialfall — nämlich für Polynome 0. Ordnung — enthält.

5.4.1 Approximation durch Polynome 0. Ordnung

Die Approximation durch ein bivariates Polynom 0. Ordnung bedeutet die Anpassung einer nicht geneigten Ebene.⁶ Dies entspricht jedoch genau unserer Annahme einer stückweise konstanten Bildfunktion, die wir bei der

⁶im Falle unserer Farbbilder einer Hyperebene im 5D-Merkmalraum

Verallgemeinerung des Tests von Yakimovsky in Abschnitt 5.1 gemacht haben. Haben wir bei der Herleitung unseres Tests im vorigen Abschnitt keinen Fehler gemacht, so sollten wir für Polynome 0. Ordnung genau den Test aus Abschnitt 5.1.1 erhalten, und haben somit zumindest eine notwendige Bedingung für die Korrektheit unserer Herleitung.⁷

Eine Beobachtung \mathbf{x}_i aus der Probe \mathbf{X} läßt sich nach Gleichung 5.20 dann beschreiben durch

$$\begin{aligned}\mathbf{x}_i &= \boldsymbol{\beta} z_i + \boldsymbol{\varepsilon}_i \quad \text{mit } z_i = 1 \\ &= \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i.\end{aligned}$$

Unsere Parametermatrix \mathbf{B} ist damit eine $d \times 1$ -Matrix, besteht also nur aus dem Vektor $\boldsymbol{\beta}$. Die Designmatrix \mathbf{Z} ist eine $1 \times n$ -Matrix bestehend aus lauter Einsen. Damit erhalten wir für die Schätzungen der Parametermatrizen nach den Gleichungen 5.24, 5.25 und 5.26

$$\begin{aligned}\hat{\mathbf{B}}_0 &= \frac{1}{m+n} \left(\sum_{i=1}^m \mathbf{x}_{1i} + \sum_{j=1}^n \mathbf{x}_{2j} \right), \\ \hat{\mathbf{B}}_1 &= \frac{1}{m} \sum_{i=1}^m \mathbf{x}_{1i}, \\ \hat{\mathbf{B}}_2 &= \frac{1}{n} \sum_{j=1}^n \mathbf{x}_{2j}.\end{aligned}$$

Dies sind genau die Schätzungen der Mittelwertvektoren $\hat{\boldsymbol{\mu}}_0$, $\hat{\boldsymbol{\mu}}_1$ sowie $\hat{\boldsymbol{\mu}}_2$ aus den Gleichungen 5.3–5.5. Setzen wir diese Mittelwertvektoren in die Gleichungen 5.21–5.23 ein, so erhalten wir die Kovarianzmatrizen aus den Gleichungen 5.6–5.8. Das Likelihood-Verhältnis aus Gleichung 5.27 ist damit dann gleich dem Likelihood-Verhältnis aus Gleichung 5.9.

Wie erwartet, erhalten wir für Polynome 0. Ordnung also den Test aus Abschnitt 5.1.1.

5.4.2 Approximation durch Polynome 1. Ordnung

Die Approximation durch ein bivariates Polynom 1. Ordnung in den Bildkoordinaten (r_i, c_i) bedeutet die Anpassung einer geeigneten Ebene.⁸

⁷Auch unsere Implementation können wir hiermit prüfen: Wir müssen die gleichen Ergebnisse erhalten wie beim Yakimovsky-Kriterium.

⁸genauer: Hyperebene im 5D-Merkmalraum

Der Farbvektor $\mathbf{x}_i = (R_i, G_i, B_i)^T$ eines Pixels aus einer bestimmten Bildfläche läßt sich hier nach Gleichung 5.20 beschreiben durch

$$\mathbf{x}_i = \beta_1 + \beta_2 r_i + \beta_3 c_i + \epsilon_i.$$

Ein \mathbf{z}_i ist also gegeben durch $\mathbf{z}_i = (1, r_i, c_i)^T$. Die Designmatrix \mathbf{Z} ist somit eine $3 \times n$ -Matrix mit

$$\mathbf{Z} = \begin{pmatrix} 1 & \cdots & 1 \\ r_1 & \cdots & r_n \\ c_1 & \cdots & c_n \end{pmatrix},$$

und die Parametermatrix \mathbf{B} ist eine 3×3 -Matrix mit

$$\mathbf{B} = (\beta_1 \beta_2 \beta_3) = \begin{pmatrix} \beta_{R1} & \beta_{R2} & \beta_{R3} \\ \beta_{G1} & \beta_{G2} & \beta_{G3} \\ \beta_{B1} & \beta_{B2} & \beta_{B3} \end{pmatrix}.$$

Wir kommen somit zu folgendem

Zerlegungskriterium 8

Die Analogie zu Abschnitt 5.1 haben wir schon erwähnt. Wir verwenden daher die Zerlegungskriterien 4 und 5, berechnen die Verteilungsparameter aber nach den Gleichungen 5.21–5.23 sowie 5.24–5.26 und verwenden die hier angegebene Form der Design- und der Parametermatrix.

Mit den Parametern $n = 15$, $\theta = 50$ sowie $\theta_g = 0$ segmentieren wir unser synthetisches Bild dann in 67 Bereiche. Bild 6 zeigt, daß die konstanten Flächen in mehrere Bereiche zerlegt, gleichzeitig unterschiedliche Flächen aber miteinander verschmolzen wurden (z.B. \mathcal{O}_1 und \mathcal{O}_2).

Da wir in diesem Bild bis auf die Rampe \mathcal{O}_4 nur Flächen mit im Mittel konstanter Farbe haben, werden die linearen Terme des approximierten Polynoms jeweils dazu verwendet, das Rauschen zu beschreiben. In natürlichen Bildern werden wir solche Flächen kaum antreffen, und so soll uns dieses schlechte Segmentationsergebnis hier nicht weiter stören.

Bild 24 zeigt die Segmentation des Blockbildes in 59 Bereiche ($n = 15$, $\theta = 150$ und $\theta_g = 30$). Im Gegensatz zu der Segmentation mit Hilfe des Yakimovsky-Kriteriums (Bild 22) können wir jetzt die rechte Seite von der Frontseite des Pyramidenstumpfes trennen. Auch der Schatten der Pyramide auf dem Papier des Hintergrundes konnte jetzt als ein zusammenhängender Bereich erkannt werden, da seine Farbe nicht konstant ist,

sondern nach oben hin heller wird. Dies läßt sich dann durch ein Polynom 1. Ordnung besser beschreiben.

Die Segmentation des Hausbildes in 392 Bereiche ($n = 15$, $\theta = 300$ und $\theta_g = 20$), die wir in Bild 16 sehen, ist jedoch eher schlechter als mit dem Yakimovsky-Kriterium (Bild 15). Das Verfahren reagiert jetzt wesentlich sensibler auf kleinste Variationen in der Farbe. Hierdurch wird das Bild in mehr Bereiche zerlegt. Gleichzeitig sind jedoch andere Bereiche verschmolzen worden: der Balkon ist mit seinem Schatten verschmolzen und ein Erker des rechten Hauses ist mit dem Dach verschmolzen worden. Ein Erhöhen des Schwellwertes θ verringert zwar die Anzahl der Bereiche, dabei verschmelzen aber weitere interessante Bereiche.

5.4.3 Approximation durch Polynome 2. Ordnung

Approximieren wir die Bildfunktion stückweise durch ein bivariates Polynom 2. Ordnung in den Bildkoordinaten (r_i, c_i) , so läßt sich der Farbvektor $\mathbf{x}_i = (R_i, G_i, B_i)^T$ eines Pixels der zu approximierenden Bildfläche nach Gleichung 5.20 beschreiben durch

$$\mathbf{x}_i = \beta_1 + \beta_2 r_i + \beta_3 c_i + \beta_4 r_i c_i + \beta_5 r_i^2 + \beta_6 c_i^2.$$

Ein \mathbf{z}_i ist damit gegeben durch $\mathbf{z}_i = (1, r_i, c_i, r_i c_i, r_i^2, c_i^2)^T$. Die Designmatrix \mathbf{Z} ist dann eine $6 \times n$ -Matrix mit

$$\mathbf{Z} = \begin{pmatrix} 1 & \cdots & 1 \\ r_1 & \cdots & r_n \\ c_1 & \cdots & c_n \\ r_1 c_1 & \cdots & r_n c_n \\ r_1^2 & \cdots & r_n^2 \\ c_1^2 & \cdots & c_n^2 \end{pmatrix}$$

und die Parametermatrix \mathbf{B} ist eine 3×6 -Matrix mit

$$\mathbf{B} = (\beta_1 \beta_2 \beta_3 \beta_4 \beta_5 \beta_6) = \begin{pmatrix} \beta_{R1} & \beta_{R2} & \beta_{R3} & \beta_{R4} & \beta_{R5} & \beta_{R6} \\ \beta_{G1} & \beta_{G2} & \beta_{G3} & \beta_{G4} & \beta_{G5} & \beta_{G6} \\ \beta_{B1} & \beta_{B2} & \beta_{B3} & \beta_{B4} & \beta_{B5} & \beta_{B6} \end{pmatrix}.$$

Hiermit erhalten wir nun unser letztes

Zerlegungskriterium 9

Wir modifizieren Zerlegungskriterium 8, indem wir die hier dargestellte Designmatrix \mathbf{Z} und Parametermatrix \mathbf{B} verwenden.

Bild 25 zeigt die Segmentation des Blockbildes in 174 Bereiche ($n = 15$, $\theta_g = 30$ und $\theta = 150$). Bild 17 zeigt die Segmentation des Hausbildes in 369 Bereiche ($n = 15$, $\theta_g = 20$ und $\theta = 1500$).

Die Ergebnisse sind nun deutlich schlechter als bei der Verwendung des Yakimovsky-Kriteriums. Das Verfahren paßt sich noch sensibler lokalen Änderungen der Farbe an. Dies wird besonders deutlich bei unserem Blockbild: Hintergrund und Drehteller, die in ihrer Farbe relativ gering variieren, sind zu einem großen Bereich verschmolzen worden; die Holzmaserung des Tisches führt jedoch dazu, daß dieser in sehr viele recht kleine Bereiche zerlegt wurde. Ist der Übergang zwischen zwei Bereichen sehr 'sanft', so kann dieser auch schon durch ein Polynom 2. Ordnung beschrieben werden. Dies ist genau der Grund, warum die rechte Seite des Blockes mit dem Papier des Drehtisches verschmolzen ist.

Bei unserem Hausbild können wir den gleichen Effekt beobachten. Einerseits reagiert das Verfahren sensibler; beispielsweise wurden die beiden kleinen Giebel im Dach des rechten Hauses in kleinere Bereiche zerlegt als mit dem Yakimovsky-Kriterium. Andererseits werden 'sanfte' Übergänge verschmolzen; hier ist ein Teil des Daches des rechten Hauses mit dem Himmel verschmolzen worden. Wir werden hierauf in Abschnitt 6.1 des folgenden Kapitels noch zurückkommen.

5.5 Zusammenfassung

In Kapitel 4 haben wir versucht, den MST in Teilbäume und damit das Bild in Bereiche zu zerlegen, indem wir die Länge der einzelnen MST-Kanten für unser Zerlegungskriterium verwendet haben.

Eine globale Schwelle für diese Kantenlänge führte zu schlechten Segmentationsergebnissen, besonders bei unterschiedlichen Kontrasten im Bild. Wir konnten jedoch zeigen, daß sich dieses Verfahren nicht von einem primitiven Verschmelzungsverfahren unterscheidet. Die schlechten Ergebnisse sind somit nicht weiter verwunderlich.

Die Verwendung einer lokalen Schwelle, die mit Hilfe unseres MST aus der Verteilung der Kantenlängen in einer lokalen Umgebung einer zu betrachtenden Kante ermittelt werden kann, führte zu wesentlich besseren Ergebnissen. Die Untersuchung der Segmentationsergebnisse für unser synthetisches Bild zeigte jedoch, daß sich Schwierigkeiten ergeben, wenn die Länge einer MST-Kante, die eigentlich entfernt werden sollte, durch Rauschen stark verändert (verkürzt) wird. Da die Entscheidung, ob eine Kante

entfernt werden soll, ganz wesentlich von ihrer Länge abhängt, ist eine korrekte Entscheidung dann nicht mehr möglich.

In diesem Kapitel haben wir dann statistische Testverfahren hergeleitet und für die Entscheidung zum Entfernen einer MST-Kante verwendet. Dabei haben wir nicht mehr die Kantenlängen, sondern die Farbvektoren der einzelnen Pixel betrachtet.

Als erstes haben wir ein Verfahren von Yakimovsky 76 für die Verwendung auf unseren Farbbildern verallgemeinert. Hierbei wird angenommen, daß sich ein Bild stückweise durch Flächen konstanter Farbe beschreiben läßt. Es wird dann getestet, ob sich zwei benachbarte Flächen des Bildes als eine zusammenhängende konstante Fläche beschreiben lassen. Diese Flächen brauchen wir jedoch nicht wie Yakimovsky durch Masken fester Form und Größe explizit vorzugeben, sondern können sie durch unseren MST definieren als eine vorgegebene Anzahl von Pixeln, die, ausgehend von einer zu betrachtenden Kante, in einem Teilbaum unseres MST zusammenhängen. Hierdurch können wir uns besser an die lokalen Eigenschaften des Bildes anpassen. Liefert uns der Test die Antwort, daß zwei so definierte Flächen nicht als eine gemeinsame Fläche konstanter Farbe beschrieben werden können, so wird die sie verbindende Kante des MST entfernt. Die Segmentationsergebnisse zeigen, daß wir mit diesem Verfahren die angesprochenen Probleme durch Rauschen vermeiden können.

Bei diesem Verfahren fließt ein Schwellwert in die Entscheidung mit ein, ob der Test positiv oder negativ ausfällt. Motiviert durch die Arbeiten von Hsu *et al.* 82 haben wir dann versucht, das Likelihood-Verhältnis, mit dessen Hilfe wir uns den Test konstruiert haben, auf eine bekannte und tabellierte Verteilung zurückzuführen, so daß wir keinen Schwellwert mehr explizit vorgeben müssen. Es wird vielmehr ein Konfidenzniveau angegeben, das ausdrückt, wie sicher wir sein wollen, daß die durch den Test getroffene Entscheidung richtig ist. Es zeigte sich jedoch, daß die beiden hergeleiteten Testverfahren keine guten Segmentationsergebnisse liefern, da die bei ihrer Herleitung gemachten Annahmen bei uns in der Regel nicht erfüllt sind.

Wieder motiviert durch die Arbeiten von Hsu *et al.* 82 haben wir dann einen Test hergeleitet, bei dem angenommen wird, ein Bild ließe sich stückweise durch ein bivariates Polynom n . Ordnung in den Bildkoordinaten beschreiben. Hierbei wird dann analog zu dem Test von Yakimovsky getestet, ob sich zwei benachbarte Flächen des Bildes besser durch ein oder durch zwei Polynome beschreiben lassen. Für ein Polynom 0. Ordnung geht dieser Test in den Test von Yakimovsky über. Verwendet haben wir dieses Verfahren für Polynome 1. und 2. Ordnung. Es zeigte sich jedoch, daß dieses

Verfahren mit steigender Ordnung des Polynoms immer sensibler auf lokale Farbvariationen des Bildes reagiert. Die Bilder werden hierdurch in zu viele Bereiche segmentiert. Außerdem können sanfte Übergänge zwischen zwei Bereichen durch Polynome höherer Ordnung besser beschrieben werden, so daß hierdurch einige Bereiche miteinander verschmolzen werden.

Kapitel 6

Versuch einer Bewertung

Wir wollen in diesem Kapitel nun den Versuch unternehmen, unser Segmentationsverfahren zu bewerten. In Abschnitt 1.1 haben wir schon gesehen, daß sich dabei Schwierigkeiten ergeben werden, da wir aufgrund einer mangelnden Definition des Begriffs *Bereich* kein objektives Gütekriterium angeben können. Auch uns bleibt hier kaum etwas anderes übrig, als ikonische Darstellungen der Segmentationsergebnisse mit den Originalbildern zu vergleichen, und wir hoffen dabei, daß der Leser beim Betrachten der Ergebnisse in Anhang A unserer subjektiven Beurteilung zustimmen wird.

6.1 Vergleich der einzelnen Zerlegungsverfahren

In den beiden vorangegangenen Kapiteln haben wir mehrere Verfahren vorgestellt, mit denen wir unseren minimal spannenden Baum zerlegen können, um so zu einer Segmentation des Bildes zu gelangen. Welches dieser Verfahren ist nun am besten für diese Aufgabe geeignet?

Die Zerlegung des MST mit Hilfe einer globalen Schwelle über die Kantenlängen, wie wir sie in Abschnitt 4.1 vorgeschlagen haben, lieferte keine guten Ergebnisse. Unser kleiner Beweis auf Seite 32 hat uns auch gezeigt, daß die Berechnung des MST hierbei gar keinen Vorteil bringt, sondern vielmehr überflüssig ist; die gleichen Ergebnisse können auch mit einem einfachen Schwellwertverfahren erreicht werden. Diese Zerlegungsverfahren scheidet daher von vornherein aus.

Im Anschluß daran haben wir die Verwendung einer lokalen Schwelle vorgeschlagen, die jeweils aus der Verteilung der Kantenlängen in einer 'Umgebung' einer zu betrachtenden Kante im MST ermittelt wird. Mit diesem

Verfahren haben wir für alle Bilder recht gute Ergebnisse erzielt. Wir wollen die entgeltige Bewertung dieses Verfahrens jedoch etwas aufschieben und uns zunächst den statistischen Testverfahren aus Kapitel 5 zuwenden.

Diesem Verfahren liegt jeweils ein bestimmtes *Modell der Bildfunktion* zugrunde: Es wird angenommen, die Bildfunktion ließe sich stückweise durch ein *bivariates Polynom n . Ordnung in den Bildkoordinaten* beschreiben. Wir haben dies mit drei verschiedenen Ordnungen probiert: Die Verallgemeinerung des Verfahrens von Yakimovsky nimmt an, die Bildfunktion ließe sich stückweise durch Flecken konstanter Farbe beschreiben (dies entspricht einem Polynom 0. Ordnung), in Abschnitt 5.4 haben wir dann versucht, die Bildfunktion linear oder quadratisch zu approximieren.

Entspricht ein Bild dem angenommenen Modell, so können mit dem entsprechenden Verfahren sicherlich optimale Ergebnisse erzielt werden. Dies konnten wir z.B. bei unserem synthetischen Bild beobachten. Die einzelnen Flächen wurden jeweils mit einer konstanten Farbe eingefärbt und dann mit Rauschen überlagert.¹ Dieses Bild entspricht also genau der Modellvorstellung des Yakimovsky-Verfahrens, mit dem wir hier ein (fast) perfektes Ergebnis erzielen konnten (vergleiche Bild 1 und 3).

Entspricht ein Bild dem zugrunde liegenden Modell jedoch nicht, so führt dies zu Segmentationsfehlern. Wir können hier die beiden folgenden Fälle unterscheiden:

- Das Bild variiert mit höherer Ordnung als angenommen. Dies hat eine *Übersegmentierung* zur Folge. Eine Fläche wird in kleinere Bereiche zerlegt, die sich durch Polynome der angenommenen Ordnung beschreiben lassen.
- Das Bild variiert mit niedrigerer Ordnung als angenommen. Innerhalb einer Fläche, die sich schon durch ein Polynom niedrigerer Ordnung beschreiben ließe, hat dies ebenfalls eine *Übersegmentierung* zur Folge, da die zusätzlichen Polynomterme dann dazu verwendet werden, das Rauschen zu beschreiben. Dies wird in zwei kleineren Flächen jedoch besser gelingen als in einer größeren, so daß in der Segmentierung dann zwei Bereiche entstehen.

Die Verwendung von Polynomen zu hoher Ordnung kann jedoch auch genau den entgegengesetzten Effekt haben. Stellen wir uns hierzu

¹Genau genommen hat die Fläche \mathcal{O}_4 eine linear variierende Farbe. Die Steigung der Fläche ist verglichen mit dem Rauschen jedoch so gering, daß man hier lokal von Konstanz sprechen kann.

zwei Bildflächen vor, die in ihrer Farbe jeweils linear variieren. Wenn diese Flächen an ihrem Übergang die gleiche Farbe haben, so kann dieser Übergang eventuell mit einem Polynom 2. Ordnung beschrieben werden, wenn das Rauschen stark genug ist und die Steigungen der beiden Flächen nicht allzu sehr differieren. Hierdurch würden die beiden Flächen dann zu einem Bereich *verschmolzen*.

Haben wir also Bilder zu segmentieren, die tatsächlich *einem* Modell entsprechen, und kennen wir dieses Modell zudem auch noch, so können wir das dazu passende Testverfahren verwenden und werden gute Segmentationsergebnisse erhalten.

Bei unseren Bildern von natürlichen Szenen kennen wir das Modell jedoch nicht. Es gibt zwar statistische Testverfahren, mit denen wir die optimale Ordnung eines zu approximierenden Polynoms für eine bestimmte Bildfläche bestimmen können [Hille 82], dies wird uns jedoch nicht weiterhelfen, da bei Bildern natürlicher Szenen ein Modell nicht ausreichen wird, um alle Bildstrukturen optimal beschreiben zu können.

Welches Verfahren sollen wir nun zur Segmentation verwenden? Um hier zu einer Entscheidung zu kommen, kann es helfen, ein wenig Wissen über die Bildentstehung zu verwenden. Nehmen wir einmal an, die Oberflächen der aufgenommenen Szene sind matte Oberflächen und diese Oberflächen sind nicht gekrümmt. Wenn wir unsere Szene im Freien bei direktem Sonnenlicht aufgenommen haben, so werden die Oberflächen jeweils konstante Helligkeit haben, da bei einer perfekten matten Oberfläche — dem Lambert'schen Strahler — die Helligkeit nur vom Einfallswinkel des Lichtes abhängt. Diesen können wir bei der großen Entfernung der Sonne verglichen mit der Ausdehnung unserer Objektoberflächen jedoch als konstant voraussetzen. Haben wir unsere Szene dagegen bei künstlichem Licht, z.B. durch einen Scheinwerfer, aufgenommen, so ist der Einfallswinkel auch bei kleinen Objektoberflächen nicht mehr konstant, und somit wird auch die Helligkeit nicht mehr konstant sein. Dies zeigt sich z.B. deutlich bei unserem Blockbild. Hier konnten wir mit unserem Testverfahren, das Polynome 1. Ordnung approximiert, recht gute Ergebnisse erzielen (siehe Bild 24).

Solche Überlegungen können jedoch nur in beschränktem Umfang eine Hilfe leisten, da hierzu doch recht starke Einschränkungen gemacht werden müssen, um überhaupt zu Ergebnissen zu kommen (eine nicht gekrümmte Lambert'sche Oberfläche werden wir sicherlich nicht allzu häufig vorfinden).

Ein Vergleich unserer Segmentationsergebnisse in Anhang A zeigt, daß wir mit der Verallgemeinerung des Verfahrens von Yakimovsky bei allen Bil-

dern recht gute Ergebnisse erhalten. Können wir kein bestimmtes Modell der Bildfunktion zugrunde legen, so erscheint uns dieses Verfahren denjenigen Verfahren, die Polynome höherer Ordnung an die Bildfunktion approximieren, überlegen zu sein.

Kommen wir nun auf das Verfahren aus Abschnitt 4.2 zurück, bei dem wir den MST mit Hilfe einer lokal berechneten Schwelle zerlegen. Hierbei ist es äußerst schwierig, anzugeben, welches Modell der Bildfunktion diesem Verfahren zugrunde liegt. Da bei diesem Verfahren die Kantenlängen des MST miteinander verglichen werden, wird eine Bildfläche, in der die Kantenlängen alle gleich sind oder innerhalb der vorgegebenen Varianz liegen, in einen zusammenhängenden Bereich segmentiert werden. Wie ein Bild aussehen darf, daß diese zulässige Varianz der Kantenlängen voll ausnutzt, aber noch perfekt segmentiert wird, kann man sich jedoch kaum vorstellen.

Die mit diesem Verfahren erzielten Ergebnisse sind durchaus mit den Ergebnissen des Yakimovsky-Verfahrens vergleichbar. Beide zeigen an unterschiedlichen Stellen der Bildern jeweils bessere Ergebnisse als das andere Verfahren. Vergleichen wird hierzu die Segmentationen unseres Hausbildes (Bild 14 und 15). Das Verfahren mit einer lokalen Schwelle über die Kantenlängen konnte im Gegensatz zum Yakimovsky-Verfahren die Strukturen des Fensters unterhalb des Balkons von dessen Schatten trennen und der weiße Giebel des rechten Hauses wurde nicht in viele kleine Bereiche zerlegt. Mit Hilfe des Yakimovsky-Verfahrens wurden jedoch die beiden mittleren Fenster in der oberen Reihe des rechten Hauses nicht mit der Hauswand verschmolzen. Außerdem konnte die Frontseite des linken Hauses sowohl vom Dach als auch von der Frontseite des rechten Hauses getrennt werden. Die Ergebnisse der beiden Verfahren unterscheiden sich in ihrer Qualität jedoch nicht grundlegend voneinander.

Aufgrund der Schwierigkeiten des lokalen Verfahrens mit starkem Rauschen, die wir in Abschnitt 4.2.4 schon ausführlich diskutiert haben, halten wir die Verallgemeinerung des Verfahrens von Yakimovsky im allgemeinen jedoch für besser geeignet. Wir werden uns in den folgenden Abschnitten daher nur noch auf dieses Verfahren beziehen. Die Diskussion in diesem Abschnitt hat jedoch gezeigt, daß mit anderen Verfahren durchaus bessere Resultate erzielt werden können, wenn sich die Bildfunktion durch ein anderes Modell besser beschreiben läßt, als das hier mit der Annahme stückweise konstanter Farbe geschieht.

6.2 Robustheit

Eine Anforderung an ein Segmentationsverfahren ist, daß es sich gutmütig gegenüber der Veränderung von Parametern des Verfahrens verhält. Bei kleiner Änderung der Parameter soll sich das Ergebnis nicht drastisch verändern. Außerdem ist es wünschenswert, daß ein fester Satz von Parametern für alle zu segmentierenden Bilder gute Ergebnisse liefert, so daß diese Parameter nicht bei jedem Bild wieder interaktiv ermittelt werden müssen.

Bei der Entscheidung zum Entfernen einer MST-Kante müssen wir ausgehend von dieser Kante in jedem der beiden Teilbäume, die an dieser Kante hängen, eine vorgegebene Anzahl von Pixeln betrachten, aus denen dann die jeweiligen Verteilungsparameter berechnet werden (siehe Abschnitt 5.1.1). Bei unseren Experimenten hat sich gezeigt, daß eine Zahl von $n = 15$ Pixeln je Teilbaum für alle Bilder gute Resultate liefert. Bei einer kleineren Anzahl haben die berechneten Parameter keine statistische Aussagekraft mehr, bei einer größeren Anzahl erhöht sich die Gefahr, daß wir mit zunehmender Entfernung von der zu betrachtenden Kante in ein Gebiet des Bildes 'hineinlaufen', das nicht mehr zum selben Bereich gehört. Hierdurch werden dann aber die Verteilungsparameter verfälscht, oder genauer, die einzelnen Komponenten der Kovarianzmatrizen werden größer.

In Abschnitt 5.2 hatten wir versucht, das statistische Testverfahren, das wir zur Zerlegung unseres MST verwenden, so zu modifizieren, daß wir den Schwellwert θ nicht mehr explizit vorgeben müssen. Vielmehr sollte nur noch ein Konfidenzniveau vorgegeben werden, das ausdrückt, wie sicher wir sein wollen, daß die durch den Test getroffene Entscheidung richtig ist. Dies war jedoch nicht möglich, da bestimmte Voraussetzungen, die wir hierzu hätten machen müssen, bei uns nicht erfüllt waren.

Es hat sich dann auch gezeigt, daß die Segmentationsergebnisse wesentlich von diesem Schwellwert θ abhängen. Es konnte kein Wert für diesen Parameter gefunden werden, der für alle Bilder 'optimale' Ergebnisse liefert (siehe Bild 3, 15 und 22). Allerdings liegen diese Werte doch in einem relativ eng begrenzten Bereich von 120 bis 160.

Wir wollen nun einmal untersuchen, wie stark sich die Segmentationsergebnisse bei Variation des Schwellwertes θ verändern. Hierzu haben wir in Abbildung 6.1 die Anzahl der Bereiche, in die unser Blockbild segmentiert wurde, in Abhängigkeit vom verwendeten Schwellwert aufgetragen.

Es zeigt sich hierbei, daß diese Kurve gerade in dem Bereich des Schwellwertes von $\theta = 130$, der für unser Blockbild die besten Ergebnisse lieferte (Bild 22), einen sehr steilen Verlauf hat. Dies bedeutet jedoch, daß schon

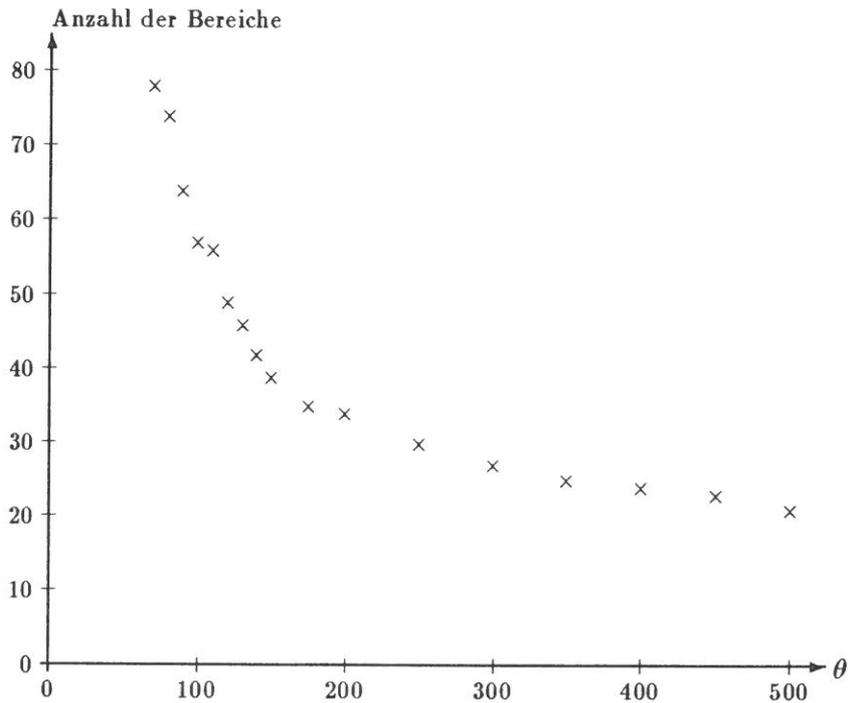


Abbildung 6.1: Anzahl der Bereiche in Abhängigkeit von der Schwelle θ für Bild 18 mit Zerlegungskriterium 5 ($n = 15$, $\theta_g = 30$).

geringe Variationen des Schwellwertes große Veränderungen der Segmentationsergebnisse verursachen. Diese äußerst unerwünschte Eigenschaft konnte auch für unser Hausbild bestätigt werden.

6.3 Reproduzierbarkeit

Aufgabe unseres Segmentationsverfahrens ist es, neben einem Punktefinder und einem Kantenfinder Merkmale für die Korrespondenzanalyse aus den Bildern zu extrahieren (siehe Kapitel 1). Eine Anforderung, die an die extrahierten Merkmale gestellt wird, um für die Korrespondenzanalyse brauchbar zu sein, ist, daß sie *reproduzierbar* sind. Dies bedeutet, daß Merkmale,

die in einem Bild gefunden werden konnten, im darauffolgenden Bild einer Bildfolge oder dem zweiten Bild eines Stereobildpaares ebenfalls gefunden werden.

Im Falle markanter Punkte wird dies relativ einfach zu messen sein. Markante Punkte, die irrtümlich aufgrund von Rauschen detektiert wurden, werden im nächsten Bild nicht reproduziert werden. Dies ist auch wünschenswert, da sie bei der Korrespondenzanalyse nicht von Interesse sind. Entspricht ein detektierter markanter Punkt einem Objektpunkt, so kann leicht festgestellt werden, ob er im Folgebild ebenfalls detektiert wird. Dies ist deshalb so einfach, weil sich ein Objektpunkt als ein nulldimensionales Gebilde von Bild zu Bild nicht verändert (vorausgesetzt natürlich es handelt sich um einem körperfesten Punkt der im Folgebild überhaupt noch sichtbar und nicht verdeckt ist). Man kann also einfach nachschauen: entweder er wurde gefunden oder nicht. Bei Bereichen ist der Begriff der Reproduzierbarkeit jedoch wesentlich schwieriger zu fassen, da sich die Oberflächen der Objekte von Bild zu Bild ändern, wenn sie z.B durch Bewegung aus einer anderen Ansicht zu sehen sind.

Betrachten wir hierzu einmal die einzelnen Segmentationen für eine Folge von Bildern, zu der auch das uns schon bekannte Blockbild gehört (Bild 28 und 29). Beispielsweise wurde die schwarze Samtfläche auf der rechten Seite in allen vier Bildern der Folge korrekt in einen Bereich segmentiert. In jedem Bild entspricht der Bereich jedoch einer anderen Objektoberfläche, da durch die Drehung des Tisches jeweils ein anderer Teil der Samtfläche sichtbar wird. Kann man dann noch von reproduzierten Bereichen sprechen?

Als zweites Beispiel soll uns der vordere Pyramidenstumpf dienen. Abgesehen davon, daß die linke Seite des Pyramidenstumpfes im letzten Bild nicht mehr sichtbar ist, bleiben hier im Gegensatz zu dem vorherigen Beispiel die gleichen Objektoberflächen sichtbar. Im ersten Bild wurde die linke Seite des Pyramidenstumpfes als ein Bereich segmentiert. Die Frontseite bildet zusammen mit der rechten Seite einen gemeinsamen Bereich, da diese beiden Oberflächen in ihrer Farbe kaum zu unterscheiden sind. Im dritten Bild wurde der Pyramidenstumpf dann mit dem Drehteller soweit gedreht, daß aufgrund der geänderten Beleuchtungsverhältnisse die Frontseite nun kaum mehr von der linken, sehr wohl aber von der rechten Seite zu unterscheiden ist. Hierdurch wurde dann die rechte Seite als ein Bereich segmentiert, die Frontseite nun aber mit der *linken* Seite zu einem Bereich verschmolzen. Auch hier fällt es schwer, von reproduzierten Bereichen zu reden.

In den beiden bisherigen Beispielen haben wir bewußt Bereiche verwendet, für die die Segmentierung sehr gut ist. Wir wollen nun noch ein Beispiel

wählen, wo wir — verursacht durch Rauschen — eine ‘falsche’ Segmentierung erhalten haben. Betrachten wir hierfür die rote Frontseite der hinteren Pyramide. In der zweiten und vierten Ansicht konnte diese Oberfläche korrekt in jeweils einen Bereich segmentiert werden, in der ersten und dritten Ansicht wurde die Frontseite jedoch in jeweils zwei Bereiche zerlegt. Hier ist ebenfalls nicht klar, was reproduziert werden konnte.

Der aufmerksame Leser wird bemerkt haben, daß wir hier zwei Begriffe miteinander vermischt haben: den der *Reproduzierbarkeit* und den der *Bedeutung* von Bereichen. Dies ist jedoch ganz bewußt geschehen. Unsere Beispiele sollten zeigen, daß der Begriff Reproduzierbarkeit sehr eng mit der Bedeutung der detektierten Merkmale verbunden ist und ohne sie gar keinen Sinn hat. Verfahren zur Korrespondenzanalyse fordern von den Merkmalen Reproduzierbarkeit, damit die Merkmale aus verschiedenen Bildern einander zugeordnet werden können. Dabei sollen aber nur solche Merkmale einander zugeordnet werden, die auch die gleiche Bedeutung haben. Bei markanten Punkten wird dies besonders deutlich: Es sollen nur markante Punkte einander zugeordnet werden, die den gleichen Objektpunkten entsprechen; nur dann ist z.B. bei Stereobildern die Rekonstruktion von Tiefeninformationen mittels Triangulation möglich (siehe Kapitel 1).

Der Begriff *Reproduzierbarkeit* hat hier also folgende Bedeutung: Wurde in einem Bild ein Merkmal detektiert, das eine bestimmte Bedeutung hat, so soll in einem Folgebild ebenfalls ein Merkmal detektiert werden, und zwar mit der *gleichen* Bedeutung.

Unsere Beispiele haben uns gezeigt, daß Bereiche diese Anforderung im allgemeinen nicht erfüllen. Zum einen natürlich durch Segmentationsfehler, zum anderen jedoch bedingt durch teilweise Verdeckungen bei unterschiedlichen Objektansichten entsprechen Bereiche eben oftmals *nicht* den gleichen Objektflächen. Die Verwendung von Bereichen als Merkmale für die Korrespondenzanalyse erweist sich somit als äußerst problematisch. Zumindest können Bereiche nicht ohne weiteres dazu verwendet werden, sie wie markante Punkte oder Kanten einander zuzuordnen. Es ist jedoch denkbar, daß sie für die Zuordnung anderer Merkmale, wie markanter Punkte oder Kanten, eine wertvolle Hilfe darstellen, indem sie z.B. den räumlichen Zusammenhang dieser Merkmale im Bild klären. Als Beispiel hierfür soll noch einmal die schwarze Samtfläche auf der rechten Seite dienen: es erscheint einfacher, die diese Fläche umrandenden Kanten einander zuzuordnen, wenn genau diese Information, daß sie eine schwarze Fläche umranden, ausgenutzt wird.

Als Fazit dieses Abschnittes sehen wir, daß die Forderung der Reproduzierbarkeit von Segmentationsergebnissen in einer Bildfolge nur selten uneingeschränkt erfüllbar ist, die Ergebnisse der Bereichsegmentation aber gleichwohl hilfreich bei der Korrespondenzanalyse sein können.

6.4 Vergleich mit Ergebnissen des Kantensfinders

Im Projekt SISSY werden die Ergebnisse der Bereichsegmentation, der wir uns in dieser Arbeit gewidmet haben, neben den Ergebnissen eines Kantendetektors verwendet [Drewniok]. Wir hatten in Kapitel 1 schon auf die Dualität der beiden Ansätze hingewiesen: Bei perfekten Ergebnissen beider Verfahren umranden die gefundenen Kanten jeweils die Bereiche, und Bereichsgrenzen entsprechen den Kanten. Es ist jedoch zu erwarten, daß die beiden Verfahren aufgrund ihrer unterschiedlichen Vorgehensweise nicht die gleichen Resultate liefern werden. In diesem Abschnitt wollen wir nun untersuchen, inwieweit die Ergebnisse übereinstimmen, und ob sich die beiden Verfahren gegenseitig ergänzen.

Bild 26 zeigt das Ergebnis des Kantendetektors für unser Blockbild, und in Bild 27 haben wir die gefundenen Kanten unserem Segmentationsergebnis aus Bild 22 überlagert. Bei diesen Ergebnissen wurden noch keine Farbinformationen ausgenutzt; der Kantendetektor arbeitet bisher auf Intensitätsbildern. Das verwendete Verfahren beruht auf [Canny 83].

Es zeigt sich, daß viele der gefundenen Kanten mit den Bereichsgrenzen übereinstimmen. Außerdem scheint in den Fällen, in denen beide Verfahren eine Diskontinuität im Bild detektieren konnten, die Lokalisation ebenfalls übereinzustimmen. Dies ist eine sehr wünschenswerte Eigenschaft der beiden Verfahren, da hierdurch den Verfahren der Korrespondenzanalyse erleichtert wird, festzustellen, ob eine Bereichsgrenze mit einer Kante übereinstimmt.

Wir wollen die Übereinstimmung nun quantitativ messen. Uns interessiert, zu wievielen Kantenelementen an der gleichen Stelle im Bild ein entsprechendes Element einer Bereichsgrenze existiert, und umgekehrt, zu wievielen Elementen der Bereichsgrenzen ein entsprechendes Kantenelement existiert. Hierbei ergibt sich jedoch folgendes Problem: Die Kantenelemente, die von dem Kantendetektor geliefert werden, liegen jeweils direkt auf den Pixeln des ursprünglichen Bildes, Bereichsgrenzen liegen jedoch zwischen den Pixeln des Bildes. Wir haben die Übereinstimmung deshalb folgendermaßen definiert: Ein Kantenelement hat ein entsprechendes Grenzele-

ment, wenn mindestens eines der vier Nachbarpixel (4-Nachbarschaft) zu einem anderen Bereich gehört als das Pixel des Kantenelementes. Dies bedeutet, daß zwischen diesen beiden Pixeln eine Bereichsgrenze verläuft. Ein Element einer Bereichsgrenze hat ein entsprechendes Kantenelement, wenn mindestens eines der beiden Pixel, zwischen denen dieses Grenzelement liegt, ein Kantenelement ist. Der Abstand zwischen Kantenelement und Grenzelement beträgt bei Übereinstimmung also genau ein halbes Pixel. Man beachte, daß diese Definition der Übereinstimmung etwas ungleichmäßig ist: Ein Kantenelement kann von vier Grenzelementen umgeben sein, ein Grenzelement hat jedoch nur zwei benachbarte Pixel, die ein Kantenelement sein könnten.

Für unser Blockbild ergibt sich dann folgende Übereinstimmung: Zu 70 Prozent aller Kantenelemente existiert ein entsprechendes Grenzelement, und zu 77 Prozent aller Grenzelemente existiert ein entsprechendes Kantenelement. Trotz des geringen zugelassenen Abstandes von einem halben Pixel ist die Übereinstimmung also recht hoch.

Wir wollen nun betrachten, inwieweit sich die Ergebnisse gegenseitig ergänzen können. In Bild 26 sehen wir, daß der Kantenfinder einige Informationen liefert, die in der Bereichssegmentation nicht enthalten sind. Beispielsweise konnte sowohl auf der linken wie auch auf der rechten Seite der roten Pyramide der hintere Rand des Drehtellers gefunden werden. In der Bereichssegmentation sind Hintergrund und Drehteller miteinander verschmolzen worden. Weiterhin konnte die Kante zwischen der Frontseite und der rechten Seite des Pyramidenstumpfes gefunden werden, und auch die Kante des Schattens, die der Pyramidenstumpf auf den Drehteller wirft, konnte vollständig detektiert werden. In der Bereichssegmentation konnte dagegen die rote Frontseite der hinteren Pyramide von deren grüner linker Seite getrennt werden, und bei dem Block wurde der Deckel von der rechten Seitenwand getrennt. Beides ist in dem Ergebnis des Kantenfinders durch kurze Kantenzüge angedeutet, so daß die Ergebnisse beider Verfahren sich hier unterstützen. Die Ergebnisse der Bereichssegmentation könnten auch dazu verwendet werden, bestimmte Kantenzüge des Kantendetektors zu schließen. Kantenfinder haben prinzipiell an solchen Stellen Schwierigkeiten, an denen mehrere Kanten aufeinander treffen. Hier entsteht dann in der Regel eine Lücke. Beispielsweise ist die Frontseite des Blocks auf der linken Seite nicht von einem geschlossenen Linienzug umrandet.

In Abschnitt 1.2 hatten wir erwähnt, daß im Projekt SISSY die Ergebnisse der verschiedenen Verfahren nicht zu einer gemeinsamen Beschreibung der Bilder verschmolzen werden sollen. Wir sehen hier beim Vergleich der

Ergebnisse der Bereichssegmentation und des Kantenfinders, daß dies auf dieser Ebene der Verarbeitung auch nur schwer möglich ist. Beide Verfahren liefern sowohl korrekte also auch falsche Ergebnisse. Wie soll hier aber entschieden werden, was richtig und was falsch ist? *Hanson + Riseman 78* verwenden in ihrem System VISIONS ebenfalls Kanten und Bereiche, verschmelzen die Ergebnisse aber zu einer Darstellung, indem sie nur einen sogenannten k -Durchschnitt verwenden. Das heißt, sie verwenden nur den Schnitt der beiden Ergebnismengen, wobei eine Kante und eine Bereichsgrenze jedoch als identisch betrachtet werden, wenn sie nicht weiter als k Pixel voneinander entfernt sind. Bei unserem Beispiel würden hierdurch jedoch wertvolle Informationen verlorengehen. Das Vorgehen in SISSY, die Ergebnisse beider Verfahren nebeneinander bestehen zu lassen und die Entscheidung auf eine höhere Ebene der Verarbeitung zu verlegen, ist für unsere Anwendung daher sicherlich der richtige Weg und entspricht dem "Prinzip der kleinsten Verpflichtung".

6.5 Nachteile des minimal spannenden Baumes

Nachdem wir bisher stets die Vorteile unseres minimal spannenden Baumes hervorgehoben haben, sollen einige Nachteile jedoch nicht unerwähnt bleiben.

Bei unserem Segmentationsverfahren berechnen wir zunächst den MST für ein Bild, und zerlegen ihn anschließend in Teilbäume und damit das Bild in Bereiche. Hierdurch können jedoch nur noch Flächen des Bildes zu Bereichen zusammengefaßt werden, die innerhalb unseres Baumes direkt über eine Kante zusammenhängen. Existiert zwischen diesen Flächen jedoch keine direkte Verbindung, so können sie auf keinen Fall mehr zu einem Bereich verschmolzen werden. Wir haben bereits einen Fall gesehen, in dem sich dies negativ bemerkbar macht. Bei dem MST für unser synthetisches Bild mit einem Rauschen von $\sigma = 12$, den wir in Abbildung 4.2 gezeigt haben, ist die dünne Linie unterhalb der Fläche \mathcal{O}_2 weder mit dem mittleren noch mit dem rechten Teil der Fläche \mathcal{O}_1 über eine Kante des MST verbunden. Hierdurch ist es dann nicht mehr möglich, daß die Fläche \mathcal{O}_1 als ein zusammenhängender Bereich detektiert wird.

Ein weiterer Nachteil unseres Baumes ist, daß zwei Flächen des Bildes nur durch *eine* Kante des MST miteinander verbunden sind. Wir haben dies in Abschnitt 4.2.4 schon erwähnt. Bei unseren statistischen Testverfahren verwenden wir zwar nicht mehr nur die Länge dieser Kante als einzige Infor-

mation, prinzipiell ändert sich an der Situation jedoch nichts; machen wir an dieser einen Stelle einen Fehler, so werden die beiden Flächen irrtümlich verschmolzen (oder getrennt). Da unser Baum die Eigenschaft hat, die Summe der Kantenlängen zu minimieren, werden zwei Flächen des Bildes zudem noch an genau der Stelle durch eine MST-Kante miteinander verbunden, an der sie sich am ähnlichsten sind. Dies erhöht die Gefahr, die beiden Flächen irrtümlich zu verschmelzen.

Außerdem scheint es sehr schwierig zu sein, Gestaltinformationen auszunutzen, die in der Struktur des MST enthalten sind. Wir möchten hierfür ein Beispiel geben. In der Segmentation unseres Blockbildes sehen wir, daß die untere Bereichsgrenze des Schattens, den die hintere Pyramide auf den Hintergrund wirft, genau dem Rand des Drehtellers entspricht, der ein Stück weiter links mit dem Hintergrund verschmolzen ist (Bild 22). Bei einem Blick auf den für dieses Bild berechneten MST stellt der menschliche Betrachter jedoch sofort fest, daß zwischen dem Block auf der linken Seite und der roten Pyramide eine horizontale Linie existiert, in der neben der einen MST-Kante, die den Drehteller mit dem Hintergrund verbindet, keine einzige weitere MST-Kante zu finden ist. Der menschliche Betrachter sieht hier in der Darstellung des MST quasi eine 'Kante', die dem Rand des Drehtellers entspricht. Diese Gestaltinformation, daß hier eine Linie ohne MST-Kanten existiert, können wir jedoch bisher nicht ausnutzen, um das Verschmelzen des Drehtellers mit dem Hintergrund zu vermeiden.

Abschließend möchten wir noch bemerken, daß es zur Beurteilung von Segmentationsverfahren sicherlich hilfreich wäre, wenn ein Satz von Referenzbildern zur Verfügung stünde, der dann einheitlich bei der Veröffentlichung von Segmentationsverfahren in der Literatur verwendet werden sollte. Da sich ein objektives Gütemaß zum Vergleich verschiedener Verfahren nicht angeben läßt, und die Implementation aller anderen bekannten Verfahren mit einem erheblichen Aufwand verbunden wäre, bestünde hierdurch die Möglichkeit, die Leistungsfähigkeit der einzelnen Verfahren auf den *gleichen* Bildern direkt mit Hilfe der Literatur zu vergleichen (siehe hierzu auch [Price 86]).

Kapitel 7

Zusammenfassung und Ausblick

Diese Arbeit beschäftigte sich mit Verfahren zur Segmentation von Farbbildern in Bereiche, wie sie im Projekt SISSY durchgeführt wird, um Merkmale für eine Korrespondenzanalyse zu erhalten.

Bei einem kurzen Überblick über die in der Literatur bekannten Verfahren stellten wir fest, daß die derzeit am häufigsten benutzten Segmentationsverfahren, die Bilder durch Klassifikation der einzelnen Pixel mit Hilfe einer Ballungsanalyse in einem Merkmalsraum segmentieren, Probleme mit verrauschten Bildern haben. Der Grund hierfür liegt darin, daß diesen Verfahren die nicht immer gerechtfertigte Annahme zugrunde liegt, daß eine Ballung im Merkmalsraum durch einen *zusammenhängenden* Bildbereich zustande kommt. Diesen Verfahren fehlt dabei die lokale Information, ob die einer Ballung entsprechenden Pixel im Bild tatsächlich zusammenhängen. Der Versuch, die dadurch entstehenden Segmentationsfehler durch Anwendung eines Relaxationsverfahrens nachträglich zu korrigieren, bringt neue Probleme mit sich, wenn in den Bildern feine, linienartige Strukturen vorhanden sind; diese werden durch ein Relaxationsverfahren dann entfernt.

In dieser Arbeit haben wir ein Segmentationsverfahren vorgeschlagen, das diejenigen Fehler, die die Anwendung eines Relaxationsverfahrens erst notwendig machen, von vornherein vermeidet, indem neben den drei Farbkomponenten eines Pixels auch noch seine Ortskoordinaten als Merkmale verwendet werden. In dem dadurch entstehenden fünfdimensionalen Farben-Orts-Merkmalsraum steht die Information über den räumlichen Zusammenhang der Pixel im Bild direkt bei der Ballungsanalyse zur Verfügung. Die

Annahme, daß eine Ballung im Merkmalsraum durch einen zusammenhängenden Bereich im Bild erzeugt wird, ist hierdurch dann gerechtfertigt.

In unserem Verfahren führen wir in diesem fünfdimensionalen Merkmalsraum eine hierarchische Ballungsanalyse mittels des "minimal spannenden Baumes" (MST) durch. Durch unsere Wahl der Abstandsfunktion in diesem Merkmalsraum erreichen wir, daß der Graph, für den wir unseren MST berechnen, ein *planarer* Graph ist. Jedes Pixel des Bildes entspricht dabei einem Knoten des Graphen und ist mit jedem seiner 4 Nachbarn über eine Kante des Graphen verbunden. Die Länge dieser Kante ist gleich dem Euklidischen Farbabstand der Pixel.

Da in dem MST nur benachbarte Pixel durch eine Kante verbunden sind, entspricht ein Teilbaum des MST einem zusammenhängenden Bildbereich. Entfernen wir nun die 'richtigen' Kanten des MST, so erhalten wir durch die entstehenden Teilbäume schon die Segmentation unseres Bildes in die gewünschten Bereiche.

Der Hauptschwerpunkt dieser Arbeit lag dann auf der Entwicklung eines geeigneten Kriteriums, mit dessen Hilfe entschieden werden kann, ob eine Kante des MST entfernt werden soll. Neben Versuchen, die Längen der einzelnen MST-Kanten für ein solches Kriterium zu nutzen, haben wir insbesondere statistische Testverfahren auf ihre Eignung hin untersucht. Alle untersuchten Testverfahren mußten für ihre Verwendung auf Farbbildern zunächst auf multivariate Normalverteilungen verallgemeinert werden.

Das allgemeinste dieser Verfahren approximiert die Farbwerte zweier beliebig geformter, benachbarter und durch eine Kante des MST verbundener Bildflächen durch bivariate Polynome einer vorgegebenen Ordnung, und testet dann, ob sich die Farbwerte der beiden Flächen besser durch *zwei* Polynome approximieren lassen als die Farbwerte der Gesamtfläche durch *ein* Polynom. Ist dies der Fall, so wird die MST-Kante entfernt.

Die beiden Bildflächen, die für den Test verwendet werden, brauchen nicht durch rechteckige Masken bestimmter Größe fest vorgegeben zu werden, sondern können bei unserem Verfahren für jede Kante individuell durch den MST definiert werden. Hierzu werden die beiden Teilbäume des MST, die an der gerade betrachteten Kante hängen, durchlaufen, bis eine vorgegebene Anzahl von Pixeln erreicht ist. Dies hat den großen Vorteil, daß sich das Verfahren an die lokalen Strukturen des Bildes besser anpassen kann, denn es werden dadurch nur Pixel betrachtet, die durch den MST in Teilbäumen zusammengruppiert werden und sich dadurch sehr ähnlich sind. Dies zeigte sich besonders deutlich bei den feinen linienartigen Strukturen in unserem synthetischen Bild.

Es besteht hierbei jedoch immer noch die Möglichkeit, daß wir über eine bestimmte Kante des MST in eine Bildfläche 'hineinlaufen', die zu einem anderen Bereich gehört. Hierdurch werden wir beim Entfernen einer MST-Kante dann eine falsche Entscheidung treffen.

Dieses Problem können wir eventuell dadurch vermindern, daß wir unser Verfahren mehrstufig auslegen: In einem ersten Durchlauf würden alle diejenigen Kanten beseitigt, für die die Entscheidung zum Entfernen sehr sicher ist; in einem zweiten Durchlauf würden dann auch Kanten entfernt, für die die Entscheidung weniger sicher ist. Hierbei könnte dann nicht mehr über die bereits entfernten Kanten, die sehr sicher einen Übergang zwischen zwei Bereichen bilden, in einen Nachbarbereich hineingelaufen werden.

Der Spezialfall der Approximation durch Polynome 0. Ordnung entspricht der Verallgemeinerung eines von Yakimovsky 76 zur Detektion von Diskontinuitätskanten vorgeschlagenen Testverfahrens. Hiermit konnten wir für unsere Bilder die besten Resultate erzielen.

Dieses Verfahren hängt recht sensibel von einem Schwellwert ab, der für jedes Bild interaktiv ermittelt werden muß. Motiviert durch die Arbeit von Hsu et al. 82 haben wir aus diesem Grunde einen Test hergeleitet, bei dem nur noch ein Konfidenzniveau vorgegeben werden muß. Die hierbei gemachte Annahme, die Kovarianzmatrizen der Farbwerte beider Flächen seien gleich (Behrens-Fisher Problem), war bei uns jedoch nicht erfüllt. Ein Testverfahren zur Lösung dieses Problems erwies sich wegen anderer prinzipiell ungerechtfertigter Annahmen als ungeeignet.

Es sollte jedoch möglich sein, ein anderes Verfahren zur Lösung des Behrens-Fisher Problems [Scheffé 70] auf multivariate Normalverteilungen und damit zur Verwendung auf Farbbildern zu verallgemeinern, wenn auch mit erheblichem mathematischem Aufwand.

Die von uns hergeleiteten Testverfahren sind neben unserer speziellen Anwendung auch zur Änderungsdetektion in Farbbildfolgen einsetzbar und stellen damit eine Verallgemeinerung der von Hsu et al. 82 für Schwarzweißbildfolgen vorgeschlagenen Verfahren dar.

Bei dieser Anwendung wird auch die Annahme gleicher Kovarianzmatrizen gerechtfertigt sein, so daß die investierte Arbeit zur Herleitung eines Test zu einem vorgegebenen Konfidenzniveau nicht umsonst gewesen ist. Da bei der Anwendung der Tests zur Änderungsdetektion quadratische Bildflächen verwendet werden, müßte es hierfür auch möglich sein, die Tests, die die Farbwerte durch Polynome höherer Ordnung approximieren, so weiterzuentwickeln, daß bei ihnen anstelle eines Schwellwertes ebenfalls nur noch ein Konfidenzniveau vorgegeben zu werden braucht (siehe auch [James 54]).

Die vordringlichste Aufgabe wird jedoch die Entwicklung des Verfahrens zur Korrespondenzanalyse sein, so daß wir feststellen können, inwieweit die Ergebnisse unseres Segmentationsverfahrens hierfür geeignet sind. Wir werden den Schwerpunkt unserer Arbeit in Zukunft daher zunächst auf dieses Gebiet verlagern.

Anhang A

Bilder

Unsere Bilder wurden aufgenommen mit einer RGB-Farbkamera der Firma BOSCH, die mit drei Plumbiconröhren ausgestattet ist. Nach der Analog-Digital-Wandlung des Kamerasignals erhalten wir ein Bild mit einer Auflösung von 512×576 Pixeln. Ein solches Bild besteht aus zwei Halbbildern: zwei aufeinanderfolgende Zeilen haben einen zeitlichen Abstand von 20ms. Da dies bei bewegten Objekten zu störenden Effekten führt, haben wir ein solches Bild in die beiden Halbbilder von jeweils 512×288 Pixeln aufgeteilt. Hiervon wurde jeweils nur das erste Halbbild weiter verwendet. Anschließend wurde noch eine Korrektur der Kamerageometrie vorgenommen, um auf das Seitenverhältnis von 4 : 3 der CCIR-Norm B zu kommen. Hierdurch erhalten wir dann ein sogenanntes GEO-Bild im Format von 256×192 Pixeln.

Aus drucktechnischen Gründen sind die Originalbilder auf den folgenden Seiten nicht als Farb- sondern als Grauwertbilder dargestellt. Unsere Segmentationsergebnisse werden ebenfalls als Grauwertbilder dargestellt. Hierbei sind die einzelnen Bereiche mit einem von 4, 5 bzw. 6 unterschiedlichen Grauwerten so eingefärbt worden, daß keine zwei benachbarten Bereiche den gleichen Grauwert haben.¹

¹Hierzu wurde ein Programm von Rainer Sprengel benutzt.

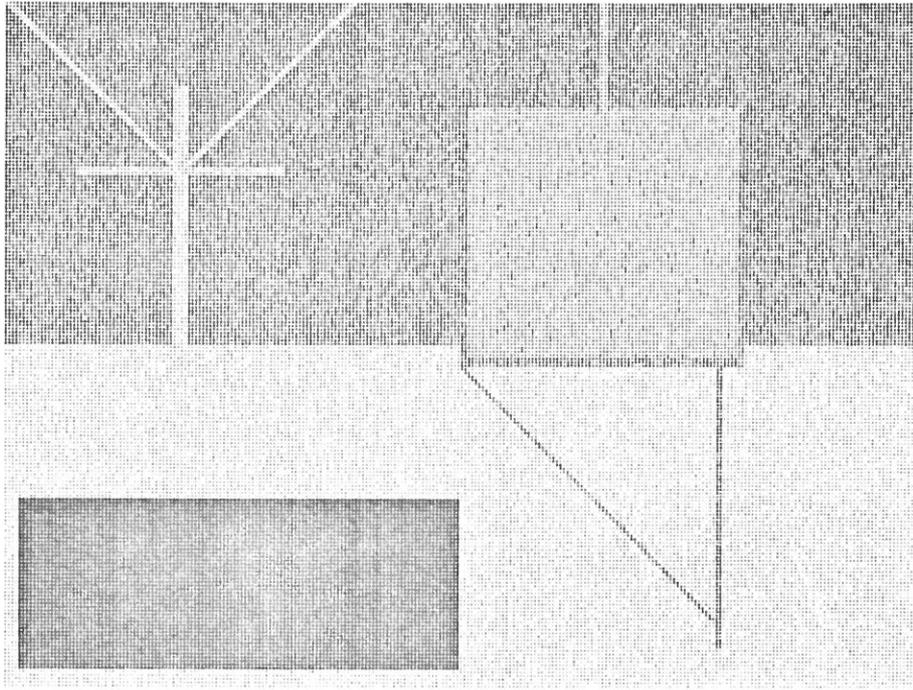


Bild 1: Synthetisches Bild mit normalverteiltem Rauschen ($\sigma = 12$). Vergleiche hierzu die Angaben auf Seite 36.



Bild 2: Segmentation von Bild 1 mit Zerlegungskriterium 3 (lokale Schwelle). 111 Bereiche, $\tau = 7$, $\theta = 1.5$, $\theta_g = 0$, $\theta_{min} = 0$.

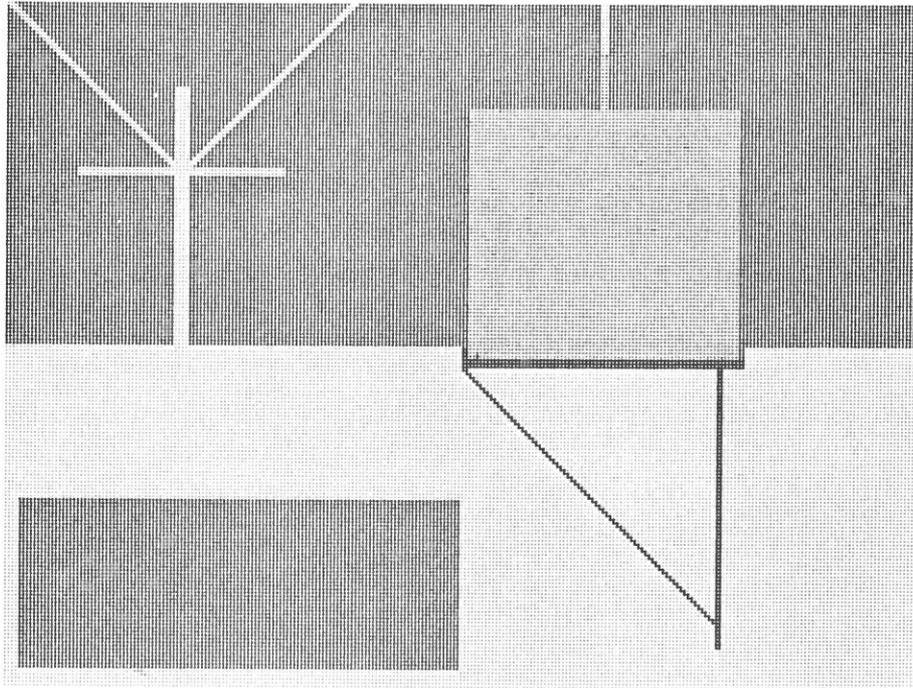


Bild 3: Segmentation von Bild 1 mit Zerlegungskriterium 5 (Yaki-movsky-Kriterium). 11 Bereiche, $n = 15$, $\theta = 120$, $\theta_g = 0$.

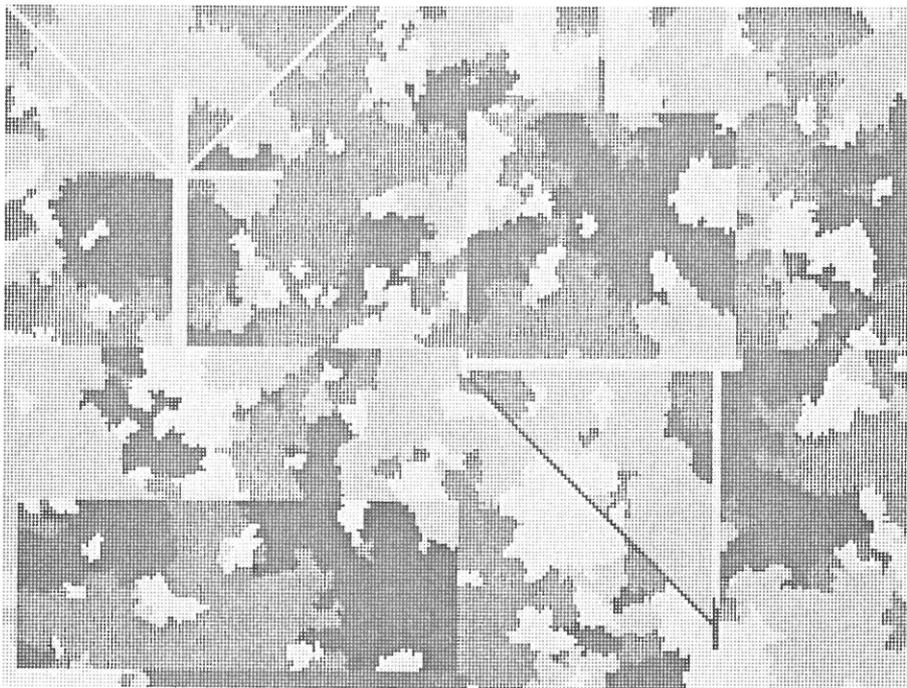


Bild 4: Segmentation von Bild 1 mit Zerlegungskriterium 6 (Hotellings T^2 -Test). 232 Bereiche, $n = 15$, $\alpha = 0.001$, $\theta_g = 0$.

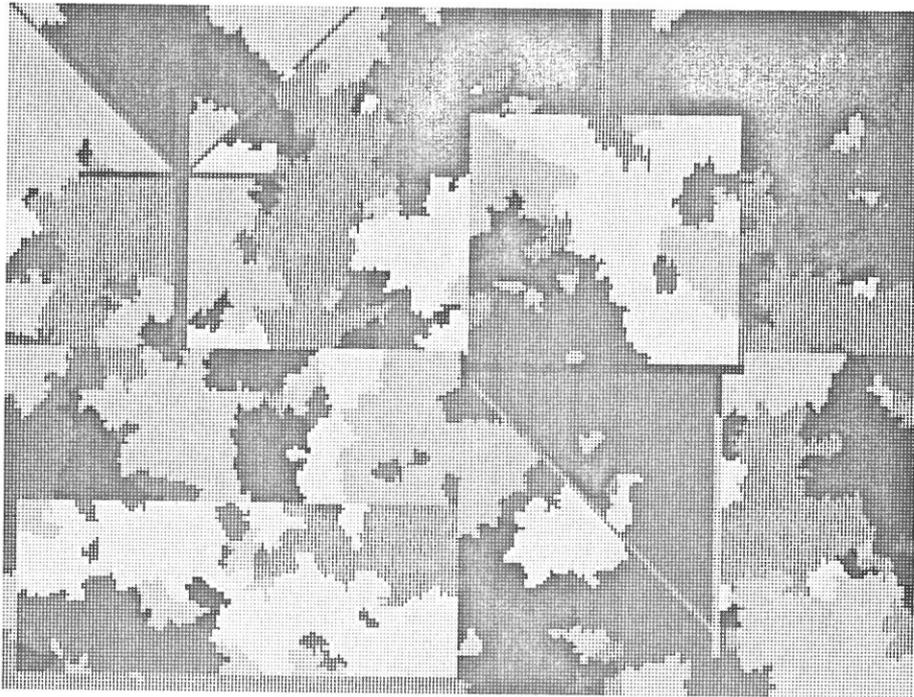


Bild 5: Segmentation von Bild 1 mit Zerlegungskriterium 7 (Behrens-Fisher Problem). 189 Bereiche, $n = 15$, $\alpha = 0.001$, $\theta_g = 0$.

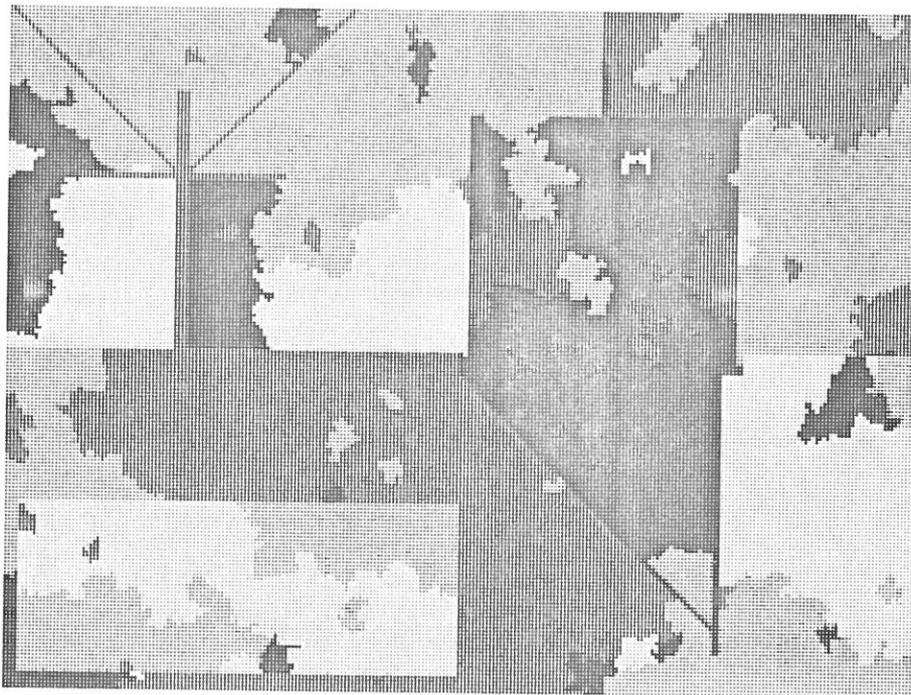


Bild 6: Segmentation von Bild 1 mit Zerlegungskriterium 8 (Polynom 1. Ordnung). 67 Bereiche, $n = 15$, $\theta = 50$, $\theta_g = 0$.

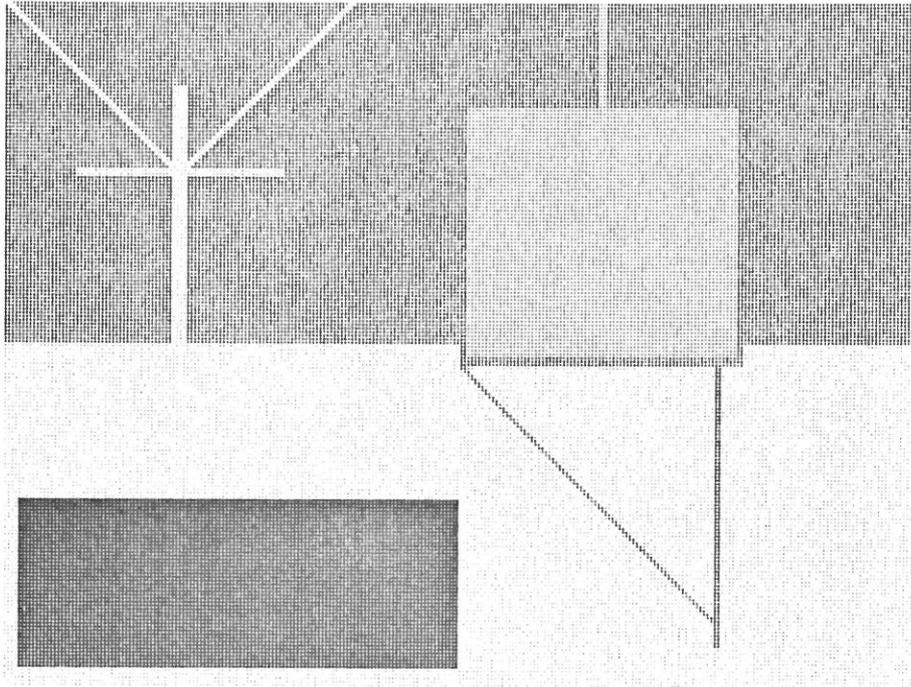


Bild 7: Synthetisches Bild mit normalverteiltem Rauschen ($\sigma = 6$).

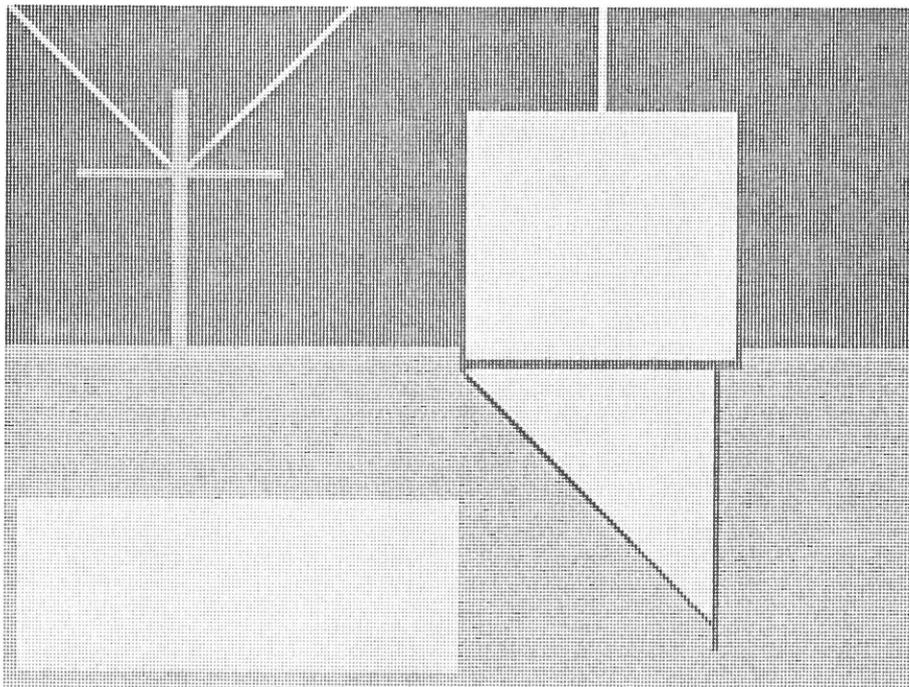


Bild 8: Segmentation von Bild 7 mit Zerlegungskriterium 3 (lokale Schwelle). 10 Bereiche, $\tau = 7$, $\theta = 3.0$, $\theta_g = 0$, $\theta_{min} = 0$.



Bild 9: Blick auf die Schlüterstraße.



Bild 10: Segmentation von Bild 9 mit Zerlegungskriterium 1 (globale Schwelle). 6562 Bereiche, $\theta_1 = 350$. Wegen der großen Anzahl wurden die Bereiche hier mit ihrem mittleren Grauwert eingefärbt.



Bild 11: Segmentation von Bild 9 mit Zerlegungskriterium 2 (lokale Schwelle). 637 Bereiche, $\tau = 7$, $\theta = 1.0$. Zur besseren Vergleichbarkeit wurden die Bereiche hier ebenfalls mit ihrem mittleren Grauwert eingefärbt.

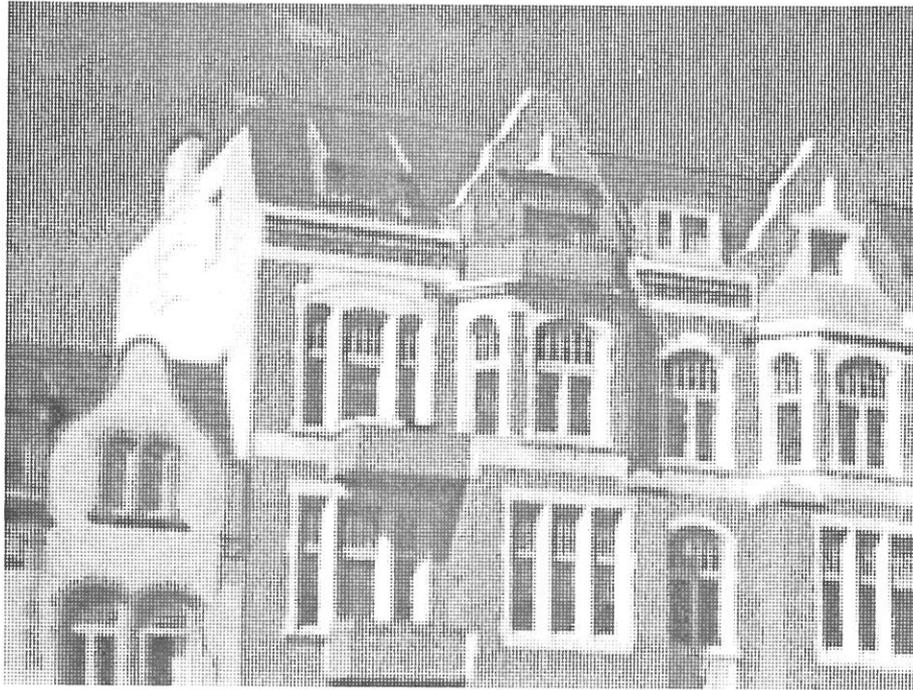


Bild 12: Häuserfassade.



Bild 13: Segmentation von Bild 12 mit Zerlegungskriterium 2 (lokale Schwelle). 650 Bereiche, $\tau = 7$, $\theta = 1.5$.



Bild 14: Segmentation von Bild 12 mit Zerlegungskriterium 2 (lokale Schwelle und Größenkriterium). 256 Bereiche, $\tau = 7$, $\theta = 1.5$, $\theta_g = 20$.



Bild 15: Segmentation von Bild 12 mit Zerlegungskriterium 5 (Yakhimovskiy-Kriterium). 329 Bereiche, $n = 15$, $\theta = 160$, $\theta_g = 20$.



Bild 16: Segmentation von Bild 12 mit Zerlegungskriterium 8 (Polynom 1. Ordnung). 392 Bereiche, $n = 15$, $\theta = 300$, $\theta_g = 20$.



Bild 17: Segmentation von Bild 12 mit Zerlegungskriterium 9 (Polynom 2. Ordnung). 369 Bereiche, $n = 15$, $\theta = 1500$, $\theta_g = 20$.

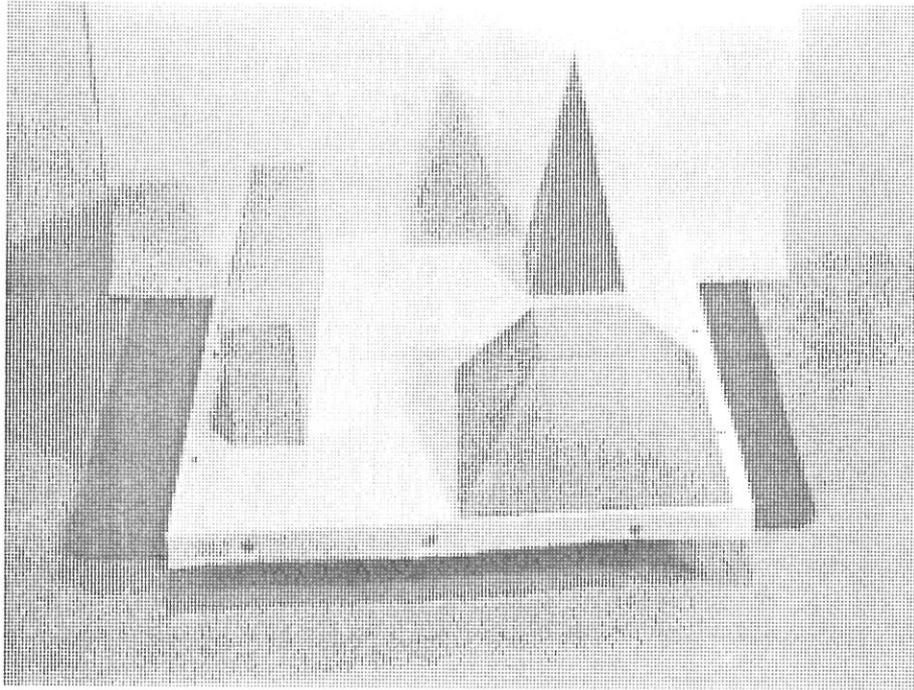


Bild 18: Blockbild.

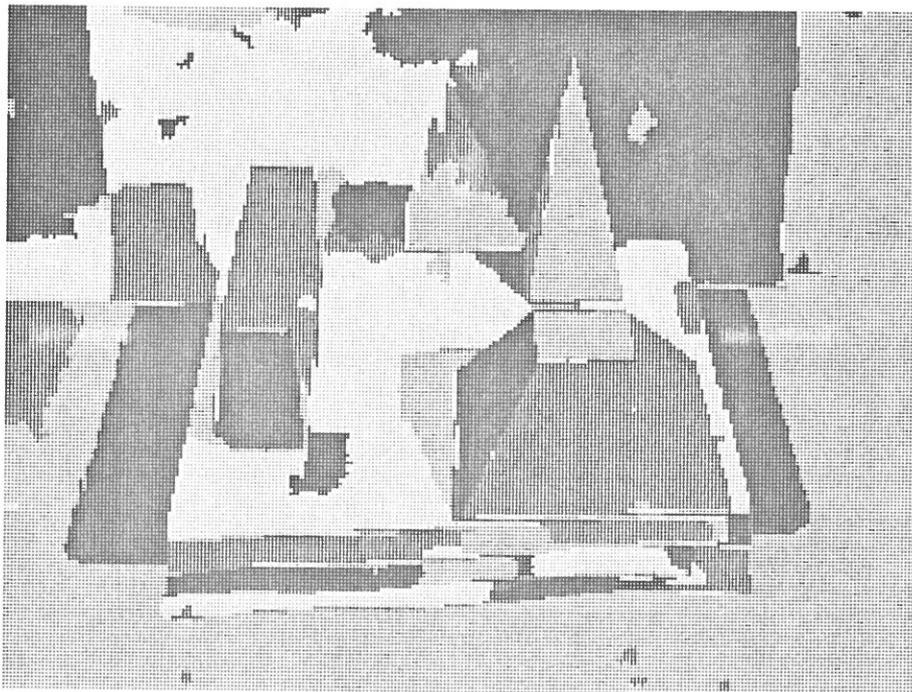


Bild 19: Segmentation von Bild 18 mit Zerlegungskriterium 2 (lokale Schwelle). 171 Bereiche, $\tau = 7$, $\theta = 1.5$.

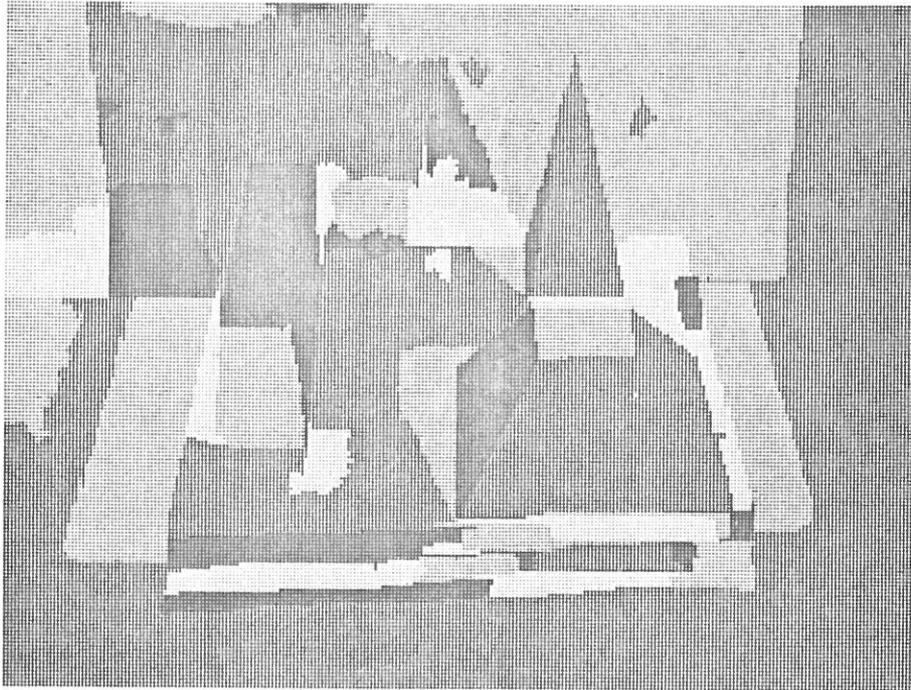


Bild 20: Segmentation von Bild 18 mit Zerlegungskriterium 2 (lokale Schwelle und Größenkriterium). 49 Bereiche, $\tau = 7$, $\theta = 1.5$, $\theta_g = 30$.

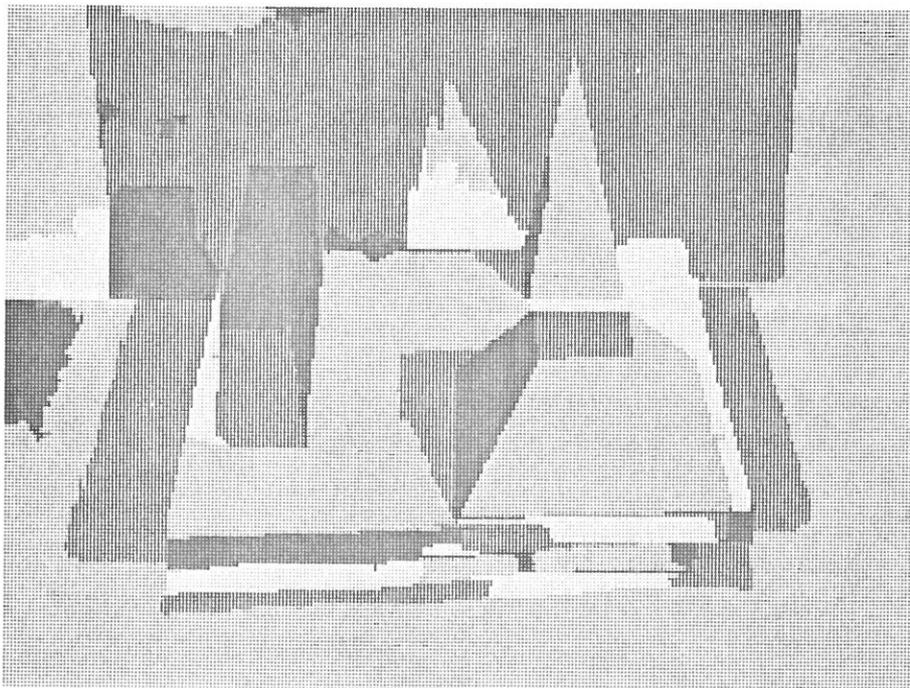


Bild 21: Segmentation von Bild 18 mit Zerlegungskriterium 3 (lokale Schwelle, Größenkriterium und Mindestkantenlänge). 41 Bereiche, $\tau = 7$, $\theta = 1.5$, $\theta_g = 30$, $\theta_{min} = 10$.

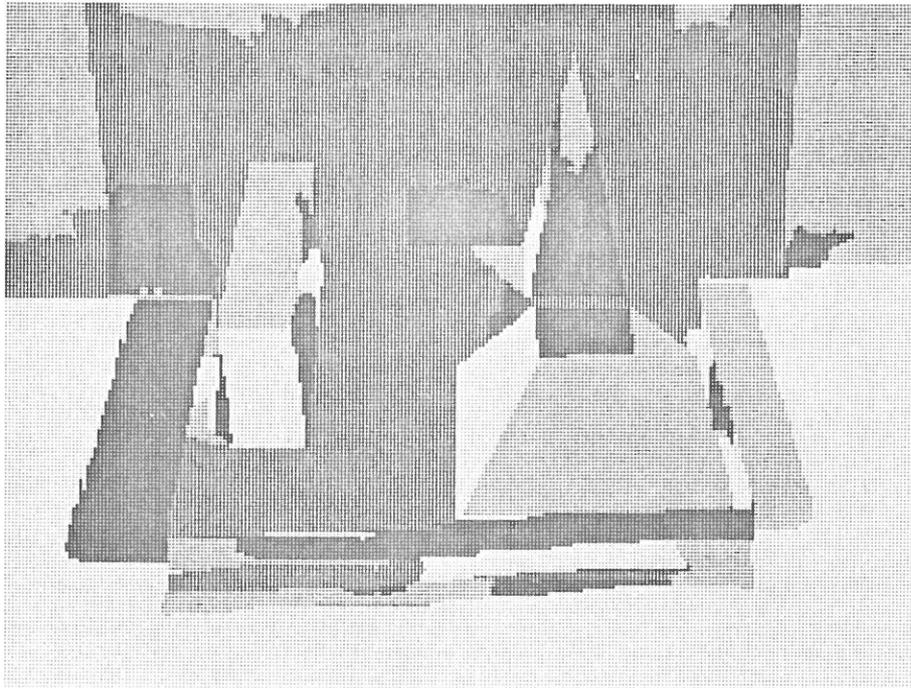


Bild 22: Segmentation von Bild 18 mit Zerlegungskriterium 5 (Yaki-movsky-Kriterium). 46 Bereiche, $n = 15$, $\theta = 130$, $\theta_y = 30$.

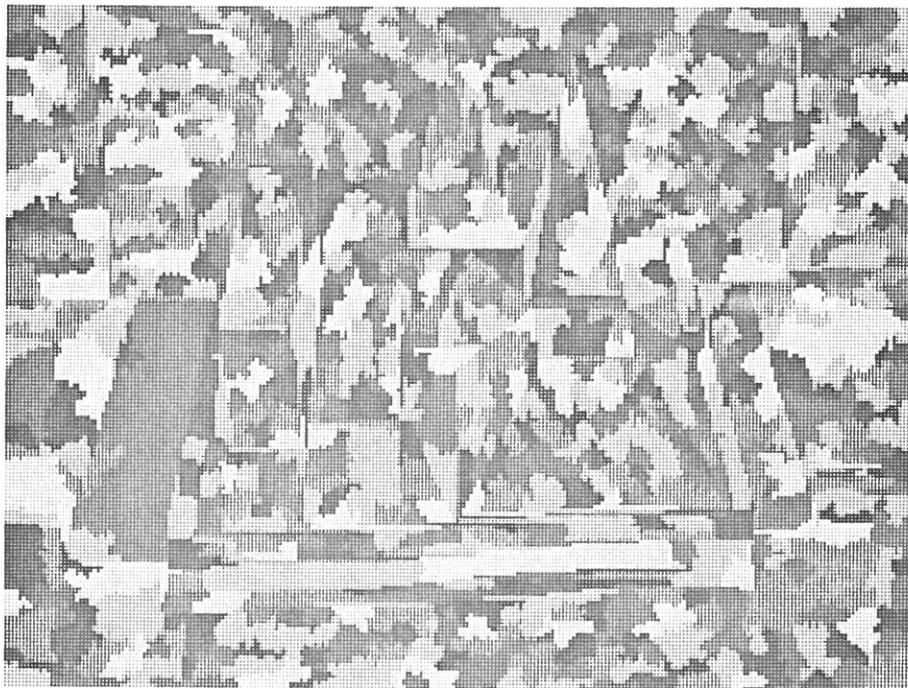


Bild 23: Segmentation von Bild 18 mit Zerlegungskriterium 6 (Hotellings T^2 -Test). 604 Bereiche, $n = 15$, $\alpha = 0.001$, $\theta_y = 30$.

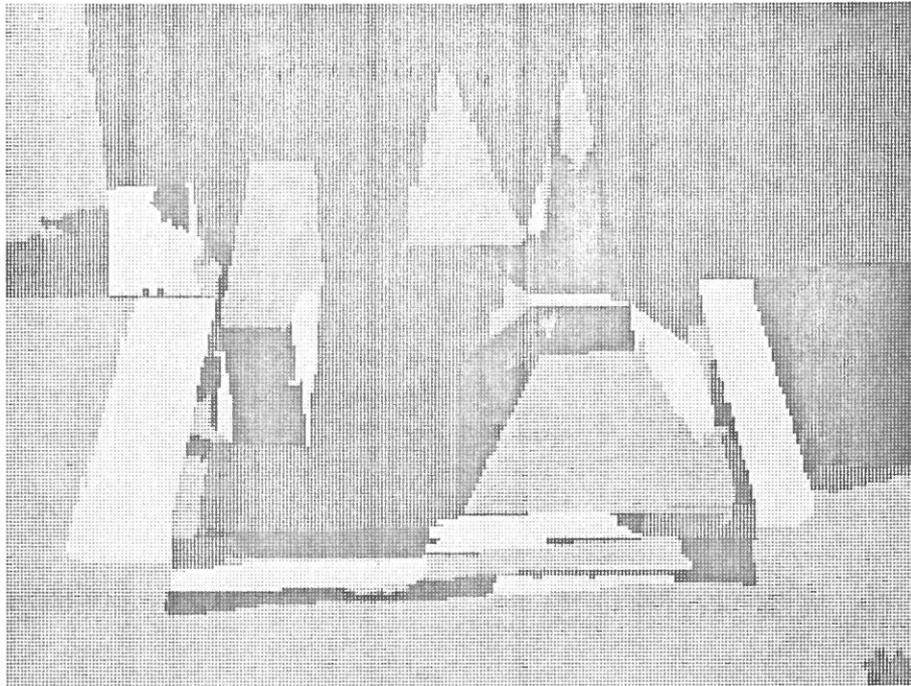


Bild 24: Segmentation von Bild 18 mit Zerlegungskriterium 8 (Polynom 1. Ordnung). 59 Bereiche, $n = 15$, $\theta = 150$, $\theta_g = 30$.



Bild 25: Segmentation von Bild 18 mit Zerlegungskriterium 9 (Polynom 2. Ordnung). 174 Bereiche, $n = 15$, $\theta = 150$, $\theta_g = 30$.

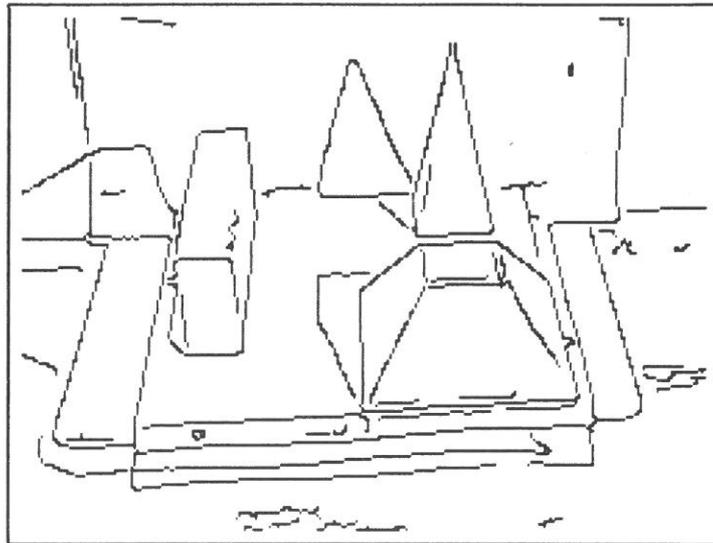


Bild 26: Ergebnis des Kantendetektors für Bild 18 [Drewniok].

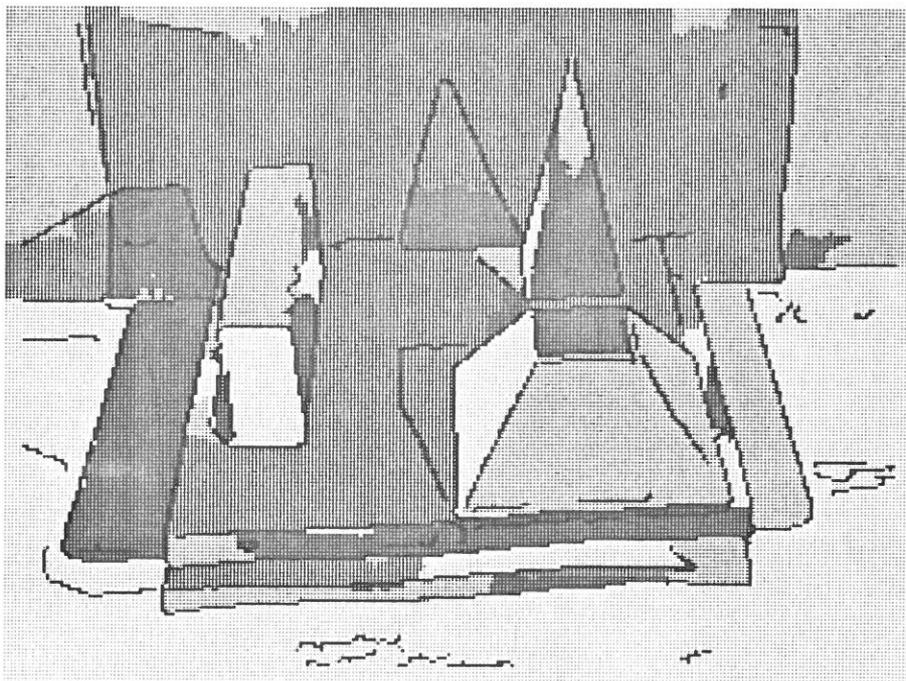


Bild 27: Überlagerung unseres Segmentationsergebnisses aus Bild 22 mit den Ergebnissen des Kantendetektors.

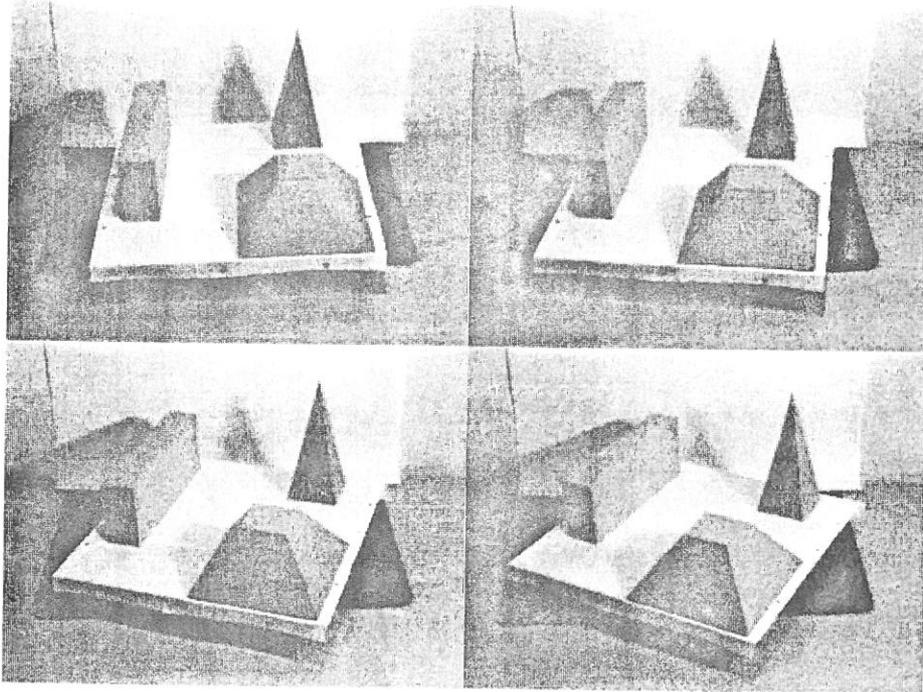


Bild 28: Folge von 4 Bildern. Das erste Bild ist das gleiche wie Bild 18. Der Drehteller wurde von Aufnahme zu Aufnahme jeweils um 10° weiter gedreht.

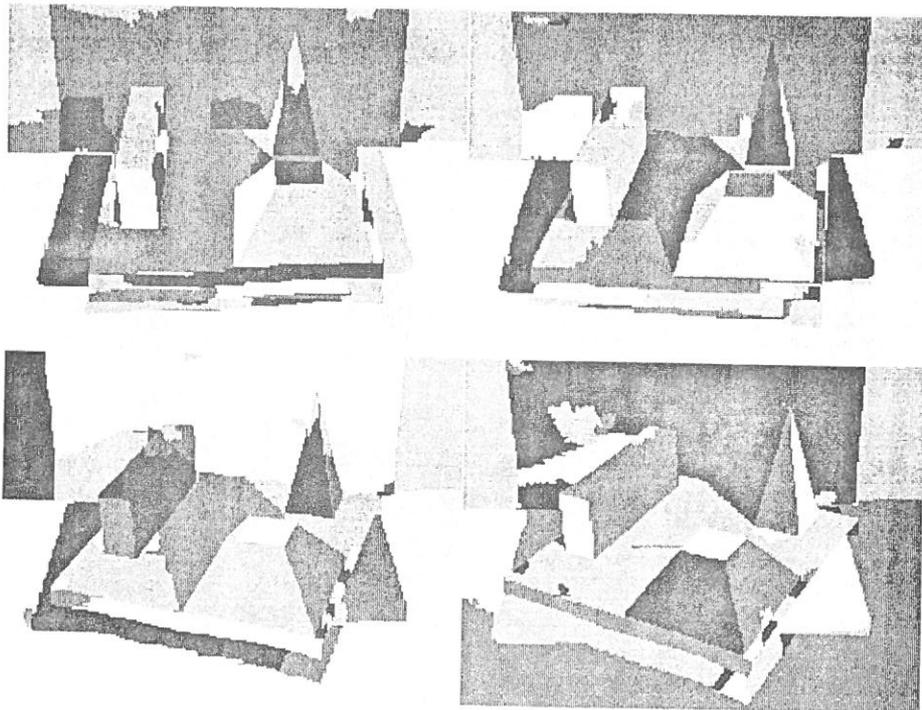


Bild 29: Segmentation der 4 Bilder der Folge von Bild 28. Mit gleichem Kriterium und gleichen Parametern wie in Bild 22 wurden 46, 49, 41 bzw. 45 Bereiche erzeugt.

Anhang B

Schätzung der Verteilungsparameter

Wenn der Logarithmus der Likelihood-Funktion $L(\boldsymbol{\theta})$ nach θ_α differenzierbar ist und der Maximumlikelihood-Schätzer $\hat{\theta}_\alpha$ existiert, dann muß er die folgende Likelihood-Gleichung erfüllen:

$$\frac{\partial}{\partial \theta_\alpha} \ln L(\boldsymbol{\theta}) = 0.$$

Sei die Likelihood-Funktion gegeben durch

$$\begin{aligned} L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \left(\frac{1}{(2\pi)^d \det(\boldsymbol{\Sigma})} \right)^{\frac{n}{2}} \exp \left[-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right] \\ &= \left(\frac{\det(\boldsymbol{\Sigma}^{-1})}{(2\pi)^d} \right)^{\frac{n}{2}} \exp \left[-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right], \end{aligned}$$

dann ist der Logarithmus davon

$$\ln L = -\frac{n}{2} \ln(2\pi)^d + \frac{n}{2} \ln(\det \boldsymbol{\Sigma}^{-1}) - \frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}).$$

B.1 Schätzung des Mittelwertvektors $\boldsymbol{\mu}$

Für die Ableitung von $\ln L$ nach $\boldsymbol{\mu}$ gilt

$$\frac{\partial \ln L}{\partial \boldsymbol{\mu}} = \frac{\partial}{\partial \boldsymbol{\mu}} \left(-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right).$$

Für die Ableitung nach Vektoren verwenden wir folgende Regeln [Horn 86]:

Sei \mathbf{x} ein Vektor und \mathbf{M} eine Matrix, dann gilt

$$\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^T \mathbf{M} \mathbf{x} = (\mathbf{M} + \mathbf{M}^T) \mathbf{x}.$$

Die Ableitung eines Vektors nach einem Vektor ergibt eine Matrix, und speziell die Ableitung eines Vektors nach sich selbst, die Einheitsmatrix \mathbf{I} .

Damit erhalten wir

$$\begin{aligned} \frac{\partial \ln L}{\partial \boldsymbol{\mu}} &= \frac{1}{2} \sum_{i=1}^n \mathbf{I} (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-1T}) (\mathbf{x}_i - \boldsymbol{\mu}) \\ &= \boldsymbol{\Sigma}^{-1} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu}). \end{aligned}$$

Durch Nullsetzen erhalten wir dann die Schätzung

$$\hat{\boldsymbol{\mu}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i. \quad (\text{B.1})$$

B.2 Schätzung der Kovarianzmatrix $\boldsymbol{\Sigma}$

Das Maximum von $\ln L$ bezüglich $\boldsymbol{\Sigma}$ ist gleich dem Maximum von $\ln L$ bezüglich $\boldsymbol{\Sigma}^{-1}$, und der Wert von $\boldsymbol{\Sigma}$ der $\ln L$ maximiert, ist gleich dem Inversen des Wertes von $\boldsymbol{\Sigma}^{-1}$, der $\ln L$ maximiert [Anderson 58]. Wir können daher einfach nach $\boldsymbol{\Sigma}^{-1}$ differenzieren. Dabei verwenden wir noch folgende Regeln für die Ableitung nach Matrizen [Horn 86]:

Sei \mathbf{M} eine Matrix und \mathbf{x} ein Vektor, dann gilt

$$\frac{\partial}{\partial \mathbf{M}} \mathbf{x}^T \mathbf{M} \mathbf{x} = \mathbf{x} \mathbf{x}^T$$

$$\frac{\partial}{\partial \mathbf{M}} \det(\mathbf{M}) = \det(\mathbf{M}) (\mathbf{M}^{-1})^T.$$

Hiermit erhalten wir dann

$$\frac{\partial \ln L}{\partial \boldsymbol{\Sigma}^{-1}} = \frac{n}{2} \frac{1}{\det(\boldsymbol{\Sigma}^{-1})} \det(\boldsymbol{\Sigma}^{-1}) \boldsymbol{\Sigma}^T - \frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T.$$

Nullsetzen ergibt

$$\frac{n}{2}\boldsymbol{\Sigma}^T - \frac{1}{2}\sum_{i=1}^n(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T = 0.$$

Da die Kovarianzmatrix symmetrisch ist ($\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^T$), erhalten wir somit als Schätzung

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^n(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T. \quad (\text{B.2})$$

Anhang C

Schätzung der Parametermatrix \mathbf{B}

Wenn der Logarithmus der Likelihood-Funktion L nach \mathbf{B} differenzierbar ist und der Maximumlikelihood-Schätzer $\hat{\mathbf{B}}$ existiert, dann muß er die folgende Likelihood-Gleichung erfüllen:

$$\frac{\partial}{\partial \mathbf{B}} \ln L = \mathbf{0}.$$

Sei die Likelihood-Funktion gegeben durch

$$L(\mathbf{B}\mathbf{Z}, \Sigma) = \left(\frac{1}{(2\pi)^d \det(\Sigma)} \right)^{\frac{n}{2}} \exp \left[-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{B}\mathbf{z}_i)^T \Sigma^{-1} (\mathbf{x}_i - \mathbf{B}\mathbf{z}_i) \right].$$

Der Logarithmus hiervon ist

$$\ln L = -\frac{n}{2} \ln(2\pi)^d - \frac{n}{2} \ln(\det \Sigma) - \frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{B}\mathbf{z}_i)^T \Sigma^{-1} (\mathbf{x}_i - \mathbf{B}\mathbf{z}_i).$$

Bei der Ableitung nach \mathbf{B} werden die beiden ersten Terme zu Null; uns interessiert also nur noch der dritte, den wir im folgenden mit L_3 bezeichnen wollen. Da wir keine Regel für die Ableitung von Matrizen nach Matrizen kennen, wandeln wir den Term L_3 um in Komponentenschreibweise (wobei σ_{rs}^- die einzelnen Komponenten der inversen Kovarianzmatrix Σ^{-1} sind):

$$\begin{aligned} L_3 &= -\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{B}\mathbf{z}_i)^T \Sigma^{-1} (\mathbf{x}_i - \mathbf{B}\mathbf{z}_i) \\ &= -\frac{1}{2} \sum_{i=1}^n \sum_{s=1}^d \sum_{r=1}^d (x_{ri} - \sum_{g=1}^q \beta_{rg} z_{gi}) \sigma_{rs}^- (x_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}). \end{aligned}$$

Eine geeignete Zerlegung der Summe ist

$$L_3 = -\frac{1}{2} \sum_{i=1}^n \left(\sum_{\substack{s=1 \\ s \neq t}}^d \sum_{\substack{r=1 \\ r \neq t}}^d (x_{ri} - \sum_{g=1}^q \beta_{rg} z_{gi}) \sigma_{rs}^- (x_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}) \right. \\ + \sum_{\substack{r=1 \\ r \neq t}}^d (x_{ri} - \sum_{g=1}^q \beta_{rg} z_{gi}) \sigma_{rt}^- (x_{ti} - \sum_{g=1}^q \beta_{tg} z_{gi}) \\ + \sum_{\substack{s=1 \\ s \neq t}}^d (x_{ti} - \sum_{g=1}^q \beta_{tg} z_{gi}) \sigma_{ts}^- (x_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}) \\ \left. + (x_{ti} - \sum_{g=1}^q \beta_{tg} z_{gi}) \sigma_{tt}^- (x_{ti} - \sum_{g=1}^q \beta_{tg} z_{gi}) \right).$$

Wie man hieran leicht sieht, ist die Ableitung nach einer Komponente β_{th} von **B**

$$\begin{aligned} \frac{\partial L_3}{\partial \beta_{th}} &= \sum_{i=1}^n \sum_{s=1}^d z_{hi} \sigma_{ts}^- (x_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}) \\ &= \sum_{i=1}^n z_{hi} \sum_{s=1}^d \sigma_{ts}^- (x_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}). \end{aligned}$$

Um die Ableitung nach **B** zu bekommen, verwenden wir folgende Ableitungsregel [Horn 86]:

Sei f ein Skalar und **M** eine Matrix mit

$$\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{dann gilt} \quad \frac{\partial f}{\partial \mathbf{M}} = \begin{pmatrix} \frac{\partial f}{\partial a} & \frac{\partial f}{\partial b} \\ \frac{\partial f}{\partial c} & \frac{\partial f}{\partial d} \end{pmatrix}.$$

Damit erhalten wir dann

$$\begin{aligned}
\frac{\partial L_3}{\partial \mathbf{B}} &= \sum_{i=1}^n \begin{pmatrix} z_{1i} \sum_{s=1}^d \sigma_{1s}^- (\mathbf{x}_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}) & \cdots & z_{qi} \sum_{s=1}^d \sigma_{1s}^- (\mathbf{x}_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}) \\ \vdots & \ddots & \vdots \\ z_{1i} \sum_{s=1}^d \sigma_{ds}^- (\mathbf{x}_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}) & \cdots & z_{qi} \sum_{s=1}^d \sigma_{ds}^- (\mathbf{x}_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}) \end{pmatrix} \\
&= \sum_{i=1}^n \begin{pmatrix} \sum_{s=1}^d \sigma_{1s}^- (\mathbf{x}_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}) \\ \vdots \\ \sum_{s=1}^d \sigma_{ds}^- (\mathbf{x}_{si} - \sum_{g=1}^q \beta_{sg} z_{gi}) \end{pmatrix} \mathbf{z}_i^T \\
&= \sum_{i=1}^n \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \mathbf{B} \mathbf{z}_i) \mathbf{z}_i^T.
\end{aligned}$$

Wir setzen die Ableitung gleich Null

$$\boldsymbol{\Sigma}^{-1} \sum_{i=1}^n (\mathbf{x}_i \mathbf{z}_i^T - \mathbf{B} \mathbf{z}_i \mathbf{z}_i^T) = \mathbf{0}$$

und erhalten als Schätzung für die Parametermatrix \mathbf{B}

$$\begin{aligned}
\hat{\mathbf{B}} &= \left(\sum_{i=1}^n \mathbf{x}_i \mathbf{z}_i^T \right) \left(\sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i^T \right)^{-1} \\
\hat{\mathbf{B}} &= \mathbf{C} \mathbf{A}^{-1}
\end{aligned}$$

oder

$$\begin{aligned}
\hat{\mathbf{B}} &= \left(\sum_{i=1}^n \mathbf{x}_i \mathbf{z}_i^T \right) \left(\sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i^T \right)^{-1} \\
\hat{\mathbf{B}} &= \mathbf{X} \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \\
\hat{\mathbf{B}} &= \mathbf{X} \mathbf{Z}^\dagger,
\end{aligned} \tag{C.1}$$

wobei \mathbf{Z}^\dagger die Pseudoinverse von \mathbf{Z} ist.

Zum Abschluß sei noch bemerkt, daß diese Lösung durch Berechnung der Pseudoinversen die Standardlösung bei der Anpassung von Funktionen an Daten mit der Methode der kleinsten Fehlerquadrate ist. Dies liegt daran, daß wir hier eine Normalverteilung verwenden, bei der ein Maximumlikelihood-Schätzer gleich einem Schätzer der Methode der kleinsten Fehlerquadrate ist.

Anhang D

Exponentenvereinfachung

Die Wahrscheinlichkeitsdichte einer multivariaten Normalverteilung ist gegeben durch

$$p(\mathbf{x}_i) = \frac{1}{\sqrt{(2\pi)^d \det(\boldsymbol{\Sigma})}} \exp \left[-\frac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right]$$

und ihre Likelihood-Funktion durch

$$\begin{aligned} L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \prod_{i=1}^n p(\mathbf{x}_i) \\ &= \frac{1}{\left(\sqrt{(2\pi)^d \det(\boldsymbol{\Sigma})} \right)^n} \exp \left[-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right]. \end{aligned}$$

Der Exponent der Likelihood-Funktion lässt sich nun wie folgt vereinfachen: Mit $\mathbf{x}^T \mathbf{M} \mathbf{x} = \text{Spur}(\mathbf{M} \mathbf{x} \mathbf{x}^T)$ erhalten wir

$$\exp \left[-\frac{1}{2} \sum_{i=1}^n \text{Spur} \left(\boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) (\mathbf{x}_i - \boldsymbol{\mu})^T \right) \right].$$

Verwenden wir $\text{Spur} \mathbf{A} + \text{Spur} \mathbf{B} = \text{Spur}(\mathbf{A} + \mathbf{B})$, so ergibt sich

$$\exp \left[-\frac{1}{2} \text{Spur} \left(\sum_{i=1}^n \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) (\mathbf{x}_i - \boldsymbol{\mu})^T \right) \right].$$

Die Distributivität gilt auch für Matrizenrechnung; es folgt also

$$\exp \left[-\frac{1}{2} \text{Spur} \left(\boldsymbol{\Sigma}^{-1} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu}) (\mathbf{x}_i - \boldsymbol{\mu})^T \right) \right].$$

Mit Hilfe von Gleichung B.2 erhalten wir

$$\exp \left[-\frac{1}{2} \text{Spur} (n \Sigma^{-1} \Sigma) \right]$$
$$\exp \left[-\frac{n}{2} \text{Spur} \mathbf{I} \right].$$

Mit $\text{Spur} \mathbf{I} = \dim \mathbf{I} = d$ ergibt sich schließlich

$$\exp \left[-\frac{nd}{2} \right].$$

Literaturverzeichnis

Abramowitz + Stegun 68

Handbook of Mathematical Functions.
Milton Abramowitz, Irene A. Stegun.
Dover Publications, New York 1968.

Aho et al. 83

Data Structures and Algorithms.
Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman.
Addison-Wesley, Reading, Mass. 1983.

Anderson 58

An Introduction to Multivariate Statistical Analysis.
T.W. Anderson.
John Wiley & Sons, New York 1958.

Ballard + Brown 82

Computer Vision.
Dana H. Ballard, Christopher M. Brown.
Prentice-Hall, Englewood Cliffs, New Jersey 1982.

Bartsch et al. 86

Merkmalsdetektion in Farbbildern als Grundlage zur Korrespondenzanalyse in Stereo-Bildfolgen.
Thomas Bartsch, Leonie Dreschler-Fischer, Carsten Schröder.
"Mustererkennung 1986", 8. DAGM-Symposium, Paderborn, 30. Sept. - 2. Okt. 1986, G. Hartmann (Hrsg.), Informatik-Fachberichte Bd. 125, Springer Verlag, Berlin-Heidelberg-New York 1986, 94-98.

Bartsch 87

Untersuchungen zur Detektion und Lokalisation markanter Punkte

in Farbbildern.

Thomas Bartsch.

Diplomarbeit, Fachbereich Informatik, Universität Hamburg, 1987.

Bickel + Doksum 77

Mathematical Statistics: Basic Ideas and Selected Topics.

Peter J. Bickel, Kjell A. Doksum.

Holden-Day, San Francisco 1977.

Bovik et al. 87

The Effect of Median Filtering on Edge Estimation and Detection.

Alan Conrad Bovik, Thomas S. Huang, David C. Munson, Jr.

IEEE Transactions on Pattern Analysis and Machine Intelligence

PAMI-9 (1987) 181-194.

Brice + Fennema 70

Scene Analysis Using Regions.

Claude R. Brice, Claude L. Fennema.

Artificial Intelligence **1** (1970) 205-226.

Burr + Chien 76

The Minimal Spanning Tree in Visual Data Segmentation.

D.J. Burr, R.T. Chien.

Proc. 3. Int. Joint Conf. Patt. Recogn. IJCPR-76, Nov. 8-11, 1976,
Coronado/CA, 519-523.

Canny 83

Finding Edges and Lines in Images.

John F. Canny.

Technical Report TR-720, Artificial Intelligence Laboratory, MIT,
1983.

Coleman + Andrews 79

Image Segmentation by Clustering.

Guy B. Coleman, Harry C. Andrews.

Proc. of the IEEE **67** (1979) 773-785.

Dreschler + Nagel 82

*Volumetric Model and 3D-Trajectory of a Moving Car Derived from
Monocular TV-Frame Sequences of a Street Scene.*

Leonie Dreschler, Hans-Hellmut Nagel.

Computer Graphics and Image Processing **20** (1982) 199-228.

Dreschler-Fischer 86

A Blackboard System for Dynamic Stereo Matching.

Leonie S. Dreschler-Fischer.

Proc. 1. Int. Conf. on Autonomous Mobile Systems, Amsterdam, Dezember 1986, (im Druck).

Dreschler-Fischer 87

Das "Bootstrap-Problem" bei der geometrischen Szenenrekonstruktion — Eine Übersicht.

Leonie S. Dreschler-Fischer.

"GWAI-87", 11. Fachtagung über Künstliche Intelligenz, Schloß Eringerfeld bei Geseke, 28. Sept. – 2. Okt. 1987, Katharina Morik (Hrsg.), Informatik Fachberichte Bd. 152, Springer Verlag, Berlin-Heidelberg-New York 1987, 1–15.

Drewniok

Untersuchungen zur Detektion von Kanten in digitalen Bildern unter Ausnutzung von Farbinformation.

Christian Drewniok.

Diplomarbeit, Fachbereich Informatik, Universität Hamburg, (in Vorbereitung).

Duda + Hart 73

Pattern Classification and Scene Analysis.

Richard O. Duda, Peter E. Hart.

John Wiley & Sons, New York 1973.

Freuder 76

Affinity: A Relative Approach to Region Finding.

Eugene C. Freuder.

Computer Graphics and Image Processing 5 (1976) 254–264.

Hanson + Riseman 78

Segmentation of Natural Scenes.

Allen R. Hanson, Edward M. Riseman.

in: *Computer Vision Systems*, A.R. Hanson, E.M. Riseman (Hrsg.), Academic Press, Orlando 1978, 129–162.

Haralick + Shapiro 85

Image Segmentation Techniques.

Robert M. Haralick, Linda G. Shapiro.

Computer Vision, Graphics, and Image Processing **29** (1985) 100–132.

Hille 80

Methoden und Modelle in der Bildsegmentation. Eine Übersicht.
Gunter Hille.

Bericht IFI-HH-B-80/81, Fachbereich Informatik, Universität Hamburg, 1980.

Hille 82

Polynomiale Approximation von digitalen Bildfunktionen.
Gunter Hille.

Mitteilung IFI-HH-M-98/82, Fachbereich Informatik, Universität Hamburg, 1982.

Horn 86

Robot Vision.

Berthold Klaus Paul Horn.

MIT Press, Cambridge, Mass. 1986.

Horowitz + Pavlides 74

Picture Segmentation by a directed Split-and-Merge Procedure.

Steven L. Horowitz, Theodosios Pavlidis.

Proc. 2. Int. Joint Conf. Patt. Recogn. IJCPR-74, Aug. 1974, Kopenhagen, Dänemark.

Hsu et al. 82

New Likelihood Test Methods for Change Detection in Image Sequences.

Y.Z. Hsu, Hans-Hellmut Nagel, Gudrun Rekers.

Mitteilung IFI-HH-M-104/82, Fachbereich Informatik, Universität Hamburg, 1982

siehe auch: *Computer Vision, Graphics, and Image Processing* **26** (1984) 73–106.

James 54

Tests of Linear Hypotheses in Univariate and Multivariate Analysis when the Ratios of the Population Variances are Unknown.

G.S. James.

Biometrika **41** (1954) 19–43.

Kanade 80

Region Segmentation: Signal vs Semantics.

Takeo Kanade.

Computer Graphics and Image Processing **13** (1980) 279–297.

Kender 76

Saturation, Hue, and Normalized Color: Calculation, Digitization Effects, and Use.

John R. Kender.

Technical Report, Dep. of Computer Science, Carnegie-Mellon University, 1976.

Kruskal 56

On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem.

Joseph B. Kruskal, Jr.

Proc. Amer. Math. Soc. **7** (1956) 48–50.

Mohnhaupt 87

Private Korrespondenz.

Michael Mohnhaupt.

Fachbereich Informatik, Universität Hamburg, 1987.

Morrison 67

Multivariate Statistical Methods.

Donald F. Morrison.

McGraw Hill, New York 1967.

Mottschall 85

Segmentation als Bewertung der Eignung von bivariaten Polynomen 2. Ordnung zur Approximation der Bildfunktion.

Thomas Mottschall.

Diplomarbeit, Fachbereich Informatik, Universität Hamburg, April 1985.

Muerle + Allen 68

Experimental Evaluation of Techniques for Automatic Segmentation of Objects in a Complex Scene.

J.L. Muerle, D.C. Allen.

in: *Pictorial Pattern Recognition*, G.C. Cheng et al. (Hrsg.), Thompson, Washington 1968, 3–13.

Nagel 76

Experiences with Yakimovsky's Algorithm for Boundary and Object Detection in Real World Images.

Hans-Hellmut Nagel.

Bericht IFI-HH-B-23/76, Fachbereich Informatik, Universität Hamburg, 1976.

Nagin et al. 82

Studies in Global and Local Histogram-Guided Relaxation Algorithms.

Paul A. Nagin, Allen R. Hanson, Edward M. Riseman.

IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-4 (1982) 263-277.

Ohlander et al. 78

Picture Segmentation Using a Recursive Splitting Method.

Ron Ohlander, Keith E. Price, D. Raj Reddy.

Computer Graphics and Image Processing **8** (1978) 313-333.

Ohta et al. 80

Color Information for Region Segmentation.

Yuichi Ohta, Takeo Kanade, Toshiyuki Sakai.

Computer Graphics and Image Processing **13** (1980) 222-241.

Price 86

Anything You Can Do, I Can Do Better (No You Can't)

Keith E. Price.

Computer Vision, Graphics, and Image Processing **36** (1986) 387-391.

Prim 57

Shortest Connection Networks and some Generalizations.

R.C. Prim.

Bell Syst. Tech. Journ. **36** (1957) 1389-1401.

Rosenfeld + Kak 82

Digital Picture Processing. Second Edition.

Azriel Rosenfeld, Avinash C. Kak.

Academic Press, Orlando 1982, 2 Bände.

Schachter et al. 75

Scene Segmentation by Cluster Detection in Color Spaces.

Bruce J. Schachter, Larry S. Davis, Azriel Rosenfeld.
Technical Report TR-424, Computer Science Center, University of
Maryland, 1975.
Kurzfassung in: *ACM SIGART Newsletter* **58** (1976) 16–17.

Schachter et al. 79

Some Experiments in Image Segmentation by Clustering of Local Feature Values.

Bruce J. Schachter, Larry S. Davis, Azriel Rosenfeld.
Pattern Recognition **11** (1979) 19–28.

Shafer 80

MOOSE: Users' Manual, Implementation Guide, Evaluation.

Steven A. Shafer.

Bericht IFI-HH-B-70/80, Fachbereich Informatik, Universität Hamburg, 1980.

Shafer + Kanade 82

Recursive Region Segmentation by Analysis of Histograms.

Steven A. Shafer, Takeo Kanade.

Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing,
Vol. 2, Paris, May 1982, 1166–1171.

Scheffé 70

Practical Solutions of the Behrens-Fisher Problem.

Henry Scheffé.

Journal of the American Stat. Assoc. **65** (1970) 1501–1508.

Sedgewick 83

Algorithms.

Robert Sedgewick.

Addison-Wesley, Reading, Mass. 1983.

Sprengel 86

DESIRE: Entwurf einer Datenstruktur zur symbolischen Beschreibung von Bildern.

Rainer Sprengel.

Mitteilung IFI-HH-M-142/86, Fachbereich Informatik, Universität Hamburg, 1986.

Tarjan 75

On the Efficiency of a Good but not Linear Set Merging Algorithm.

Robert E. Tarjan.
Journal of the ACM **22** (1975) 215-225.

Tou + Gonzalez 74

Pattern Recognition Principles.
Julius T. Tou, Raphael C. Gonzalez.
Addison-Wesley, Reading, Mass. 1974.

Ullman 79

The Interpretation of Visual Motion.
Shimon Ullman.
MIT Press, Cambridge, Mass. 1979.

Wirth 79

Algorithmen und Datenstrukturen.
Niklaus Wirth.
Teubner Studienbücher Informatik, Stuttgart 1979.

Yakimovsky 76

Boundary and Object Detection in Real World Images.
Yoram Yakimovsky.
Journal of the ACM **23** (1976) 599-618.

Zahn 71

Graph-Theoretical Methods for Detecting and Describing Gestalt Clusters.
Charles T. Zahn.
IEEE Transactions on Computers **C-20** (1971) 68-86.

Zucker 76

Region Growing: Childhood and Adolescence.
Steven W. Zucker.
Computer Graphics and Image Processing **5** (1976) 382-399.

Danksagungen

An dieser Stelle möchte ich ganz besonders meinen Eltern danken, die mir mein Studium und damit auch die Anfertigung dieser Arbeit ermöglicht haben.

Meiner Betreuerin Dr. Leonie Dreschler-Fischer danke ich für ihre Unterstützung bei der Durchführung dieser Arbeit. Ihr unerschütterlicher Optimismus hat mir bei Rückschlägen stets neuen Mut eingeflößt.

Besonderer Dank gebührt auch Rainer Sprengel. Er hat mir durch viele Diskussionen zu tieferer Einsicht in das bearbeitete Gebiet verholfen und eine erste Fassung dieser Arbeit kritisch durchgesehen.

Den Mitgliedern der Projektgruppe SISSY danke ich für die gute Zusammenarbeit, vor allem danke ich jedoch Chrischan Drewniok für die fruchtbare und freundschaftliche Zusammenarbeit während der ganzen Zeit unseres gemeinsamen Studiums.

Michael Mohnhaupt hat durch ständiges Nachfragen und Drängen die Fertigstellung dieser Arbeit sehr beschleunigt. Ohne ihn hätte ich mein Studium nicht so rasch beenden können. Hierfür bin ich ihm sehr dankbar.

Ingeborg Heer-Mück und Hartmuth Krüger waren sehr hilfsbereit, wenn es Probleme mit der verwendeten Bildverarbeitungsapparatur gab, und Helmut Faasch und Volker Haarslev standen bei Problemen mit der Software stets hilfreich zur Seite.

Die Mitarbeiter des Projektes TEX-K haben mir für meine recht aufwendigen Berechnungen ihre MicroVAX BIANCA zur Verfügung gestellt.

Bei der Durchführung dieser Arbeit wurde Software benutzt, die durch von der Deutschen Forschungsgemeinschaft finanzierte studentische Hilfskräfte erstellt wurde.

