

DEUTSCHE GEODÄTISCHE KOMMISSION

bei der Bayerischen Akademie der Wissenschaften

Reihe C

Dissertationen

Heft Nr. 370

**Statistische Verfahren
für die automatische Bildanalyse und ihre Bewertung
bei der Objekterkennung und -vermessung**

Von der Fakultät für
Bauingenieur- und Vermessungswesen
der Universität Stuttgart
zur Verleihung der Lehrbefähigung
für das Lehrgebiet
"Digitale Bildverarbeitung"
genehmigte Habilitationsschrift

vorgelegt von

Dr.-Ing. Wolfgang Förstner

München 1991

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission bei der C.H.Beck'schen Verlagsbuchhandlung München

ISSN 0065-5325

ISBN 3 7696 9417 1

Adresse der Deutschen Geodätischen Kommission:

DEUTSCHE GEODÄTISCHE KOMMISSION

D-8000 MÜNCHEN 22 • Marstallplatz 8 • Postfach 220 002

Telefon (089) 23 031 113 • Telefax (089) 23 031 240

Telex 5 213 550 dgfi d

Hauptberichter: Prof. Dr.-Ing. F. Ackermann

Mitberichter: Prof. Dr.-Ing. W. H. Bloss
Prof. Dr.-Ing.habil. E. W. Grafarend

Tag der Einreichung: 24. März 1988

Tag des Habilitationskolloquiums: 23. Januar 1989

© 1991 Deutsche Geodätische Kommission, München

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

<u>Inhalt</u>	3
Liste der Figuren	4
1. EINLEITUNG	7
1.1 Allgemeines	7
1.2 Probleme der Bildanalyse	8
1.3 Zielsetzung der Arbeit	13
2. VERFAHREN DER EINBILDANALYSE	16
2.1 Vorverarbeitung und Extraktion von Merkmalskarten	16
2.1.1 Das Bildmodell	18
2.1.1.1 Die Gebietseinteilung	18
2.1.1.2 Die Textur	18
2.1.1.3 Das Abtauschen	19
2.1.2 Schätzung der Kovarianzfunktion C	19
2.1.3 Merkmalskarten	20
2.1.3.1 Ikonische Merkmale	20
2.1.3.2 Beispiele für Merkmalskarten	23
2.1.3.3 Vergleich mit anderen Verfahren der lokalen Bildanalyse	25
2.1.4 Fehleranalyse und Filterung der Schätzungen	27
2.1.4.1 Verteilung der geschätzten Merkmale	27
2.1.4.2 Filterung der geschätzten Signalvarianz und der Hessematrix	29
2.1.5 Informationserhaltende Filterung	30
2.1.5.1 Schätzung des Rauschanteils	30
2.1.5.2 Das adaptive Filter	36
2.2 Extraktion von markanten Punkten und Bildkanten	40
2.2.1 Kriterien für die Auswahl markanter Punkte	41
2.2.2 Lokalisierung von Punkten	42
2.2.2.1 Mathematisches Modell	42
2.2.2.2 Normalgleichungen	45
2.2.2.3 Genauigkeitseigenschaften der Punkte	48
2.2.2.4 Markante Punkte in Farbbildern	48
2.2.3 Klassifizierung markanter Punkte	49
2.2.4 Selektion optimaler Fenster für die Punktauswahl	52
2.2.5 Extraktion von Bildkanten	54
2.2.5.1 Selektion optimaler Fenster für die Kantenextraktion	55
2.2.5.2 Lokalisierung von Kantenelementen	55
2.2.5.3 Verknüpfung der Kantenelemente zu geraden Kantenstücken	57
2.2.5.4 Beispiele zur Kantenextraktion	60
2.3 Schlußfolgerungen	61
3. ÄHNLICHKEIT UND SELTENHEIT SYMBOLISCHER UND NUMERISCHER MERKMALE	63
3.1 Das Zuordnungsproblem	64
3.2 Grundlagen aus der Informationstheorie	65
3.3 Ähnlichkeit symbolischer und numerischer Attribute	69
3.3.1 Das informationstheoretische Ähnlichkeitsmaß nach BOYER und KAK	69
3.3.2 Relationen zu Prüfgrößen aus der Statistik	71
3.4 Beurteilung von Attributen	73
3.4.1 Überwachte Beurteilung von Attributen auf Invarianz	75
3.4.2 Unüberwachte Beurteilung von Attributen auf Seltenheit	77
3.4.3 Beispiele für die unüberwachte Bewertung	79
3.4.3.1 Isolierte und seltene Punkte	80
3.4.3.2 Auffällige Symbolfolgen	82
4. Zusammenfassung und Ausblick	87
Verwendete Literatur	92

Anhang 1: Beweis von Gl. 215-15 und 215-20	97
Anhang 2: Ableitung der Verteilung von \underline{g} nach Gl. 215-16	98
Anhang 3: Figuren zu Teil 2	99

Liste der Figuren:

1- 1	Eine Struktur zur Beschreibung des Prozesses der automatischen Bildanalyse	8
1- 2	Komponenten eines Systems zur Analyse von Mustern (nach NIEMANN 1981) . .	12
1- 3	Schema für die Bildzuordnung und die Objekterkennung	14
2- 1	Prozeß zur Erstellung von Merkmalskarten	17
2- 2	Dichte- und Verteilungsfunktion der X_2^2 -Verteilung	33
2- 3	Bedingte Dichte- und Verteilungsfunktion der X_2^2 -Verteilung	34
2- 4	Kantenelement und Gefällelement	46
2- 5	Markante Punkte in selektierten Fenstern mit Kantenelementen und Konfidenzellipse, simulierte Daten	47
2- 6	Position der Mitten selektierter Fenster und markanter Punkte	54
2- 7	Für die Position von Ecken und Kanten maßgebliche Stellen innerhalb eines Fensters	57
2- 8	Repräsentation eines geraden Kantenstücks während der Verknüpfung der Kantenelemente	58
3- 1	Übertragungskanal für das Attribut "Kantenrichtung" und Kanalmatrix . . .	70
3- 2	Punktgruppe mit einem isolierten Punkt	80
3- 3	Teil eines Schachbretts, 4 Punkte mit Seltenheitsmaß	81
3- 4	Teil eines Schachbretts, 9 Punkte mit Seltenheitsmaß	81
3- 5	Teil eines Schachbretts, 16 Punkte mit Seltenheitsmaß	81
3- 6	Teil eines Schachbretts, 25 Punkte mit Seltenheitsmaß	81
3- 7	Bestimmung der Levenshtein-Distanz als kürzestem Weg in einem gerichteten, bewerteten Graphen	83
3- 8	Zuordnung der Symbolfolgen WESTEN und OSTEN	83
3- 9	Zuordnung zweier Symbolfolgen	84
<u>Figuren im Anhang 3</u>		99

A2- 1	Merkmalskarten eines unverrauschten Siemenssterns	
A2- 2	Merkmalskarten eines verrauschten Siemenssterns	
A2- 3	Merkmalskarten eines schwach verrauschten Schachbretts	
A2- 4	Merkmalskarten eines verrauschten Schachbretts	
A2- 5	Merkmalskarten eines schwach verrauschten Fischgrätenmusters	
A2- 6	Merkmalskarten eines schwach verrauschten Wellenmusters	
A2- 7	Merkmalskarten eines Spielzeugbildes	
A2- 8	Merkmalskarten eines Gebäudeteils in einem Luftbild 1 : 12 000	
A2- 9	Streifenmuster mit Histogramm der Grauwertgradientenquadrate	
A2-10	Verrauschtes Rosettenmuster mit Histogramm der Grauwertgradientenquadrate	
A2-11	Verrauschtes Wellenmuster mit Histogramm der Grauwertgradientenquadrate	
A2-12	Luftbild mit Histogramm der Grauwertgradientenquadrate	
A2-13	Verrauschter Siemensstern, informationserhaltend gefiltert mit Merkmalskarten	
A2-14	Verrauschtes Schachbrett, informationserhaltend gefiltert mit Merkmalskarten	
A2-15	Stark verrauschtes Fischgrätenmuster, zweimal informationserhaltend gefiltert, mit Merkmalskarten	
A2-16	Luftbild "Antarktis I", informationserhaltend gefiltert, mit Merkmalskarten	
A2-17	Luftbild "Antarktis II", informationserhaltend gefiltert, mit Merkmalskarten	
A2-18	Baum, informationserhaltend gefiltert, mit Merkmalskarten	
A2-19	Schwach verrauschtes Streifenmuster, informationserhaltend gefiltert, Merkmalskarten	

- A2-20 Schwach verrauschtes Streifenmuster, 3x3 Median gefiltert, mit Merkmalskarten
- A2-21 Stark verrauschtes Streifenmuster, informationserhaltend gefiltert, mit Merkmalskarten
- A2-22 Stark verrauschtes Streifenmuster, 3x3 Median gefiltert, mit Merkmalskarten
- A2-23 Stark verrauschtes Streifenmuster, zweimal informationserhaltend gefiltert, mit Merkmalskarten
- A2-24 Stark verrauschtes Streifenmuster, 5x5 Median gefiltert, mit Merkmalskarten
- A2-25 Stark verrauschtes Rosettenmuster, informationserhaltend gefiltert, mit Merkmalskarten
- A2-26 Stark verrauschtes Rosettenmuster, 3x3 Median gefiltert, mit Merkmalskarten
- A2-27 Stark verrauschtes Rosettenmuster, zweimal informationserhaltend gefiltert, mit Merkmalskarten
- A2-28 Stark verrauschtes Rosettenmuster, 5x5 Median gefiltert, mit Merkmalskarten
- A2-29 Luftbild, informationserhaltend gefiltert, mit Merkmalskarten
- A2-30 Luftbild, 3x3 Median gefiltert, mit Merkmalskarten
- A2-31 Schwach verrauschtes Schachbrett, mit Merkmalskarte, Kantenbereich, Kantenpixel
- A2-32 Schwach verrauschtes Schachbrett mit Merkmalskarten, ungefiltert, gerade Kantenstücke, Kantenelemente
- A2-33 Schwach verrauschtes Schachbrett mit Merkmalskarten, informationserhaltend gefiltert, gerade Kantenstücke, Kantenelemente
- A2-34 Verrauschtes Schachbrett mit Merkmalskarten ungefiltert, gerade Kantenstücke, Kantenelemente
- A2-35 Verrauschtes Schachbrett mit Merkmalskarten informationserhaltend gefiltert, gerade Kantenstücke, Kantenelemente
- A2-36 Stark verrauschtes Schachbrett mit Merkmalskarten ungefiltert, gerade Kantenstücke, Kantenelemente
- A2-37 Stark verrauschtes Schachbrett mit Merkmalskarten informationserhaltend gefiltert, gerade Kantenstücke, Kantenelemente
- A2-38 Schwach verrauschtes Streifenmuster mit Merkmalskarten informationserhaltend gefiltert, gerade Kantenstücke, Kantenelemente
- A2-39 Schwach verrauschtes Streifenmuster mit Merkmalskarten 3x3 Median gefiltert, gerade Kantenstücke, Kantenelemente
- A2-40 Verrauschtes Streifenmuster mit Merkmalskarten informationserhaltend gefiltert, gerade Kantenstücke, Kantenelemente
- A2-41 Verrauschtes Streifenmuster mit Merkmalskarten 3x3 Median gefiltert, gerade Kantenstücke, Kantenelemente
- A2-42 Stark verrauschtes Streifenmuster mit Merkmalskarten zweimal informationserhaltend gefiltert, gerade Kantenstücke, Kantenelemente
- A2-43 Stark verrauschtes Streifenmuster mit Merkmalskarten 5x5 Median gefiltert, gerade Kantenstücke, Kantenelemente

1. EINLEITUNG

1.1 Allgemeines

Eines der faszinierendsten Wissenschaftsgebiete ist die Analyse des menschlichen Sehsystems, beziehen wir doch den größten Teil der Information über unser Auge. Wesentliches Hilfsmittel dabei ist die Nachbildung des Sehvorgangs auf dem Rechner zum Zweck des Vergleichs mit psychologischen Kontrollversuchen. Die Fähigkeit, auf Rechnern automatisch Bilder zu analysieren und zu interpretieren, ist aber als eigenständiges Ziel von mindestens gleichem Interesse. Automatische Bildanalyse-Systeme werden in verschiedenen Gebieten erfolgreich eingesetzt: etwa in der Industrie für die kameragestützte Führung von Robotern oder der Fertigungskontrolle, in der Medizin bei der Blutanalyse oder der Computertomographie, oder in der Fernerkundung zur Umweltüberwachung oder Kartenherstellung.

Die anfängliche Euphorie, sehende Maschinen herstellen zu können, wurde trotz dieser Erfolge bald gedämpft, als man erkannte, daß beim Menschen das Gehirn, nicht das Auge, den Hauptanteil an der Analyse der optischen Reize trägt, wie etwa der Versuch am MIT belegt, bei dem einem von Geburt an Blinden die Intensitätswerte einer Digitalkamera als Nadeldruck-Raster über den Rücken übermittelt wurden und der mit Hilfe dieses künstlichen "Auges" den Henkel einer Kaffeekanne oder die Form einer Kerzenflamme - eine völlig neue Wahrnehmung - erkennen konnte. Die Einsicht, daß das Auge als optischer Sensor wohl notwendig, aber keineswegs hinreichend ist für die Fähigkeit, erkennen zu können, machte den Bereich "Computer Vision" zum Teilgebiet der künstlichen Intelligenz. Es ist wegen seiner Breite und Bedeutung heute ein eigenes Gebiet wissenschaftlicher Forschung. Neben dieser Einsicht in die Komplexität der Aufgabe, die vom menschlichen Sehsystem scheinbar ohne Aufwand gelöst wird, hat auch die Überschätzung der Leistungsfähigkeit und Flexibilität realisierter Systeme dazu geführt, daß andere, konkurrierende Verfahren, die für spezielle Aufgaben zugeschnitten sind, stärkere Verbreitung gefunden haben, wie etwa Laser-Ranging- oder Lichtschnitt-Verfahren gegenüber Stereoverfahren zur Tiefenmessung.

Im Bereich der Photogrammetrie haben die Verfahren der digitalen Bildanalyse trotz beachtlicher Einzelentwicklungen, wie etwa dem Gestalt-Photomapper, erst recht spät Fuß gefaßt. Der wesentliche Grund liegt wohl in der im Vergleich zum menschlichen Operateur von der Effizienz und der Genauigkeit her nicht konkurrenzfähigen Verfahren für die Lösung des Stereo- bzw. Bildzuordnungsproblems. Die Leistungssteigerung in der Rechentechnik, vor allem aber die theoretische Aufarbeitung des Bildzuordnungsproblems haben die Entwicklung von meßtechnischen Verfahren auf der Basis digitaler bzw. digitalisierter Bilder in der Photogrammetrie vorange-trieben. Heute sind automatische, photogrammetrische wie geodätische Meßsysteme in der Anwendung, deren Sensor eine CCD-Camera ist und die in Geschwindigkeit, Zuverlässigkeit und Präzision manuellen Messungen wenigstens gleichwertig sind. Wieder liegt der Grund für die Konkurrenzfähigkeit der Verfahren in der Spezialisierung, etwa der Beschränkung auf retroreflektierende Signale (vgl. z. B. FRASER/BROWN 1986, HAGGREN 1985) oder künstliche, aufprojizierte Texturen (SCHEWE und FÖRSTNER 1986) und in der gesicherten Abgrenzung des automatischen vom interaktiven Meßteil. Einer breiteren Anwendung von Bildanalysetechniken steht auch hier das Fehlen zuverlässiger Interpretationsverfahren entgegen, etwa zur automatischen Erkennung natürlicher, nicht signalisierter Meßpunkte oder zu kartierender Objekte.

Wir wollen den Ursachen für die Schwierigkeiten der automatischen Bildanalyse etwas näher auf den Grund gehen und dabei die Vielfalt der Probleme auffächern. Eine vollständige Darstellung des Bereichs ist wohl kaum möglich. Die Schlußfolgerungen aus den folgenden Darlegungen bleiben dennoch gültig und sollen zur Motivation der vorliegenden Arbeit dienen.

1.2 Probleme der Bildanalyse

Als Aufgabe der automatischen Bildanalyse gilt derzeit, mit Hilfe von Rechnern aus Bildern Informationen über die abgebildeten Objekte zu erhalten. Wir wollen in unserer Darstellung alle flächenhaft abbildenden Systeme als mögliche Sensoren einbeziehen, also neben optischen auch Infrarot-, Röntgen- oder Radarbilder, aber auch akustische, wie Ultraschall oder durch Ranging-Techniken erzeugte Entfernungsbilder.

Die Aufgabe läßt sich offensichtlich nicht in einem Schritt lösen. Eine mögliche Struktur zur Beschreibung des Analyseprozesses ist in Fig. 1-1 dargestellt. Danach unterscheiden wir drei Prozeßebenen:

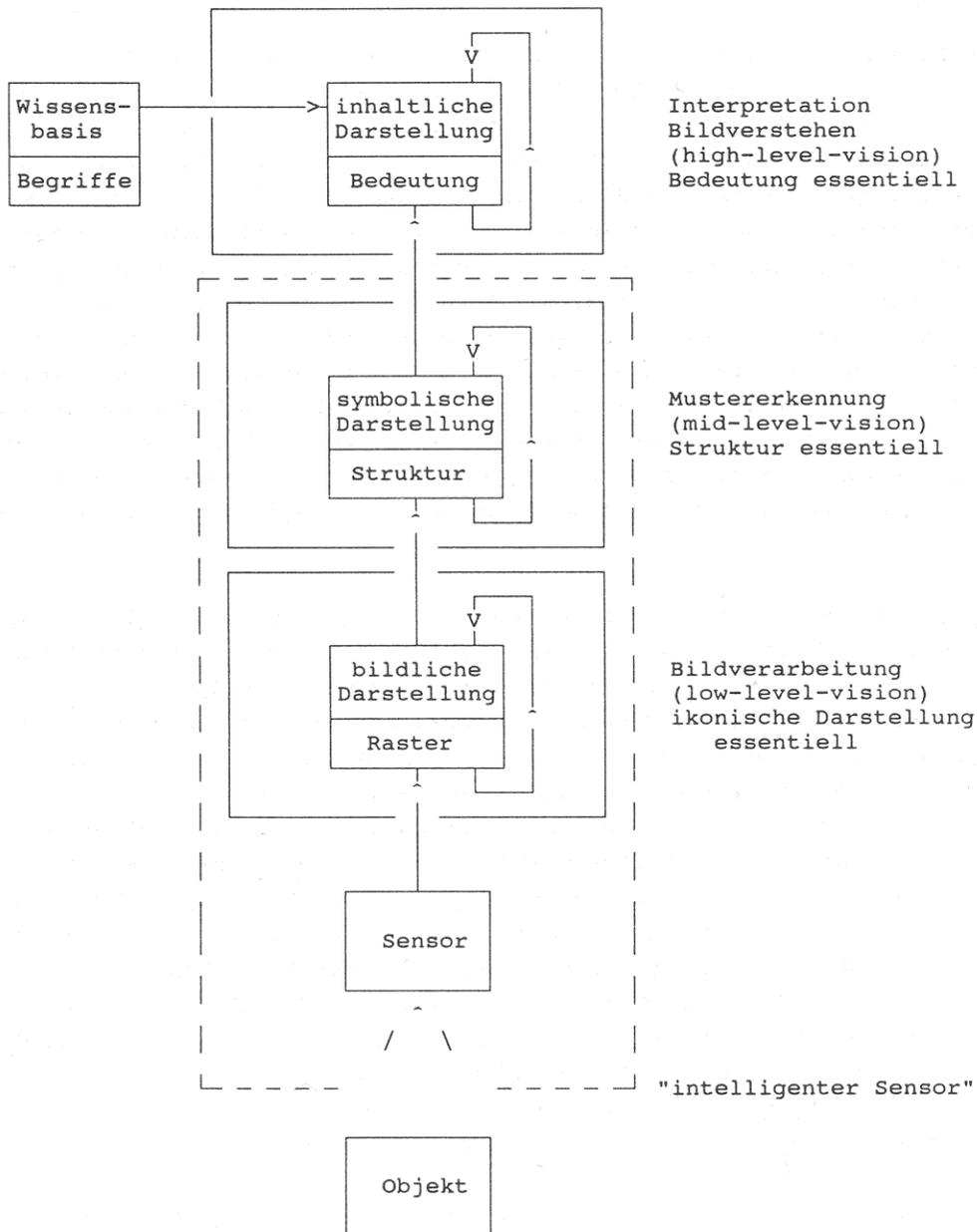


Fig. 1-1 Eine Struktur zur Beschreibung des Prozesses der automatischen Bildanalyse. Sensor, low-level- und mid-level-vision können zu einem "intelligenten Sensor" zusammen gefaßt werden, der die Vorverarbeitung für die Interpretation der Bilder übernimmt

1. In der untersten Stufe der Auswertung geht es um die Vorverarbeitung der Bilder mit dem Ziel, die für die Analyse wesentlichen Bildmerkmale zu extrahieren. Wesentliches Kennzeichen dieses Prozesses ist, daß das Ergebnis wieder eine ikonische Darstellung ist, also die gleiche Rasterform aufweist wie das Bild selbst. Beispiele für diesen, englisch "low-level-vision" genannten Verarbeitungsschritt sind, abgesehen von allen Arten von Bildverbesserungsverfahren, die eventuelle Fehler oder Unzulänglichkeiten des Abbildungsprozesses rückgängig machen, insbesondere Kantenfilter oder Korrelationsverfahren zur Ableitung dichter Disparitäts- oder Parallaxenfelder.

Diese klassischen Aufgaben der Digitalen Bildverarbeitung sind unter allen Schritten der Bildanalyse diejenigen, die derzeit am besten verstanden und theoretisch durchgearbeitet sind und für die es daher auch leistungsfähige Hardware gibt.

2. In der zweiten Stufe der Auswertung geht es um die Extraktion der Bildstrukturen und damit um den wichtigen Schritt der Datenreduktion. Wesentliches Merkmal dieses Schrittes ist, daß das Ergebnis Strukturen, d. h. symbolische Beschreibungen des Bildes sind, also z. B. Listen, Graphen oder Relationen. Typische Aufgaben dieser Stufe (engl. "mid-level-vision") sind etwa die Extraktion von markanten Punkten, geraden Kanten, Linien, eine Segmentierung oder die Suche eines Polygons vorgegebener Geometrie. Gegenüber der dritten Stufe spielt die Bedeutung der Strukturen, im Sinne einer Anwendung, noch keine Rolle. Auch kann der Prozess selbst in mehreren Stufen ablaufen, wenn z. B. eine Segmentierung auf einer vorherigen Kantenextraktion beruht oder die Lokalisierung von Rechtecken (für die spätere Interpretation von Gebäuden) auf eine zuvor erstellte Liste anti-paralleler gerader Kanten zurückgreift.

Für diese Aufgaben steht eine Vielzahl von Verfahren der Mustererkennung zur Verfügung.

3. In der obersten Stufe (engl. "high-level-vision") geht es um die Bedeutung der extrahierten Strukturen, also um die Interpretation des Bildinhaltes. Notwendige Voraussetzung für diese Aufgabe ist eine Wissensbasis mit einer symbolischen Beschreibung der Objekte, die für die jeweilige Aufgabe von Interesse sind. Es werden zwar zur inhaltlichen Beschreibung des Bildes die gleichen Strukturen benötigt wie in der vorigen Stufe, die Bedeutung der Elemente ist aber wesentlich für diesen Schritt. Beispiele sind die Erkennung von Fehlern in einem gefertigten Produkt, die Erkennung und Positionierung von Teilen auf einem Fließband oder die Erkennung und Lokalisierung von Gebäuden in einem Luftbild.

Dieser letzte Schritt innerhalb der Bildanalyse fällt in den Aufgabenbereich der Künstlichen Intelligenz. Hier sind die theoretischen Grundlagen für das Design und die Analyse der Methoden noch am wenigsten weit entwickelt.

Wir wollen die Problemkreise bei der Bildanalyse in diese Struktur einzuordnen, sie aber später selbst - als weitere Ursache der Schwierigkeit bei der Automatisierung des Auswerteprozesses - kritisieren. Die Analyse von Daten in der Meßtechnik soll uns im folgenden zum Vergleich dienen.

a. Das erste Problem wird durch das i. a. niedrige Signal-Rausch-Verhältnis der Bildsignale verursacht. Gegenüber heute ausgereiften Meßtechniken ist die Relativgenauigkeit der gemessenen Intensitätswerte bestenfalls 1:100 bis 1:1000, im Durchschnitt 1:10 und kann in ungünstigen Fällen um 1:1 oder gar darunter liegen. Das drückt sich in der bei digitalen Bildern üblichen bit-Zahl von 6-8 pro Grauwert aus. Das niedrige Signal-Rausch-Verhältnis weist auf physikalischen Grenzen des Meßprozesses selbst hin, obwohl hier im Einzelfall sicher eine deutliche Verbesserung möglich wäre, kann aber auch als Grenze der Genauigkeit der Modelle interpretiert werden.

Die statistischen Eigenschaften der Ausgangsdaten, etwa der Intensitätswerte, müssen daher untersucht und unter allen Umständen bei der Bildauswertung berücksichtigt werden. Dies beginnt schon etwa bei der Überprüfung, ob linearisierte Modelle ausreichen. Die Auswirkung der stochastischen Eigenschaften auf die Interpretation muß analysiert werden, um eine Qualitätsaussage über das Ergebnis zu ermöglichen. Da beim Übergang auf die symbolische Ebene, etwa durch Schwellwertoperationen, Entscheidungen getroffen werden, die einen Informationsverlust bedeuten, und alle folgenden Entscheidungen beeinflussen, sollte diesem Übergang besondere Aufmerksamkeit bei der Entwicklung gewidmet werden.

Dies ist zunächst nur eines unter mehreren Problemen bei der Bildanalyse, bildet aber den Kern der vorliegenden Arbeit.

- b. Gegenüber klassischen Meßaufgaben haben wir es in der Bildauswertung mit Datenmengen zu tun, die um 3-4 Größenordnungen, bei vergleichsweise umfangreichen Aufgaben, ohne weiteres um 6-8 Größenordnungen höher liegen können, wenn man etwa an die Auswertung von Bildfolgen denkt. Auch wenn die Entwicklung leistungsfähiger Rechner uns in die Lage versetzte, eine Million mal schneller zu rechnen als heute, würde Parallelität das Problem der Bildanalyse nicht lösen. Das wird eindrücklich von TSOTSOS (1987) belegt. Er zeigt, daß nur eine Struktur, wie man sie beim menschlichen Sehvorgang vermutet, geeignet ist, die Komplexität der Bildinterpretation auf ein realistisches Maß zu reduzieren. Neben einer Hierarchie der Bildmerkmale, die einer Bildpyramide (vgl. z. B. ZIMMER 1986) entsprechen, und einer hierarchischen Ordnung der zu erkennenden Objekte ist dafür vor allem der möglichst frühe Übergang auf eine symbolische Repräsentation des Bildes auf der Basis weniger Merkmale notwendig.

Aus diesen Gründen sollte man digitale Bilder allein zur Ableitung von Meßwerten nicht verwenden, sondern nur dann einsetzen, wenn gleichzeitig der Bildinhalt von Interesse ist, die Gewinnung der Meßwerte also nicht im Vordergrund steht.

- c. Ein weiteres für die praktische Anwendung wie für die theoretische Durchdringung des Bildanalyseprozesses nachteiliges Kennzeichen ist die Vielzahl der, selbst für die gleiche Detailaufgabe, vorhandenen Verfahren, typisches Indiz für das Frühstadium einer sich rasch entwickelnden Wissenschaft. Zudem fehlt für viele Verfahren eine Methode zur Beurteilung ihrer Leistungsfähigkeit, geschweige ein für alle Methoden gemeinsames Modell, das einen Vergleich gestatten würde (vgl. ALOIMONOS et. al. 1987, BARNARD und THOMPSON 1980, BURNS et. al. 1986). Aber selbst, wenn die Verfahren im einzelnen auf einer starken Theorie aufgebaut sind, fehlt oft der Bezug zu den nächsten Schritten in der Abstraktionshierarchie der Interpretation, wie etwa bei BLOSTEIN und HUANG (1987). Eine Ausnahme hiervon bildet etwa das Modell von WITKIN et. al. (1987), die ein hierarchisches Verfahren zur Zuordnung beliebig vieler Signale beliebiger Dimension geschlossen darstellen und überzeugende Beispiele für seine Anwendung geben.

Die Schwäche fast aller Verfahren liegt schon beim Übergang von der ikonischen zur symbolischen Darstellung, da die statistischen Eigenschaften der symbolischen Repräsentation nicht oder nur sehr unzulänglich aus den statistischen Eigenschaften der ikonischen Bildmerkmale abgeleitet werden können. Dies ist aber - wie oben schon bemerkt - entscheidend für die Beurteilung des Gesamtergebnisses. Am besten noch ist die Situation bei der Extraktion von Kantenelementen oder Punkten (vgl. HARALICK 1984, TORRE und POGGIO 1986, DORST und DUIN 1984. FÖRSTNER 1986b) und bei der Zuordnung kleiner Bildflächen durch kleinste Quadrate Anpassung der Grauwertfunktionen (vgl. FÖRSTNER 1982, 1984), wo z. T. auch der Einfluß der Vorverarbeitung empirisch und theoretisch bekannt ist. Über die empirische oder theoretische Genauigkeit etwa von Linien oder gar die Qualität von Segmentierungen ist so gut wie nichts bekannt.

d. Trotzdem sind die meisten Modelle für eine theoretische Analyse unvollständig oder unzulänglich. Denn sie greifen auf ein kontinuierliches Modell für die Bildfunktion zurück, wie etwa CANNY 1986, FÖRSTNER 1982/1984, TORRE und POGGIO 1986. Dafür gibt es gute Gründe: die Kontinuität des abgebildeten Objektes und die Leistungsfähigkeit der linearen Systemtheorie zur Beschreibung des Einflusses von Filtern auf die statistischen Eigenschaften von Schätzgrößen, die aus den Signalen abgeleitet sind. Viele Ergebnisse, die aus kontinuierlichen Modellen gewonnen sind, lassen sich etwa über eine Finite Elemente Approximation auf die diskrete Realisierung im Rechner übertragen, vor allem wenn die Rundungsfehler bei der Quantisierung unterhalb des Rauschpegels liegen.

Grenzen dieser Modelle sind aber einerseits die in keiner Weise intuitive Repräsentation einfacher Funktionen im Spektralbereich, wie etwa der Sprungfunktion. Dies spiegelt sich mathematisch in der Unmöglichkeit endlicher und gleichzeitig bandbegrenzter Signale wider. Andererseits ist die Modellierung der Quantisierung, etwa als weißes Rauschen, bei langwelligen Signalen fragwürdig, da benachbarte Rundungsfehler stark korreliert sind, aber eigentlich eine stochastische Modellierung nicht trifft.

Z. B. läßt sich die Unsicherheit, die bei der Lokalisierung einer geraden Kante in einem rauschfreien Binärbild besteht, mit Hilfe der linearen Systemtheorie wohl kaum adäquat beschreiben. Die dafür von DORST und DUIN (1984) entwickelte Spirographentheorie verdeutlicht, daß dafür starke Konzepte aus der Zahlentheorie benötigt werden. Sie läßt aber, und dies ist m. E. symptomatisch, keine unmittelbare Verallgemeinerung auf verrauschte Daten zu. Der Vergleich läßt sich leicht kritisieren, da bereits bei der Binärisierung des Signals die wesentliche Information verloren ging. Die Kritik trifft aber nicht den Kernpunkt, den das Beispiel, als Extrem gewählt, veranschaulichen sollte: Es fehlen mathematische Hilfsmittel, die Operationen auf stochastischen, diskreten, ganzzahligen Funktionen analysieren helfen und ähnlich durchsichtig und elegant zu handhaben sind wie die Methoden der Analysis.

Die beiden folgenden Problemkreise sind für die automatische Bildanalyse spezifisch und treten bei der Auswertung von Daten in der Meßtechnik deshalb nicht so offen zutage, da dort der Experte die Ergebnisse der Berechnungen selbst interpretiert.

e. Die Repräsentation von geometrischen Formen und von Textur ist weitgehend ungelöst. Wie bei den Algorithmen, gibt es zahlreiche, für abgegrenzte Anwendungsbereiche hervorragende Repräsentationsformen, wie etwa Superquadricken (Normalform: $(x/a)^p + (y/b)^q + (z/c)^r = 1$), generalisierte Zylinder (freie Achsen und Querschnitte), Fraktale oder Finite Elemente Darstellungen von Oberflächenformen (vgl. z. B. BAJCSY und SOLINA 1987, PENTLAND 1984, 1987, PONCE et. al. 1987, TERZOPOULOS 1986a, 1987, 1988, KASS et. al. 1987, und BALLARD und BROWN 1983, Kap. 3) oder Spektren, Korrelationsfunktionen, Autoregressive Prozesse, Fraktale oder Strukturelle Modelle zur Repräsentation von Textur (vgl. z. B. HARALICK 1979, ROSENFELD 1980/1, PENTLAND 1984, SZELISKI 1987).

Für einen flexiblen Auswerteprozess sind allerdings meistens mehrere Repräsentationen bereitzuhalten bzw. bei unterschiedlichen Operationen auf den Formen bzw. Texturen ggf. von einer Repräsentation auf eine andere überzugehen. Auch hier müßten zusätzlich die statistischen Eigenschaften, d. h. die Unsicherheit der Parametrisierung oder eventuelle Transformationen mit in die Repräsentation und die Operationen auf bzw. zwischen den Formen oder Texturen einbezogen werden, etwa bei der Analyse der Topologie des Schnitts zweier Flächen oder der Erkennung einer Texturgrenze. Daneben ist nicht allgemein geklärt, in welcher Struktur die symbolische Beschreibung der Bilder in der mittleren Ebene repräsentiert werden sollen, um innerhalb einer bestimmten Klasse von Aufgaben kurze Zugriffszei-

ten zu erhalten. Hier berühren sich Methoden aus der Statistik, dem rechnergestützten Entwerfen und der Informatik.

- f. Schließlich ist eine Theorie erforderlich, die die Nutzung der Begriffe, die in der Wissensbasis enthalten sind, beurteilen kann und damit erst eine kontrollierte, d. h. auch in ihrer Leistungsfähigkeit vorhersehbare automatische Deutung der Bildinhalte ermöglicht. Hier entsteht die grundsätzliche Frage, ob ein Rechner Begriffe wie etwa "Wald" oder "Stuhl" speichern und richtig nutzen kann, ohne die Erfahrung etwa des Gehens, Riechens oder Sitzens. Von einer solchen Theorie sind wir wohl noch weit entfernt.

Aus der Sicht einer Ingenieurwissenschaft mögen solche Fragen von untergeordneter Bedeutung sein. Partikuläre Lösungen, etwa aus der Farbe allein auf den Begriff zu schließen, umgehen gezielt diese Schwierigkeiten und machen den schnellen Erfolg spezieller Bildanalyse-systeme verständlich. Diese kennen aber selten ihre eigenen Begrenzungen und reagieren daher in unvorhergesehenen Situationen u. U. unkontrolliert bzw. falsch, verlangen daher eine Interaktion mit dem Spezialisten und sind deshalb nur bedingt für die Automatisierung geeignet.

Nach dieser Übersicht über die Probleme der Einzelschritte bei der automatischen Bildanalyse wollen wir nun die lineare Struktur des Auswerteprozesses, wie er in Fig. 1-1 dargestellt ist, und der uns bei der begrifflichen Einordnung der Einzelschritte nützlich war, selbst kritisieren.

- g. Bei der Realisierung eines Systems ist selbstverständlich eine klare Zuordnung der Einzelprozesse zu den drei Stufen nicht immer eindeutig möglich. Sie greifen aus praktischen Gründen, etwa der Effizienz, und aus prinzipiellen Gründen, etwa der Abhängigkeit untereinander, über die gezogenen Grenzlinien hinweg. Umgekehrt sind die Prozesse, wie bei der Diskussion der mittleren Ebene schon angesprochen, in den einzelnen Stufen wesentlich stärker gegliedert. Dies könnte man als Modifikationen abtun, denn auch ein Zugriff auf die unterste Ebene während der Interpretationsphase würde die Struktur nicht wesentlich verändern.

- h. Entscheidende Schwäche des Schemas ist die Einseitigkeit des Informationsflusses. "Spontanes Sehen" (pre-attentive vision), bei dem ein sofortiger Eindruck über den Bildinhalt entsteht, könnte zwar mit dieser Struktur modelliert werden. Eine zielorientierte Interpretation, die etwa die Erkennung und Anordnung von Objekten aus einer Wissensbasis zum Ziel hat, benötigt Information und Kontrollfluß in mehreren Richtungen.

Dies kann z. B. auf die Struktur eines Musteranalyse-systems führen, wie sie NIEMANN (1981) skizziert hat.

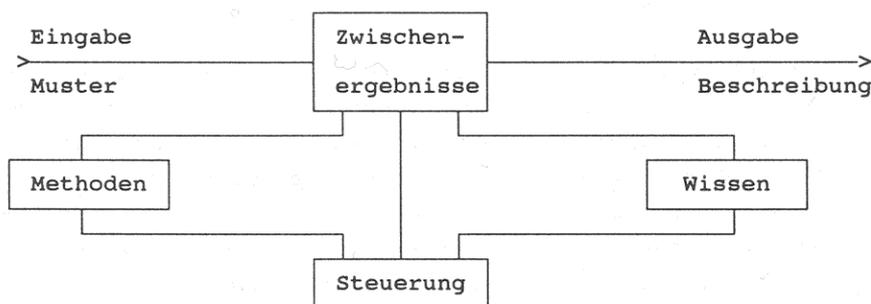


Fig. 1-2 Komponenten eines Systems zur Analyse von Mustern nach NIEMANN (1981)

Danach kann das Kontroll- oder Steuermodul in Abhängigkeit von dem Wissen, das auf Datenträger zur Verfügung gestellt oder auf Grund der Analyse bisheriger Interpretationen gewonnenen wurde, mit den Methoden, die in einer Bibliothek bereitstehen, seine Aufgaben datenabhängig, etwa regelgesteuert durchführen. Die Freiheit bei dieser Konzeption liegt dabei in den Regeln, in der Flexibilität des Einsatzes von Algorithmen und in dem im Prinzip frei wählbaren Zugriff auf Ergebnisse von Zwischenschritten. Bei realisierten Systemen, etwa zur Segmentation von Bildern (vgl. 1-2) etwa NAZIF und LEVINE 1984) oder zur Analyse des linken Herzventrikels (NIEMANN et. al. 1985), die regelgesteuert arbeiten, weisen sowohl die Struktur der intermediären Bildrepräsentationen als auch die Struktur der Regeln selbst eine der Fig. 1-1 verwandte Struktur auf. Allerdings stehen in den einzelnen Auswertestufen Repräsentationsformen verschiedenen Typs nebeneinander, so daß gegebenenfalls auch ein Rückgriff auf die Originaldaten zu einem späteren Zeitpunkt der Analyse möglich bleibt, wie etwa bei dem System von FUA und HANSON (1987) zur Erkennung von Gebäuden in Luftbildern, bei dem nach einer vorläufigen Analyse des Bildes mit einem anderen Verfahren die Segmentierung verfeinert wird.

Insgesamt legen die Darlegungen den Mangel an einer Theorie der automatischen Bildanalyse offen (vgl. HARALICK 1985). Trotz hervorragender einzelner theoretischer und praktischer Ergebnisse fehlt bisher ein Rahmen zur Verknüpfung der Verfahren und zur Beurteilung der Einzelschritte in Bezug auf die Qualität der Interpretationsergebnisse.

1.3 Zielsetzung der Arbeit

Die Arbeit möchte exemplarisch Wege zur Vereinheitlichung der theoretischen Grundlagen der automatischen Bildanalyse zeigen. Auch wenn eine einheitliche Theorie, die alle Aspekte der Bildanalyse umfaßt, wohl noch auf sich warten läßt, so ist doch für die Anwendung der Verfahren jeder Schritt auf eine theoretische Durchdringung der Verfahren von unmittelbarem Vorteil, selbst wenn sie in Einzelfällen ihre Aufgabe noch nicht lösen. PREGIBON (1986) gibt im Zusammenhang mit der Entwicklung von Strategien ein Kriterium dafür an, wann man mit der Entwicklung von Verfahren aufhören kann, und die auch hier Gültigkeit haben:

- a. Das Verfahren kann $XX\%$ aller neuen Fälle lösen,
- b. In den $(100-XX)\%$ anderen Fällen, die es nicht löst, weiß es dies
- c. man ist mit dem Prozentsatz XX zufrieden.

Die unter b. angesprochene Selbstdiagnose scheint daher unerlässlich beim Entwurf von Verfahren und kann nur theoretisch, d. h. auf der Basis von Modellen, entwickelt werden. Sie kann zunächst, wenn anders nicht möglich, für verschiedene Verfahren unterschiedlich sein. In einem System wie der Bildanalyse mit mehreren Einzelkomponenten, dessen Gesamtleistung im Sinne einer Selbstdiagnose zu beurteilen ist, benötigt man aber, wenn nicht einheitliche Modelle für die Einzelverfahren, so doch Modelle mit einer einheitlichen Sprache zur Beschreibung der Qualität ihrer Ergebnisse.

Gemeinsam ist allen im folgenden diskutierten Verfahren die Nutzung der Wahrscheinlichkeitstheorie und der Statistik, einschließlich der daraus von SHANNON (1949) entwickelten Informationstheorie. M. E. sind dies die einzigen Hilfsmittel, die uns heute für eine Qualitätsbeurteilung zur Verfügung stehen. Denn sie erlauben eine konsistente Beschreibung zufälliger oder scheinbar zufälliger Komponenten in den Experimenten, besitzen mit der Wahrscheinlichkeitstheorie ein Instrumentarium zur Vorhersage der Qualität der Verfahren und sind mit der Statistik in der Lage, die Modelle auf der Basis von Experimenten zu prüfen und zu verfeinern. Sie sind damit anderen Techniken, die etwa mit Toleranzen, Maximalfehlern oder Kenngrößen aus der

Approximationstheorie arbeiten, überlegen bzw. enthalten diese Techniken meist als Spezialfälle. Selbst wenn der mathematische Formalismus dazu benutzt wird, um Wahrscheinlichkeiten als Maß für den Grad der Überzeugung oder der Sicherheit zu verwenden, also nicht im Sinne einer relativen Häufigkeit, bietet das wahrscheinlichkeitstheoretische Konzept wegen seiner Geschlossenheit Vorteile (vgl. CHEESEMAN 1985). In unserem Zusammenhang ist die Informationstheorie als wichtiges Bindeglied zwischen symbolischen und numerischen Bildbeschreibungen von besonderem Interesse.

Unter diesen Gesichtspunkten sind nun unter den oben genannten Problemkreisen drei von besonderer Bedeutung:

1. Die Verschmelzung von Bildanalyseverfahren unterschiedlicher theoretischer Konzeption.
2. Die Mitführung der Unsicherheit der Bildrepräsentation beim Übergang von der ikonischen zu einer symbolischen Darstellung.
3. Die Bewertung numerischer und symbolischer Attribute auf der Basis eines einheitlichen Modells.

Als Beispiel diskutieren wir Querverbindungen zwischen der Interpretation elementarer Texturmerkmale und der Extraktion von markanten Punkten und Kanten. Dabei gehen wir ausführlich darauf ein, wie man aus den Verfahren selbst abschätzen kann, mit welcher Unsicherheit sich markante Punkte und gerade Kantenstücke aus der Bildfunktion ableiten lassen. Schließlich behandeln wir die bei der Bildzuordnung und Objekterkennung wichtige Bemessung der Ähnlichkeit, der Seltenheit und der Invarianz von numerischen und symbolischen Attributen

Die gewählten Anwendungsbeispiele sind Verfahrensschritte aus der Zuordnung von Bildern bzw. aus der Objekterkennung, und waren Anlaß für diese Arbeit. Da sie noch kein auch nur vorläufig abgeschlossenes System bilden, sollen sie an Hand des Schemas der Fig. 1-3 eingeordnet werden.

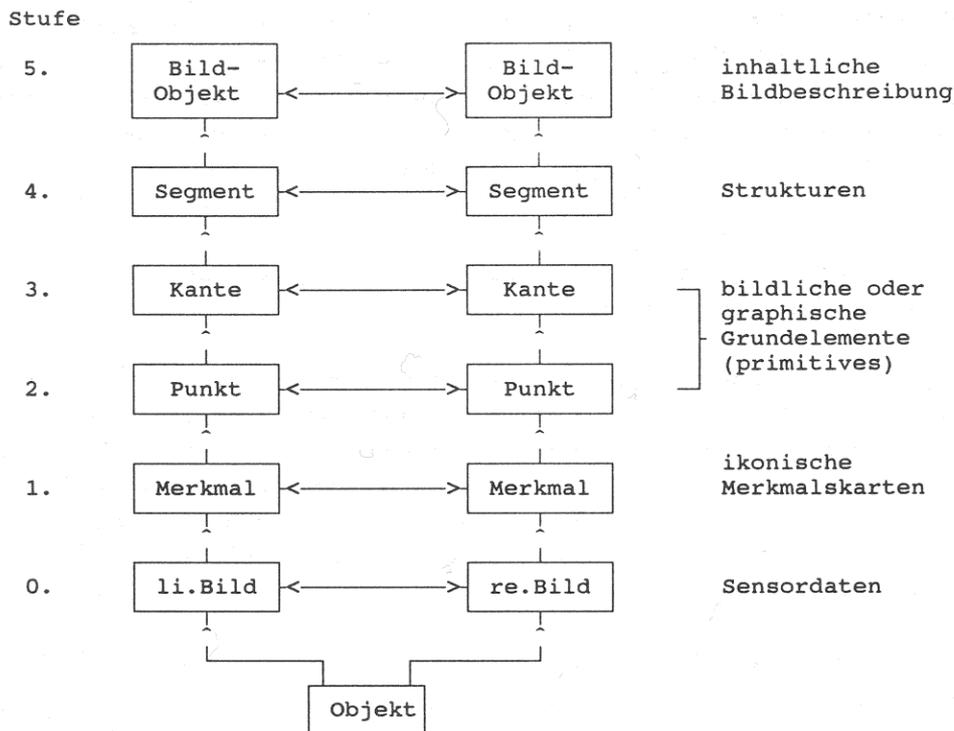


Fig. 1-3 Schema für die Bildzuordnung und die Objekterkennung

Der rechte vertikale Ast kann auch ab Stufe 2 eine Karte oder ein Objektmodell enthalten

Fig. 1-3 stellt das Grundmuster der meisten Zuordnungsverfahren dar. Die Zuordnung bezieht sich dabei entweder auf ein Bildpaar, etwa zum Zweck der Rekonstruktion eines dreidimensionalen Objektes. Es kann sich dabei aber auch um die Zuordnung eines Bildes zu einer Karte, etwa für die Navigation, oder zu einem Objektmodell für die Objekterkennung und -positionierung, etwa bei der Zielverfolgung, handeln. Wir können zwei Aufgaben unterscheiden: die Herstellung einer Bildbeschreibung gewählter Abstraktionsstufe (in vertikaler Richtung) und die Zuordnung der so gewonnenen Beschreibungen (in horizontaler Richtung). Die Bildzuordnung kann in jeder Abstraktionsstufe erfolgen, sich etwa - wie bei der klassischen Korrelation unmittelbar auf die gemessenen Intensitätswerte beziehen, oder auf daraus abgeleitete "Merkmalskarten" (nach TSOTSOS 1987, vgl. Abschnitt 2.1), d. h. auf das Ergebnis eines Bildverarbeitungsprozesses, der etwa ein Kantenbild erzeugt. Aber auch abstraktere Bildbeschreibungen, wie in der Figur angedeutet, können für die Zuordnung verwendet werden. Wesentlich ist, daß die Beschreibungen der Bilder, der Karte oder des Objektmodells zugeordnet werden. Dagegen ist unerheblich, ob die Herleitung der Beschreibung auf beiden Seiten des Prozesses parallel abläuft, oder die Beschreibung des rechten Bildes, der Karte oder des Objektmodells für die Analyse des linken Bildes genutzt wird.

In der vorliegenden Arbeit werden eine Reihe von Verfahren behandelt, die bei der Bildzuordnung und der Objekterkennung auftreten und die sich in das Schema der Fig. 1-3 einordnen lassen.

Kapitel 2 behandelt Fragen der Einbildanalyse. Insbesondere geht es um die Bestimmung von Texturmerkmalen, ein neues informationserhaltendes Filter und um die Extraktion von markanten Punkten und Kantenelementen. Wir werden zeigen, daß diese Aufgaben sich in einem einheitlichen Modell für die Bildfunktion formulieren lassen und die vorgeschlagenen Lösungen bekannten Verfahren vergleichbar oder überlegen sind. Darüberhinaus lassen sich für diese Primitive sowie für daraus abgeleitete Elemente, wie gerade Kantenstücke, die Genauigkeitseigenschaften aus den Extraktionsverfahren selbst ableiten.

Kapitel 3 befaßt sich mit dem Problem eines für numerische und symbolische Merkmale geeigneten Ähnlichkeitsmaßes. Mit Hilfe informationstheoretischer Konzepte gelingt es, insbesondere für die Invarianz und die Seltenheit oder Auffälligkeit von Merkmalen ein gemeinsames Maß zu finden. Die dabei anliegenden Querverbindungen zur Theorie der Trennbarkeit in der Parameterschätzung (FÖRSTNER 1983) zeigen, daß auch die in Kapitel 2.2 vorgeschlagene Selektion markanter Punkte sich informationstheoretisch begründen läßt, womit eine Verbindung zwischen horizontalen und vertikalen Prozessen bei der Bildzuordnung hergestellt wird.

Insgesamt soll die Arbeit einen ersten Schritt zur Vereinheitlichung der Qualitätsbeurteilung der einzelnen Verfahrensschritte in der Bildanalyse darstellen. Die Repräsentation der Unsicherheit der symbolischen Bildbeschreibungen durch die zweiten Momente der Verteilung der beteiligten geometrischen Grundelemente, die aus den Analyseschritten selbst ableitbar ist, erlaubt eine Verbindung der Verfahren der mathematischen Statistik mit den heuristischen Verfahren der Künstlichen Intelligenz. Erste Anwendungen finden sich bei der Suche von groben Fehlern (SARJAKOSKI 1986) oder der Objektlokalisierung (FÖRSTNER 1988).

2. VERFAHREN DER EINBILDANALYSE

Das vorliegende Kapitel behandelt zwei wichtige Fragestellungen der Analyse einzelner Bilder: die Filterung der beobachteten Bildfunktion zur Elimination des Rauschens und die Extraktion von markanten Punkten und Kanten. Gemeinsam ist beiden Aufgaben die Nutzung der gleichen Texturmerkmale: Texturstärke, Texturanisotropie und Texturrichtung.

Das adaptive Filter greift auf diese Texturmerkmale zu und kann so in Gebieten mit flacher Bildfunktion stark glätten, während es an Kanten nur entlang der Kante und an isolierten Punkten oder Ecken nicht glättet. Es ist ein informationserhaltendes Filter, das für homogene Gebiete ein Wiener Filter darstellt. Es benötigt eine Näherung für die Varianz des Rauschanteils. Die dafür entwickelte Schätzung stützt sich zwar nicht unmittelbar auf das Modell der Bildfunktion, ist aber auch in schwach texturiertem Gebiet robust und genau genug, wie die Analyse und Tests an künstlichen Mustern zeigen.

Die Extraktion von Punkten nutzt die Texturmerkmale, um solche Stellen im Bild zu lokalisieren, bei denen die Textur nicht anisotrop ist und relativ zur Umgebung am stärksten ist. Die so gefundenen Bildfenster kann man weiter analysieren und klassifizieren. Bei der einen Klasse schneiden sich die Gradientenvektoren, bei der anderen die dazu senkrecht stehenden Vektoren in einem Punkt. Typische Repräsentanten für die erste Klasse sind runde Punkte, Kreise oder Ringe, für die zweite Klasse Ecken. Auf entsprechende Weise lassen sich Kantenelemente extrahieren.

Die Texturanalyse und das daraus abgeleitete informationserhaltende Filter (Abschn. 2.1) greifen auf ein stochastisches Modell der Bildfunktion zurück, während die Extraktion von Punkten und Kantenelementen (Abschn. 2.2) auf einem geometrischen Bildmodell beruht. Insofern repräsentieren sie gerade in der Verschiedenheit ihrer Modellvorstellung typische Bildanalyseverfahren. Wir werden sie daher getrennt darstellen und auf die Querverbindungen an geeigneter Stelle hinweisen. Sie zeigen, daß sich die Extraktion von markanten Punkten und von Kantenelementen als Interpretation der Merkmalskarten deuten läßt. Wir werden bei allen Einzelschritten die Frage nach den statistischen Eigenschaften der geschätzten Größen behandeln. Dies ist für ihre kontrollierte Weiterverwendung notwendig und wird abschließend am Beispiel von geraden Kantenstücken (Abschn. 2.2.5), die sich aus Kantenelementen zusammensetzen, diskutiert.

2.1 Vorverarbeitung und Extraktion von Merkmalskarten

Im ersten Schritt der Bildanalyse werden abgeleitete Bilder erzeugt, flächenhaft über das Bild verteilte Merkmale, die lokale Eigenschaften des Bildes darstellen. Wir wollen diese Bilder, die Merkmale enthalten, in Anlehnung an TSOTSOS (1987) als Merkmalskarten bezeichnen. Sie sind i. a., wie auch das ursprüngliche Bild mehrkanalig, also vektorwertige diskrete Funktionen.

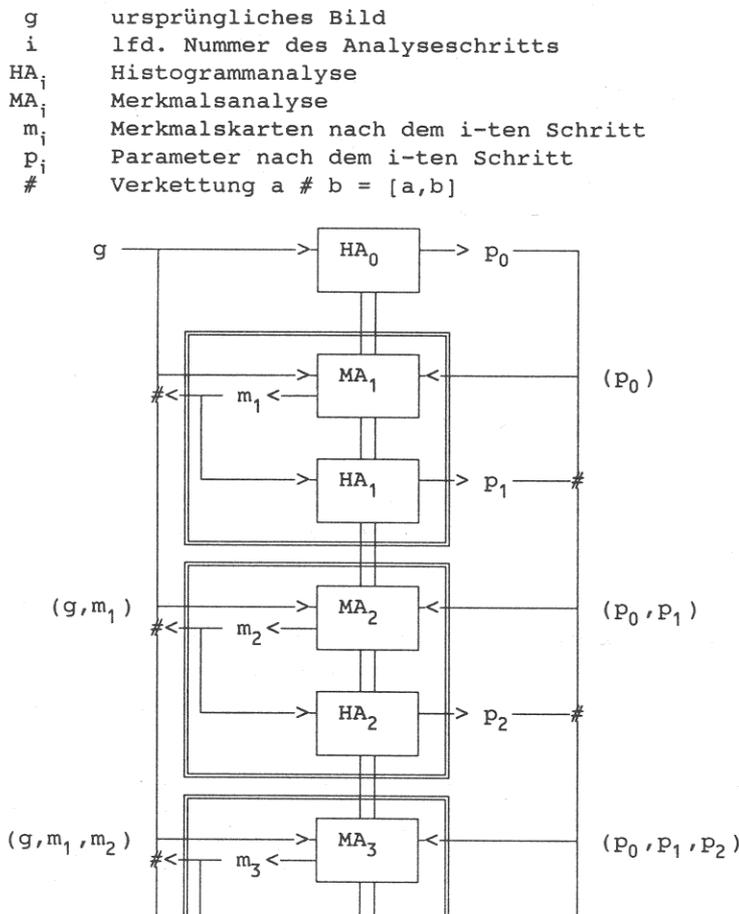
Bei diesem Schritt wird unterstellt, daß unabhängig vom Bildinhalt die extrahierten Merkmale für die Bildanalyse ausreichen und daher auch das Bild bezüglich dieser Merkmale homogen ist. Damit ist sichergestellt, daß dieser erste Prozeß der Merkmalsextraktion parallel bzw. unabhängig sowohl innerhalb des Bildes als auch für mehrere Bilder ablaufen kann. Es ist aber nicht notwendig, daß die Merkmalsextraktion mit einem einzigen Durchgang durch das Bild abgeschlossen ist. Vielmehr werden u. U. zur Steuerung der parallel ablaufenden Prozesse globale, d. h. das ganze Bild betreffende Parameter gebraucht, z. B. aus den Histogrammen der Merkmale ableitbare Kenngrößen, die dann als Schwellwerte dienen.

Damit läßt sich der Prozeß zur Ableitung der Merkmalskarten wie folgt darstellen (vgl. Fig. 2-1):

Ausgehend vom Bild g werden zunächst über die Analyse des Histogramms (HA_0) Parameter der Grauwertverteilung p_0 ermittelt. Das Bild wird dann, als erste Merkmalskarte, zusammen mit den Parametern p_0 einer ersten Merkmalsanalyse (MA_1) zugeführt, aus der die Merkmalskarte m_1 entsteht, deren Histogramm analysiert wird (HA_1) und die Parameter p_1 der Verteilung des ersten Merkmals ergibt. Damit steht ein erweiterter Satz von Merkmalskarten (g, m_1) und ein erweiterter Satz von globalen Parametern (p_0, p_1) zur Verfügung, die im nächsten Schritt der Analyse verwendet werden. Welche Merkmale extrahiert werden und auf welche Weise die globalen Parameter der Verteilung der Merkmale ermittelt werden, nicht notwendigerweise aus dem Histogramm, bleibt offen.

Die modulare Struktur erlaubt, über die Histogrammanalysen globale und über die Analyse bereits abgeleiteter Merkmalskarten lokale Schwellwertoperationen, Punktoperationen, ortsabhängige Filter u. ä. zu realisieren.

Fig. 2-1 Prozeß zur Erstellung von Merkmalskarten



Das Ziel dieses Abschnitts ist es nun, einen Satz von Merkmalen vorzustellen, der für eine Reihe von sehr verschiedenartigen Bildanalyseaufgaben geeignet ist und der sich unter der genannten Struktur aus einer kleinen Umgebung um jedes Bildelement berechnen läßt. Die algorithmische Lösung wird in Abschnitt 2.2.4 dargestellt. Von dem zugrunde gelegten 2-dimensionalen stochastischen Modell für das Bild ausgehend lassen sich mit den vorgeschlagenen Merkmalen klassische Aufgaben der Bildanalyse lösen, wie u. a.:

- Analyse der Textur auf Homogenität und Isotropie
- Analyse von Kanten auf Lage, Orientierung und Glattheit
- Extraktion markanter Punkte
- Schätzung des Rauschanteils in Bildern
- informationserhaltende Filterung von Bildern.

Damit ist diese Gruppe von Bildanalyseaufgaben aufeinander beziehbar, und Beurteilungskriterien der einen Aufgabe lassen sich für die jeweilig anderen heranziehen.

2.1.1 Das Bildmodell

Wir wollen ein stochastisches Modell für das Bild, d. h. die Funktion $f(r,c)$ ($r = \text{row}$, $c = \text{column}$, manchmal auch mit t_1 und t_2 bezeichnet) zugrunde legen, das für die meisten Verfahren zur Ableitung von Merkmalskarten hinreichend allgemein ist. Es stellt in keiner Weise den Anspruch, die Entstehung von Texturen oder den Abbildungsvorgang zu modellieren. Es dient lediglich als Modell zur Beurteilung der verwendeten Algorithmen. Es lehnt sich an die Darstellung von McLURE (1980,1981, p. 259 ff.) an und enthält drei Komponenten:

1. Eine Gebietseinteilung der Bildfläche
2. Eine Textur, hier einen stochastischen Prozess zur Beschreibung der "wahren" Bildfunktion für jedes Gebiet unabhängig und
3. Die Abtastung, modelliert durch einen additiven Rauschprozess.

2.1.1.1 Die Gebietseinteilung

Die Gebietseinteilung der Bildfläche

$$g = \{s^1, s^2, \dots, s^s \mid s^i \cap s^j = \emptyset \text{ für } i \neq j, \bigcup_i s^i = S\} \quad (211-1)$$

Die Gebiete sollen zusammenhängend sein. An die Grenze zwischen den Gebieten werden gewisse Regularitätsanforderungen gestellt. Insbesondere sollen sie stückweise stetig sein.

Segmentierungsalgorithmen wird oft dieses Modell zugrunde gelegt. Dabei wird allerdings meistens zusätzlich postuliert, daß die Prozesse in einzelnen Gebieten voneinander mit hinreichender Sicherheit trennbar sind - eine Forderung, die nicht realistisch ist. Wir wollen hier das Modell auch nur dazu verwenden, um diejenigen Bildpunkte zu kennzeichnen, für die wir uns bei unserer Analyse interessieren: Punkte im Inneren der Gebiete, Randpunkte oder Punkte auf Kanten und Knotenpunkte, an denen mehrere Gebiete zusammenkommen. Bei einer Repräsentation der Gebietseinteilung durch einen planaren Graphen sind dies Punkte in Flächen, auf Kanten und die Knoten des Graphen.

2.1.1.2 Die Textur

Die "wahre" Bildfunktion $\underline{f}(t_1, t_2)$, das Signal, wird innerhalb der Gebiete als schwach stationärer stochastischer Prozeß modelliert:

$$\underline{f}^i = \{ \underline{f}^i(t_1, t_2) \mid (t_1, t_2) \in S^i \}$$

$$\underline{f}^i \sim N(\mu^i, C^i) \quad (211-2)$$

$$E(\underline{f}^i) = \mu^i = \text{const. pro Gebiet}$$

$$\text{Cov}(\underline{f}^i(t_1, t_2), \underline{f}^i(u_1, u_2)) = \text{Cov}(\underline{f}^i(t_1 + \tau_1, t_2 + \tau_2), \underline{f}^i(u_1 + \tau_1, u_2 + \tau_2))$$

für alle $t, u, t + \tau, u + \tau \in S^i$

Die Prozesse \underline{f}^i werden als unabhängig angenommen. Wir wollen weiterhin annehmen, daß die Prozesse ergodisch sind, d. h. die Mittelwerte über Stichproben und über den Ort identisch sind. Dahinter verbirgt sich die Möglichkeit, aus einer Realisierung der Prozesse Informationen über ihre statistischen Eigenschaften zu erhalten. Wir benötigen diese Annahme für die Schätzungen in den folgenden Abschnitten.

Die meisten der im folgenden angegebenen Schätzungen stützen sich auf kleine Bildfenster. Die Forderung nach Stationarität und Homogenität bei der Schätzung kann genau dann fallengelassen werden, wenn Erwartungswert und Kovarianzfunktion langsam innerhalb des Gebietes variieren, so daß die Schätzung durch die geringe Nichtstationarität oder Inhomogenität innerhalb der Fenster nicht wesentlich beeinflußt werden. Wir werden auf diese Verallgemeinerung gelegentlich zurückgreifen, um der Lesbarkeit willen aber das oben angegebene Modell verwenden.

2.1.1.3 Das Abtastrauschen

Das Modell des Abtastprozesses wird extrem einfach angenommen. Wir sehen lediglich ein additives, weißes und, falls es um die Schätzung und Beurteilung geht, Gaußsches Rauschen vor, das unabhängig von der Gebietseinteilung und dem Signal \underline{f} ist:

$$g = \underline{f} + \underline{n}$$

(211-3)

$$\underline{n} \sim N(0, \sigma_n^2)$$

Wir nehmen an, daß eventuelle Unschärfen bei der Abbildung durch einen Restaurierungsprozeß (im wesentlichen) rückgängig gemacht wurden.

In Wirklichkeit sind die Intensitätswerte eines Bildes in erster Näherung Poisson-verteilt und daher die Varianz linear von den Intensitäten g abhängig: $\sigma_g^2 = a + b g$. Denn die Intensitäten sind im wesentlichen proportional zu den aufgenommenen Photonen. Wir wollen dies im folgenden vernachlässigen. Ggf. kann man die Intensitäten durch eine Wurzeltransformation $k_1 + k_2 \sqrt{a + b g} \rightarrow g$, bei freier Wahl von k_1 und k_2 , so modifizieren, daß nach der Transformation g konstante Varianz aufweist. Die Schätzung der Rauschvarianz (vgl. Abschn. 2.1.5.1) muß dann entsprechend modifiziert werden, um gleichzeitig a und b zu bestimmen.

Im folgenden wollen wir zunächst die Schätzung der Modellparameter diskutieren. Dies sind insbesondere der Mittelwert des Rauschprozesses, die Varianz und die wichtigsten Parameter des Signals. Wir werden dabei in umgekehrter Reihenfolge vorgehen. Das hat den Vorteil, daß die Schätzungen in ihrer Darstellung einfacher sind, bzw. auch auf vereinfachende Annahmen zurückgreifen können. Wir werden diese Schätzungen dann für verschiedene Bildanalyseaufgaben verwenden.

2.1.2 Schätzung der Korrelationsfunktion C

Die Schätzung der Korrelationsfunktion aus den Signalen g oder \underline{f} kann auf die Definitionsgleichung zurückgreifen, etwa

$$C_{gg}(\tau_1, \tau_1) = \iint g(t_1, t_2) g(t_1 + \tau_1, t_2 + \tau_1) dt_1 dt_2 \quad (212-1)$$

und nutzt dabei die Annahme der Ergodizität.

Hier und im folgenden unterdrücken wir die Kennzeichnung des betroffenen Gebietes, falls dadurch keine Mißverständnisse entstehen, nehmen also an, daß die Fenster zur Schätzung der Parameter ganz in einem Gebiet liegen.

Für diskrete Werte t_k und τ_k innerhalb eines Fensters F erhält man die Schätzung

$$\hat{C}_{gg}(\tau_1, \tau_2) = \frac{1}{f-1} \sum_F \sum \{g(t_1, t_2) - \hat{\mu}_g\} \cdot \{g(t_1+\tau_1, t_2+\tau_2) - \hat{\mu}_g\} \quad (212-2)$$

wobei $f = |F|$ die Zahl der Summanden angibt und

$$\hat{\mu}_g = \frac{1}{f} \sum_F \sum g(t_1, t_2) \quad (212-3)$$

eine lokale Schätzung des Mittelwertes ist. μ_g ist unverzerrt, falls die verwendeten Werte ganz in einem Gebiet liegen. Für $\tau_1 = \tau_2 = 0$ erhalten wir die Varianz $\hat{\sigma}_g^2$ des Signals g .

Aus einer Schätzung für C_{gg} wollen wir später eine Schätzung für C_{ff} ableiten (vgl. Abschnitt 2.1.4.2). Im nächsten Abschnitt wollen wir nun verschiedene lokale Merkmale aus diesen Schätzungen ableiten.

2.1.3 Merkmalskarten

2.1.3.1 Ikonische Merkmale

Auf der Basis des Bildmodells wollen wir nun einen Minimalatz von ikonischen Merkmalen entwickeln. Wichtigste Forderung an die Merkmale ist ihre Interpretierbarkeit unter möglichst allgemeinen Randbedingungen. Wir werden die Merkmale schrittweise entwickeln und dann ausführlich diskutieren.

Falls wir uns auf die Punkte $(0,0)$, $(0,1)$, $(1,0)$ und $(1,1)$ der Kovarianzfunktion beschränken, erhalten wir 5 lokale Merkmale:

Merkmalsatz A

1. den Mittelwert des Signals
2. die Varianz des Signals
3. Die Kovarianz in Zeilenrichtung zwischen benachbarten Bildpunkten
4. Die Kovarianz in Spaltenrichtung zwischen benachbarten Bildpunkten
5. Die Kovarianz in Richtung der ersten Diagonalen zwischen benachbarten Bildpunkten.

Die ersten beiden Merkmale sind gegenseitig unabhängig, unabhängig von Koordinatensystem und unmittelbar interpretierbar. Dies gilt nicht für die 3 Parameter der Kovarianzfunktion. Wir wollen sie daher umformen. Dabei wird auch deutlich werden, warum wir uns auf diese 3 Parameter der Kovarianzfunktion beschränkt haben.

Die Kovarianzfunktion $C_{gg}^{(i)}(\tau_1, \tau_2) = C(\tau)$ kann im Ursprung durch (vgl. z. B. O'NEILL 1966)

$$C(\tau) = C(0) - \frac{1}{2} \tau' H \tau + O(\tau^4) \quad (213-1)$$

approximiert werden, worin

$$H = H_{gg} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = - \begin{bmatrix} \frac{d^2C(\tau)}{d\tau_1^2} & \frac{d^2C(\tau)}{d\tau_1 d\tau_2} \\ \frac{d^2C(\tau)}{d\tau_2 d\tau_1} & \frac{d^2C(\tau)}{d\tau_2^2} \end{bmatrix} = - \frac{d^2C(\tau)}{d\tau^2} \quad (213-2)$$

die (negative) Hesse-Matrix von $C(\tau)$ am Ursprung ist.

Die Elemente dieser Matrix sind identisch mit (vgl. PAPOULIS 1965, S. 317)

$$\begin{aligned} H_{11} &= E(\underline{g}_{\tau_1}^2) = - C_{\tau_1 \tau_1}(0,0) \\ H_{12} &= E(\underline{g}_{\tau_1} \cdot \underline{g}_{\tau_2}) = - C_{\tau_1 \tau_2}(0,0) \\ H_{22} &= E(\underline{g}_{\tau_2}^2) = - C_{\tau_2 \tau_2}(0,0) \end{aligned} \quad (213-3)$$

worin $\underline{g}_{\tau_k} = (d\underline{g}(\tau)/d\tau_k)$ die partiellen Ableitungen von \underline{g} sind.

Das bedeutet, daß die Kovarianzen $C(1,0)$, $C(0,1)$ und $C(1,1)$ zwischen den Grauwerten \underline{g} auch durch die Varianzen bzw. die Kovarianz der Gradienten der Bildfunktion dargestellt werden können, wir also auch folgenden Merkmalssatz verwenden können:

Merkmalssatz B

1. Der Mittelwert des Signals
2. Die Varianz des Signals
3. Die Varianz der Gradienten in Zeilenrichtung
4. Die Varianz der Gradienten in Spaltenrichtung
5. Die Kovarianz zwischen den Gradienten in Zeilen und Spaltenrichtung

Damit ist auch begründet, warum wir die Kovarianz $C(-1,0)$ in der zweiten Diagonalrichtung nicht benötigen.

Schätzungen für die letzten 3 Merkmale sind ($\tau_1 \rightarrow r, \tau_2 \rightarrow c$)

$$\begin{aligned} \hat{H}_{11} &= \hat{\underline{g}}_{g_r}^2 = \frac{1}{f} \sum_{(r,c) \in F} g_r^2(r,c) \\ \hat{H}_{12} &= \hat{\underline{g}}_{g_r g_c} = \frac{1}{f} \sum_{(r,c) \in F} g_r(r,c) \cdot g_c(r,c) \\ \hat{H}_{22} &= \hat{\underline{g}}_{g_c}^2 = \frac{1}{f} \sum_{(r,c) \in F} g_c^2(r,c) \end{aligned} \quad (213-4)$$

oder mit der Matrix

$$\underline{vq}' = \underline{G}' = \begin{bmatrix} g_{r_1} & g_{r_2} & \dots & g_{r_f} \\ g_{c_1} & g_{c_2} & \dots & g_{c_f} \end{bmatrix}$$

$$\underline{H} = \frac{1}{f} \underline{vq}' \underline{vq} = \frac{1}{f} \underline{G}' \underline{G}$$

Der Nenner f ist identisch mit der Anzahl der Summanden. Die Merkmalssätze A und B sind dann äquivalent, d. h. funktional abhängig, wenn

1. die Gradienten etwa aus (modifizierter Robertsgradient, Gitterweite 1)

$$g_r\left(r+\frac{1}{2}, c+\frac{1}{2}\right) = \frac{1}{2} \sum_{k=0}^1 g(r+1, c+k) - g(r, c+k) \quad (213-5)$$

$$g_c\left(r+\frac{1}{2}, c+\frac{1}{2}\right) = \frac{1}{2} \sum_{k=0}^1 g(r+k, c+1) - g(r+k, c)$$

berechnet werden und

2. die für die Berechnung verwendeten Fenster F die gleichen Elemente enthalten, d. h.

- a. μ und σ_g^2 aus $m_1 \times m_2$ Elementen und
- b. die Varianzen und Kovarianzen der Gradienten aus $(m_1-1) \times (m_2-1)$ Elementen berechnet werden.

Damit wird gewährleistet, daß sich die Merkmale auf die gleiche Stelle, nämlich $(r+1/2, c+1/2)$, beziehen und aus den gleichen Beobachtungen abgeleitet werden.

Die drei Elemente der Hesse-Matrix H sind nicht Koordinatensystem invariant. Aus der Eigenwertzerlegung von H kann man drei invariante Maße ableiten, z. B.:

1. Die Spur der Matrix

$$\text{tr } H = \delta_1 + \delta_2 \quad (213-6)$$

ist als Summe der Eigenwerte invariant gegen Drehungen des Koordinatensystems. Ihr Wert ist identisch mit der zweifachen mittleren Krümmung der Kovarianzfunktion im Ursprung und mit der Gesamtvarianz der Gradienten im Fenster.

Um Unabhängigkeit von der Varianz σ_g^2 zu erreichen, bilden wir das Verhältnis

$$b = \frac{1}{2\pi} \frac{\text{tr } H}{\sigma_g^2} \quad (213-7)$$

Es ist wegen $b_r = \frac{\sigma_r^2}{g_r} / (2\pi \frac{\sigma_g^2}{g})$, der effektiven Bandbreite des Signals in Zeilenrichtung, die gesamte effektive Bandbreite des Signals. Es ist gleichzeitig die zweifache (negative) Krümmung der Korrelationsfunktion $R_{gg}(\tau) = C_{gg}(\tau) / C_{gg}(0)$.

2. Die Richtung des zum größten Eigenwert gehörenden Eigenvektors

$$\Phi = \frac{1}{2} \arctan \frac{2 H_{12}}{H_{11} - H_{22}} \quad (213-8)$$

Der Winkel Φ gibt die Richtung der größten Krümmung von C_{gg} an. Er ist gleichzeitig die Richtung, in der die Varianz der Richtungsableitung von g am größten ist. Falls es sich um ein texturiertes Gebiet handelt und die Eigenwerte von H unterschiedlich sind, ist $\Phi+90^\circ$ die lokale Richtung der Textur.

3. Das Verhältnis der Eigenwerte

$$a = \delta_1 / \delta_2 \quad (213-9)$$

gibt über die Anisotropie des stochastischen Prozesses Auskunft. Es ist gleichzeitig das Verhältnis der größten und der kleinsten Varianz der Richtungsableitungen.

Mit diesen 3, von der Hesse-Matrix abgeleiteten Parametern, stehen uns nun ein Satz von 5 unabhängigen Merkmalen zur Verfügung

Merkmalssatz C:

1. Der Mittelwert des Signals
2. Die Varianz des Signals
3. Die Stärke der Textur
4. Die Hauptrichtung der Textur
5. Die Anisotropie der Textur

2.1.3.2 Beispiele für Merkmalskarten

Die folgenden Beispiele wollen die Merkmalskarten, insbesondere die für die Textur wichtigen Parameter der Hesse-Matrix darstellen. Dabei haben wir als erstes Texturmerkmal statt der Bandbreite immer die Gesamtvarianz der Gradienten, also die Spur der Hesse-Matrix verwendet.

Bilder A2-1 bis A2-6 (s. Anhang 3) zeigen zunächst einen Siemensstern, ein Schachbrett, ein Wellen- und ein Fischgrätenmuster, z. T. mit verschieden starkem Rauschen und - darunter - die zugehörigen Merkmalskarten für Texturstärke, für Anisotropie und, farbig codiert, die Richtung der Textur. Die Merkmale sind aus den Schätzungen 213-4 hervorgegangen, enthalten also noch den Anteil des Rauschens.

- a. Das erste Merkmal ist die Spur der Hesse-Matrix. Sie mißt die Schärfe oder die Dichte der Textur, eine Unterscheidung zwischen beiden ist nicht möglich.

Wie man an dem Schachbrett erkennen kann, gibt dieses Merkmal offensichtlich die Kantenbereiche an. Die weniger kontrastreichen Kanten am Rand des Schachbrettes führen zu einer deutlich schwächeren Ausprägung dieses Merkmals am Rand verglichen mit den inneren Kanten. Weniger deutlich ist auch die gegenüber den Rändern größere Varianz der Gradienten in der Ecke, da hier in der Spur der Hesse-Matrix Gradienten von zwei Richtungen beitragen.

Die Kantenbereiche sind 4 Bildelemente breit, da mit einem 4x4 Boxfilter die Elemente der Hesse Matrix bestimmt wurden. Bessere Ergebnisse würde man mit einem Gaußfilter erreichen, dann wären die Übergänge fließender.

Der Kontrast und die Schärfe der Kanten ist bei Wellen- und Fischgrätenmuster über das ganze Bild konstant. Wegen der gleichen Dichte der Textur beim Fischgrätenmuster weist das erste Merkmal kaum Schwankungen auf. Durch das Box-Filter werden scheinbare Auflösungsdefekte erzeugt. Dagegen zeigt das Wellenmuster eine deutliche Zunahme der Texturdichte zur rechten oberen Ecke hin.

Beim Siemensstern vermischen sich Kontrastschwankungen und Texturdichte. Der Kontrast beim Siemensstern nimmt nach außen hin bei gleicher Schärfe, d. h. in diesem Fall Breite der

Kanten zu, so daß die Werte g_x und g_y linear nach außen hin zunehmen, während die Dichte der Textur linear abnimmt. Da die Gradienten quadratisch in die Spur der Hesse-Matrix eingehen, ist insgesamt ein Zuwachs des ersten Merkmals nach außen hin zu beobachten. Wollte man Dichte und Schärfe der Textur unterscheiden, müßte man ein zusätzliches Merkmal einführen, etwa die Summe der Absolutbeträge der Gradienten, gemittelt über ein kleines Fenster, was in diesem Fall einen konstanten Betrag über das ganze Bild wohl mit Ausnahme des Innern des Sterns - ergeben würde. Wir wollen diese Erweiterung hier nicht weiter verfolgen.

Zum Rand des Siemenssterns werden wegen der Breite der Strahlen die einzelnen Kanten aufgelöst. Ein ähnlicher Effekt ist beim Wellenmuster zu beobachten.

Sowohl beim Schachbrett wie beim Siemensstern ist der Einfluß stärkeren Rauschens auf das Merkmal deutlich als Fluktuationen zu sehen.

Für eine Bildanalyse genügt diese Merkmalskarte offensichtlich nur in Bildern mit texturlosen Gebieten wie dem Schachbrett, wenn man sich auf die Erkennung von Kanten bzw. Gebietsgrenzen beschränkt. Eine Unterscheidung von Gebietsgrenzen, -knoten und -ecken ist allerdings auf der Basis dieses Merkmals nicht möglich.

- b. Das zweite Merkmal, die Anisotropie der Textur, gibt an, inwieweit die Textur eine dominante Richtung aufweist. Wir haben hier statt des Verhältnisses der Eigenwerte eine monotone Funktion

$$p = \left[\frac{a - 1}{a + 1} \right]^2 = \left[\frac{\delta_1 - \delta_2}{\delta_1 + \delta_2} \right]^2 = 1 - \frac{4 \det H}{\text{tr}^2 H} \quad (213-10)$$

dieses Verhältnisses dargestellt. Sie liegt zwischen 0 und 1 und läßt sich ohne Eigenwertberechnung aus der Spur und der Determinante der Hesse-Matrix ableiten. In den Bildern entsprechen helle Gebiete Stellen mit hohen Werten für p , d. h. für Stellen mit stark gerichteter Textur.

Dieses Merkmal ermöglicht uns, in Gebieten mit starker Textur zwischen Kanten und Ecken oder zwischen Linien und Punkten, bzw. linienartigen und punktartigen Texturen zu unterscheiden. Eine Unterscheidung zwischen Linien und Kanten in anisotropen Gebieten, bzw. zwischen Ecken und Punkten bei isotroper Textur ist dagegen auf der Basis dieses Merkmals nicht möglich. Wir werden dieses Problem aber im Abschnitt 2.2 wieder aufgreifen.

Beim Siemensstern wird die stark gerichtete Struktur außerhalb des Zentrums richtig erfaßt. Die mangelnde Auflösung der Strahlen im Zentrum führt zu einer unregelmäßigen Textur, die nicht speziell ausgerichtet ist. Ähnlich gut ist die Erfassung der Anisotropie beim Wellenmuster.

Bei beiden Mustern fällt allerdings auf, daß der Anisotropiewert bei den vertikal bzw. diagonal ausgerichteten Kantenbereichen größer ist als in den anderen Gebieten. Dies ist ein Diskretisierungseffekt, der bei künstlich generierten Mustern besonders deutlich in Erscheinung tritt (vgl. die Kritik an MORAVEC's Operator in DRESCHLER 1981).

Beim Schachbrett kann man deutlich die Eck- und Kantenbereiche unterscheiden. Offensichtlich ist es nicht möglich, Ecken an Gebietsrändern und Gebietsknoten, aber auch untexturierte Gebiete zu unterscheiden.

Die anisotropen Stellen, nämlich an den Stielen des Fischgrätenmusters, treten gegenüber den Gräten deutlich als dunkle Linien hervor.

c. Das dritte Merkmal, die Richtung der Textur, ist nur in rauschfreien bzw. rauscharmen Gebieten mit großem Anisotropiewert nützlich. Da Rauschen den Anisotropiewert verringert, genügt es, für Gebiete, in denen die beiden ersten Merkmale eine Schwelle überschreiten, die Richtung der Textur anzugeben und darzustellen. Im Zentrum des Siemenssterns, in der Mitte oder an den Ecken der Schachbrettfelder, ist der Richtungswinkel nicht bestimmt.

Sowohl beim Siemensstern als auch beim Wellenmuster wird die Richtung der Textur richtig erfaßt, insbesondere die gleitenden Übergänge.

Beim Schachbrett erkennt man deutlich die Unterscheidung der horizontalen (gelben) und der vertikalen (blauen) Kanten. Beim Fischgrätenmuster ergibt sich eine klare Basis für eine Segmentierung allein in Abhängigkeit von der Richtung. Man beachte, daß diese Struktur wieder gerichtet ist, und Teil einer Makrotextur sein kann, die auf einer geringeren Auflösungsstufe auf die gleiche Weise analysiert werden kann (vgl. KASS und WITKIN 1987).

Fig. A2-7 und A2-8 zeigen die drei Merkmalskarten für zwei reale Bilder. Das Spielzeugbild, auf das wir im Abschnitt 2.2 zurückgreifen werden, zeigt uns die für die Deutung des Bildes wichtigen Bildbereiche: im Texturstärkenbild die Kanten, im Anisotropiebild die dunklen Stellen unter den Kantenbereichen mit geringer Anisotropie, die wir später zur Suche optimaler Fenster für markante Punkte wählen, und im Richtungsbild die für die Linienverfolgung, bzw. die Extraktion gerader Kanten wichtige Richtungsinformation. Man beachte, daß im Anisotropiebild die homogenen Flächen der Quader sich deutlich von der gekrümmten Zylinderfläche unterscheiden. Ähnlich klar zeigen die Merkmalskarten des in Fig. A2-8 teilweise abgebildeten Gebäudes in einem Luftbild 1 : 12 000 die für die Erfassung etwa des Daches wichtigen Strukturen. Insbesondere wird die schwächere Bildkante des Giebels durch ihre Anisotropie deutlich hervorgehoben.

Insgesamt zeigen die Beispiele, daß bereits diese drei Merkmale die Textur gut repräsentieren und darüber hinaus in texturarmen, aber gegliederten Bildern Informationen über die Lage und die Art der Gebietseinteilung geben.

2.1.3.3 Vergleich mit anderen Verfahren der lokalen Bildanalyse

Textur als stochastischen Prozess zu modellieren und lokale Schätzungen der Texturparameter zur Bildverbesserung, Segmentierung oder Datenkompression zu verwenden, ist nicht neu. Eine große Klasse von Verfahren modelliert dabei unmittelbar die Bildfunktion. Bei stationären Prozessen werden die Elemente der Kovarianzfunktion, Parameter von Autoregressiven Prozessen, das Leistungsspektrum oder Parameter eines fraktalen Modells zur Charakterisierung benutzt. Sie lassen sich unter bestimmten Bedingungen ineinander überführen, so daß dann ein Vergleich der Repräsentation möglich ist (vgl. HARALICK 1979, ROSENFELD 1980/1). Die Modelle sind, was die Zahl der Parameter angeht, unterschiedlich allgemein. Fraktale bzw. Leistungsspektrum stellen dabei Extreme dar. In unserem Zusammenhang ist die Bedeutung der Parameter von besonderem Interesse. Wir wollen im folgenden unsere Repräsentation mit denjenigen vergleichen, die entweder eine ähnliche Fragestellung verfolgen oder eine ähnliche Repräsentationsform verwenden.

1. Wenn man über den ersten Parameter, die Varianz der Grauwerte, als Texturparameter hinausgeht, erhalten wir eine Gruppe von 3 zusätzlichen Parametern, wie wir es im Abschnitt 2.1.3 diskutiert haben. Dieselben Parameter verwenden u. a. TOU (1980/1), McLURE (1980/1), KASS und WITKIN (1987), LONGUET-HIGGINS (1958, vgl. SCHACHTER 1980/1) und BIGÜN und GRANLUND (1987) für z. T. sehr verschiedene Aufgabenstellungen.

TOU (1980/1) schätzt die Parameter eines ARMA-Prozesses mit Hilfe der Yule-Walker-Gleichungen. Sein Ansatz gilt für eine beliebige Anzahl von Parametern. Wegen des Bezugs zum Bildgitter sind es aber keine rotationsinvarianten Merkmale und so nicht unmittelbar für eine Klassifizierung geeignet.

McLURE (1980/1) zielt auf eine optimale räumliche Abtastung, mit lokal angepaßten Bildelementen. Für isotrope, aber inhomogene Texturen in denen also $H(x,y) = t(x,y) \cdot I_2$ gilt, müssen die Flächen der Bildelemente proportional zu $t^{1/2}(x,y)$ sein. Für anisotrope Texturen ergibt sich eine entsprechende Schlußfolgerung.

KASS und WITKIN (1987) verfolgen neben einer Richtungsanalyse ein ähnliches Ziel, wenn sie die Hesse-Matrix als Spannungsmatrix interpretieren und Linienscharen gleicher Spannung bzw. dazu orthogonale Linienscharen gleicher Spannungsänderung zur Repräsentation gerichteter Texturen vorschlagen.

Schließlich greift SCHACHTER (1980/1) für die Texturmodellierung das Modell auf, das LONGUET-HIGGINS (1958) für die Modellierung von Ozean-Wellen vorschlug. Dort werden die Momente des Leistungsspektrums zur Charakterisierung der Wellentextur verwendet. Die zweiten Momente sind aber gerade mit den Elementen der Hesse-Matrix identisch, so daß ihre Analyse bzgl. Rauigkeit, Richtung und Anisotropie mit unserer übereinstimmt. Dort werden zusätzlich noch Momente 4. Ordnung verwendet.

2. In allen Fällen kommt der Anisotropie der Textur eine besondere Bedeutung zu.

Nur bei SCHACHTER (1980/1) wird das Verhältnis der Eigenwerte der Hesse-Matrix als Anisotropie bzw. "Langgekämtheit" (long-crestedness) des Wellenmusters interpretiert.

KASS und WITKIN (1987) bzw. BIGÜN und GRANLUND (1987) verwenden die Anisotropie, zusammen mit der Texturstärke lediglich als Maß für die Qualität der Richtungen. Tatsächlich ist die Varianz des Richtungswinkels genähert proportional zu $\sigma_n^2 / (\delta_1 - \delta_2) = \sigma_n^2 / \delta_1 / (1 - \delta_2 / \delta_1)$. Bei genähert isotroper Textur ist die Anisotropie, bei anisotroper Textur ihre Stärke entscheidend für die Genauigkeit der Richtung, womit die "certainty-measures" Stärke und Anisotropie begründet sind. Sie lassen sich aber offensichtlich zur Beurteilung der Genauigkeit der Richtung integrieren.

Entscheidend ist aber die Eignung der Anisotropie für die Klassifizierung, wie die Beispiele und die Anwendung im nächsten Kapitel zeigen.

3. Die Schätzung von Texturparametern muß sich definitionsgemäß immer auf ein Fenster stützen, das für die Erfassung der Textur genügend groß ist. Die vorgeschlagene Schätzung der Hesse-Matrix, wie sie auch KASS und WITKIN bzw. BIGÜN und GRANLUND angeben, hat gegenüber anderen Verfahren zur Approximation der Kovarianzfunktion den Vorteil, daß bereits mit sehr kleinen Fenstern zuverlässige Schätzungen möglich sind. Wir verwendeten in unseren Beispielen immer 5x5 Fenster. Alle übrigen Verfahren benötigen entweder größere Fenster oder schätzen die Parameter nicht unmittelbar aus der Bildfunktion. Sie erreichen damit nur eine geringere Auflösung der Textur oder sind entsprechend ungenauer. Auch erhält man aus den Elementen der Hesse-Matrix ohne Umwege rotationsinvariante Merkmale.

Eine Alternative zur Schätzung von Texturparametern aus den Gradienten wurde von PEET und SAHOTA (1985) angegeben. Sie verwenden die Krümmung der Bildfunktion selbst zur Charakterisierung der Textur. Allerdings extrahieren sie Parameter, die nicht unmittelbar vergleichbar mit unseren sind, obwohl dies möglich wäre (vgl. auch FRITSCH und DÜSEDAU 1987).

Die Schätzung der Elemente der Taylorentwicklung der Kovarianzfunktion aus Momenten des Leistungsspektrums, wie wir es für das erste Glied gezeigt haben, läßt sich auf höhere Momente ausdehnen. Etwa könnte man zusätzlich wie aus den Gradienten, aus den Krümmungen Momente 4. Ordnung bestimmen und daraus rotations- oder auch affininvariante Merkmale extrahieren.

2.1.4 Fehleranalyse und Filterung der Schätzungen

Im folgenden wollen wir die statistischen Eigenschaften der geschätzten Merkmale zusammenstellen und für den Fall verrauschter Daten unverzerrte Schätzungen für die Varianz des Signals und der Hesse-Matrix bereitstellen.

Da für diese Schätzung die Varianz des Rauschanteils benötigt wird, werden wir die in texturierten Bildern erschwerte Schätzung der Rauschvarianz gesondert in Abschnitt 2.1.5 behandeln.

In allen Fällen gehen wir von einem normal-verteilten Zufallsprozeß aus.

2.1.4.1 Verteilung der geschätzten Merkmale

1. Mittelwert des Signals

Wegen $\underline{g} = \underline{f} + \underline{n}$ ist die Schätzung

$$\hat{\underline{\mu}} = \frac{1}{f_m} \sum_F \underline{g}(r, c) \quad (214-1)$$

für den Mittelwert des Signals \underline{f} normal verteilt

$$\hat{\underline{\mu}} \sim N(\underline{\mu}, (\sigma_f^2 + \sigma_n^2)/f_m). \quad (214-2)$$

f_m ist die Zahl der Elemente zur Bestimmung des Mittelwertes.

2. Varianz $\hat{\sigma}_g^2$ des beobachteten Signals \underline{g}

Da $\underline{g} \sim N(\underline{\mu}, \sigma_n^2 + \sigma_f^2) = N(\underline{\mu}, \sigma_g^2)$ ist die normierte quadratische Form

$$\hat{\underline{\Omega}} = (f_m - 1) \cdot \hat{\sigma}_g^2 / \sigma_g^2 = \frac{1}{\sigma_g^2} \sum_F (\underline{g}(r, c) - \hat{\underline{\mu}})^2 \quad (214-3)$$

χ^2 -verteilt mit $f_m - 1$ Freiheitsgraden. Damit ist die Varianz von $\hat{\underline{\Omega}}$

$$V(\hat{\underline{\Omega}}) = 2 \cdot (f_m - 1), \quad (214-4)$$

und daher die Varianz der Schätzung

$$\hat{\sigma}^2 = \frac{1}{f_m - 1} \sum_F (\underline{g}(r, c) - \hat{\underline{\mu}})^2 \quad (214-5)$$

gleich

$$V(\hat{\sigma}_g^2) = \frac{2 \cdot \sigma_g^4}{(f_m - 1)} \quad (214-6)$$

3. Verteilung der geschätzten Hesse-Matrix \hat{H}_{gg}

In diesem Abschnitt diskutieren wir die Verteilung der Hesse-Matrix für das beobachtete Signal \underline{g} , um daraus - analog zur Schätzung der Varianz des Signals - mit Hilfe der Rauschvarianz eine Schätzung für die Hesse-Matrix des unverrauschten Signals \underline{f} abzuleiten. Alle Größen beziehen sich in diesem Abschnitt, wenn nicht anders vermerkt, auf \underline{g} .

Die Verteilung von

$$\hat{V}_{gg} = f_h \cdot \hat{H}_{gg} = \underline{G}'_g \underline{G}_g \quad (214-7)$$

mit

$$\underline{G}_g = \begin{bmatrix} \underline{v}q_1 & \underline{v}q_2 & \dots & \underline{v}q_{f_h} \end{bmatrix}$$

ist unter der Bedingung, daß die Gradienten $\underline{v}q_i$ unabhängig sind, Wishart-verteilt

$$\hat{V} \sim W(f_h, C_{g'g'}) \quad (214-8)$$

d. h. $E(\hat{V}) = f_h \cdot C_{g'g'}$. Hierin ist $C_{g'g'}$ die Kovarianzmatrix der Gradienten

$$C_{g'g'} = D(\underline{v}q) = D \begin{bmatrix} \underline{q}_r \\ \underline{q}_c \end{bmatrix} = \begin{bmatrix} \sigma^2 g_r & \sigma g_r g_c \\ \sigma g_c g_r & \sigma^2 g_c \end{bmatrix} = H_{gg} \quad (214-9)$$

$$= D \begin{bmatrix} \underline{f}_r + \underline{n}_r \\ \underline{f}_c + \underline{n}_c \end{bmatrix} = C_{f'f'} + \sigma_n^2 \cdot I = H_{ff} + C_{n'n'}$$

Wir nehmen dabei an, daß die Rauschteile bei der Schätzung der Gradienten unabhängig sind und gleiche Varianz σ_n^2 besitzen.

Bei der praktischen Berechnung sind allerdings auf dem Bildgitter benachbarte Gradienten wegen gemeinsamer Rauschteile korreliert, so daß die für die Herleitung der Verteilung von \hat{H} notwendige Unabhängigkeit der Summanden (hier \underline{q}_r und \underline{q}_c) nicht erfüllt ist. Wir wollen aber die Eigenschaft der Wishart-Verteilung, Verallgemeinerung der X^2 -Verteilung zu sein, benutzen, um - wenigstens genäherte - Varianzen der Schätzungen zur Verfügung zu haben.

Wegen $C_{g'g'} = C_{f'f'} + \sigma_n^2 \cdot I$ sind die Eigenvektoren von $C_{g'g'}$ und $C_{f'f'}$ identisch und die Eigenwerte unterscheiden sich um σ_n^2 ,

$$\delta_l^g = \delta_l^f + \sigma_n^2, \quad l = 1, 2 \quad (214-10)$$

Daher gilt mit der orthogonalen Matrix U und den Matrizen $D^g = \text{diag}(\delta_l^g)$ bzw. $D^f = \text{diag}(\delta_l^f)$ der Eigenwerte

$$C_{g'g'} = U D^g U' = U (D^f + \sigma_n^2, I) U' = U D^f U' + \sigma_n^2, I \quad (214-11)$$

Daher ist die Diagonalmatrix

$$\hat{K} = U' \hat{V} U = f_h U' \hat{H}_{gg} U \quad (214-12)$$

genähert Wishart-verteilt gemäß

$$\hat{K} \sim W(f, D^f + \sigma_n^2, I) = W(f, \text{diag}(\delta_l^f + \sigma_n^2)) \quad (214-13)$$

Hierin leitet sich die Zahl der Freiheitsgrade aus der Zahl der unabhängigen Komponenten bei der Schätzung von V ab und kann ungefähr mit $f_h/2$ abgeschätzt werden (s. u.).

Dann gilt aber für die Diagonalelemente von $\hat{K} = \text{diag}(\hat{k}_l)$ nach KOCH (1987, p. 163) genähert

$$\hat{k}_l \sim (\delta_l^f + \sigma_n^2) \cdot \chi^2(f) \quad (214-14)$$

woraus wir für die Eigenwerte $\hat{\sigma}_l^g$ der geschätzten Matrix \hat{H}_{gg} schließlich folgende Näherungen erhalten

$$E(\hat{\sigma}_l^g) = \delta_l^f + \sigma_n^2 \quad (214-15)$$

$$V(\hat{\sigma}_l^g) = \frac{2 \cdot (\delta_l^f + \sigma_n^2)^2}{f}$$

2.1.4.2 Filterung der geschätzten Signalvarianz und der Hesse-Matrix

Wir sind nun in der Lage, aus $\hat{\sigma}_g^2$, \hat{H}_{gg} und Schätzungen $\hat{\sigma}_n^2$ bzw. $\hat{\sigma}_f^2$ mit $\sigma_g^2 = \sigma_f^2 + \sigma_n^2$ und $H_{gg} = H_{ff} + \sigma_n^2, I$, Schätzungen für die Varianz σ_f^2 und die Hesse-Matrix der Kovarianzfunktion des nicht beobachtbaren Signals zu erhalten.

Für die Varianz σ_f^2 des Signals ergibt sich unmittelbar die folgende unverzerrte Schätzung

$$\hat{\sigma}_f^2 = \begin{cases} \hat{\sigma}_g^2 - \hat{\sigma}_n^2, & \text{falls } \hat{\sigma}_g^2 > \hat{\sigma}_n^2 \\ 0 & \text{sonst.} \end{cases} \quad (214-16)$$

Die Varianz $V(\hat{\sigma}_f^2)$ ist genähert gleich der Varianz von $\hat{\sigma}_g^2$, wenn $\hat{\sigma}_n^2$ aus dem ganzen Bild abgeleitet wird und deswegen eine gegenüber $\hat{\sigma}_f^2$ vernachlässigbare Varianz aufweist.

Ganz analog ergeben sich unverzerrte Schätzer für die Eigenwerte der Hesse-Matrix \hat{H}_{ff} :

$$\hat{\underline{\delta}}_l^f = \begin{cases} \hat{\underline{\delta}}_l^g - \hat{\sigma}_n^2, & \text{falls } \hat{\underline{\delta}}_l^f > \hat{\sigma}_n^2, \quad l = 1, 2 \\ 0 & \text{sonst.} \end{cases} \quad (214-17)$$

und daraus die Schätzung für die Matrix H_{ff}

$$\hat{H}_{ff} = \hat{C}_{f,f} = U \text{diag}(\hat{\underline{\delta}}_l^f) U' \quad (214-18)$$

Wir werden diese Schätzung für eine informationserhaltende Filterung verwenden.

2.1.5 Informationserhaltende Filterung

In diesem Abschnitt wollen wir eine mögliche Anwendung der Merkmalskarten diskutieren. Es handelt sich um eine Filterung der Bildfunktion, die, gestützt auf die lokal extrahierten Merkmale, das Signal prädiziert. Gegenüber linearen Filtern, bei denen die Filtermatrix über das ganze Bild konstant ist, wird hier in Abhängigkeit von der Textur die Filtermatrix angepaßt. Damit wird erreicht, daß Kanten, Ecken und einzelne Flecken erhalten bleiben und nur in Gebieten mit theoretisch konstanter Bildfunktion stark geglättet wird.

Wir werden das Filter mit dem Medianfilter vergleichen. Da die Unterschiede zum Medianfilter i. a. groß genug sind, genügt zur Beurteilung ein visueller Vergleich der Wirkung der Filter auf verschiedene Testmuster.

Für die Filterung benötigen wir eine zuverlässige Schätzung des Rauschanteils in den Bildern, die wir daher zuerst entwickeln.

2.1.5.1 Schätzung des Rauschanteils

1. Problematik der Schätzung des Rauschanteils in beobachteten Signalen

Bei der Bestimmung der Rauschvarianz handelt es sich streng genommen um eine Varianzkomponentenschätzung. Das beobachtete Signal \underline{g} setzt sich additiv aus den unabhängigen stochastischen Anteilen \underline{f} und \underline{n} mit den Varianzen σ_f^2 und σ_n^2 zusammen. Die Problematik der Schätzung des Rauschanteils aus \underline{g} ohne zusätzliche Modellannahmen wird transparent, wenn wir sie im Spektralbereich durchführen. Dort gilt mit den Fourier-Transformierten oder Amplitudenspektren $\underline{G}(u)$, $\underline{F}(u)$ und $\underline{N}(u)$

$$\underline{G}(u) = \underline{F}(u) + \underline{N}(u). \quad (215-1)$$

Ohne weitere Annahmen über \underline{F} und \underline{N} ist offensichtlich eine Trennung von Signal und Rauschen nicht möglich.

Falls wir aber über die Abhängigkeit der Varianzverhältnisse von $\underline{F}(u)$ bzw. $\underline{N}(u)$ von der Frequenz u Annahmen treffen können, lassen sich die beiden Komponenten trennen. Für die Leistungsspektren gilt in Analogie zu Gl. 215-1

$$P_g(u) = P_f(u) + P_n(u) \quad (215-2)$$

Wir wollen nun zunächst annehmen, daß die Leistungsspektren von \underline{f} und \underline{n} bis auf konstante Faktoren bekannt sind:

$$P_f(u) = Q_f(u) * \sigma_f^2 \tag{215-3}$$

$$P_n(u) = Q_n(u) * \sigma_n^2$$

dann ergibt sich

$$P_g(u) = Q_f(u) * \sigma_f^2 + Q_n(u) * \sigma_n^2 \tag{215-4}$$

Die Fourier-Transformation bewirkt nun zweierlei:

1. Falls f und n normalverteilt sind, sind, wegen der Linearität der Fourier-Transformation auch F und N und damit auch G normal verteilt. Real- und Imaginäranteile sind gegenseitig unabhängige normalverteilte Zufallsvariable mit gleicher Varianz. Daher sind Betrag und Phase ebenfalls unabhängig und Rayleigh- bzw. gleichverteilt.
2. Die einzelnen Frequenzanteile sind gegenseitig unabhängig, falls f und n periodisch angenommen werden können (vgl. PAPOULIS 1965, S. 368).

Bei beobachteten $g(x)$, und damit ableitbarem $P_g(u)$ ($u = 0, 1, \dots$) lassen sich deshalb genau dann Kleinste Quadrate Schätzungen für die Varianzfaktoren σ_f^2 und σ_n^2 schätzen, wenn die bekannt angenommenen Funktionen $Q_f(u)$ und $Q_n(u)$ nicht proportional sind. Im allgemeinen wird $Q_n(u)$ als konstant, und damit weißes Rauschen angenommen, während $Q_f(u)$ normalerweise eine abfallende Funktion ist. Man kann zeigen, daß sich diese Überlegungen in den Ortsbereich übertragen lassen und dort - unter den genannten Bedingungen - auf die wesentlich allgemeinere Lösung der klassischen Varianzkomponentenschätzung führen (vgl. FÖRSTNER 1983). Dabei entfallen dann auch die Forderungen nach Periodizität und Gleichabständigkeit der Meßwerte.

Von den einschränkenden Bedingungen läßt sich diejenige der Kenntnis über den funktionalen Zusammenhang zwischen Leistungsspektrum und Frequenz des Signals am leichtesten beheben. Für beobachtete autoregressive Prozesse, gleiche Abstände der Meßwerte voraussetzend, lassen sich die Prozeßparameter und die Varianzen simultan (PAGANO 1974, SCHULTE 1987) oder iterativ (FÖRSTNER 1983, LINDENBERGER 1987) bestimmen. Die Schätzung des Rauschanteils etwa bei Geländeprofilen (LINDENBERGER 1987) oder Zeitreihen von Orientierungsgrößen bewegter Sensorplattformen ergaben auch unter kontrollierten Versuchsbedingungen realistische Ergebnisse. Die Homogenität der untersuchten Profile ist der Grund für den Erfolg, aus einzelnen Profilen das Signal-Rausch-Verhältnis zuverlässig ableiten zu können.

Genau diese Bedingung ist aber verletzt, wenn wir das Verfahren, ohne vorherige Segmentierung auf Bilder anwenden wollen. Unser Bildmodell postuliert nur gebietsweise homogene stochastische Prozesse. Für eine Anwendung des Verfahrens müßten wir daher eine Bestimmung, d. h. Schätzung der Gebietsgrenzen vornehmen, die aber selbst die Schätzung des Rauschanteils enthalten oder zumindest eine gute Näherung für den Rauschanteil als bekannt voraussetzen müßte.

Dies ist unser Ziel in diesem Abschnitt. Wir wollen eine Näherung für die Rauschvarianz aus der Verteilung einer Funktion der beobachteten Bildfunktion herleiten.

2. Bestimmung der Rauschvarianz aus der Verteilung der Gradienten

Da die einzelnen Komponenten von g und damit g selbst normalverteilt angenommen werden, läßt sich aus der Verteilung von g oder auch von linearen Funktionen von g wegen der reproduzierenden Eigenschaft der Normalverteilung nicht auf die einzelnen Komponenten schließen.

Nichtlineare Funktionen von q sind dagegen i. a. nicht mehr normal verteilt. Man kann daher erwarten, aus der empirischen Verteilung Kenngrößen zumindest über den Rauschanteil zu erhalten. Insbesondere geeignet sind Beträge oder Quadrate der Ableitungen der Bildfunktion, denn sie eliminieren etwaige konstante oder lineare Änderungen der Funktion und ermöglichen eine robuste Schätzung der Rauschvarianz ohne vorherige Segmentierung.

Wir wollen hier zunächst die Betrachtung auf die erste Ableitung einer eindimensionalen Funktion beschränken. Die Herleitung für höhere Ableitungen und für zweidimensionale Funktionen erfolgt analog. Sie wollen wir später bei den Beispielen behandeln.

Sei $\underline{f}(x)$ das Signal und $\underline{n}(x)$ das additive Rauschen, beide normalverteilt:

$$\begin{aligned}\underline{f} &= \{\underline{f}(x)\} \sim N(\mu_f, \sigma_f^2 C) \\ \underline{n} &= \{\underline{n}(x)\} \sim N(0, \sigma_n^2 I)\end{aligned}\tag{215-5}$$

mit einer Toeplitzmatrix C , entsprechend unserer Annahme über die Homogenität des Signals.

Dann folgt die Ableitung des diskreten Signals etwa aus

$$\underline{q}'(x) = d\underline{q}(x)/dx = \underline{f}'(x) + \underline{n}'(x) = [\underline{q}(x+1) - \underline{q}(x-1)]/2\tag{215-6}$$

ebenfalls einer Normalverteilung

$$\underline{q}'(x) \sim N(0, 2(1-\rho) \sigma_f^2 + 2 \sigma_n^2) = N(0, 2 \sigma_g^2)\tag{215-7}$$

wobei ρ die Korrelation zum Abstand 2 bedeutet.

Daher gilt für das Quadrat eines Gradienten

$$\underline{q}'^2(x) / \sigma_{g'}^2 \sim X_1^2\tag{215-8}$$

mit

$$E(\underline{q}'^2(x)) = 2(1-\rho) \sigma_f^2 + 2 \sigma_n^2\tag{215-9}$$

Die Verteilung von $\underline{q}'^2(x)$ wird dann vom Rauschen dominiert, wenn

$$(1-\rho) \sigma_f^2 \ll \sigma_n^2\tag{215-10}$$

Dies tritt in zwei Fällen auf:

- a. $(1-\rho) \ll 1$: In diesem Fall handelt es sich um ein stark korreliertes, d. h. langwelliges, glattes Signal
- b. $\sigma_f^2 \ll \sigma_n^2$: In diesem Fall ist die mittlere Abweichung des Signals vom Mittelwert wesentlich kleiner als die Streuung des Rauschens.

Der wesentliche Punkt der im folgenden abgeleiteten Schätzung ist, daß wenigstens ein kleiner Prozentsatz des betrachteten Bildes einer der beiden Bedingungen genügt.

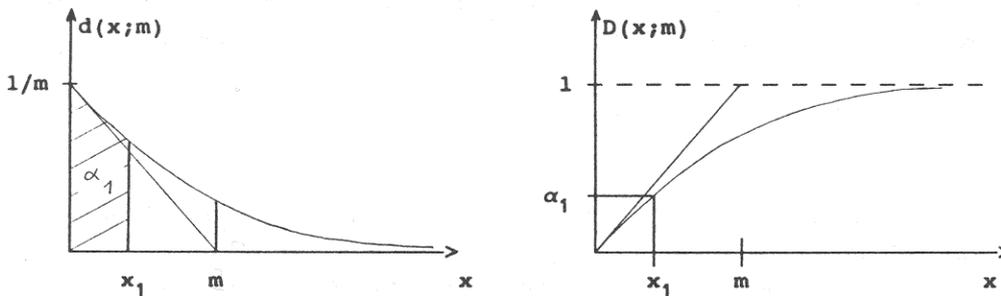
Wir wollen daher eine Schätzung von σ_n^2 aus der Verteilung von \underline{q}'^2 für den Fall herleiten, daß

die Bedingung Gl. 215-10 erfüllt ist. Dabei werden wir die Ableitung für die Verteilung einer beliebigen Funktion \underline{g} von \underline{g}'^2 formulieren, da wir damit formal auch den Fall anderer, nichtlinearer Funktionen, insbesondere auch mehrdimensionaler Signale erfassen. Wesentlich ist nur, daß sich die unterschiedlichen Eigenschaften von Signal und Rauschen in der Verteilung von \underline{g} widerspiegeln, d. h. eine Bedingung ähnlich der Gl. 215-10 gilt.

Sei $d(x;m)$ die Dichtefunktion und $D(x;m)$ die Verteilungsfunktion einer nichtlinearen Funktion \underline{g} des Signals. Sie hänge nur von dem zu bestimmenden Parameter m ab, der den Mittelwert von \underline{g} darstellt (vgl. Fig. 2-2). Also sei

$$E(\underline{g}) = m \tag{215-11}$$

Fig. 2-2 Dichte und Verteilungsfunktion der X_2^2 -Verteilung
 α_1 -Punkt $x_1 = x(\alpha_1)$,
 Mittelwert $m = E(\underline{g}) =$ Tangentenabschnitt



Im speziellen Fall für $\underline{g} = \underline{g}'^2$ gilt für signalfreies \underline{g} : $E(\underline{g}) = 2 * \sigma_n^2$. Wir nehmen an, daß auch im Allgemeinfall eine einfache Beziehung zwischen m und der gesuchten Rauschvarianz besteht, so daß wir aus einer Schätzung für m eine Schätzung für σ_n^2 ermitteln können

Wir schätzen nun in einer 1. Iteration m aus aus dem α_1 -Punkt $x(\alpha_1) = D^{-1}(\alpha_1)$ der empirischen Verteilung und erhalten

$$\hat{m}^{(1)} = m_0 \frac{\hat{x}(\alpha_1)}{x_0(\alpha_1)}, \tag{215-12}$$

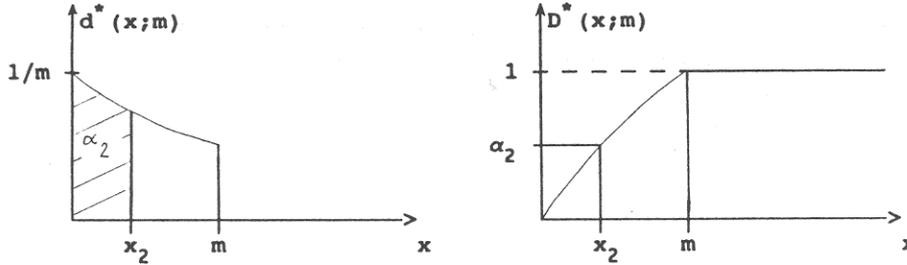
worin x_0 und m_0 sich auf die theoretische Verteilung von \underline{g} beziehen. Für kleine α , etwa 0.05 oder 0.1, sei die Schätzung aus der empirischen Verteilung nur unwesentlich durch "Ausreißer" in \underline{g} beeinflusst, die etwa an Gebietsgrenzen entstehen. Damit läßt sich aus dieser Schätzung eine gute 1. Näherung für die Rauschvarianz ableiten, im speziellen Fall etwa

$$\hat{\sigma}_n^2 = \hat{m} / 2.$$

Falls die Zahl der Sprungstellen groß oder die Bereiche mit Textur umfangreich sind, aber unter 50 % bleiben, ist die Schätzung nach Gl. 215-12 zu pessimistisch, da der α -Punkt der Verteilung von \underline{g} gegenüber der theoretischen Verteilung bei signalfreiem Bild nach rechts verschoben wird. Wenn man aber annehmen kann, daß die Form der Verteilung links von $\hat{m}^{(1)}$ nicht oder nur unwesentlich durch Sprungstellen oder Textur beeinflusst ist, kann man aus der bedingten empirischen Verteilung $\underline{D}^* = \underline{D}(x;m | x < \hat{m}^{(1)})$, d. h. nach Ausschluß aller Werte $s > \hat{m}^{(1)}$ eine verbesserte Schätzung für m erhalten. Die bedingte Verteilung \underline{D}^* hängt von dem gleichen Parameter m ab wie D , so daß bei Wahl eines geeigneten α_2 die Schätzung gewonnen wird:

$$\hat{m}^{(2)} = m_0 \frac{\hat{x}^*(\alpha_2)}{x_0^*(\alpha_2)} \quad (215-13)$$

Fig. 2-3 Bedingte Dichte und Verteilungsfunktion der X_2^2 Verteilung für $\underline{g} < m$



Hierin sind wieder m_0 und $x_0^*(\alpha_2)$ aus der theoretischen bedingten Verteilung. Man beachte, daß m nicht der Mittelwert der bedingten Verteilung ist.

Diese Schätzung ist robust sowohl gegen Kanten bei Verwendung von Daten aus verschiedenen Gebieten als auch bei Texturen, die genügend deutlich von weißem Rauschen abweichen.

Für die Wahl der α -Punkte sind zwei Kriterien gegeneinander abzuwägen:

1. α sollte klein genug sein, damit die Schätzung nicht durch "Ausreißer" beeinflusst wird, also die gewünschte Robustheit aufweist.
2. α sollte groß genug sein, um eine genügend genaue Schätzung der Rauschvarianz zu erhalten.

Die Varianz der Schätzung von m ist (vgl. Anhang 1)

$$\sigma_m^2 = \frac{m_0^2}{x_0^2} \frac{1 - \alpha(1 - \alpha)}{d^2(x) n} \quad (215-15)$$

und hängt damit wesentlich ab von

- dem gewählten Punkt α
- der zugehörigen Stelle $x = x_0$ der Verteilungsfunktion
- dem Wert $d(x)$ der Dichtefunktion
- dem Stichprobenumfang n und
- dem relativen Abstand $m_0/x_0 = m/x$ des Mittelwertes m vom Punkt $x(\alpha)$

Ohne eine Spezifizierung der Funktion \underline{g} ist eine Diskussion der Genauigkeit nicht möglich. Wir wollen daher an einem repräsentativen Beispiel die Schätzung und ihre Eigenschaften diskutieren.

3. Beispiel:

Wir wollen die Rauschvarianz in digitalen Bildern aus der Verteilung der Quadrate der Gradienten schätzen. Dazu müssen wir zunächst die Berechnung der Gradienten spezifizieren. Wir verwenden den auf Zeilen und Spalten bezogenen Roberts-Gradient, der sich aus Gl. (213-5) ergibt. Damit ist der Erwartungswert des Betragsquadrats des Gradienten,

$$\underline{g} = g_r^2 + g_c^2 \quad (215-16)$$

falls keine Textur vorliegt

$$E(\underline{g}) = m = 2 \sigma_n^2 \quad (215-17)$$

und $\underline{g}/(2 \sigma_n^2)$ folgt einer X_2^2 -Verteilung mit 2 Freiheitsgraden (vgl. Anhang 2). Damit hat \underline{g} die Dichte

$$d_s(x;m) = \frac{1}{m} e^{-\frac{x}{m}}, \quad x \geq 0 \quad (215-18)$$

und die Verteilung

$$D_s(x;m) = 1 - e^{-\frac{x}{m}}, \quad x \geq 0. \quad (215-19)$$

(vgl. Fig. 2-2)

Damit erhalten wir mit Gl. 215-15 die Streuung für σ_m die Schätzung von m aus dem α -Punkt der empirischen Verteilung mit $n = n_\alpha$ für kleine α genähert zu

$$\sigma_{\hat{m}} = \frac{m}{\sqrt{n_\alpha}} \quad (215-20)$$

n_α ist die Zahl der 100 α % kleinsten Werte von \underline{g} . Für konstant gewähltes α wird die Genauigkeit der Schätzung besser, wenn m aus einer größeren Stichprobe abgeleitet wird. Dagegen muß n_α konstant gewählt werden, wenn man eine vorgegebene Relativgenauigkeit für m erreichen will. Wir haben in unseren Experimenten $n_\alpha = 150$ gewählt, d. h. die Punkte α_1 und α_2 vom Stichprobenumfang abhängig gemacht und erhalten so die Genauigkeit der geschätzten Rauschvarianz bei Bildern mit mehr als 40x40 Bildelementen besser als 8 %, und daher die Streuung besser als 4 %.

VOORHEES und POGGIO (1987) schätzen die Varianz des Rauschens aus dem differenzierten Histogramm der Beträge der Gradienten, das der Ableitung einer Rayleigh-Verteilung folgt. Ihr Verfahren nutzt ebenfalls den linken Teil des Histogramms, um den Einfluß der Textur auf die Schätzung zu vermindern, und führt wie das obige Verfahren in stark texturierten Gebieten zu einer verzerrten Schätzung. Die Verzerrung wird aber bei VOORHEES und POGGIO größer sein, da sie einen größeren Anteil des Histogramms nutzen. Auch dürfte die Genauigkeit wegen der Diskretisierungsfehler und der Gradientenbildung aus dem Histogramm geringer sein als bei unserem Verfahren, das über die integrierte, nicht die differenzierte Dichtefunktion arbeitet.

Fig. A2-9 bis -12 (s. Anhang 3) zeigen für drei künstliche und ein natürliches Bild das Histogramm der Gradientenquadrate und die angepaßte Dichte. Die ebenfalls eingezeichnete Tangente im Ursprung schneidet die Abszissenachse im geschätzten Erwartungswert der Verteilung.

Bei dem Streifenmuster in Figur A2-9 liegen für die Schätzung ideale Verhältnisse vor. Bei einem Kontrast von 20 (Graustufen) und einer theoretischen Streuung des Rauschens von 5 ist das Signal-Rausch-Verhältnis von $(20/2) / 5 = 2$. Die erste Näherung für die Standardabweichung nach Gl. 215-12 ist 5.17, die zweite Näherung 5.07, ein verglichen mit dem theoretischen Wert gutes Ergebnis. 86.7 % der Gradienten liegen unter dem 3-fachen des mittleren Gradienten. Das empirische Signal-Rausch-Verhältnis ist 2.1. Das Histogramm zeigt klar den exponentiellen Verlauf der Dichtefunktion.

Das künstliche Muster in Fig. A2-10 zeigt in der linken oberen Ecke wegen der mangelnden Auflösung eine rauschähnliche Struktur. Sie macht etwa 10 % der Fläche des Bildes aus. Bei einem empirische Signal-Rausch-Verhältnis von 2.1 war die erste Näherung für die Streuung des

Rauschens 2.87, die zweite 2.60. Die Textur im linken oberen Teil konnte offensichtlich nicht gut von dem additiven Rauschen unterschieden werden und führte zu einer um 30 % zu großen Streuung.

Das Histogramm des Wellenmusters in Fig. A2-11 zeigt anschaulich die Bereiche des Rauschens und der Kanten. Offensichtlich tritt auch hier eine gewisse Überlappung ein. Die erste Schätzung ist mit 3.83 fast doppelt so hoch wie der theoretische Wert. Die zweite Schätzung ist dagegen nur noch um 40 % zu pessimistisch. Die Abweichung der angepaßten Dichte an die empirische ist deutlich erkennbar.

Daß solche Histogramme in natürlichen Bildern vorkommen, in denen noch nicht einmal ausgesprochen texturierte Gebiete enthalten sind, zeigt der Luftbildausschnitt in Fig. A2-12. Bei einem Signal-Rausch-Verhältnis von 1.8 ergab die erste Schätzung 6.00, die zweite 4.47. Die angepaßte Exponentialfunktion gibt den Histogrammverlauf im linken Teil recht gut wieder.

Nachdem wir nun eine zuverlässige Schätzung für den Rauschanteil in digitalen Bildern zur Verfügung haben, sind wir in der Lage, ein entsprechendes Filter zur Unterdrückung des Rauschens zu entwickeln.

2.1.5.2 Das adaptive Filter

Die Qualität jeder Bildanalyse hängt wesentlich davon ab, wie gut das durch Rauschen überlagerte Bild rekonstruiert werden kann. Die meisten dazu verwendeten Filter sind lediglich von der globalen Signalcharakteristik, etwa der Korrelationsfunktion und dem Signal-Rauschverhältnis abhängig, wie z. B. das Wiener-Filter. Der Nachteil dieser Verfahren ist, daß sie auf lokale Änderungen der Signalstruktur nicht reagieren können und so entweder Kanten und Ecken glätten oder das Rauschen zu wenig unterdrücken. Unter den kantenerhaltenden Filtern ist das Median-Filter das bekannteste und insbesondere bei nicht-normalverteiltem Rauschen äußerst wirksam. Es schneidet aber Ecken ab und eliminiert schmale Linien. Neben diesen Eigenschaften, die das Median-Filter als generelles Vorfilter für Bildanalyseaufgaben untauglich erscheinen lassen, führt es auch gegenüber linearen Filtern zu geometrisch schlechteren Ergebnissen (AKEY und MITCHELL 1984, GALLAGHER 1982). Andere nichtlineare Filter, wie etwa das KUWAHARA-Filter, das Ecken, die von horizontalen und vertikalen Kanten begrenzt sind, erhält, sind für spezielle Aufgaben entwickelt und, was die Qualität des gefilterten Bildes betrifft, theoretisch weniger durchsichtig (vgl. etwa auch CHIN und YEN 1983, HARWOOD et. al. 1987, LEE 1981, 1983, NAGAO und MATSUYAMA 1979, WANG et. al. 1981).

Das hier vorgeschlagene informationserhaltende Filter ist ein Wiener Filter, das sich adaptiv auf die lokal gültige Kovarianzstruktur einstellt.

Wir wollen zunächst das Filter für den Fall homogener Bildbereiche ableiten und dann das Verhalten des Filters an Gebietsrändern, -ecken und anderen singulären Stellen diskutieren. Trotz der oben genannten ungünstigen Eigenschaften soll uns das Median-Filter als Referenz dienen.

1. Das Rechenmodell des adaptiven Filters und eine Approximation

In statistisch homogenen Gebieten, d. h. in Gebieten mit konstanter Kovarianzstruktur ist das Filter ein Wiener Filter. Wäre uns die Kovarianzmatrix C_{ff} der Bildpunkte f in einem kleinen m_1, m_2 -Fenster (3×3 oder größer) und die Rauschvarianz bekannt, so könnten wir mit

$$\hat{\underline{f}} = C_{fg} C_{gg}^{-1} \underline{g} = C_{ff} (C_{ff} + \sigma_n^2 I)^{-1} \underline{g} = P \underline{g} \quad (215-21)$$

eine Schätzung für \underline{f} erhalten. Wie wir aber in den Abschnitten 2.1.4.2 und 2.1.5.1 gesehen haben, sind wir in der Lage, aus der beobachteten Bildfunktion \underline{g} Schätzungen für C_{ff} und σ_n^2 zu erhalten. Damit können wir die Prädiktionsgl. 215-21 an jedem Bildpunkt mit Hilfe der Werte in seiner unmittelbaren Umgebung anwenden. Denn die Werte der Kovarianzfunktion können wir aus der Approximation Gl. 213-1 für jeden Bildpunkt des betreffenden Fenster bestimmen. Wir erhalten dann durch Lösung der Gleichung $C_{ff} = P C_{gg}$ die Gewichte P für die einzelnen Funktionswerte \underline{g} .

Der Aufwand für diese Filterung ist wegen der Gleichungslösung recht hoch. Wir wollen hier eine rechen-technisch weniger aufwendige Näherung entwickeln. Ziel ist die direkte Bestimmung der Gewichte im Vektor P in unmittelbarer Abhängigkeit von der Hesse-Matrix und der geschätzten Rauschvarianz, d. h. die Bestimmung der Gewichte über eine Gewichtsfunktion analog zur Bestimmung der Elemente der Kovarianzmatrix durch eine Kovarianzfunktion.

Dazu wollen wir zunächst die Abhängigkeit der Kovarianzen und der Gewichte vom Abstand zum Ursprung miteinander vergleichen, um damit die verwendete Struktur der Gewichtsfunktion zu motivieren.

Wir nehmen an, daß ein gedrehtes Koordinatensystem (u,v) existiert, in dem die Kovarianzfunktion C_{ff} separierbar ist:

$$C_{ff}(\tau_1, \tau_2) = C_{ff}^{(u)}(\tau_1) \cdot C_{ff}^{(v)}(\tau_2)$$

bzw. diskret als Kroneker-Produkt der Kovarianzmatrizen

$$C_{ff} = C_{ff}^{(u)} \otimes C_{ff}^{(v)} .$$

Die Koordinatenunterschiede τ_1 und τ_2 seien parallel zu u und v .

Das Kronekerprodukt setzt eine entsprechende Anordnung der Funktionswerte des betrachteten rechteckigen Fensters voraus. Die Krümmungen der beiden Kovarianzfunktionen im Ursprung stimmen mit den Eigenwerten der Hesse-Matrix von C_{ff} überein. Falls C_{gg} bezüglich des gleichen Koordinatensystems separierbar wäre, gälte wegen $(A \otimes B)(C \otimes D) = AC \otimes BD$ dies auch für P , so daß wir die Filterung in u - und v - Richtung getrennt nacheinander vornehmen könnten. Da wir die Kovarianz durch eine abfallende Funktion approximieren, können wir auch die zugehörige Gewichtsfunktion in erster Näherung durch eine abfallende Funktion annähern, wobei starke Krümmungen der Kovarianzfunktionen mit starken Krümmungen der Gewichtsfunktionen einhergehen. Außerdem stimmen die Hauptkrümmungsrichtungen der Funktionen C und P überein.

Wir wollen aus diesen Überlegungen heraus die $m_1 \times m_2$ Werte des Vektors $P(\tau) = P(\tau_1, \tau_2)$ mit folgender Gewichtsfunktion bestimmen

$$P(\tau) = \frac{c}{1 + \frac{1}{2} \tau' \frac{H_{ff}}{\sigma_n^2} \tau} \quad (215-22)$$

Die Konstante c dient dazu, die Summe der Gewichte auf 1 zu normieren. Die Normierung mit σ_n^2 sorgt dafür, daß eine geringfügige Filterung stattfindet, wenn Signal und Rauschen gleiche Varianz im Gradienten aufweisen. Etwa wird, falls $H_{ff} = \sigma_n^2 I = 2 \sigma_n^2 I$ ist, das beobachtete Signal mit

$$\frac{6}{26} \begin{bmatrix} \frac{1}{3} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{3} \end{bmatrix} = \frac{1}{26} \begin{bmatrix} 2 & 3 & 2 \\ 3 & 6 & 3 \\ 2 & 3 & 2 \end{bmatrix}$$

gefaltet, ein Kern, der einem 3x3-Gauß-Filter sehr ähnlich ist. Für $H_{ff} = 0$, d. h. für ein signalfreies Fenster, wird der Kern zu einem Box-Filter. Das ist die stärkste Filterung, die wir bei vorgegebener Größe des Kerns erreichen können. In diesem Extremfall treten Verflechtungseffekte auf, die die Berechnung der Gewichte, vor allem aber die Begrenzung auf ein 3x3 Fenster als Näherung ausweisen.

Für H verwenden wir die nach Gl. 214-18 gewonnene Schätzung.

2. Das informationserhaltende Filter an Kanten und Ecken

Nachdem wir nun ein von der Texturcharakteristik abhängiges und bei festgehaltenem H in homogenen Gebieten lineares Filter entwickelt haben, wollen wir seine adaptiven Eigenschaften an Gebietsgrenzen, -ecken und -knoten für den Fall diskutieren, daß die Gebiete selbst keine Textur aufweisen.

Zunächst betrachten wir Kanten, d. h. Gebietsgrenzen, die glatt sind, also lokal als geradlinig angenommen werden können. Hier sind die Gradienten parallel und daher die Ableitungen in Zeilen- und Spaltenrichtung linear abhängig. So wird, zumindest theoretisch, die Matrix H singular, wenn kein Rauschen vorliegt. Der zum größten Eigenwert gehörige Eigenvektor gibt die Richtung des Maximalgefälles an und steht daher senkrecht auf der Kante. Wegen der Diskretisierung und des Abtaustrauschens wird die nach Gl. 213-4 geschätzte Matrix i. a. aber regulär. Falls aber der kleinste Eigenwert von H, der nur durch das Rauschen entlang der Kante verursacht wird, identisch mit der aus dem gesamten Bild abgeleiteten Varianz σ_n^2 des Rauschgradienten wird, dann ist die nach Gl. 214-18 geschätzte Matrix H singular und gibt damit die Eigenschaften des Signals richtig wieder (vgl. die Diskussion der Merkmalskarten im Abschnitt 2.1.4).

Bei der vorgeschlagenen Filterung wird aber in diesem Fall lediglich entlang der - lokal geradlinig angenommenen - Kante gefiltert und zwar im gesamten Kantenbereich, der von dem Fenster überdeckt wird, mit dem H berechnet wird. Der Grund liegt darin, daß die Funktionswerte g von Bildelementen, die - bezogen auf die Kantenrichtung - rechts und links vom zu filternden Bildelement liegen, wegen der großen Gradienten quer zur Kante ein kleines Gewicht erhalten, während die Funktionswerte von Bildelementen, die - bezogen auf die Kantenrichtung - vor bzw. hinter dem zu filternden Bildelement liegen für die gewogene Mittelung ein ähnliches Gewicht wie das zu filternde Bildelement erhalten und daher zur Glättung beitragen.

An gekrümmten Gebietsgrenzen und an Knoten wird dagegen H nie singular. H hat dort etwa gleich große Eigenwerte, weist also auf eine ungefähr isotrope Textur hin. Im Extrem kann H zur Diagonalmatrix werden, etwa bei rechtwinkligen Ecken. Da H in diesen Fällen - gegenüber σ_n^2 - groß ist, wird nicht oder kaum gefiltert, so daß Ecken bzw. Knickpunkte erhalten bleiben. Allerdings wird das Rauschen an diesen Stellen auch nicht oder nur sehr wenig unterdrückt. Im Gegensatz zum Median bleiben deshalb auch einzelne helle oder dunkle Flecken erhalten.

Die folgenden Beispiele sollen das informationserhaltende Filter demonstrieren.

3. Beispiele

Zunächst wollen wir die Wirkungsweise des Filters an dem Siemensstern und dem Schachbrett verdeutlichen.

Fig. A2-13 zeigt links oben denselben Siemensstern wie Fig. A2-2. In der Mitte oben ist der mit dem informationserhaltenden Filter bearbeitete Stern und darunter die zugehörigen Merkmalskarten (Texturstärke, - anisotropie und -richtung) dargestellt.

Man erkennt deutlich, daß nur innerhalb der Strahlen geglättet wurde und die Kanten in ihrer Schärfe erhalten blieben. Das geschätzte Rauschen ging von 3.3 auf 1.0 zurück. Damit wurde das Signal-Rausch-Verhältnis von 1.6 auf 2.5 erhöht. Das Anisotropiebild in Fig. A2-13 zeigt, daß die Anisotropie, die vor allem um den inneren Kern durch das Rauschen abgeschwächt wurde, nun stark angestiegen ist. Daher ist auch das Richtungsbild vollständiger.

Fig. A2-14 zeigt die Wirkung des Filters bei dem Schachbrett von Fig. A2-4. Die Schachbrettfelder wurden stark geglättet und dabei die Kanten erhalten. Außerdem wurden die Ecken nicht abgerundet. Insbesondere wurde die Anisotropie entlang der Kanten verbessert und das Richtungsbild vervollständigt.

Das Filter läßt sich auch iterativ anwenden. Fig. A2-15 zeigt das Fischgrätenmuster diesmal mit einer 3.5 mal so großen Streuung des Rauschens und das 2-mal gefilterte Bild. Die Gräten wurden in ihrem Kontrast durch die Filterung nicht beeinflusst aber wesentlich klarer durch die Filterung. Die Stiele sind durch die Filterung kaum beeinflusst, so daß auch hier die Ecken unverfälscht wiedergegeben werden.

Besonders fällt auf, daß - hier durch Zufall entstandene - Kleinstrukturen bzw. Unregelmäßigkeiten, die über das Rauschen herausragen, durch die Filterung deutlicher sichtbar sind. Die Stellen lassen sich, soweit sie im Grätenbereich liegen, im Anisotropiebild leicht lokalisieren, da sie dort als dunkle Flecken erscheinen.

Die drei folgenden Bilder zeigen die Wirkung des Filters auf reale Bilder. In allen Fällen werden flache Gebiete geglättet, während Bereiche mit Kanten oder Flecken nicht sehr stark verändert werden.

Wir wollen nun an weiteren Beispielen die Leistungsfähigkeit dieses Filters mit dem Median-Filter vergleichen, das oft als kantenerhaltendes Filter verwendet wird.

Fig. A2-19 und A2-20 zeigen die Filterung eines schwach verrauschten Streifenmusters mit dem informationserhaltenden und dem Median-Filter zusammen mit den drei Merkmalskarten nach der jeweiligen Filterung. Der Unterschied ist gering. Das geschätzte Rauschen ist allerdings beim Median gefilterten Bild mit 0.27 gegenüber 1.5 deutlich geringer. Dies zeigt sich in der größeren Glattheit der flachen Gebiete, wo beim Median auch Kleinstrukturen weggefiltert wurden. Dagegen erscheinen die Kanten beim informationserhaltend gefilterten Bild etwas glatter.

Bei diesem Beispiel wird die Schwierigkeit der Beurteilung von Vorverarbeitungsverfahren deutlich. Ihr Wert kann eigentlich nur in Bezug auf eine spezifische Aufgabenstellung bestimmt werden, etwa einer Objektlokalisierung. Wir werden hierauf bei der Kantenextraktion zurückkommen.

Bei starkem Rauschen wird allerdings der Unterschied beider Filter offensichtlich. Fig. A2-21 bis A2-24 zeigen für ein Ausgangs Signal-Rausch-Verhältnis von empirisch 1.6 die Wirkung des informationserhaltenden Filters (Fig. A2-21), die des 3x3 Median-Filters (Fig. A2-22), des 2-mal angewandten informationserhaltenden Filters (Fig. A2-23) und des 5x5 Medians (Fig. A2-24). Die Median gefilterten Bilder zeigen in beiden Fällen einen deutlich unruhigeren Kantenverlauf. Die zweimalige Anwendung des informationserhaltenden Filters bewirkt noch einmal eine Straffung der Richtungen entlang der Kanten.

Wieder fällt auf, daß Details innerhalb der glatten Flächen wenig gefiltert werden, oder anders gesehen, daß die gegenüber der Modellvorstellung unvollkommene Darstellung des Streifenmusters durch das Original begründet ist. Ob diese Details Information darstellen, oder nicht - wie hier - kann und sollte auch nicht auf dieser Ebene der Bildauswertung entschieden werden, sondern muß späteren, die Aufgabe und das Wissen berücksichtigenden Analyseschritten vorbehalten bleiben.

In Fig. A2-25 bis A2-28 finden wir dieselben Effekte bei einem anderen Muster. Wieder wird die Schwäche des Median-Filters, insbesondere wegen des Ausrundens der Ecken und des Eliminierens feiner Linien gegenüber dem informationserhaltenden Filter deutlich. Der 2. Durchgang mit dem informationserhaltenden Filter führt hier aber zu kaum einer sichtbaren Verbesserung mehr. Bei dem Luftbild in Fig. A2-29 und A2-30 wird an einem realen Beispiel der Unterschied beider Filter insgesamt wirksam demonstriert.

Das Ergebnis dieses Vergleichs wird durch die Kantenextraktion im Abschnitt 2.2.5, als nächstem Schritt der Bildanalyse, bestätigt.

2.2 Extraktion von markanten Punkten und Bildkanten

Ein entscheidender Schritt der Bildanalyse besteht im Übergang von der ikonischen zur symbolischen Repräsentation des Bildinhalts. Wir wollen die bildlichen oder graphischen Grundelemente in Anlehnung an den englischen Sprachgebrauch "Primitive" nennen (engl. "primitives", vgl. etwa GUDE, MÄRGNER, PAULUS 1985). Dies sind im wesentlichen punkt-, linien- und flächenartige Elemente die durch einige Attribute, etwa Helligkeit, Länge oder Fläche näher gekennzeichnet werden. Sie bilden eine symbolische Repräsentation des Bildes.

Wir wollen in diesem Abschnitt einen Operator zur Extraktion markanter Punkte diskutieren, der in abgewandelter Form auch für die Extraktion von Kantenelementen geeignet ist. Unter markanten Punkten verstehen wir solche, die sich eindeutig lokalisieren lassen, wie helle oder dunkle Flecken, Ecken, die Mitte von rotationssymmetrischen Figuren oder auch Stellen, an denen die Textur gegenüber der Umgebung besonders stark ist. Solche Punkte werden bei verschiedenen Bildanalyseaufgaben benötigt, wie der Bildzuordnung für die 3D-Rekonstruktion der abgebildeten Objekte, für die Verknüpfung von Bildern innerhalb einer Bildfolge oder für die Lokalisierung von Objekten in einem Bild auf der Grundlage eines Objektmodells. Bei der Einbildanalyse spielen daneben auch die Zentren rotationssymmetrischer Figuren eine zentrale Rolle, etwa bei der Lokalisierung von markierten Punkten, der Bestimmung der Mittelpunkte von Ringen, Löchern oder Scheiben.

Der Operator wurde für die Zuordnung von Bildern entwickelt (PADERES et. al. 1984, vgl. Förstner 1986a). In seiner ursprünglichen Form war er lediglich zur Suche markanter Punkte gedacht, wie der MORAVEC-Operator, der z. B. dem Bildzuordnungsverfahren von BARNARD und THOMPSON (1980) zugrundeliegt. Er stellte sich aber als wesentlich leistungsfähiger heraus.

Ein Grund dafür ist, daß der Operator in seiner ursprünglichen Form lediglich optimale Bildfenster selektierte, nicht Punkte. Es war eine Fehlinterpretation des Algorithmus, die Mitten der Fenster als markante Punkte zu deklarieren. Sie erklärt die systematischen Fehler des ursprünglichen Operators wie die des MORAVEC-Operators. Der wesentliche zweite Schritt des Verfahrens, der im ursprünglichen Algorithmus fehlte, besteht in der Bestimmung des optimalen Punktes innerhalb der selektierten Fenster.

Der andere Grund für die Eignung des Operators liegt darin, daß er optimale Fenster für die Korrelation oder Kleinste Quadrate Zuordnung, für die Erkennung von Ecken, von Zentren rotationssymmetrischer oder auch spiralförmiger Figuren findet, ohne daß man vorher den Typ des zu findenden Punktes spezifizieren muß. Die Klassifizierung des selektierten Fensters kann später erfolgen, falls sie benötigt wird.

In unserem Zusammenhang ist nun folgendes von besonderem Interesse: Das Modell ist geometrisch, und die für die Punktauswahl wesentlichen Merkmale ergeben sich aus der Gewichts- oder Kovarianzmatrix des zu bestimmenden Punktes. Die Gewichtsmatrix, bzw. die daraus abgeleiteten Parameter sind aber identisch mit der Hesse-Matrix, bzw. mit den aus ihr abgeleiteten Merkmalen, insbesondere mit der Spur und dem Anisotropiewert. Damit ist eine enge Verbindung mit den Merkmalskarten hergestellt. Man kann so die Suche von markanten Punkten oder von Kantenpunkten als eine Interpretation der Merkmalskarten verstehen.

Das Ziel dieses Kapitels ist es, einerseits während der Darstellung des Operators die Querverbindungen unter den verschiedenen Aufgabenstellungen, für die er geeignet ist, offenzulegen. Andererseits soll gezeigt werden, wie mit Hilfe der Analyse der selektierten Fenster eine verfeinerte Deutung der Merkmalskarten möglich ist, die ja auf einem stochastischen Modell der Bildfunktion basieren.

Das Kapitel stellt zunächst die Bestimmung des optimalen Punktes innerhalb eines Fensters vor (2.2.1/2) und behandelt die Klassifizierung des Punktes (2.2.3), die sich auf die Qualität der Punktlokalisierung stützt. Die Auswahl optimaler Fenster benötigt allein die Texturstärke und -anisotropie und wird in Abschnitt 2.2.4 diskutiert. Schließlich wird das Verfahren auf die Extraktion von Kantenelementen übertragen (2.2.5) und erlaubt so eine vollständige geometrische Analyse texturfreier Bilder.

Der Inhalt der Abschnitte 2.2.1, 2.2.2 und 2.2.4 stützt sich im wesentlichen auf den Beitrag von FÖRSTNER und GÜLCH (1987).

2.2.1 Kriterien für die Auswahl markanter Punkte

Bei der Auswahl markanter Punkte sind eine Reihe von Anforderungen zu erfüllen:

- Deutlichkeit:

Die Punkte sollten markant sein, d. h. sich deutlich von ihrer Umgebung unterscheiden. So sollten z. B. Punkte auf Kanten nicht ausgewählt werden, solange sie nicht für die Zuordnung von Bildern verwendet werden, bei denen die Epipolargeometrie verwendet wird.

- Invarianz:

Die Auswahl wie auch die Position des selektierten Punktes sollte invariant gegen die zu erwartenden geometrischen und radiometrischen Verzerrungen des Bildes sein. Neben der Deutlichkeit ist diese Forderung wohl die wichtigste. Denn die Invarianz beeinflusst unmittelbar die Genauigkeit und die Zuverlässigkeit der folgenden Bildanalyseschritte.

- Stabilität:

Die Auswahl sollte unempfindlich gegen Störungen sein. Dies soll bei der Bildzuordnung u. a. dafür garantieren, daß die ausgewählten Punkte mit großer Wahrscheinlichkeit auch im anderen Bild gefunden werden.

- Seltenheit:

Während Deutlichkeit lokale Unterscheidbarkeit garantiert, zielt Seltenheit auf Unterscheidbarkeit der selektierten Punkte innerhalb des Bildes. Dies ist besonders wichtig, wenn Bilder periodische Muster enthalten und die selektierten Punkte für die Zuordnung verwendet werden sollen. Solche Punkte sollten aber ein geringes Gewicht bei der Zuordnung erhalten. Denn die Auswahl seltener oder interessanter Punkte führt zu einer zuverlässigeren Zuordnung. Daher werden diese Punktesucher im Englischen "interest operators" genannt. Wir werden auf die Anforderung der Seltenheit im Kapitel 3 ausführlich zurückkommen.

- Interpretierbarkeit:

Das Auswahlprinzip sollte auf Punkte führen, die eine Bedeutung i. S. der Interpretation des Bildes haben, etwa Ecken oder Kreise. Für die Objektidentifizierung ist dies eine wesentliche Forderung. Bei der Bildzuordnung spielt sie dagegen keine Rolle.

Der im folgenden diskutierte Operator erfüllt im wesentlichen diese Anforderungen. Es ist ein zweistufiges Verfahren:

1. Selektion optimaler Fenster
2. Lokalisierung der optimalen Position innerhalb der selektierten Fenster

Da die Selektion der Fenster auf die Genauigkeit der Punkte zurückgreift, behandeln wir den Positionierungsschritt als erstes.

2.2.2 Lokalisierung von Punkten

2.2.2.1 Mathematisches Modell

Wir behandeln vier verschiedene Aufgabenstellungen, die sich als Gauß-Markoff-Modell formulieren lassen. In allen vier Fällen sind n beobachtete Werte l_i gegeben, die in dem Vektor \underline{l} zusammengefaßt sind. Für die u unbekannt Parameter \underline{x} gilt dann:

$$\underline{l} + \underline{e} = A \hat{\underline{x}}, \quad D(\underline{l}) = C = \sigma_0^2 W^{-1} \quad (222-1)$$

Die $n \times u$ Designmatrix A sei regulär. Die Genauigkeit der Beobachtungen sei entweder durch die Kovarianzmatrix C oder die Gewichtsmatrix W gegeben. Wir erhalten dann Schätzungen für die Unbekannten \underline{x} aus den Normalgleichungen

$$N \hat{\underline{x}} = \underline{h} \quad (222-2)$$

mit der Normalgleichungsmatrix und den rechten Seiten

$$N = A' W A, \quad \underline{h} = A' W \underline{l} \quad (222-3)$$

und eine Schätzung für den frei wählbaren Varianzfaktor σ_0^2

$$\hat{\sigma}_0^2 = \underline{e}' W \underline{e} / r, \quad (222-4)$$

der sich aus der Residuen \underline{e} ableiten läßt. Hierin ist $r = n - u$ die Redundanz des Systems (cf. MIKHAIL/ACKERMANN 1976).

Die gewählten Aufgabenstellungen haben gemeinsam, daß lediglich $u=2$ Unbekannte, nämlich die Position x_0, y_0 (r_0, c_0) eines Punktes zu bestimmen ist und daß jedes Bildelement innerhalb eines kleinen, etwa 5×5 bis 16×16 Bildelemente enthaltenden Fensters auf die gleiche Weise zur Lösung beiträgt:

a. Kleinste Quadrate Zuordnung

Wir nehmen an, daß die Funktionswerte (Grauwerte) $g(r,c)$ in einem Fenster eine verrauschte und um r_0 bzw. c_0 verschobene Kopie eines Signals $f(r,c)$ darstellen. Das nichtlineare Modell lautet daher

$$g(r,c) = f(r + r_0, c + c_0) + \underline{n}(r,c) \quad (222-5)$$

Nach der Linearisierung an Näherungswerten Null für beide Unbekannte erhalten wir

$$\underline{d}g(r,c) - \underline{n}(r,c) = f_r(r,c) \cdot \underline{r}_0 + f_c(r,c) \cdot \underline{c}_0 \quad (222-6a)$$

Die Varianz der Grauwerte g ist identisch mit der des Rauschens \underline{n} . Wir nehmen weißes Rauschen an. Wenn wir dann als Varianzfaktor σ_0^2 die Varianz σ_n^2 des Rauschens wählen, erhalten wir die Gewichte

$$w_{dg}(r,c) = 1 \quad (222-6b)$$

für alle Bildelemente, also $W = I$.

b. Schnitt von Kantenelementen, Ecken

Das Kantenelement an jedem Bildelement sei definiert durch eine Gerade, die durch die Mitte des Bildelementes in der Richtung senkrecht zum Gradienten der Bildfunktion verläuft. Eine Ecke (r_0, c_0) kann man als Schnittpunkt aller Kantenelemente des betrachteten Fensters erhalten. Die Gerade kann durch $r \cos \Phi + c \sin \Phi - l = 0$ repräsentiert werden, worin l der Abstand der Geraden vom Ursprung und Φ die Richtung des Gradienten darstellt. Es gilt also $\underline{v}f' = |\underline{v}f| \cdot (\cos \Phi, \sin \Phi)$. Das lineare Modell für den Schnittpunkt der Kantenelemente lautet demnach

$$\underline{l}(r,c) + \underline{e}_l(r,c) = \cos \Phi(r,c) \cdot \underline{r}_0 + \sin \Phi(r,c) \cdot \underline{c}_0$$

mit $\underline{l}(r,c) = r \cdot \cos \Phi(r,c) + c \cdot \sin \Phi(r,c)$. Das Gewicht jedes Kantenelements ist proportional zum Quadrat des Absolutbetrags des Gradienten.

$$w_l(r,c) = |\underline{v}f|^2. \quad (222-7b)$$

Dies läßt sich zeigen, indem man die Rauschvarianz zu σ_n^2 , also konstant annimmt und beachtet, daß $|\underline{v}f| = df/dl$ also $\sigma_l = \sigma_n / |\underline{v}f|$ gilt.

c. Gewogener Schwerpunkt

Wenn wir annehmen, daß jedes Bildelement zu einem gewogenen Schwerpunkt beiträgt, wobei die Gewichte wieder proportional zum Quadrat des Betrags des Gradienten sind, erhalten wir unmittelbar

$$\underline{r} + \underline{e}_r = \hat{\underline{r}}_0$$

$$\underline{c} + \underline{e}_c = \hat{\underline{c}}_0$$

Dabei muß aber die Richtung des lokalen Gradienten berücksichtigt werden. Durch Rotation des Vektors $(|\nabla f|, 0)$ nach $\nabla f(r, c)$ mit einer lokalen Rotationsmatrix erhalten wir für das Bildelement (r, c)

$$W_{rc}(r, c) = |\nabla f|^2 \begin{bmatrix} \cos^2 \Phi & \cos \Phi \cdot \sin \Phi \\ \cos \Phi \cdot \sin \Phi & \sin^2 \Phi \end{bmatrix} \quad (222-8b)$$

$$= \nabla f \cdot \nabla f' = \begin{bmatrix} f_r^2(r, c) & f_r(r, c) \cdot f_c(r, c) \\ f_r(r, c) \cdot f_c(r, c) & f_c^2(r, c) \end{bmatrix}$$

Falls z. B. das Kantenelement horizontal verläuft, gilt $\Phi = 0$, und es trägt die Zeilen-Koordinate zum Schwerpunkt bei.

Bemerkung:

Die Gewichtsmatrix in Gl 222-8b ist singulär. Für ihre Herleitung verwenden wir das Fehlerfortpflanzungsgesetz in der Form für Gewichtsmatrizen. Die Kovarianzmatrix von $\underline{y} = A \underline{x}$ ist $C_{yy} = A C_{xx} A'$, wenn C_{xx} die Kovarianzmatrix von \underline{x} ist. Falls die Inverse von A existiert, gilt demnach $W_{yy} = (A')^{-1} W_{xx} A^{-1}$. Der Gradient ∇f ergibt sich aus $e = (|\nabla f|, 0)$ durch Rotation um

$$R_{\Phi} = \begin{bmatrix} \cos \Phi & -\sin \Phi \\ \sin \Phi & \cos \Phi \end{bmatrix} \quad (222-8c)$$

Also gilt $\nabla f = R_{\Phi} \cdot (|\nabla f|, 0)$. Wenn nun die Zeilenkomponente in e das Gewicht $|\nabla f|^2$ und die Spaltenkomponente das Gewicht 0 haben, also

$$W_{ee} = \begin{bmatrix} |\nabla f|^2 & 0 \\ 0 & 0 \end{bmatrix}$$

gilt, so erhalten wir mit $A = R_{\Phi}$ schließlich $W_{rc} = (R_{\Phi}')^{-1} W_{ee} (R_{\Phi})^{-1}$ in Gl. (222-8b). PADERES (PADERES et. al. 1984) verwendet diese Repräsentation von Kantenelementen für die Entzerrung von Satellitenaufnahmen.

d. Schnitt von Gefäll-Elementen, Mittelpunkt rotationssymmetrischer Figuren

Im Gegensatz zum Schnittpunkt von Kantenelementen können wir auch lokale Gefäll-Elemente ver-schneiden. Im Falle rotationssymmetrischer Figuren erhalten wir dann den gemeinsamen Mittel-punkt, etwa eines Rings. Gefäll-Elemente können wir analog als Geraden durch das Zentrum von Bildelementen definieren, aber mit der Richtung des lokalen Gradienten, also mit einer um $\beta = 90^\circ$ gegenüber dem Kantenelement gedrehten Richtung. Analog zu Gl 222-7 erhalten wir so das lineare Modell für den Mittelpunkt (r_0, c_0)

$$\underline{l}(r, c) + \underline{e}_l(r, c) = -\sin \Phi(r, c) \cdot \hat{\underline{r}}_0 + \cos \Phi(r, c) \cdot \hat{\underline{c}}_0$$

$$w_l(r, c) = |\nabla f|^2.$$

(222-9a)

Wenn wir
Gewichte
als

2.2.2.2 Normalgleichungen

Der Grund, weshalb wir diese vier Aufgabenstellungen gewählt haben, wird klar, wenn wir die dazu gehörigen Normalgleichungen betrachten. Wir erhalten nacheinander

a: Kleinste Quadrate Zuordnung

$$\begin{bmatrix} \Sigma f_r^2 & \Sigma f_r f_c \\ \Sigma f_r f_c & \Sigma f_c^2 \end{bmatrix} \begin{bmatrix} \underline{x}_0 \\ \underline{c}_0 \end{bmatrix} = \begin{bmatrix} \Sigma f_r \underline{d}g \\ \Sigma f_c \underline{d}g \end{bmatrix} \quad (222-10)$$

b and c: Ecke und gewogener Schwerpunkt

$$\begin{bmatrix} \Sigma f_r^2 & \Sigma f_r f_c \\ \Sigma f_r f_c & \Sigma f_c^2 \end{bmatrix} \begin{bmatrix} \underline{x}_0 \\ \underline{c}_0 \end{bmatrix} = \begin{bmatrix} \Sigma f_r^2 \underline{x} + \Sigma f_r f_c \underline{c} \\ \Sigma f_r f_c \underline{x} + \Sigma f_c^2 \underline{c} \end{bmatrix} \quad (222-11)$$

d: Mittelpunkt rotationssymmetrischer Figuren

$$\begin{bmatrix} \Sigma f_c^2 & -\Sigma f_r f_c \\ -\Sigma f_r f_c & \Sigma f_r^2 \end{bmatrix} \begin{bmatrix} \underline{x}_0 \\ \underline{c}_0 \end{bmatrix} = \begin{bmatrix} \Sigma f_c^2 \underline{x} - \Sigma f_r f_c \underline{c} \\ -\Sigma f_r f_c \underline{x} + \Sigma f_r^2 \underline{c} \end{bmatrix} \quad (222-12)$$

Die Summen sind dabei über alle beteiligten Bildelemente zu nehmen.

Diskussion:

1. Die Normalgleichungsmatrizen der ersten drei Aufgabenstellungen sind identisch. Mehr noch, die Eigenwerte der Normalgleichungsmatrizen aller vier Aufgaben sind gleich. Da sich die Selektion der optimalen Fenster lediglich auf die Eigenwerte dieser Normalgleichungen stützt, kann man optimale Fenster für alle vier Aufgaben gleichzeitig selektieren. Man benötigt keine Vorkenntnis über die Art der zu suchenden Punkte.
2. Die Normalgleichungen von b. und c. sind identisch: Der Eckpunkt in einem Fenster ist identisch mit dem gewogenen Schwerpunkt. Der Schnittpunkt der Kantenelemente ist leicht interpretierbar, während die Formulierung der Aufgabe als gewogener Schwerpunkt einfacher ist, da die Designmatrix nur aus 2x2 Einheitsmatrizen besteht.
3. Die Bestimmung des Schnittpunktes der Kantenelemente ist invariant gegen Rotationen eines Polyeders im Raum um die betreffende Ecke. Darüber hinaus benötigt man keine Vorkenntnis über die Richtung und die Zahl der beteiligten Kanten. Dies ist wohl die für die Einbildanalyse wichtigste Eigenschaft des Verfahrens.
4. Analog dazu muß man für eine Bestimmung des Mittelpunktes die Zahl der kreisförmig angeordneten Kanten nicht spezifizieren.
5. Den Schnitt von Kantenelementen kann man auch als lineare Regression im Hough-Raum interpretieren. Jedes Kantenelement (r,c) entspricht dem Punkt $(\tan \Phi(r,c), l(r,c)/\cos \Phi(r,c))$ im Hough-Raum. Die Kantenelemente einer Kante bilden ein Cluster im Hough-Raum. Falls sich mehrere Kanten in einer Ecke schneiden, so liegen die zugehörigen Cluster auf einer Geraden, deren Koordinaten dem Schnittpunkt der Kanten entsprechen. Das Modell, das wir hier verwenden, nimmt an, daß die Richtung der Gerade fest ist und der Achsabschnitt $l/\cos \Phi$ die Beobachtung darstellt (vgl. Fig. 2-4). Ihre Standardabweichung ist demnach $\sigma_a = \sigma_l / \cos \Phi = 1/|f_r|$. Dann gilt das lineare Modell für die ausgleichende Gerade im Hough-Raum

$$\underline{a}(r,c) + \underline{e}_a(r,c) = \hat{\underline{x}}_0 + \tan \Phi(r,c) \cdot \hat{\underline{c}}_0 \quad (222-13a)$$

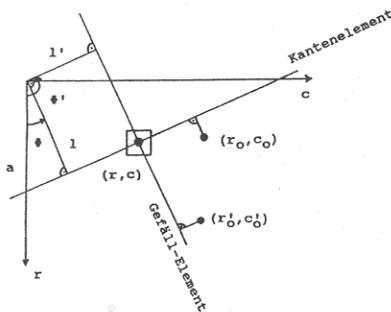
mit den Gewichten

$$w_a(r, c) = f_r^2(r, c) \quad (222-13b)$$

welches auf das Normalgleichungssystem Gl. 222-11 führt.

Fig. 2-4 Kantenelement an der Position (r, c) zur Bestimmung des Eckpunktes (r_0, c_0)

Gefäll-Element an der Position (r, c) zur Bestimmung des Kreismittelpunktes (r_0', c_0')



6. Auch der Schnitt der Gefäll-Elemente läßt sich als gewogener Schwerpunkt darstellen, aber mit einer anderen Gewichtsmatrix. Wenn tf den Kantenvektor senkrecht zum Gradienten darstellt, also $tf \cdot \nabla f = 0$ und $|tf| = |\nabla f|$ gilt, dann führt das lineare Modell Gl. 222-8a mit der Gewichtsmatrix

$$W_{rc}(r, c) = tf \cdot tf' = \begin{bmatrix} f_c^2(r, c) & -f_r(r, c) \cdot f_c(r, c) \\ -f_r(r, c) \cdot f_c(r, c) & f_r^2(r, c) \end{bmatrix} \quad (222-14)$$

auf das Normalgleichungssystem Gl. 222-12. Diese Parallele kann wieder zur Interpretation oder zur vereinfachten Herleitung von Beziehungen herangezogen werden.

Die empirische Genauigkeit des Punktes (\hat{r}_0, \hat{c}_0) kann man nun aus

$$D \begin{bmatrix} \hat{r}_0 \\ \hat{c}_0 \end{bmatrix} = \hat{\sigma}_n^2 \cdot N^{-1} \quad (222-15)$$

ableiten, worin $\hat{\sigma}_n^2$ aus Gl. 222-4 mit der Relation

$$\underline{e}' W \underline{e} = \underline{l}' W \underline{l} - \hat{\underline{x}}' \underline{h}, \quad (222-16)$$

$\hat{\underline{x}}' = (\hat{r}_0, \hat{c}_0)$ und der rechten Seite \underline{h} der Normalgleichungen bestimmt ist.

Für die Modelle b. und c. bzw. d. erhalten wir

$$\underline{e}' W \underline{e} = \Sigma f_r^2 \underline{r}^2 + 2 \Sigma f_r f_c \underline{r} \underline{c} + \Sigma f_c^2 \underline{c}^2 - \hat{\underline{x}}' \underline{h} \quad (222-17)$$

bzw.

$$\underline{e}' W \underline{e} = \Sigma f_r^2 \underline{c}^2 - 2 \Sigma f_r f_c \underline{r} \underline{c} + \Sigma f_c^2 \underline{r}^2 - \hat{\underline{x}}' \underline{h} \quad (222-18)$$

Falls die Beobachtungen als normalverteilt angenommen werden können, sind auch die geschätzten Koordinaten gemeinsam normalverteilt mit der Kovarianzmatrix aus Gl. 222-15. Die Dichtefunktion kann dann durch die Fehler- oder Konfidenzellipse repräsentiert werden, welche unmittelbar Auskunft über die Struktur der Normalgleichungsmatrix gibt.

Fig. 2-5 zeigt das Resultat des Schätzprozesses für verschiedene künstliche Bildausschnitte. Die Kanten- oder Gefäll-Elemente sind zusammen mit dem optimalen Punkt und seiner 99% Konfidenzellipse dargestellt.

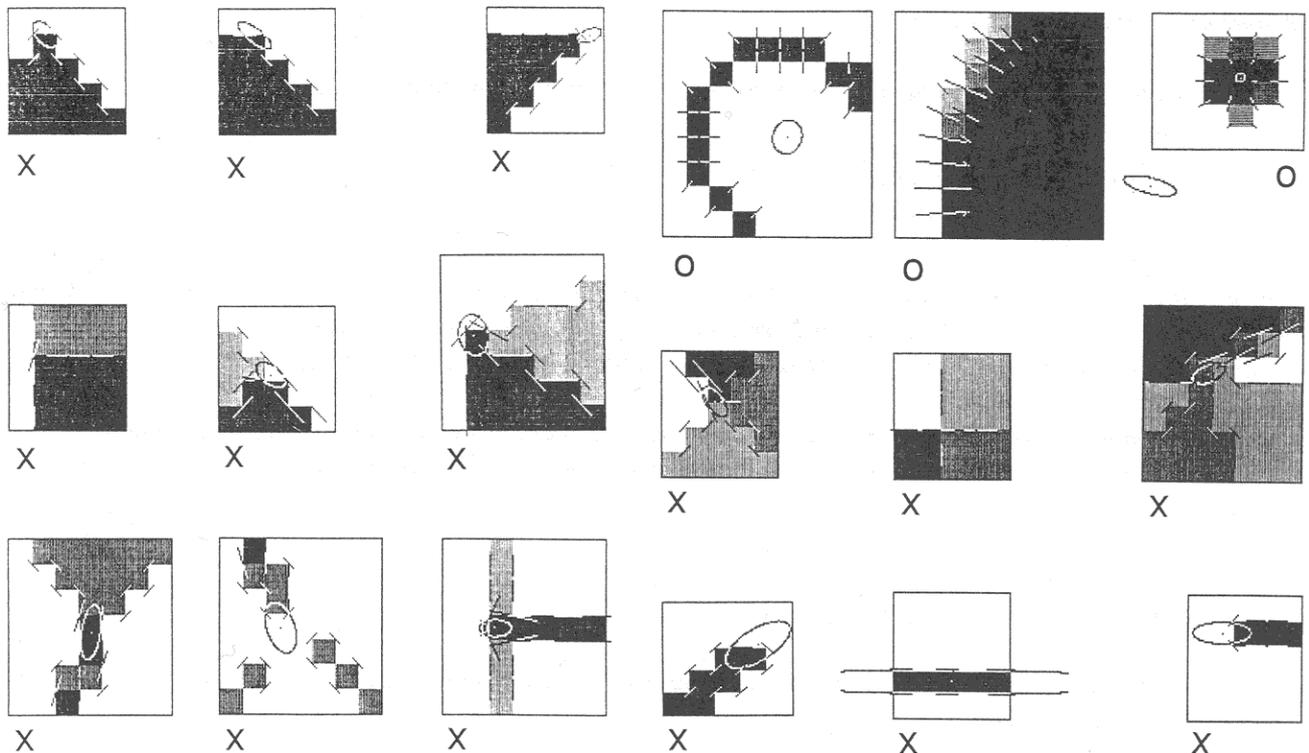
Bevor wir die Genauigkeit der Punktlokalisierung diskutieren, wollen wir die Ergebnisse mit denen des vorigen Kapitels vergleichen. Hier ist die Identität der Normalgleichungsmatrix der Aufgaben a., b. und c. mit der Hesse-Matrix bei der Texturanalyse von besonderer Bedeutung.

Wie die Kennzeichnung der Elemente der Normalgleichungsmatrix verdeutlicht, nehmen wir hier an, daß uns die Gradienten des Signals bekannt sind. Wir sind also nicht an einer stochastischen Modellierung der Bildfunktion f interessiert. Selbstverständlich benötigen wir bei der praktischen Bestimmung der Punktkoordinaten eine Schätzung für das Signal, welche wir mit Hilfe eines beliebigen Restaurationsfilters gewinnen können, etwa auch mit dem in Abschnitt 2.1.5 vorgeschlagenen informationserhaltenden Filters. Danach postulieren wir aber die Funktionswerte f als fehlerfrei.

Damit sind alle vier genannten Modelle Näherungen, denn sie berücksichtigen nicht die stochastischen Eigenschaften der Gradienten ∇f . Vielmehr wird statt dessen die Position des Bildelementes als stochastisch angenommen, wie etwa die Modellgleichungen für den gewogenen Schwerpunkt zeigen. Welchen Einfluß diese Approximation auf das Ergebnis hat, ist noch zu untersuchen. Wir werden aber auf dieses Rechenmodell bei der Klassifizierung der selektierten Fenster zurückkommen.

Davon unabhängig wird die Interpretation der Hesse-Matrix durch die Interpretation der Eigenschaften der geschätzten Punktkoordinaten gestützt.

Fig 2-5 Markante Punkte in selektierten Fenstern mit Kantenelementen und 99 %-Konfidenzellipse, simulierte Daten (aus FÖRSTNER 1986)
x = Ecken aus Kantenelementen,
o = Kreise aus Gefällelementen



2.2.2.3 Genauigkeitseigenschaften der Punkte

Wir wollen nun an Hand von Gl. 222-15 die Genauigkeitseigenschaften des geschätzten Punktes unter der Bedingung diskutieren, daß wir es nur mit einer der 4 Aufgabenstellungen zu tun haben, etwa der Schwerpunktbildung. Da wir annehmen können, daß die Rauschvarianz im Bild konstant ist, enthält die Normalgleichungsmatrix die für die Genauigkeit entscheidende Information. Sie kann, wie die Hesse-Matrix durch drei Kenngrößen charakterisiert werden:

1. Die Größe der Fehlerellipse

Die Länge der Halbachsen der Fehlerellipse sind umgekehrt proportional zu den Eigenwerten von N. Das Gewicht des Punktes kann durch

$$w = 1/\text{tr } N^{-1} = \det N / \text{tr } N \quad (222-19)$$

angegeben werden. Daher kann das Gewicht eines Punktes auf einfache Weise, insbesondere ohne Inversion, aus der Spur und der Determinante von N ermittelt werden

2. Die Richtung der großen Halbachse

Die Richtung der großen Halbachse der Fehlerellipse kann man aus

$$\tan 2\phi = 2 N_{12} / (N_{11} - N_{22}) \quad (222-20)$$

erhalten. Wenn durch das Fenster eine Bildkante verläuft, ist ϕ die Richtung dieser Kante (s. die Diskussion in Abschn. 2.1.3.)

3. Die Form der Fehlerellipse

Wie oben können wir wenigstens zwei Maße für die Form der Fehlerellipse verwenden. Das Verhältnis der Halbachsen der Fehlerellipse ist identisch mit der Wurzel aus dem Verhältnis der Eigenwerte von N oder von N^{-1} .

Um eine Eigenwertberechnung zu umgehen, kann man auch das Rundheitsmaß

$$\begin{aligned} q &= 1 - p \\ &= 1 - [(\delta_1 - \delta_2)/(\delta_1 + \delta_2)]^2 \\ &= (4 \det N) / \text{tr}^2 N \\ &= 1 - [(\text{SNR}^2 - 1)/(\text{SNR}^2 + 1)]^2 \end{aligned} \quad (222-21)$$

verwenden. Es ist die Ergänzung vom Anisotropiemaß p, das wir in Gl. 213-10 definiert hatten. Beide Maße lassen sich ohne Eigenwertberechnung aus der Spur und der Determinante von N bestimmen. Die Inversion von Gl. 222-21 ergibt $\text{SNR}^2 = (1+\sqrt{1-q})/(1-\sqrt{1-q})$.

Wie oben schon erwähnt, werden wir es bei der Selektion von Punkten dazu verwenden, Punkte auf Kanten auszuschließen, bzw. bei der Kantendetektion dazu benutzen, um markante Punkte zu umgehen.

2.2.2.4 Markante Punkte in Farbbildern

Die Erweiterung der Aufgaben auf Farbbilder ist einfach möglich. Wir erwähnen sie hier um der Vollständigkeit willen.

Statt der skalarwertigen Funktion $f(r,c)$ haben wir nun eine vektorwertige Funktion $f(r,c) = [f_1(r,c), f_2(r,c), \dots, f_k(r,c)]$. Der Gradient ist dann eine Matrix mit den partiellen Ableitungen

$$\nabla f = \begin{bmatrix} f_{1r} & f_{2r} & \dots & f_{kr} \\ f_{1c} & f_{2c} & \dots & f_{kc} \end{bmatrix} \quad (222-22)$$

Jeder Kanal $1, \dots, k$ trägt zur Lösung gemäß seinem Rausch- und seinem Gradientenanteil bei. Die Kovarianzmatrix der Rauschanteile in den k Kanälen sei C_{nn} . Dann müssen wir zur Bestimmung des gewogenen Schwerpunkts die Gewichtsmatrix

$$W = \nabla f (C_{nn})^{-1} \nabla f' \quad (222-23)$$

verwenden. Dies ist eine unmittelbare Verallgemeinerung von Gl. 222-8b. Für $g = f + n$ mit zwei unabhängigen Kanälen, die Rauschvarianzen σ_1^2 und σ_2^2 aufweisen, erhalten wir die Gewichtsmatrix

$$W = \begin{bmatrix} f_{1r}^2/\sigma_1^2 + f_{2r}^2/\sigma_2^2 & f_{1r} \cdot f_{1c}/\sigma_1^2 + f_{2r} \cdot f_{2c}/\sigma_2^2 \\ * & f_{1c}^2/\sigma_1^2 + f_{2c}^2/\sigma_2^2 \end{bmatrix} \\ = (W_1/\sigma_1^2 + W_2/\sigma_2^2) \quad (222-24)$$

Das bedeutet, daß der Kanal mit der größeren Rauschvarianz geringer zur Punktlokalisierung beiträgt. Falls $\sigma_1 = \sigma_2$ gilt, ein Kanal aber kein Signal aufweist, reduziert sich das Gewicht des Punktes auf die Hälfte. Falls es sich um ein Farbbild mit zwei Kanälen handelt, nämlich Intensität und Farbe, tragen Farbänderungen insbesondere dann zur Lokalisierung bei, wenn die Intensität keine Struktur aufweist.

Falls beide Kanäle die gleichen Gradienten aufweisen und die gleiche Rauschvarianz besitzen, aber eine Korrelation von r haben, ist das Gewicht

$$W = W_0 \cdot 2/(1 + r) \quad (222-25)$$

worin W_0 das Gewicht aus Gl. 222-8b ist. Nur wenn die Korrelation der Kanäle gering ist, kann man mit einer Genauigkeitssteigerung durch die Hinzunahme eines Kanals rechnen.

Die Interpretation von Gl. 222-24 führt also auf die erwarteten Schlußfolgerungen.

Wir haben damit die notwendigen Hilfsmittel zur Bestimmung der optimalen Position eines markanten Punktes, wenn die Aufgabe spezifiziert und das Fenster, in dem der Punkt liegen soll selektiert ist.

2.2.3 Klassifizierung markanter Punkte

Für die Lokalisierung des markanten Punktes müssen wir entscheiden, ob es sich um eine Ecke, einen Kreis oder eine zufällige, isotrope Textur handelt. Wir schließen beim Kreis auch alle anderen rotationssymmetrischen Figuren mit ein. Als Kriterium steht uns die Anpassung der Bildfunktion an die Modelle Gl. 222-7 und 222-9 zur Verfügung.

Wir wollen daher die Nullhypothese

$$H_0: \text{allgemeine Textur}$$

gegen die Alternativen

$$H_A: \text{Ecke, Modell Gl. 222-7}$$

$$H_B: \text{Kreis, Modell Gl. 222-9}$$

prüfen.

Dazu zeigen wir, daß die Testgröße

$$\underline{T} = \frac{\sigma_A^2}{\sigma_B^2} = \frac{\Omega_A}{\Omega_B} \sim F_{r,r} \quad (223-1)$$

Fisher-verteilt ist. Hierin sind

$$\Omega_A = \underline{e}_A' W \underline{e}_A = \Sigma f_r^2 \underline{r}^2 + 2 \Sigma f_r f_c \underline{r} \underline{c} + \Sigma f_c^2 \underline{c}^2 - \hat{\underline{x}}_A' \underline{h}_A \quad (223-2)$$

und

$$\Omega_B = \underline{e}_B' W \underline{e}_B = \Sigma f_r^2 \underline{c}^2 - 2 \Sigma f_r f_c \underline{r} \underline{c} + \Sigma f_c^2 \underline{r}^2 - \hat{\underline{x}}_B' \underline{h}_B \quad (223-3)$$

von Gl. 222-17 und 222-19 und r die Redundanz.

Dazu vereinheitlichen wir das Rechenmodell und nehmen an, daß die Zentren $(\underline{r}_i, \underline{c}_i)$ der Bildelemente gemäß

$$\begin{bmatrix} \underline{r}_i \\ \underline{c}_i \end{bmatrix} \sim N \left[E \begin{bmatrix} \underline{r}_i \\ \underline{c}_i \end{bmatrix}, \frac{\sigma_n^2}{|\nabla g|^2} I_2 \right] \quad (223-4)$$

gemeinsam normalverteilt sind. Dann sind die in den Gl. 222-17 und 222-19 auftretenden Beobachtungen $\underline{l}_A(r_i, c_i)$ und $\underline{l}_B(r_i, c_i)$ ebenfalls normalverteilt und genügen der Verteilung von Gl. 223-4. Denn sie gehen als senkrechte Abstände vom Ursprung von der Kanten-Gerade bzw. der Gefäll-Gerade aus den Koordinaten hervor. Da die beiden Geraden aufeinander senkrecht stehen, sind wegen der Isotropie der Fehlerellipsen der Bildpunkte auch die Abstände unabhängig und gleich genau (vgl. Fig. 2-34) und haben mit $\sigma_0^2 = \sigma_n^2$ das in Gl. 222-7b und 222-9b repräsentierte Gewicht.

Nun lassen sich die quadratischen Formen Ω_A und Ω_B auch in Abhängigkeit von den Beobachtungen darstellen. Z. B. gilt mit der Projektionsmatrix

$$P_A = I - A (A' W_A A)^{-1} A' W_A \quad (223-5)$$

$$\underline{e}_A = - P_A \underline{l}_A, \quad (223-6)$$

daher

$$\underline{\Omega}_A = \underline{e}_A' W_A \underline{e}_A = \underline{1}_A' P_A' W_A P_A \underline{1}_A = \underline{1}_A' M_A \underline{1}_A \quad (223-7)$$

und daher analog

$$\underline{\Omega}_B = \underline{e}_B' W_B \underline{e}_B = \underline{1}_B' P_B' W_B P_B \underline{1}_B = \underline{1}_B' M_B \underline{1}_B \quad (223-7)$$

Da aber beide quadratische Formen X_r^2 -verteilt und wegen der Orthogonalität von $\underline{1}_A$ und $\underline{1}_B$ selbst unabhängig sind, ist die Testgröße T in Gl. 223-1 Fisher-verteilt.

Mit zwei kritischen Werten k_1 und k_2 können wir also folgende Klassifizierung vornehmen

$$\begin{aligned} T > k_1 &\longrightarrow \text{Kreis} \\ T < k_2 &\longrightarrow \text{Ecke} \\ \text{sonst} &\longrightarrow \text{isotrope Textur} \end{aligned} \quad (223-8)$$

Für eine sichere Entscheidung sollten die Werte $1/k_1$ und k_2 bzw. das dazugehörige Signifikanzniveau groß gewählt werden, z. B. 99.9 %.

Wir wollen nun zeigen, daß das Verfahren sich auf Fenster verallgemeinern läßt, die eine spiralförmige Textur enthalten. Die beiden Modelle A und B haben gemeinsam, daß der Winkel zwischen der Kantenrichtung und der Richtung zum gesuchten Punkte konstant ist; für die Ecke ist dieser Winkel 0, für die rotationssymmetrische Textur 90°. Damit erfüllen die Kantenelemente bei A die Differentialgleichung

$$\frac{dr}{dc} = \frac{r}{c} \quad (223-9)$$

während für den Fall B gilt

$$\frac{dr}{dc} = -\frac{c}{r} \quad (223-10)$$

Mit der Rotationsmatrix R_{β} (wie in Gl. 222-8c) können wir beide Differentialgleichungen auch als

$$\left| \begin{array}{c} dr \\ dc \end{array} R_{\beta_i} \cdot \begin{bmatrix} r \\ c \end{bmatrix} \right| = 0 \quad (223-11)$$

mit $\beta_A = 0$ und $\beta_B = 90^\circ$ darstellen.

Damit können wir aber den Anwendungsbereich des Operators auf Fenster erweitern, in denen Kantenelemente der Differentialgleichung 223-11 mit beliebigem Winkel β genügen. Dies sind im Allgemeinfall logarithmische Spiralen, von denen Kreise und Ecken einen Spezialfall darstellen. Auch die Eigenwerte der Normalgleichungsmatrix sind unabhängig vom Winkel β . So können wir auch zwischen zwei spiralförmig texturierten Fenstern mit Hilfe der o. g. Testgröße entscheiden, wenn die zugehörigen Winkel sich um 90° unterscheiden. Schließlich können wir statt der Klassifizierung eine Schätzung des Winkels β aus $\underline{\Omega}(\beta)$ vornehmen, wozu 3 Werte für $\underline{\Omega}(\beta)$ genügen. Eine kompakte, unabhängig entwickelte Schätzung von Spiralparametern hat kürzlich BIGÜN (1990) angegeben. Eine mögliche Anwendung, in der diese Erweiterung benötigt werden könnte, wäre etwa die Lokalisierung der Zentren von Wirbelfeldern (vgl. BOUTHEMY und BENVENISTE 1985) ohne vorherige Analyse ihrer Struktur, etwa ihres Drehsinns.

Damit haben wir die Hilfsmittel für die Analyse der extrahierten Fenster bereitgestellt. Im nächsten Abschnitt diskutieren wir die Selektion der Fenster, die unmittelbar auf die Merkmalskarten zurückgreift.

2.2.4 Selektion optimaler Fenster für die Punktauswahl

Der Operator soll Punkte finden, die in einer spezifischen Weise optimal sind. Die Hauptanforderung ist Deutlichkeit oder Markantheit der Punkte. Die Punkte sollen sich leicht von ihrer Umgebung unterscheiden lassen. Die folgenden beiden Kriterien stützen sich auf die erwartete Präzision der Zuordnung, der Ecken- oder Kreispositionierung. Wie wir im vorletzten Abschnitt zeigten, sind die Halbachsen der Fehlerellipsen identisch für alle der angesprochenen Aufgaben, auch der Lokalisierung des Zentrums einer Spirale und proportional zu der Wurzeln aus den Eigenwerten der Hesse-Matrix aus denen wir die Merkmalskarten ableiteten. Ohne die Aufgabe spezifizieren zu müssen, können wir daher fordern:

C1: Die Fehlerellipse sollte nahezu kreisförmig sein.

C2: Die Fehlerellipse sollte klein sein, insbesondere kleiner als diejenige, die aus Nachbarbildfenstern abgeleitet wird.

Die Maße für beide Kriterien sollten sich einfach aus der Bildfunktion ableiten lassen, da sie für jede Stelle im Bild berechnet werden müssen. Die Maße q (Gl. 222-21) und w (Gl. 222-19) für die Rundheit und die Größe der Fehlerellipse, hier das Gewicht, des Punktes, erfüllen die Anforderung, da sie auf einfache Weise aus den drei Elementen der Normalgleichungs- bzw. Hesse-Matrix ableitbar sind. Bis auf die Determinante bei der Berechnung des Gewichts stimmen sie mit den beiden ersten Merkmalen der in Kapitel 2.1 diskutierten Merkmalskarten überein.

Die Selektion der optimalen Fenster läßt sich daher in folgenden Schritten erreichen:

1. Bestimmung der Elemente von N

Dies sind drei Faltungen, nämlich der drei abgeleiteten Bilder $g_r^2(r,c)$, $g_c^2(r,c)$ and $g_r(r,c) \cdot g_c(r,c)$.

In den Gleichungen der Normalgleichungs- bzw. der Hesse-Matrix haben wir ein Boxfilter für diese Faltungen formuliert. Die Größe dieses Faltungskerns ist eine freier Parameter. Wir haben bei der Bildzuordnung gute Erfahrungen mit Fenstern von 5x5 bzw. 7x7 Bildelementen gemacht. Für eine bessere Lokalisierung der Fenster, insbesondere in der Nähe von stufenartigen Kanten, etwa an den Ecken eines Schachbretts, ist ein Dreiecksfilter oder ein Gauß-Filter günstiger. In diesem Fall ist neben der Größe des Faltungskerns die Streuung des Gauß-Kerns zu spezifizieren.

2. Bestimmung der abgeleiteten Bilder $q(r,c)$ und $w(r,c)$ aus Gl. 222-19 und 222-21. Um Rechenzeit zu sparen, kann man statt der Gewichte auch nur die Spur der Hesse-Matrix verwenden.

3. Bestimmung der Punkt-Gebiete durch Schwellwertbildung

Dies sind alle Stellen im Bild, an denen potentiell ein optimales Fenster liegen kann: .

$$w^*(r,c) = \begin{cases} w(r,c) , & \text{falls } q(r,c) > q_{\min} \\ & \text{und } w(r,c) > w_{\min} \\ 0 & , \text{sonst} \end{cases} \quad (224-1)$$

Die untere Grenze q_{\min} für die Rundheit dient dazu, Fenster auszuschließen, in denen eine gerichtete Textur liegt, etwa Fenster auf Kanten. Es ist ein freier Parameter in diesem Schritt. Das Ergebnis ist nicht sehr empfindlich gegen Veränderungen dieses Schwellwerts. Die Experimente legen einen Schwellwert zwischen 0.5 und 0.75 nahe; das entspricht Verhältnissen von $\sqrt{3}$ und 2 der Halbachsen der Fehlerellipse.

Der Schwellwert w_{\min} für das Gewicht dient dazu, Fenster in flachen Teilen des Bildes auszuschließen. Daher ist es günstig, diesen Schwellwert an das Rauschen im Bild anzupassen. Eine robuste Schätzung für die Rauschvarianz σ_n^2 haben wir in Abschn. 2.1.5 im Zusammenhang mit dem informationserhaltenden Filter entwickelt. Falls ein $m \times m$ Boxfilter im ersten Schritt verwendet wurde, ist ein Schwellwert von $k \cdot 4 \cdot m^2 / \sigma_n^2$ mit einem geeigneten kritischen Wert k , etwa $k = 3$, günstig. Liegt die Schätzung der Rauschvarianz nicht vor, kann man auch unmittelbar aus dem Gewichtsbild $w(r,c)$, etwa über das Median w_{med} , einen gegen Textur robusten Schwellwert ableiten, z. B. $w_{\min} = k \cdot w_{\text{med}}$. Mit kleineren Schwellwerten q_{\min} und w_{\min} erhält man insgesamt mehr selektierte Fenster.

4. Unterdrückung aller Nicht-Maxima in w^*

Hier werden alle Bildelemente des Punkt-Gebiets-Bildes $w^*(r,c)$ Null gesetzt, an denen kein relatives Maximum in w^* vorliegt.

Die Größe des Fenster, das für die Bestimmung der relativen Maxima verwendet wird ist ein freier Parameter. Bei dem kleinstmöglichen Fenster von 3×3 liegt zwischen den Mittelpunkten selektierter Fenster ein Abstand von einem Bildelement. Größere Abstände erreicht man durch die Wahl eines größeren Fensters bei dieser Nicht-Maximum-Unterdrückung.

5. Selektion aller Fenster, deren Mittelpunkt $w^*(r,c)$ nicht Null sind.

Die algorithmische Lösung und die mögliche Realisierung auf einem Pipelineprozessor, die eine Realtimeselektion erlaubt, findet sich in (FÖRSTNER und GÜLCH 1986).

Bevor wir an einem Beispiel die Selektion zeigen, wollen wir auf die Beziehung zum Eckenfinder von DRESCHLER (1981) und NAGEL/ENKELMANN (1986) hinweisen. Die wesentliche Eigenschaft der hier vorgeschlagenen Methode zur Fensterselektion folgt aus der Maximierung von $g_r^2 + g_c^2$, wenn wir vorübergehend die Faltung und die Division durch die Determinante vernachlässigen. Diese Minimierung entspricht der Lösung des homogenen Gleichungssystems:

$$\begin{bmatrix} g_{rr} & g_{rc} \\ g_{cr} & g_{cc} \end{bmatrix} \cdot \begin{bmatrix} g_r \\ g_c \end{bmatrix} = K \cdot v_g = 0$$

worin K die Hesse-Matrix der Bildfunktion darstellt. Für $g_r^2 + g_c^2 > 0$ führt dies zur Forderung $|K| = 0$. Nach geeigneter Rotation des Koordinatensystems geht das Gleichungssystem in

$$\begin{bmatrix} g_{uu} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ g_v \end{bmatrix} = 0$$

über. Es zeigt, daß die zweite Ableitung g_{vv} in v -Richtung Null ist, also ein Wendepunkt in der Richtung größten Gefälles auftritt, während die Krümmung senkrecht dazu maximal ist. Dies sind aber gerade die Bedingungen, die DRESCHLER und NAGEL für ihren Eckensucher fordern. Für ein Maximum von $g_r^2 + g_c^2$ muß auch noch das Produkt $g_{vv} \cdot g_v$ negativ sein. Dies wird meistens bei der Anwendung von Nulldurchgängen zweiter Ableitungen zur Kantendetektion vernachlässigt.

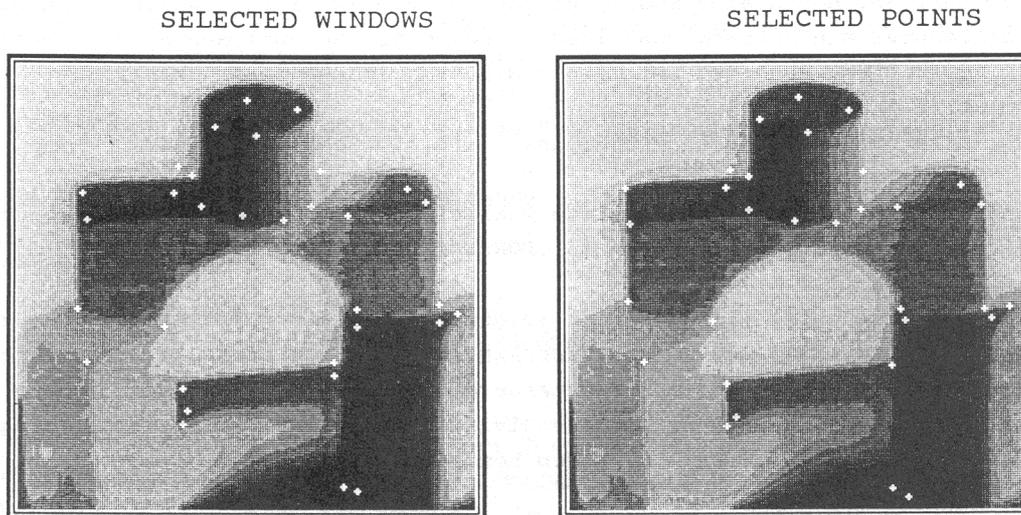
Es bestehen aber zwei wesentliche Unterschiede zwischen den beiden Operatoren:

1. Unsere Entwicklung modelliert die Bildfunktion nicht explizit, und kann daher alle Typen von markanten Punkten verarbeiten, während der Operator von DRESCHLER und NAGEL auf Ecken mit zwei Kanten beschränkt ist.

2. Die genaue Bestimmung des Eckpunktes kann sich bei unserem Operator auf größere Fenster stützen, während sie bei DRESCHLER und NAGEL die Verwendung von 3x3-Fenstern verlangt. Der Grund für den Unterschied ist die Unverzerrtheit unseres Schätzers für den markanten Punkt.

Fig. 2-6 zeigt die Mitten der 32 selektierten Fenster eines Spielzeug-Bildes und die lokalisierten Punkte. Der Vergleich der beiden Bilder zeigt, daß die lokalisierten Punkte wesentlich genauer auf den Ecken sitzen als die Mitten der Fenster. Man beachte, daß die scheinbare Ecke zwischen dem Spielzeug und der Hintergrundkante zu zwei selektierten Fenstern führte, die aber dann zu einem markanten Punkte verschmolzen wurden.

Fig. 2-6 Position der Mitten selektierter Fenster und markanter Punkte
(aus FÖRSTNER/GÜLCH 1987)



2.2.5 Extraktion von Bildkanten

Die in den vorigen Abschnitten diskutierte Interpretation der Merkmalskarten wollen wir nun in Hinblick auf die Extraktion von Kanten erweitern. Während wir bisher Fenster mit starker, isotroper Textur analysierten, geht es nun um Fenster mit starker, aber anisotroper Textur. Dabei wollen wir uns auf den Fall beschränken, daß das Bild aus untexturierten Gebieten besteht, deren Gebietsgrenzen wir lokalisieren wollen.

Das Verfahren ist dreistufig, wobei die ersten beiden Stufen, nämlich die Selektion von Fenstern und die Lokalisierung der Kantenelemente, sich durch geringfügige Modifikation aus dem Interestoperator ergeben. Im dritten Schritt müssen dann die extrahierten Kantenelemente verknüpft werden.

Für eine Reihe von Aufgaben, wie die Bildzuordnung oder die Objekterkennung, sind gerade Kantenstücke von besonderem Interesse, vor allem, wenn es sich um Bilder künstlicher Objekte handelt, wie Gebäude oder Maschinenteile. Wir werden daher die Extraktion gerader Kantenstücke gesondert behandeln.

Ziel dieses Abschnitts ist es, einerseits die Komplementarität bei der Interpretation der Merkmalskarten bzgl. der Extraktion von Punkten und Kantenelementen aufzuzeigen, andererseits die Herleitung der stochastischen Eigenschaften der symbolischen Bildbeschreibung durch gerade Kantenstücke aus dem Extraktionsprozeß selbst zu demonstrieren. Die abschließenden Beispiele sollen noch deutlicher den Unterschied des informationserhaltenden Filters und des Medianfilter bei der Kantenextraktion zeigen und so die Ergebnisse von Abschnitt 2.1.5 bestätigen.

2.2.5.1 Selektion optimaler Fenster für die Kantenextraktion

Im ersten Schritt werden die für die Kantenlokalisierung optimalen Fenster selektiert. Das in Abschnitt 2.2.4 beschriebene Verfahren muß dazu nur unwesentlich verändert werden:

1. Bestimmung der Elemente der Normalgleichungsmatrix N . Auch hier ist wieder die Art und Größe des Faltungskerns zu spezifizieren. Um eine genügend hohe Auflösung zu erhalten, d. h. benachbarte Kanten trennen zu können, sind hier i. a. kleinere Fenster als beim Interestoperator zu verwenden, d. h. mit 3×3 bis maximal 5×5 Bildelementen.
2. Bestimmung der abgeleiteten Bilder $q(r,c)$ und $w(r,c)$, auch hier evtl. statt w die Spur der Normalgleichungsmatrix.
3. Bestimmung der Kantengebiete durch Schwellwertbildung. Dies sind alle Stellen im Bild, an denen potentiell ein optimales Fenster für die Kantenextraktion liegen kann:

$$w^*(r,c) = \begin{cases} w(r,c) & , \text{ falls } q(r,c) < q_{\max} \\ & \text{ und } w(r,c) > w_{\min} \\ 0 & , \text{ sonst} \end{cases} \quad (225-1)$$

Die obere Grenze für q soll garantieren, daß keine Fenster an markanten Punkten selektiert werden. Der Schwellwert für w kann wieder in Abhängigkeit vom geschätzten Rauschen oder vom Median der Gewichte gewählt werden.

4. Unterdrückung aller Nicht-Maxima von w^* . Während beim Interestoperator alle Bildelemente unterdrückt werden, die nicht relative Maxima sind, werden hier alle Bildelemente angehalten, die in Zeilen- oder Spaltenrichtung ein relatives Maximum in $w(r,c)$ aufweisen. Auch hier ist die Fenstergröße für die Unterdrückung von Nicht-Kantenelementen zu spezifizieren. Sie sollte klein gewählt werden, z. B. 3×1 bzw. 1×3 Bildelemente, um eine genügende Auflösung zu erreichen.

5. Selektion aller Fenster, deren Mittelpunkt $w^*(r,c)$ nicht Null ist.

Als algorithmische Lösung kann diesselbe wie für den Interestoperator verwendet werden.

2.2.5.2 Lokalisierung von Kantenelementen

Im Prinzip kann man die optimale Lage der Kante innerhalb eines selektierten Fensters aus dem Normalgleichungssystem Gl. 222-11 ableiten. Das Kantenelement wird dann durch den Schwerpunkt

und die Richtung der großen Halbachse der Fehlerellipse repräsentiert. Einer unmittelbaren Übernahme von Gl. 222-11 für die Bestimmung der Lage des Kantenelements stehen allerdings zwei Schwierigkeiten im Weg:

1. Das Normalgleichungssystem kann singular sein. Die Lösung des Systems führt zwar auf die gesuchte Gerade, verlangt aber eine abweichende algorithmische Lösung.
2. Der Schwerpunkt kann weit außerhalb des Bildfensters liegen. Dies betrifft zwar nicht die Lage der Gerade, aber die Repräsentation des Kantenelements, die dadurch nicht unmittelbar interpretiert werden kann.

Beide Probleme resultieren aus der Unbestimmtheit der Kantenposition entlang der Kante. Sie kann durch Hinzufügen eines (fingierten) Gefälleelementes bei der Schätzung der Position des Kantenelementes behoben werden.

Dazu zerlegen wir die Normalgleichungsmatrix mit der orthogonalen Matrix C und der Diagonalmatrix D

$$N = C' D C = c_1 \delta_1 c_1' + c_2 \delta_2 c_2' , \quad (225-2)$$

wobei $\delta_1 \gg \delta_2$ gilt. Das Gefäll-Element führt dann auf die Fehlergleichungen

$$\begin{aligned} r_m + e_r &= \hat{r}_0 \\ c_m + e_c &= \hat{c}_0 \end{aligned} \quad W = c_2 \cdot k \delta_1 \cdot c_2' \quad (225-3)$$

worin (r_m, c_m) der Mittelpunkt des Fensters und k eine frei wählbare Konstante ist, die günstigerweise klein gewählt wird, z. B. $k = 0.1$ oder 0.01 . Damit wird das Normalgleichungssystem zu

$$(N + c_2 k \delta_1 c_2') \begin{bmatrix} \hat{r}_0 \\ \hat{c}_0 \end{bmatrix} = \underline{h} + c_2 k \delta_1 c_2' \begin{bmatrix} r_m \\ c_m \end{bmatrix} \quad (225-4)$$

Seine Kondition ist besser als $1/k$. Der auf diese Weise geschätzte Punkt (\hat{r}_0, \hat{c}_0) entspricht dem Fußpunkt des Lotes vom Mittelpunkt des Fensters auf die Kante. Die Genauigkeit dieses Lotfußpunktes, insbesondere die Präzision quer zur Kante, kann man in analoger Weise aus der Inversen N^{-1} und einer Schätzung des Varianzfaktors σ_0^2 ableiten. Für die Standardabweichung quer zur Kante ergibt sich bei gerader Kante mit $\delta_1 \approx \text{tr } N$

$$\sigma_q \approx \frac{\sigma_n}{\sqrt{\text{tr} N}} = \frac{\sigma_n}{\sqrt{\sum_i |v f_i|^2}} \quad (225-5)$$

Die Interpretation der Merkmale bei Kantenelementen erfolgte ausführlich in Abschnitt 2.1.3.

Obwohl die Konzeption der Selektionsverfahren für Punkte und Kantenelemente identisch ist, besteht doch ein wesentlicher Unterschied bei der praktischen Implementierung. Die verwendeten Kerne für die Bestimmung der Elemente der Normalgleichungsmatrix sind i. a. verschieden. Für die Punktauswahl benötigt man Fenster mit wenigstens 5×5 Bildelementen, während für die Kantenextraktion Fenster mit 3×3 Bildelementen günstiger sind. Dies ist verständlich, da z. B. die horizontale Position der linken oberen Ecke eines Rechtecks nur durch die Bildelemente unterhalb der Ecke bestimmt sind, also $m-2$, während die horizontale Position einer vertikalen Kante m Bildelemente innerhalb eines Fensters von $m \times m$ Bildelementen beeinflusst (vgl Fig. 2-7)

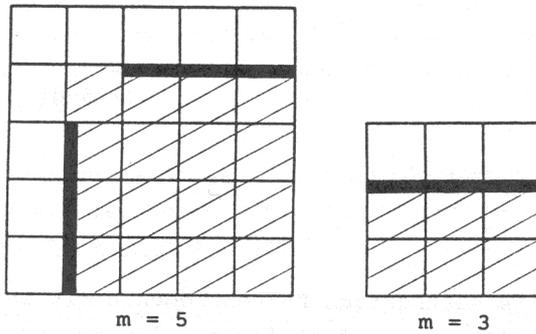


Fig. 2-7
Für die Position von Ecken und Kanten
maßgebliche Stellen innerhalb eines
Fensters

Damit stellen Punkt- und Kantenextraktion nicht Interpretation derselben Merkmalskarten dar. Dies legt ein zweistufiges Vorgehen bei der Interpretation nahe:

1. Extraktion markanter Punkte mit Kern 1 (etwa 3 x 3)
2. Extraktion von Kanten mit Kern 2 (etwa 5 x 5) unter Umgehung bereits selektierter markanter Punkte.

Wir werden diesen Problemkreis nicht weiter verfolgen (vgl. aber VOSSELMAN 1988).

2.2.5.3 Verknüpfung der Kantenelemente zu geraden Kantenstücken

Das Ergebnis der vorigen Schritte besteht in einer Liste von Kantenelementen mit Attributen, insbesondere der Kantenstärke, der Kantenrichtung und der Anisotropie. Sie stellen, wie die Liste extrahierter markanter Punkte, eine symbolische Beschreibung des Bildes dar. Die Elemente sind allerdings noch nicht untereinander verknüpft. Wir wollen nun zu einer kompakteren symbolischen Repräsentation des Bildinhaltes gelangen, indem wir die Kantenelemente zu geraden Kantenstücken zusammenfassen.

Das im folgenden beschriebene Verfahren arbeitet sequentiell auf den bezüglich einer 8-Nachbarschaft zusammenhängenden Komponenten der extrahierten Kantenelemente in der Art eines sog. Bereichswachstumsverfahrens (engl. region growing), wie es bei der Segmentierung gebräuchlich ist.

Gerade Kantenstücke werden während des Verfahrens durch den Schwerpunkt

$$S = \begin{bmatrix} r_s \\ c_s \end{bmatrix} = \begin{bmatrix} \sum_i r_i w_i / \sum_i w_i \\ \sum_i c_i w_i / \sum_i w_i \end{bmatrix}, \quad (225-6)$$

das zweite zentrale Moment

$$M = \begin{bmatrix} \sum_i (r_i - r_s)^2 w_i & \sum_i (r_i - r_s)(c_i - c_s) w_i \\ \sum_i (r_i - r_s)(c_i - c_s) w_i & \sum_i (c_i - c_s)^2 w_i \end{bmatrix} \quad (225-7)$$

und das umschließende Rechteck mit der linken oberen Ecke R_A und der rechten unteren Ecke R_E

$$R_A = \begin{bmatrix} \min_i r_i \\ \min_i c_i \end{bmatrix}$$

(225-8)

$$R_E = \begin{bmatrix} \max_i r_i \\ \max_i c_i \end{bmatrix}$$

repräsentiert. Hierin sind die Summen über die bereits verknüpften Kantenelemente mit ihren Koordinaten (r_i, c_i) zu nehmen. Diese Repräsentation läßt sich auf einfache Weise rekursiv aufbauen. Die Gewichte w_i ergeben sich aus der Kantenstärke und sind proportional zu $\text{tr}N_i$ (vgl. Fig. 2-8).

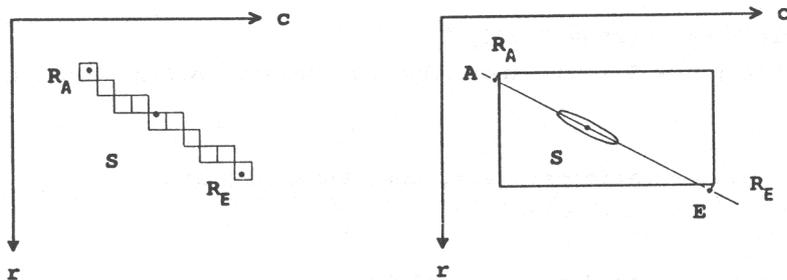


Fig. 2-8 Repräsentation eines geraden Kantenstücks während der Verknüpfung der Kantenelemente

Die Geradengleichung ergibt sich aus dem Schwerpunkt der berechneten Kantenposition und der Hauptrichtung des zweiten zentralen Moments. Die Endpunkte A und E der Kante sind die Lotfußpunkte der Ecken des umschreibenden Rechtecks auf diese Gerade.

Der Prozeß der Eingliederung neuer Kantenelemente in das Kantenstück wird abgebrochen, wenn kein benachbartes Kantenelement mehr die Abstands- und die Richtungsbedingung erfüllt. Das Verfahren wird an demjenigen Kantenelement neu begonnen, das von den übrigbleibenden die größte Kantenstärke, oder alternativ die größte Anisotropie, aufweist. Das Ergebnis ist eine Liste von Kantenstücken, die durch Anfangs- und Endpunkt repräsentiert werden.

Für eine Weiterverarbeitung benötigen wir die Genauigkeitseigenschaften der Kantenstücke, d. h. die Kovarianzmatrix der vier Koordinaten. Um sie zu bestimmen, legen wir vorübergehend das Koordinatensystem (u, v) näherungsweise in den Schwerpunkt des Kantenstücks mit der u-Achse näherungsweise entlang der Kante. Die Geradengleichung ergibt sich dann zu

$$v = a + m u$$

mit den Näherungswerten $a = m = 0$. Die Kleinste-Quadrate-Schätzungen für a und b sind unabhängig und haben die Varianzen

$$\hat{\sigma}_a^2 = \hat{\sigma}_0^2 / \Sigma w_i \quad (225-9)$$

$$\hat{\sigma}_m^2 = \hat{\sigma}_0^2 / \Sigma w_i u_i^2 \quad (225-10)$$

mit

$$\hat{\sigma}_0^2 = \Sigma w_i e_i^2 / (n - 2) \quad (225-11)$$

Nun sind die Summen $\sum w_i u_i^2$ und $\sum w_i e_i^2$ identisch mit dem größten und kleinsten Eigenwert von M , und $\sum w_i$ wird für die Berechnung der Schwerpunkte benötigt, so daß diese Werte bei der Ermittlung der Kantenstücke entstehen und daher keine erneute Schätzung der Geradenposition verlangen.

Schließlich benötigen wir noch die Genauigkeit des Anfangs- und Endpunkts in u -Richtung. Sie hängt von der Weiterverarbeitung der Kante ab, und wir können drei Fälle betrachten, die sich in der Bedeutung der geschätzten Endpunkte unterscheiden:

- a. Wenn wir annehmen, daß die geschätzten Endpunkte nur durch Rundungsfehler verfälscht sind, können wir die Varianzen zu

$$1/12 \leq \sigma_{u_A}^2 = \sigma_{u_E}^2 \leq 1/6 \quad (225-12)$$

wählen. Dies entspricht der internen Genauigkeit der Bestimmung der Maximal- bzw. Minimalwerte in R_A und R_E .

- b. Wenn wir annehmen können, daß die Endpunkte des extrahierten Kantenstücks durch einen Knick der Gebietsgrenze entstanden sind, müssen wir berücksichtigen, daß wir wegen der Prüfung auf Anisotropie den Knickpunkt nicht ganz erreichen. Daher ist das errechnete Kantenstück systematisch um einen Betrag zu kurz, der etwa zwischen 1 und 2 Bildelementen liegt:

$$\begin{aligned} \underline{u}_{A,\text{gemessen}} &= \underline{u}_{A,\text{theoretisch}} + \underline{b} + \underline{e} \\ \underline{u}_{E,\text{gemessen}} &= \underline{u}_{E,\text{theoretisch}} - \underline{b} + \underline{e} \end{aligned} \quad (225-13)$$

mit dem Bias

$$1 \leq E(\underline{b}) \leq 2 \quad (225-14)$$

Wegen der Unsicherheit des Bias, muß man die Varianzen größer als in Gl. 225-12 annehmen, also z. B.

$$1/6 \leq \sigma_{u_A}^2 = \sigma_{u_E}^2 \leq 1 \quad (225-15)$$

- ∴ Wenn man schließlich annehmen muß, daß die wahre Kante wegen Störungen sich aus mehreren Kantenstücken zusammensetzt, muß man ebenfalls mit einem Bias rechnen. Da er unbekannt ist, führt er auch zu einer großen Varianz. Damit gilt ebenfalls das Modell Gl. 225-13 mit

$$1 \leq E(\underline{b}) \leq b_{\max} \quad (225-16)$$

Die Varianzen können von b abhängig gewählt werden, also z. B.

$$\sigma_{u_A}^2 = \sigma_{u_E}^2 = k_0^2 + (k_1 b)^2 \quad (225-17)$$

mit $k_1 \geq 1$. Der maximale Betrag b_{\max} für den Bias in Gl. 225-17 ist abhängig von der Anwendung. Falls die Endpunkte vollständig unsicher sind, kann man $k_0 = \infty$ und $b = 0$ wählen.

Bei praktischen Anwendungen werden Mischformen zwischen b. und c. auftreten, so daß etwa der Punkt A den Anfangspunkt einer Kante und E einen beliebigen Punkt auf der Kante repräsentiert. Die Entscheidung hierüber kann man aber ohne zusätzliches Modell-Wissen nicht treffen. Sie fällt erst während des Zuordnungs- oder Erkennungsprozesses an und kann etwa mit Hilfe von Hypothesentests getroffen werden.

Die Kovarianzmatrix der vier Koordinaten des Kantenstücks k im (u,v) -System sind demnach

$$D(k^{(u,v)}) = D \begin{bmatrix} u_A \\ v_A \\ \hline u_E \\ v_E \end{bmatrix} = \begin{bmatrix} \sigma_{u_A}^2 & 0 & | & 0 & 0 \\ 0 & \sigma_{v_A}^2 & | & 0 & \sigma_{v_A v_E} \\ \hline 0 & 0 & | & \sigma_{u_E}^2 & 0 \\ 0 & \sigma_{v_A v_E} & | & 0 & \sigma_{v_E}^2 \end{bmatrix} \quad (225-18)$$

mit

$$\begin{aligned} \sigma_{v_A}^2 &= \sigma_a^2 + (u_A - b)^2 \sigma_m^2 \\ \sigma_{v_E}^2 &= \sigma_a^2 + (u_E + b)^2 \sigma_m^2 \\ \sigma_{v_A v_E} &= \sigma_a^2 + (u_A - b)(u_E + b) \sigma_m^2 \end{aligned}$$

Mit geeignet gewähltem Varianzfaktor erhalten wir eine Gewichtsmatrix mit derselben Struktur .

$$W(k^{(u,v)}) = W \begin{bmatrix} u_A \\ v_A \\ \hline u_E \\ v_E \end{bmatrix} = \begin{bmatrix} w_{u_A} & 0 & | & 0 & 0 \\ 0 & w_{v_A} & | & 0 & w_{v_A v_E} \\ \hline 0 & 0 & | & w_{u_E} & 0 \\ 0 & w_{v_A v_E} & | & 0 & w_{v_E} \end{bmatrix} \quad (225-19)$$

Für freie Endpunkte A und/oder E kann man hierin die Gewichte w_{u_A} und/oder w_{u_E} gleich Null setzen. Der Rang der Kovarianz- und der zugehörigen Gewichtsmatrix erniedrigt sich dann auf 2 bzw. 3. Schließlich erhalten wir durch Rotation des Koordinatensystems die Kovarianz- bzw. Gewichtsmatrix des Kantenstücks im Bildkoordinatensystem:

$$D(k^{(r,c)}) = R D(k^{(u,v)}) R' \quad (225-20)$$

bzw.

$$W(k^{(r,c)}) = R W(k^{(u,v)}) R' \quad (225-21)$$

mit der regulären Matrix

$$R = \begin{bmatrix} R_{\frac{r}{u}} & 0 \\ 0 & R_{\frac{r}{v}} \end{bmatrix} \quad (225-22)$$

mit $R_{\frac{r}{u}}$ aus Gl. 222-8c.

2.2.5.4 Beispiele zur Kantenextraktion

An zwei synthetischen Beispielen wollen wir die Wirkungsweise der Kantenextraktion zeigen und insbesondere den Einfluß verschiedener Vorfilter auf die Kantenextraktion vergleichen.

Fig. A2-31 (s. Anhang 3) zeigt ein schwach verrauschtes Schachbrett mit seinen Merkmalskarten. In der rechten oberen Ecke sind die Kantenbereiche w^* (dunkelrot) und die Mitten der selektierten Fenster von Kantenelementen (hellrot) dargestellt, die man auch als Kantenpixel bezeichnen kann. Diese Kantenpixel entsprechen dem Ergebnis der üblichen Kantendetektoren. Da

die wahre Kante zwischen je zwei Bildelementen liegt, führt bereits geringes Rauschen dazu, daß die Kantenpixel keine glatte Linie darstellen können. Darüber hinaus sind sie auch noch nicht untereinander verknüpft. In Fig. A2-32 sind rechts oben außer den Kantenbereichen und den Kantenpixeln die Kantenelemente (weiß) dargestellt. Dabei wurde die Länge der Kantenelemente proportional zur Kantenstärke gewählt, so daß sich ihre graphischen Darstellungen im Bereich der Schachbrettkanten überlappen und so die Kanten des Schachbretts recht gut wiedergeben. Die Verknüpfung zu geraden Kantenstücken ergibt die Figur in der Mitte oben. Das Schachbrett ist gut repräsentiert. Die kurzen Seiten am Rand wurden vom Algorithmus unterdrückt. Die Ecken wurden erwartungsgemäß ausgespart. Eine Kante wurde nur zum Teil erfaßt.

Fig. A2-33 zeigt das gleiche, schwach verrauschte Schachbrett. Das informationserhaltende Filter führt hier zu einer vollständigen und richtigen Erfassung aller Kanten.

Fig. A2-34 und A2-35 zeigen ein verrauschtes Schachbrett. In der Version ohne Vorfilter fehlt eine Kante ganz, da ihre Teilstücke zu kurz waren und daher unterdrückt wurden, zwei Kanten sind unterbrochen und eine nur teilweise erfaßt. Die informationserhaltende Filterung führt dazu, daß alle Kanten extrahiert werden, wobei lediglich zwei zu kurz sind. Schließlich zeigen Fig. A2-36 und A2-37 ein stark verrauschtes Schachbrett. Es ist aus den extrahierten Kanten kaum erkennbar, wenn nicht gefiltert wird. Dagegen werden nach informationserhaltender Filterung wieder alle Kanten erfaßt, z. T jedoch gestückelt oder zu kurz.

Das zweite Beispiel greift das Streifenmuster auf, das wir zum Vergleich des Median- und des informationserhaltenden Filters verwendet haben. Fig. A2-38 bis A2-41 zeigen die gleichen Muster wie Fig. A2-19-A2-24, allerdings jetzt mit den Kantenelementen (rechts oben) und den geraden Kantenstücken (Mitte oben).

Beim schwach verrauschten Streifenmuster, Fig. A2-38 und A2-39, ist der Unterschied beider Filter gering, das Median schneidet sogar etwas besser ab, da keine Kanten unterbrochen wurden. Dagegen zeigt sich der Qualitätsunterschied bei dem stark verrauschten Streifenmuster deutlich. Das informationserhaltende Filter (Fig. A2-40) führt zu besser ausgerichteten Kantenstücken, als das Medianfilter (Fig. A2-41). Eine wiederholte Anwendung des informationserhaltenden Filters (Fig. A2-42) führt zu einer weiteren Straffung der Kantenstücke, während das Ergebnis nach einer 5 x 5 Medianfilterung (Fig. A2-43) unbrauchbar scheint.

In beiden Beispielen erwartet man aber insgesamt bessere Ergebnisse. Ein Grund dafür ist sicher die Beschränkung des informationserhaltenden Filters auf ein 3 x 3-Fenster. Eine entscheidende Verbesserung würde allerdings Modellwissen bringen, etwa in der Form, daß nach einer ersten Extraktion gerader Kantenstücke Hypothesen über gerade Linien gebildet werden, die eine gezielte Filterung des Bildes erlauben und dann zu wesentlich besser ausgerichteten Kantenelementen führen können.

2.3 Schlußfolgerungen

Dieser Abschnitt hat einen wesentlichen Schritt der Einbildanalyse behandelt: die Extraktion geometrischer Grundelemente oder Primitive, wie sie für eine symbolische Repräsentation des Bildinhalts notwendig wird. Wir haben uns dabei auf markante Punkte und gerade Linienstücke beschränkt. Die Diskussion hat gezeigt, daß wir die Extraktion dieser Primitive als Interpretation von Merkmalskarten deuten können, wie man sie für die Texturanalyse verwendet. Zentrales Anliegen war in allen Fällen eine interne Schätzung der Unsicherheit der extrahierten Primitive zu erhalten. Wir konnten zeigen, daß sie durch Koordinaten und die zugehörige Kovarianzmatrix repräsentiert werden können, und daß sich diese Repräsentation aus der Bildfunktion

ableiten läßt. Damit stehen sie für Folgeaufgaben zur Verfügung, bei denen eine Qualitätsanalyse erforderlich ist, etwa für die Bildzuordnung oder die Objekterkennung.

Die vorgestellten Verfahren enthalten jedoch noch eine Reihe von Näherungen; z. B. werden bei der Extraktion markanter Punkte die Positionen der Kantenelemente stochastisch und die Richtung fest angenommen, während eine umgekehrte Zuordnung der stochastischen Eigenschaften der Realität näher kommt. Daher sind zum einen theoretische Untersuchungen notwendig, die den Einfluß dieser Näherungen auf die geometrischen Eigenschaften der Primitive untersuchen, zum anderen müssen empirische Untersuchungen zeigen, inwieweit die internen Genauigkeitsabschätzungen realistisch sind. In diesem Zusammenhang sollten auch andere Verfahren, etwa zur Kantensextraktion mituntersucht werden. Z. B. lassen die Verfahren von NEVATIA und BABU (1980) und von HANSON et. al. (1986) eine interne Schätzung der Genauigkeit der extrahierten Kantenstücke zu, die empirisch überprüft werden sollte. Darüber hinaus sind Fragen nach der Auflösung der Kantendetektion zu stellen, insbesondere in Bildern mit Kanten verschiedener Schärfe.

Auf der anderen Seite sind die vorgeschlagenen Texturmerkmale bei weitem nicht ausreichend, um komplexe Texturen repräsentieren zu können. In wenigstens drei Richtungen lassen sich die vorgeschlagenen Merkmale ergänzen bzw. erweitern:

- durch Verwendung von Bildpyramiden für die Bildfunktion (g) und das daraus abgeleitete Gradientenbild ($g_x^2, g_x g_y, g_y^2$)
- durch rekursive Texturanalyse der Merkmalskarten, wie etwa bei dem Fischgrätenmuster Fig. A2-15)
- durch Verwendung höherer Momente.

In allen drei Fällen werden dadurch Merkmale aus größeren Fenstern bestimmt. Es ist allerdings notwendig, die Merkmalsauswahl zu automatisieren, um einen vom jeweiligen Bildinhalt abhängigen Merkmalssatz zu erhalten.

Im nächsten Abschnitt werden wir Relationen unter den extrahierten Primitiven untersuchen.

3. ÄHNLICHKEIT UND SELTENHEIT SYMBOLISCHER UND NUMERISCHER MERKMALE

Unter den Verfahren der Bildanalyse kommt der Bildzuordnung und der Objekterkennung eine zentrale Rolle zu. Sie können auf die im vorigen Kapitel angesprochene Extraktion von Merkmalen zurückgreifen. Ähnlich dem Prozeß zur Erkennung gerader Kantenstücke können auch andere flächenhaft verteilte Merkmale einem Gruppierungs- oder Clusterungsprozeß unterworfen werden, um zu bildlichen bzw. graphischen Grundelementen, wie etwa Linien, Kreisen, Punkten, Flächen, zu kommen, die wir Primitive genannt haben. Sie können selbst wieder in einen Gruppierungsprozeß eingehen. Das Bild wird so durch die Struktur der Primitive repräsentiert, z. B. durch

- Listen markanter Punkte mit Koordinatenpaar und den Bildfunktionswerten in einem kleinen Fenster als Attribute
- Listen von geraden Kanten mit den Koordinaten von Anfangs- und Endpunkt, dem Vorzeichen und der Größe des Kontrasts und der Breite der Kante als Attribute
- Listen von Paaren von Kanten, die in einer bestimmten Relation stehen (parallel, anti-parallel, senkrecht, gemeinsamer Punkt) mit entsprechenden Attributen für die Relation
- Listen von Linien mit einer Liste von Zeigern zu Linienstücken bestimmter Charakteristika (geradlinig, kreisförmig, geschwungen), als Attribute (vgl. SMITH und WOLF 1984)
- Listen von zwei oder mehr Primitiven, die in einer bestimmten Relation stehen ("enthalten in", "zwischen", "bilden Rechteck").

Wichtigste Voraussetzung für eine Zuordnung von Bildern oder für die Erkennung von Objekten, die durch solche Strukturen repräsentiert werden, ist ein geeignetes Ähnlichkeits- oder Distanzmaß zur Bewertung der Zuordnung bzw. zur Suche einer optimalen Zuordnung. Nun gibt es eine Vielzahl von Ähnlichkeits- und Distanzmaßen, BOCK (1974) etwa gibt wenigstens 15 verschiedene Maße an, die bei der Klassifikation verwendet werden. Die Wahl eines geeigneten Maßes ist dadurch erheblich erschwert.

Das wesentliche Problem ist aber folgendes: Die Maße sind nicht kompatibel. Das bedeutet, Distanzen, die mit verschiedenen Maßen ermittelt wurden, können nicht addiert werden. Das Problem bei der Zuordnung - und die Objekterkennung ist damit eingeschlossen - besteht aber gerade darin, daß die Attribute von Primitiven verschiedene Typen aufweisen (Farben, Längen), daß aber ein einheitliches Maß benötigt wird, das Distanzen verschiedener Attributtypen verknüpfen kann. Wir wollen in diesem Kapitel das informationstheoretische Distanzmaß nach BOYER und KAK (1986, 1988), das für symbolische Attribute entwickelt wurde, und die klassischen Testgrößen aus der Statistik, die für den Vergleich numerischer Attribute geeignet sind, miteinander vergleichen und zeigen, daß sie unter bestimmten Bedingungen ineinander überführbar sind. Dies bildet einen ersten Schritt auf die notwendige Vereinheitlichung von Distanzmaßen von Attributen verschiedenen Typs zu.

Das zweite Problem entsteht durch den i. a. sehr großen Lösungsraum des Zuordnungsproblems, bei n Primitiven im rechten und linken Bild und einer vollständigen Zuordnung aller Primitive wächst die Zahl der möglichen Zuordnungen mit $n!$. Heuristische Suchverfahren, wie der A^* -Algorithmus, zielen einerseits darauf, den Suchraum, d. h. die Zahl der Blätter in dem zu durcharbeitenden Suchbaum, systematisch zu verkleinern, zum anderen die durchschnittliche Suchzeit, etwa durch eine Best-First-Strategie, zu reduzieren. In unserem Fall benötigen wir dazu eine Sortierung der Merkmale entweder nach ihrer Invarianz oder ihrer Seltenheit. Denn das Prüfen von Primitiven mit invarianten oder seltenen Attributen führt im Mittel früher zur Beschneidung des Suchbaums und damit zu einer im Durchschnitt geringeren Zahl der zu prüfenden Kombinationen. Nun ist Invarianz ein binärer Begriff, da ein Attribut entweder invariant bzgl. einer Transformation ist oder nicht, und Seltenheit ein Begriff, der nur für symbolische Attribute einen Sinn ergibt, da er von der Wahrscheinlichkeit abhängt, daß das Attribut einen bestimmten Wert annimmt. Eine Sortierung symbolischer und numerischer Attribute nach Invarianz

oder Seltenheit erfordert daher einen anderen Rahmen für die Definition der beiden Begriffe. Wir wollen zeigen, daß die Transinformation ein geeignetes Maß ist, um beide Begriffe, die Invarianz und die Seltenheit für symbolische und numerische Attribute zu definieren.

Zur Motivation der Konzepte wollen wir in Abschnitt 3.1 zunächst die Zuordnungsaufgabe in ihrer einfachsten Form darstellen. Die notwendigen informationstheoretischen Grundlagen stellt Abschnitt 3.2 zusammen. Die beiden folgenden Abschnitte diskutieren dann die Probleme der Ähnlichkeit, der Invarianz und der Seltenheit ausführlich.

3.1 Das Zuordnungsproblem

BOYER und KAK (1986, 1988) verwenden ihr informationstheoretisches Maß für die Zuordnung zweier Bilder auf der Basis einer strukturellen Beschreibung der beiden Bilder. Da es uns hier nicht so sehr auf ihr Zuordnungsverfahren als Ganzes, sondern auf die Analyse ihres Ähnlichkeitsmaßes ankommt, werden wir in der folgenden Darstellung das Zuordnungsproblem stark vereinfachen und so in der Lage sein, es mit Elementen des Zuordnungsverfahrens von FÖRSTNER (1984, 1986a) zu vergleichen.

Wir gehen davon aus, daß beide Bilder durch Listen $P = (p_i)$ und $Q = (q_j)$ von Primitiven p_i , $i = 1, \dots, N$ und q_j , $j = 1, \dots, M$ repräsentiert sind. Die Primitive sind durch Attribute näher gekennzeichnet. Die Attribute sind Zufallsvariable, die Werte aus einem vorgegebenem Wertebereich, einem Alphabet oder einem Intervall auf der Zahlenachse, annehmen können. Das Attribut kann man auch als den Namen des Merkmals und den Wert als Realisierung einer Zufallsvariablen verstehen. Die Zahl der Attribute ist für jedes der Primitive gleich. Falls die Attribute symbolischer Art sind, d. h. die Zufallsvariable diskret ist, können sie verschiedene Wertebereiche annehmen (z. B. Alter, Farbe, Eulerzahl). Die Zahl der Primitive im linken und rechten Bild ist nicht notwendigerweise identisch. Damit sind die Beschreibungen der beiden Bilder durch

$$P = (p_1, p_2, \dots, p_N)$$

$$Q = (q_1, q_2, \dots, q_M)$$

mit den Attributen

$$p_i = (a_{i1}, a_{i2}, \dots, a_{ik}, \dots, a_{iA})$$

$$q_j = (b_{j1}, b_{j2}, \dots, b_{jk}, \dots, b_{jA})$$

also

$$P = (a_{ik}), i = 1, \dots, N; k = 1, \dots, A$$

$$Q = (b_{jk}), j = 1, \dots, M; k = 1, \dots, A$$

gegeben. Gegenüber BOYER und KAK haben wir keine Relationen zwischen den Objekten vorgesehen.

Die Aufgabe der Zuordnung besteht darin, eine beste Abbildung

$$h: P \rightarrow Q$$

d. h. eine Liste von Paaren

$$h = \{(p_i, q_j)\} \quad \text{oder} \quad \{(h(p_i), q_j)\} \quad \text{oder} \quad h = \{(i, j)\}$$

zu finden. Zur Formulierung eines Optimalitätskriteriums benötigen wir ein Abstandsmaß zwischen den Primitiven. Es sei

$$\underline{d}_{ij} = d(h(\underline{p}_i), \underline{q}_j)$$

der Abstand zwischen den zugeordneten Primitiven \underline{p}_i und \underline{q}_j , dann kann das Zuordnungsproblem durch

$$\hat{h}: \min_h \sum_{(i,j) \in h} d(h(\underline{p}_i), \underline{q}_j)$$

beschrieben werden. Gesucht ist diejenige Abbildung \hat{h} , die die Summe der Differenzen zwischen den Bildern der linken Primitive und den dadurch zugeordneten rechten Primitiven minimiert.

Diese Formulierung ist für unseren Zweck ausreichend allgemein. Sie enthält zunächst keine Möglichkeit, Primitive, die nicht zugeordnet werden können, in der Kostenfunktion zu berücksichtigen. Dies läßt sich aber durch die Einführung eines Leer-Primitivs erreichen, auf das diejenigen Primitive zugeordnet werden, die sonst nicht zugeordnet werden können.

Falls nun die Attribute der Primitive unabhängig sind, kann man die Distanzfunktion d_p zwischen den Primitiven auf eine Distanzfunktion d_a zwischen den Attributen zurückführen und erhält z. B.

$$d_p(h(\underline{p}_i), \underline{q}_j) = \sum_k d_a(\underline{a}_{h(i),k}, \underline{b}_{jk})$$

die Distanz zwischen den Primitiven als Summe der Distanzen zwischen den Attributen.

Somit können unsere beiden Fragestellungen klar umrissen werden:

1. Unter welchen Bedingungen lassen sich Abstände zwischen Attributen symbolischer und numerischer Attribute, d. h. diskreter und kontinuierlicher Zufallsvariablen miteinander vergleichen? In wieweit kann die Abhängigkeit der Attribute untereinander berücksichtigt werden?
2. Wie kann man zur Reduktion der durchschnittlichen Suchzeit die Primitive bzw. die Attribute nach der Sicherheit ordnen, mit der man sie im (anderen) Bild unverändert wiederfindet oder nach der Auffälligkeit oder Seltenheit innerhalb der Bilder oder des Objektmodells?

Da sich beide Fragen mit Hilfe von Konzepten aus der Informationstheorie beantworten lassen, wollen wir sie im nächsten Abschnitt zusammenfassend darstellen.

3.2 Grundlagen aus der Informationstheorie

Die Theorie der Information wurde von C. SHANNON 1949 zur Analyse von Systemen zur Nachrichtenübertragung entwickelt. Insbesondere geht es dabei um die Messung des Informationsgehaltes einer Nachricht und die Effizienz ihrer Übertragung über einen Kanal, der möglicherweise verrauscht ist. Die Theorie ist statistischer Natur, insofern sie nur die statistischen Eigenschaften der Nachricht und nicht ihre Bedeutung berücksichtigt. Wir wollen hier nur die für uns notwendigen Grundbegriffe darstellen und ansonsten auf die Fachliteratur verweisen (vgl. SHANNON 1949, BERGER 1971, HÖLZLER und HOLZWARTH 1975).

Eine diskrete Nachrichtenquelle kann man nach SHANNON als Markoff-Prozeß modellieren, der zufällig Buchstaben aus einem endlichen, festgelegten Alphabet auswählt. Die Information, die pro Buchstabe übertragen wird, ist umso größer, je unwahrscheinlicher der Buchstabe gewählt wird, und kann mit dem Grad der Überraschung beim Empfänger gleichgesetzt werden, wenn er den übermittelten Buchstaben zur Kenntnis bekommt oder mit der Unsicherheit, die besteht, wenn man über den Buchstaben noch keine Kenntnis hat.

Im einfachsten Fall sind die übermittelten Buchstaben unabhängig voneinander. Die Wahrscheinlichkeit dafür, daß der Buchstabe \underline{a} (eine diskrete Zufallsvariable) den Wert w_i annimmt, sei $P(w_i) = P(\underline{a} = w_i)$, dann ist der Informationsgewinn bei Zurkenntnisnahme des Buchstaben w_i , d. i. die Information von w_i

$$I(\underline{a} = w_i) = I(w_i) = \log \frac{1}{P(w_i)} = - \log P(w_i) \quad (32-1)$$

Wenn der Logarithmus zur Basis 2 verwendet wird, ist die Einheit der Information das bit. Falls der natürliche Logarithmus verwendet wird, wie im folgenden, ist die Einheit das nat.

Entsprechend kann man die Information messen, die man erhält, wenn der Buchstabe \underline{a} mit dem Wert w_i mitgeteilt wird, aber der Wert des Buchstaben \underline{b} (eine zweite diskrete Zufallsvariable) mit w_j schon bekannt ist. Dann gilt mit der bedingten Wahrscheinlichkeit $P(w_i | w_j)$ für die bedingte Information

$$\begin{aligned} I(w_i | w_j) &= - \log P(w_i | w_j) = - \log \frac{P(w_i, w_j)}{P(w_j)} \quad (32-2) \\ &= I(w_i) - I(w_i, w_j) \end{aligned}$$

Falls die Ereignisse $\underline{a} = w_i$ und $\underline{b} = w_j$ unabhängig sind, ist die bedingte Information identisch mit der ohne Vorkenntnis. Falls dagegen die Ereignisse abhängig sind, ist die Überraschung, w_i zu erfahren, wenn man vorher w_j erfuhr, kleiner, als wenn diese Vorinformation nicht vorliegt, d. h. dann ist $I(w_i | w_j) < I(w_i)$ und man benötigt weniger bits für die Mitteilung von w_i . Die gegenseitige Information von w_i und w_j ist gerade die Differenz

$$\begin{aligned} I(w_i; w_j) &= I(w_i) - I(w_i | w_j) \\ &= - \log P(w_i) + \log P(w_i | w_j) \\ &= I(w_i) + I(w_j) - I(w_i, w_j) \\ &= \log \frac{P(w_i, w_j)}{P(w_i) P(w_j)} \quad (32-3) \end{aligned}$$

die offensichtlich symmetrisch bzgl. w_i und w_j ist (vgl. BERGER 1971).

Für kontinuierliche Nachrichtenquellen wird die Information analog definiert. Sei $p(x_i)$ die Wahrscheinlichkeitsdichte der Zufallsvariablen \underline{x}_i , die von der Nachrichtenquelle mitgeteilt wird, dann ist die differentielle Information

$$I(x_i) = - \log p(x_i). \quad (32-4)$$

Ein Vergleich zwischen diskreten und kontinuierlichen Nachrichtenquellen ist nicht unmittelbar möglich. Denn streng genommen ist die Überraschung, eine bestimmte vorgegebene reelle Zahl durch einen Zufallsgenerator mit unendlicher Stellenzahl mitgeteilt zu bekommen, unendlich, denn $P(\underline{x}_i = x_i) = 0$. Daher wird der Ausdruck Gl. 32-4 auch differentielle Information genannt. Er kann u. U. auch negativ werden, nämlich wenn $p(x_i) > 1$ wird. Außerdem ändert sich bei einer Maßstabsänderung von x_i auch die (differentielle) Information. Bei allen Anwendungen sind aber nur Differenzen oder Rangordnungen von Informationen von Bedeutung, so daß die Wahl des Maßstabs keine Rolle spielt, solange man unter kontinuierlichen Zufallsvariablen bleibt.

In unserem Zusammenhang benötigen wir die Information einer normalverteilten Zufallsvariablen $\underline{x} \sim N(\mu, \sigma^2)$. Hier gilt

$$I(x) = \frac{1}{2} \left[\frac{x-\mu}{\sigma} \right]^2 + \frac{1}{2} \log 2\pi\sigma^2 \quad (32-5a)$$

und bei mehrdimensionalen Verteilungen $\underline{x} \sim N(\underline{\mu}, C)$

$$I(x) = \frac{1}{2} (\underline{x}-\underline{\mu})' C^{-1} (\underline{x}-\underline{\mu}) + \frac{n}{2} \log 2\pi |C|^{-\frac{1}{n}} \quad (32-5b)$$

Für bedingte Verteilungen (vgl. KOCH 1987, S. 138, 244)

$$\underline{x}|y \sim N(\mu_x + C_{xy} C_{yy}^{-1} (y-\mu_y), C_{xx} - C_{xy} C_{yy}^{-1} C_{yx}) = N(\hat{x}, C_{ff}) \quad (32-6)$$

kann man auch die Information angeben, die man erhält, wenn x mitgeteilt wird, und y bereits bekannt ist. Die Überraschung, die für x bleibt, wenn y bekannt ist, ist diesselbe wie die des Prädiktionsfehlers $f = \hat{x} - x = x|y - x$

$$I(x|y) = I(f) = \frac{1}{2} f' C_{ff}^{-1} f + \frac{n}{2} \log 2\pi |C_{ff}|^{-\frac{1}{n}} \quad (32-7)$$

Mit diesen Ausdrücken läßt sich die Information messen, die eine diskrete oder kontinuierliche Nachrichtenquelle liefert.

Zur Dimensionierung eines Kanals zur Nachrichtenübertragung, bzw. zur Beurteilung einer Nachrichtenquelle benötigen wir den mittleren Wert der zu übermittelnden Information. Er wird als Entropie bezeichnet:

$$H(\underline{a}) = E(I(\underline{a})) \quad (32-8)$$

Wir erhalten nacheinander:

die durchschnittliche Information pro Buchstabe

$$H(\underline{a}) = - \sum_i P(w_i) \log P(w_i),$$

die durchschnittliche bedingte Information pro Buchstabenpaar $(\underline{a}, \underline{b})$, wenn \underline{b} bekannt ist, bevor \underline{a} mitgeteilt wird, d. h. die bedingte Entropie

$$H(\underline{a}|\underline{b}) = - \sum_{ij} P(w_i, w_j) \log P(w_i | w_j), \quad (32-9)$$

und die durchschnittliche gegenseitige Information, d. h. die gegenseitige Entropie

$$\begin{aligned} H(\underline{a}; \underline{b}) &= H(\underline{a}) + H(\underline{b}) - H(\underline{a}, \underline{b}) \\ &= - \sum_{ij} P(w_i, w_j) \log \frac{P(w_i, w_j)}{P(w_i) P(w_j)}. \end{aligned} \quad (32-10)$$

$H(\underline{a}; \underline{b})$ wird auch Transinformation genannt (vgl. HÖLZLER und HOLZWARTH 1975) und ist für viele Anwendungen, wie auch hier, die entscheidende Größe.

Die durchschnittliche Information über eine normalverteilte Zufallsvariable ist

$$H(\underline{x}) = \frac{1}{2} + \frac{1}{2} \log 2\pi\sigma^2 = \frac{1}{2} \log 2\pi e\sigma^2 = \frac{1}{2} \log 2\pi e + \frac{1}{2} \log \sigma^2 \quad (32-11)$$

und die mittlere Information bei Mitteilung eines normalverteilten Zufallsvektors

$$H(\underline{x}) = \frac{n}{2} + \frac{n}{2} \log 2\pi |C|^{\frac{1}{n}} = \frac{n}{2} \log 2\pi e |C|^{\frac{1}{n}} = \frac{n}{2} \log 2\pi e + \frac{1}{2} \log |C| \quad (32-12)$$

Die mittlere Information, die man bei Mitteilung von x erhält, falls y bekannt ist, ist gegeben durch

$$H(\underline{x}|\underline{y}) = \frac{n}{2} \log 2\pi e + \frac{1}{2} \log |C_{xx} - C_{xy} C_{yy}^{-1} C_{yx}|. \quad (32-13a)$$

Sie ist identisch mit der Entropie des Prädiktionsfehlers f

$$H(\underline{f}) = \frac{n}{2} \log 2\pi e + \frac{1}{2} \log |C_{ff}| \quad (32-13b)$$

Damit erhalten wir schließlich die mittlere gegenseitige Information von \underline{x} und \underline{y} zu

$$H(\underline{x};\underline{y}) = \frac{1}{2} \log \frac{|C_{xx}|}{|C_{xx} - C_{xy} C_{yy}^{-1} C_{yx}|}, \quad (32-14a)$$

die nach KOCH (1987, Gl. 136.12) identisch ist mit

$$H(\underline{y};\underline{x}) = \frac{1}{2} \log \frac{|C_{yy}|}{|C_{yy} - C_{yx} C_{xx}^{-1} C_{xy}|} \quad (32-14a)$$

und somit, wie die gegenseitige Information, symmetrisch ist.

Bevor wir diese Konzepte auf die Zuordnung von Bildern anwenden, wollen wir die Bedeutung der einzelnen Ausdrücke diskutieren.

1. Der Begriff der Information entspricht dem üblichen Sprachgebrauch, soweit es sich um die Statistik der übermittelten Nachricht, nicht um deren Inhalt handelt. Falls die Wahrscheinlichkeit, ein bestimmtes Zeichen übermittelt zu bekommen, sehr klein ist, d. h. für $P(\underline{a} = w_i) \ll 1$, ist die Überraschung, d. h. die Information groß; falls uns die Nachricht vorher bekannt ist, ist dagegen die verbleibende Unsicherheit = 0, und die Nachricht enthält (für uns) keine Information.
2. Für die Übertragung von Nachrichten über einen Kanal, etwa ein Magnetband, existieren optimale Codierungsschemata, wie etwa der Huffman-Code. Umgekehrt kann man bei Vorgabe einer zulässigen Störung bei der Übertragung eine Untergrenze für die notwendige Übertragungsrate ableiten (vgl. BERGER 1971).
3. Die bedingte Information $I(u|v)$ ist Null, falls \underline{u} und \underline{v} vollständig abhängig sind. Dagegen ist die bedingte Information $I(u,v)$ identisch mit der unbedingten Information $I(u)$, falls \underline{u} und \underline{v} gegenseitig unabhängig sind. Deshalb ist die gegenseitige Information gleich der Information jeder einzelnen Nachricht, wenn sie vollständig abhängig sind. Die gegenseitige Information ist Null genau dann, wenn sie unabhängig sind.
4. Die differentielle Information einer normal verteilten Zufallsvariable enthält zwei Terme. Der erste ist, abgesehen von einem Faktor $\frac{1}{2}$, der normierte Abstand vom Mittelwert. Er

dient üblicherweise als Testgröße und ist X_n^2 -verteilt. Der zweite Term ist lediglich von der Varianz bzw. der Kovarianzmatrix abhängig. Um die Überraschung bei der Mitteilung der Realisierung einer normal verteilten Zuallsvariablen gering zu halten, sollte erwartungsgemäß

- a. x nicht zu sehr vom Mittelwert abweichen, d. h. keinen Ausreißer darstellen, und
- b. die Varianz, bzw. die Determinante der Kovarianzmatrix klein sein, da wir dann mit μ und σ^2 bzw. C genügend Auskunft über x haben. Wir werden im nächsten Abschnitt auf diesen Zusammenhang zurückkommen.

- 5. Bei Bildung der Erwartungswerte bleiben für die Entropie im wesentlichen die zweiten, von der Varianz bzw. der Kovarianzmatrix abhängigen, Terme. Optimierung eines Designs bzgl. der Entropie der Unbekannten Parameter bedeutet also gleichzeitig Optimierung bzgl. der Determinante der Kovarianzmatrix dieser Parameter, worauf GRAFAREND 1971 bei der Optimierung des Vorwärtsschnittes hinweist.
- 6. Die bedingte Verteilung enthält im wesentlichen die Formel der linearen Prädiktion von x, falls y bekannt ist und die Genauigkeit des Prädiktionsfehlers f. Bei kleinen Prädiktionsfehlern sind, wegen $H(\underline{x}|\underline{y}) = 0.5 \cdot \log |C_{ff}|$ nur wenig bits notwendig, um x zu beschreiben, falls y bekannt ist. Auch hierauf werden wir später zurückkommen.

3.3 Ähnlichkeit symbolischer und numerischer Attribute

3.3.1 Das informationstheoretische Abstandsmaß nach BOYER und KAK

Nach BOYER und KAK kann man die Zuordnung zweier Bilder als Nachrichtenübertragung modellieren. Die symbolischen Attribute der Primitive, etwa im linken Bild, tragen Information, die über den Zuordnungsprozeß dem rechten Bild mitgeteilt werden. Eine Zuordnung wird genau dann als gut angesehen, wenn sie nicht überrascht, d. h. insgesamt wenig Information übertragen wird.

Die Information eines einzelnen Attributes a_{ik} des Primitivs p_i sei

$$I(a_{ik}) = -\log P(\underline{a}_{ik} = a_{ik}) = -\log P(a_{ik}) \tag{331-1}$$

Dabei ist $P(a_{ik})$ die Wahrscheinlichkeit, daß das Attribut \underline{a}_{ik} , d.h. die diskrete Zufallsvariable \underline{a}_{ik} den speziellen Wert a_{ik} aus einem Alphabet einnimmt, das für jedes Attribut verschieden sein kann. Die Zahl A der Alphabete gleicht der Zahl der Attribute.

Die Gesamtinformation, die ein Primitiv trägt, ist dann die Summe der Einzelinformationen jedes Attributs, wenn die Attribute unabhängig sind; dann gilt:

$$I(p_i) = \sum_{k=1}^A I(a_{ik}) \tag{331-2}$$

Die Gesamtinformation aller Primitive p_i , die das linke Bild beschreiben, ist dann

$$I(P) = \sum_{i=1}^N \sum_{k=1}^A I(a_{ik}) = - \sum_{i=1}^N \sum_{k=1}^A \log P(a_{ik}) \tag{331-3}$$

Jedes Attribut wird nun durch die Abbildung - den Kanal - in das rechte Bild übertragen. Zur Bestimmung der optimalen Übertragung benötigen wir für jedes Attribut die sog. Kanal-Übertragungsmatrix K. Sie gibt an, mit welcher Wahrscheinlichkeit bei richtiger Zuordnung das Attribut b_j im rechten Bild auftritt, wenn das Attribut a_i im linken Bild auftritt

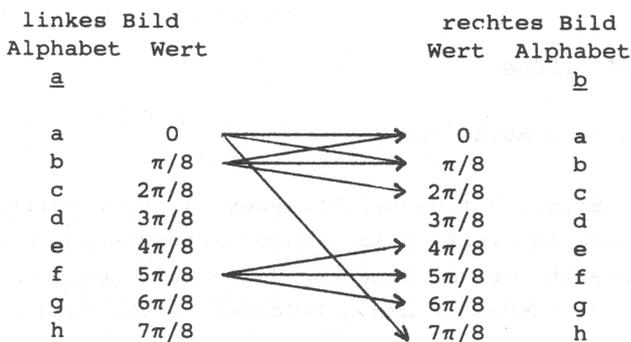
$$K = (K_{ij}) = (P(b_j | a_i)) \quad (331-4)$$

Die Dimension von K ist $r \times s$, wenn r die Größe des Alphabets im linken und s die Größe des Alphabets im rechten Bild ist. I. a. muß r nicht gleich s sein, etwa wenn 6 Farben auf Schwarz und Weiß abgebildet werden. BOYER und KAK nehmen $r = s$ an und erhalten so quadratische Kanal-matrizen.

Zur Veranschaulichung ist für das Attribut "Kantenrichtung" der Kanal und die Kanalmatrix angegeben.

Fig. 3-1 Übertragungskanal für das Attribut "Kantenrichtung" und Kanalmatrix, ideale Werte

Die diskretisierten und als Symbole gespeicherten Kantenrichtungen können bei der Abbildung nur in benachbarte Buchstaben des Alphabets übertragen werden.
(nach BOYER und KAK 1986)



$$K = \begin{bmatrix} 0.8 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0.1 \\ 0.1 & 0.8 & 0.1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0.8 & 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0.8 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0.8 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.1 & 0.8 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.1 & 0.8 & 0.1 \\ 0.1 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.8 \end{bmatrix} = (K_{ij}) = P(b_j | a_i)$$

Damit sind alle Vorbereitungen getroffen, das Maß der Informationsübertragung und somit das informationstheoretische Abstandsmaß für die Zuordnung zu definieren. Die Information, die unter einer Abbildung h vom Attribut \underline{a}_{ik} mit dem Wert a_{ik} eines linken Primitivs zum Attribut \underline{b}_{jk} mit dem Wert b_{jk} eines rechten Primitivs übertragen wird, ist

$$I_h(b_{jk} | a_{ik}) = - \log P_h(b_{jk} = b_{jk} | \underline{a}_{ik} = a_{ik}) \quad (331-5)$$

Sie ist umso kleiner, je größer die bedingte Wahrscheinlichkeit aus der Kanalmatrix ist. Die Gesamtinformation bei Zuordnung eines Primitivs ist unter der Voraussetzung der Unabhängigkeit der Attribute

$$D_h(p_i, q_j) = \sum_{k=1}^{|A|} I_h(b_{jk} | a_{ik}) \quad (331-6)$$

die Summe der Einzelinformationen. Sie kann als Distanz D_h der Objekte p_i und q_j verstanden werden. Die Gesamtinformation, die bei der Zuordnung mit der Abbildung h übertragen wird, ist

$$D_h(P, Q) = \sum_{(i,j) \in h} D_h(p_i, q_j) \quad (331-7)$$

gleich der Summe aller Distanzen durch h zugeordneter Primitive.

Das Zuordnungsproblem besteht in der Suche einer besten Zuordnung \hat{h} , die diese Distanz minimiert

$$D_{\hat{h}} = \min_{\hat{h}} D_h(P, Q) \quad (331-8)$$

Dies ist soweit die vereinfachte Darstellung des Verfahrens von BOYER und KAK.

Als Motivation für das informationstheoretische Maß nennen BOYER und KAK:

"Der informationstheoretische Ansatz zur Zuordnung von Primitiven sollte näherliegenden Techniken überlegen sein, wie der, Attribute als Vektoren in einem Attribut-Raum zu repräsentieren und Euklidische Distanzen zu rechnen, weil [hier] die Statistik der Situation berücksichtigt wird. Außerdem mag es für Attribute, die nur symbolischen Charakter haben, schwierig oder unmöglich sein, eine Ordnung in der Weise zu definieren, daß einem Euklidischen Abstand eine Bedeutung zukommt." (BOYER und KAK 1986, S. 50)

Wir können dieser Argumentation, soweit es sich um symbolische Attribute handelt, nur zustimmen. Wir wollen hier auch nicht den Lösungsweg kritisieren, insbesondere, da nur eines der insgesamt 16 von BOYER und KAK benutzten Attribute wirklich symbolischer Natur ist (das "Schleifen"-Prädikat, loop-predicate), alle anderen aber numerische Werte annehmen, wie das oben zitierte Attribut "Kantenrichtung". Im Gegenteil, wir wollen nun zeigen, daß das informationstheoretische Konzept eine Brücke zwischen symbolischen und numerischen Attributen schlagen kann.

3.3.2 Relation zu Prüfgrößen aus der Statistik

Wir wollen zunächst zeigen, daß für numerische Attribute, für die man eine Normalverteilung annehmen kann, das Prinzip von BOYER und KAK auf bekannte Optimierungsprinzipien aus der Statistik führen.

Dazu betrachten wir als Beispiel die Kanalmatrix von Fig. 3-1 für die diskreten Werte der Orientierung von Kanten, die aus Bildern extrahiert werden. Offensichtlich ist

$$P(b_j | a_i) = P(|b_i - a_j| \bmod 2\pi) \quad (332-1)$$

$$= \begin{cases} 0.8 & \text{für } a_i = b_j \\ 0.1 & \text{für } |a_i - b_j| \bmod 2\pi = \pi/8 \\ 0 & \text{sonst} \end{cases}$$

Damit wird die bedingte Information $I(b_j | a_i)$ von b_j für gegebenes a_i auf die Information der Wertedifferenz reduziert. Mit 80 % Wahrscheinlichkeit ist die Differenz 0, mit 10 % unterscheidet sie sich um eine Klasse (45 °).

Da Kantenorientierungen, wie andere numerische Merkmale, i. a. als reelle Zahlen und nicht bereits in (wenige) Klassen eingeteilt anfallen, liegt es nahe, in diesen Fällen die Kanalmatrix durch eine kontinuierliche Verteilung zu repräsentieren, z. B.

$$f_{ij} = (b_j - a_i) \bmod 2\pi \sim N(0, \sigma^2) \quad (332-2)$$

(eine hier hinreichende Approximation der v. Mises-Verteilung). Dann wäre die Information von Gl. 32-4 identisch mit

$$I(b_j | a_i) = I(f_{ij}) = \frac{1}{2} \left[\frac{f_{ij}}{\sigma} \right]^2 + \frac{1}{2} \log 2\pi\sigma^2 \quad (332-3)$$

Damit wird die Kanalmatrix zu einer Kovarianzmatrix, worin die Elemente mit Hilfe einer Gaußschen Kovarianzfunktion ermittelt werden. Die eingehenden Distanzen sind die Abstände die sich aus Gl. 332-3 ergeben. So wäre etwa in Gl. 332-2 $\sigma^2 = 45^\circ * \sqrt{0.1 + 0.1} \approx 20^\circ$ zu setzen, womit sich nach entsprechender Normierung - ohne Berücksichtigung der Intervall-Längen - ein Verhältnis der Elemente von ca. 0.86 : 0.07 statt 0.8 zu 0.1 ergäbe. Wir werden in den Beispielen darauf zurückkommen.

Tatsächlich führt auch das Prädiktionsmodell

$$b = a + f \quad (332-4)$$

mit den unabhängigen Komponenten a und f und den Varianzen σ_a^2 und σ_f^2 mit den Substitution $x=b$ und $y=a$ mit Gl. 32-7 auf Gl. 332-3.

Nun ist aber bei gegebenem Design der zweite Term in Gl. 332-3 fest. Damit führt, zumindest für Attribute dieses Typs, die Minimierung von Gl. 331-8 auf eine Zuordnung, die die Summe der normierten Abweichungen über alle Attribute aller Primitive minimiert: Das Prinzip der minimalen Informationsübertragung führt bei normalverteilten Attributen auf das Prinzip der kleinsten Quadrate.

Damit lassen sich wenigstens zwei Verallgemeinerungen unmittelbar angeben. Die Abhängigkeit der Merkmale kann auf einfache Weise durch ihre Kovarianzmatrix berücksichtigt werden. Mit $f_{ij} = b_j - a_i$ gilt

$$D_h(p_i, q_j) = \frac{1}{2} f'_{ij} C_{ff}^{-1} f_{ij} + \frac{n}{2} \log 2\pi |C_{ff}|^{-\frac{n}{2}} \quad (332-5)$$

worin C_{ff} die Kovarianzmatrix der Differenzen f_{ij} ist.

Weiterreichend ist allerdings die Möglichkeit, die Abbildung h ganz oder teilweise zu parametrisieren und damit gezielt zu erweitern. Z. B. ließe sich das Kantenorientierungs-Modell Gl. 332-2 durch

$$f_{ij} \sim N(\Phi, \sigma^2) \quad (332-6)$$

ersetzen, mit einem zusätzlichen unbekanntem Parameter, der für alle Kanten einen gleichen Richtungsfehler darstellt, verursacht etwa durch die gegenseitige Rotation der Kameras. Die bei BOYER und KAK angegebene, aus realen Daten abgeleitete, Kanalmatrix für dieses Attribut zeigt diesen Effekt sehr deutlich; dort liegt ein durchschnittlicher Richtungsunterschied von 30° vor, wobei die Varianzen von der Richtung abzuhängen scheinen.

Die Verallgemeinerung ist offensichtlich. Die Attribute des rechten Bildes gehen aus einer allgemeinen Transformation der Attribute des linken Bildes hervor

$$\underline{b}_j - T(a_i, x) \sim N(0, C) \quad \text{für alle } (i, j) \in h \quad (332-7)$$

Das bedeutet, daß die diskrete Zuordnungsfunktion h durch die mit x parametrisierte Funktion ergänzt wird. T wird etwa einem geometrisch-physikalischen Objekt-Abbildungsmodell entnommen.

Beispiel:

Wir wollen das Ähnlichkeitsmaß, das in dem merkmalsgestützten Zuordnungsalgorithmus von FÖRSTNER (1984, 1986) verwendet wird, unter diesem Blickwinkel betrachten. Dort sind die Primitive extrahierte Punkte. Ihre Attribute sind ein Koordinatenpaar zur Beschreibung ihrer Lage und die Funktionswerte in einem kleinen Fenster von z. B. 5×5 Bildelementen. Die vorläufigen Gewichte w_{ij} leiten sich dort aus dem geschätzten Abstand zwischen den vorläufig zugeordneten Punkten ab. Er errechnet sich im wesentlichen aus der Quadratsumme der Differenzen zwischen zugehörigen Funktionswerten in den Fenstern. Dieses Maß ist zunächst identisch mit Gl. 331-6, wenn man von der Berücksichtigung zweier zusätzlicher Parameter absieht, die für eine Kompensation bezüglich Helligkeit und Kontrast sorgen. Tatsächlich handelt es sich also um ein Modell der Form Gl. 332-7 mit den zwei genannten Parametern.

Eine ähnliche Überlegung gilt für das Attribut "Position". Für die Koordinaten $z_i = (x_i, y_i)'$ der Punkte im linken und $z_j = (x_j, y_j)'$ im rechten Bild wird eine lineare Transformation

$$\underline{z}_j + \underline{f}_{ij} = A z_i + a$$

mit 6 Parametern angenommen. Die 4 Parameter a_{kl} für Rotation, Maßstab und Scherungen und die 2 Parameter a_m für die Verschiebungen werden aus der Minimierung der Quadratsumme der f_{ij} gewonnen, in völliger Analogie zu Gl. 332-7 mit Gl. 332-5.

Abschließend bleibt allerdings festzuhalten, daß eine Kompatibilität zwischen Abstandsmaßen für numerische und symbolische Attribute noch nicht hergestellt ist. Dieses Problem bleibt somit vorläufig ungelöst. Inwieweit es von praktischer Bedeutung ist, bleibt offen, da sowohl bei der Bildzuordnung als auch bei der Objekterkennung geometrisch-physikalische Modelle zwischen kontinuierlich-wertigen Größen den Hauptanteil an der Modellbildung ausmachen.

Im folgenden Abschnitt wollen wir die Querverbindungen, die wir hier zusammengestellt haben, zur Beurteilung der Attribute bei der Zuordnung verwenden.

3.4 Beurteilung von Attributen

Die Aufgabe, invariante oder seltene Strukturen zu finden, bzw. Merkmale bezüglich ihrer Invarianz oder Seltenheit zu beurteilen, ist sehr allgemein und grundlegend für die Bildanalyse. Im Zusammenhang mit der Zuordnung können Invarianz und Seltenheit von Attributen oder Merkmalen dazu dienen, den Zuordnungsprozeß zu beschleunigen:

1. Invariante Attribute garantieren eine sichere Zuordnung. Daher sollten bei der Prüfung der Attribute auf Ähnlichkeit die invarianten Attribute im Sinne einer Best-First-Strategie zuerst geprüft werden.

Die Invarianz von Attributen läßt sich grundsätzlich auf zwei Weisen feststellen:

- a. analytisch, aus der Klasse der zulässigen Transformationen, oder
- b. empirisch, aus einer Stichprobe richtiger Zuordnungen

Da bei der Bildzuordnung keine einfache analytische Funktion für die Abbildung existiert, scheidet eine theoretische Beurteilung der Attribute i. a. aus. Lediglich für sehr einfache geometrische Attribute, wie "gerade", oder topologische Attribute, wie "benachbart", kann man unter der Annahme etwa zweier Zentralprojektionen die Invarianz nachweisen. Schon Strecken oder Winkel sind i. a. keine Invarianten. Invarianten wie das Doppelverhältnis setzen voraus, daß die Objekte (Punkte, Strahlen) in einer Ebene liegen, eine Annahme, die vor der Bildanalyse nicht getroffen werden kann.

Oft weichen allerdings die realen Abbildungsverhältnisse nur wenig von einer idealen, etwa einer senkrechten Parallelprojektion, ab, für die eine Reihe von Invarianten bekannt ist. Es wäre für die Zuordnung von großem Nutzen, diese Invarianten unter Berücksichtigung der unbekannt kleinen Abweichungen, verwenden zu können. Dies setzt voraus, daß wir den Grad der Invarianz eines Attributs beurteilen können.

Auf eine ähnliche Fragestellung stoßen wir, wenn wir aus einer Stichprobe richtiger Zuordnungen die Invarianz der verwendeten Attribute ableiten wollen. Schon wegen Meßfehlern, aber vor allem wegen der Komplexität der Abbildungsfunktion werden Attribute, die im linken Bild gemessen werden, sich von denen im rechten Bild unterscheiden. Man benötigt auch hier ein Maß für den Grad der Invarianz der Attribute.

Wir wollen auf der Basis des informationstheoretischen Ähnlichkeitsmaßes die Eignung von Attributen für die Zuordnung bestimmen und zu einem Maß für die Invarianz der Attribute gelangen. BOYER und KAK (1986) haben ein ähnliches Maß für die Sortierung der Attribute vorgeschlagen.

Die technischen Hilfsmittel werden z. T. auch in der überwachten Klassifikation verwendet, hier aber durch das informationstheoretische Konzept vereinheitlicht.

2. Falls keine Stichprobe richtiger Zuordnungen vorliegt, kann man aus der Analyse der Attribute innerhalb eines Bildes die auffälligen oder seltenen Attribute selektieren. Bei der Prüfung der Zuordnung von Primitiven kann man erwarten, daß die Zahl der bei der Suche notwendigen Kombinationen klein ist, und daß daher die durchschnittliche Suchzeit reduziert wird, wenn man die seltenen Attribute zuerst prüft - wieder im Sinne einer Best-First-Strategie.

Der Begriff Seltenheit ist mit der Wahrscheinlichkeit verknüpft und daher zunächst nur für symbolische Attribute sinnvoll definiert. Für numerische Attribute würde "Isoliertheit" oder "Trennbarkeit" als Eigenschaft, die man aus der Clusterung im Merkmals- oder Attributraum ableiten kann, der geeignete Begriff für die Zielgröße sein. Die Methodik entspricht hier offensichtlich der unüberwachten Klassifikation.

Auch hier wollen wir auf der Basis des informationstheoretischen Ähnlichkeitsmaßes ein für symbolische und numerische Attribute einheitliches Maß finden, das eine Sortierung der Attribute erlaubt, mit dem Ziel, die durchschnittliche Suchzeit bei der Zuordnung zu reduzieren. BLANZ (1986) und STRAUB (1986) geben ein ähnliches Kriterium für die Merkmalsextraktion bei der Klassifizierung an. Mit der gleichen Zielsetzung haben TURNEY et. al. (1985) für die Objekterkennung wichtige Merkmale extrahiert, aber einen anderen Lösungsweg vorgeschlagen.

Sowohl die Invarianz als auch die Seltenheit sollen uns dazu dienen, eine Rangordnung der Attribute, Merkmale oder Strukturen zu finden, die zu einer Beschleunigung des Zuordnungsverfahrens führt. Das gemeinsame theoretische Konzept wird die Komplementarität der beiden Kriterien offenlegen.

3.4.1 Überwachte Bewertung von Attributen auf Invarianz

Primitive, Attribute oder Merkmale sind dann gut für die Zuordnung geeignet, wenn zugeordnete Elemente viel gemeinsame Information tragen. Da es um die Bewertung der Attribute, nicht deren aktuelle Werte geht, spielt lediglich die zu erwartende gemeinsame Information, d. h. die Transinformation eine Rolle.

Seien a_i und b_j Attribute zugeordneter Primitive, dann ist die mittlere gegenseitige Information gegeben durch

$$\begin{aligned} H(\underline{a}_i; \underline{b}_j) &= H(\underline{a}_i) - H(\underline{a}_i | \underline{b}_j) \\ &= - \sum_i P(a_i) \log P(a_i) + \sum_{ij} P(a_i, b_j) \log P(a_i | b_j) \end{aligned} \quad (341-1)$$

Die mittlere gemeinsame Information kann man aus der Kanalmatrix und der Entropie der Attribute ableiten.

BOYER und KAK (1986) benutzen dieses Maß nach Division durch $H(\underline{a}_i)$ für die Sortierung der Attribute. Die Division durch $H(\underline{a}_i)$ dient dazu, für die niedrige Entropie einzelner Attribute zu kompensieren. Das gleiche Maß schlägt BLANZ (1986) vor.

Wir wollen nun dieses Konzept auf kontinuierliche Zufallsvariable, d. h. numerische Attribute übertragen.

Zunächst nehmen wir an, \underline{x} und \underline{y} seien zugehörige Attribute im linken und rechten Bild, die gemeinsam normalverteilt sind. Dann erhalten wir mit Gl. 32-14

$$H(\underline{x}; \underline{y}) = \frac{1}{2} \log \frac{\sigma_x^2}{\sigma_x^2 - \sigma_{xy} \sigma_y^{-2} \sigma_{yx}} \quad (341-2)$$

und mit dem Korrelationskoeffizient ρ_{xy}

$$H(\underline{x}; \underline{y}) = \frac{1}{2} \log \frac{1}{\sqrt{1 - \rho_{xy}^2}} \quad (341-3)$$

Mit dem Signal-Rausch-Verhältnis

$$\text{SNR} = \frac{\sigma_x}{\sigma_f} \quad (341-4)$$

wird damit wegen $\sigma_x^2 = \sigma_x^2 + \sigma_f^2$

$$H(\underline{x};\underline{y}) = \frac{1}{2} \log (1 + \text{SNR}^2) \quad (341-5)$$

Die gegenseitige Information ist erwartungsgemäß am größten, nämlich ∞ , wenn x und y 100 %-ig korreliert sind, also $\sigma_f = 0$ und damit $\text{SNR} = \infty$ ist, während $H(\underline{x};\underline{y}) = 0$ nur bei unkorrelierten, also wegen der Annahme der Normalverteilung, auch unabhängigen Attributen gilt.

Die mittlere Information hängt nur von den Varianzen und Kovarianzen ab. Damit unterscheidet sie sich wesentlich von der Divergenz als Distanzmaß zweier Verteilungen, die auch die Mittelwerte berücksichtigt (vgl. z. B. BOCK 1974) und die für gleiche Kovarianzmatrizen bei normalverteilten Zufallsvariablen in den MAHALANOBIS-Abstand übergeht. Der Grund für das Fehlen der Mittelwerte in dem Abstandsmaß der mittleren gegenseitigen Information ist die Tatsache, daß bei Kenntnis von y über lineare Prädiktion x vorausgesagt werden kann und deshalb die verbleibende Restunsicherheit $H(\underline{x}|\underline{y})$ gegenüber $H(\underline{x})$ nicht vom Mittelwert μ_x abhängt.

Ist nun \underline{x} ein Skalar und $\underline{y} = (y_i)$ ein Vektor von Attributen, fragen wir also nach demjenigen Attribut im linken Bild, welches am meisten Information mit allen Attributen y_i gemeinsam hat, erhalten wir

$$H(\underline{x};\underline{y}) = \frac{1}{2} \log \frac{\sigma_x^2}{\sigma_x^2 - c_{xy} c_{yy}^{-1} c_{yx}} \quad (341-6)$$

worin $c_{xy} = c_{yx}$ die Kovarianzen zwischen \underline{x} und \underline{y} enthält. Mit dem totalen Korrelationskoeffizient

$$\rho_{xy}^2 = \frac{c_{xy} c_{yy}^{-1} c_{yx}}{\sigma_x^2}, \quad (341-7)$$

der offensichtlich ein quadratisch gewogenes Mittel über die Korrelationen zwischen \underline{x} und den y_i darstellt, erhalten wir Gl. 341-3 mit allerdings jetzt anderer Bedeutung von ρ_{xy}^2 .

Das bedeutet, daß dasjenige Attribut x im linken Bild für die Zuordnung am wichtigsten ist, das die größte totale Korrelation mit allen Attributen im rechten Bild aufweist.

Wollen wir schließlich prüfen, ob eine Gruppe $\underline{x} = (x_i)$ von Attributen im linken Bild mit einer Gruppe $\underline{y} = (y_j)$ von Attributen im rechten Bild genügend gemeinsame Information enthält, so erhalten wir $H(\underline{x};\underline{y})$ aus Gl. 32-14. Mit der Matrix

$$M = \begin{pmatrix} c_{xx}^{-1} & c_{xy} & c_{yy}^{-1} & c_{yx} \end{pmatrix} \quad (341-8)$$

und ihren Eigenwerten μ_i^2 ergibt dies

$$H(\underline{x};\underline{y}) = -\frac{1}{2} \log \prod_i (1 - \mu_i^2) = \sum_i \log \frac{1}{\sqrt{1 - \mu_i^2}} \quad (341-9)$$

Falls die Matrix M in Gl. (341-8) wenigstens einen Eigenwert $\mu_i^2 = 1$

aufweist, ist $H(\underline{x};\underline{y}) = \infty$. Die Matrix M ist identisch mit der aus der Trennbarkeitstheorie (FÖRSTNER 1983, LI 1987). Dort dient sie dazu, um zu entscheiden, ob geschätzte Parameter \underline{x} und \underline{y} unterscheidbar sind. Die Eigenwerte von M entsprechen Quadraten von Korrelationskoeffizienten, die für eine gute Trennbarkeit deutlich kleiner als 1 sein sollten. Das bedeutet in

unserem Zusammenhang: Falls wenigstens ein Korrelationskoeffizient $\mu_i = 1$ ist, gibt es wenigstens ein Paar Linearkombinationen Ax und By , die linear abhängig sind, Information vom linken ins rechte Bild übertragen können und so die Zuordnung garantieren. Die Merkmale a_i und b_j in verschiedenen Bildern sind also gut geeignet für die Zuordnung, wenn sie sich nicht gut trennen lassen.

Damit können wir aus empirischen, d. h. aus einem Trainingssatz richtiger Zuordnungen abgeleiteten, oder aus theoretischen, d. h. durch Fehlerfortpflanzung gewonnenen Kovarianzmatrizen C_{xx} , C_{xy} , und C_{yy} die gegenseitige Information von einzelnen Attributen oder Gruppen von Attributen bestimmen. Die mittlere gemeinsame Information ist ein Maß für die Invarianz. Wie bei der Anwendung der Trennbarkeitstheorie können hier auch zusätzliche Parameter einer Abbildung zwischen den Attributen berücksichtigt werden, wie wir es bei der Verallgemeinerung des informationstheoretischen Maßes in Gl. 332-7 vorgeschlagen haben.

3.4.2 Unüberwachte Bewertung von Attributen auf Seltenheit

Die unüberwachte Bewertung zielt auf die "Seltenheit", die "Isoliertheit" oder die "Trennbarkeit" von Attributen. Seltene Attribute sollten zuerst geprüft werden, weil man von ihnen auch Seltenheit im anderen Bild erwarten kann und schnell eine sichere Zuordnung findet, oder auf kostengünstige Weise die Zuordnung des zugehörigen Primitivs verwerfen kann. Auch hier wollen wir die informationstheoretische Betrachtung der Zuordnung zugrunde legen, diesmal aber die Attribute innerhalb eines Bildes untereinander vergleichen.

Ein Attribut ist dann für die Zuordnung gut geeignet, wenn es wenig gemeinsame Information mit allen anderen Attributen hat. Sei \underline{x} das betrachtete Attribut, und $\underline{y} = (y_i)$ die übrigen, dann können wir Gl. 341-6 heranziehen, in der, wie oben, ρ_{xy} den totalen Korrelationskoeffizienten zwischen \underline{x} und \underline{y} bedeutet. Nur hier ist die Interpretation der Werte umgekehrt: \underline{x} ist ein auf das Bild bezogen seltenes Attribut, wenn es nicht mit den übrigen Attributen korreliert ist, $\rho_{xy} = 0$ ist und daher erwartungsgemäß auch $H(\underline{x};\underline{y}) = 0$ gilt. Wir können daher zur Optimierung der Zuordnung die Attribute innerhalb des Bildes nach der Seltenheit sortieren. Dabei entsprechen kleine Werte $H(\underline{x};\underline{y})$ großer Seltenheit.

Auf diesem Hintergrund kann man die in Kap. 2.2 diskutierte Selektion optimaler Fenster für die Suche markanter Punkte auch anders interpretieren: Die aus den Bildfenstern geschätzten Parameter der Autokovarianzfunktion geben Auskunft über die Korrelation mit der Bildfunktion innerhalb der Nachbarfenster. Die Nicht-Maximum-Unterdrückung, d. h. die (lokale) Suche nach dem Fenster mit der größten Spur der Hesse-Matrix, d. i. bei Vernachlässigung der Varianz die Krümmung der Autokorrelationsfunktion, entspricht der Suche nach einem Fenster, das sich möglichst gut von den Nachbarfenstern unterscheiden bzw. trennen läßt, also lokal auffällig ist. Die Suche nach lokal auffälligen Punkten kann durch die Bestimmung der globalen, d. h. auf das Bild bezogene Auffälligkeit ergänzt werden, wie es im Beispiel c. im folgenden Abschnitt für ein Schachbrett gezeigt wird.

Um den Begriff der Seltenheit mit dem der Isoliertheit zu verbinden, wollen wir hier eine weitere Querverbindung mit Verfahren zur Beurteilung von Meßanordnungen anführen.

n Beobachtungen l_i sollen zur Bestimmung von Parametern x dienen. Sie seien zusammengefaßt im Vektor $\underline{l} = (l_i)$, normalverteilt mit Kovarianzmatrix C . Die Bedingungen zwischen diesen Beobachtungen seien nach der Elimination der Parameter x

$$B' (\underline{l} + \underline{y}) = c$$

(342-1)

woraus die Verbesserungen $v = (v_i)$ der Beobachtungen und die ausgeglichenen Werte $l = l + v$ der Beobachtungen nach der Methode der Kleinsten Quadrate $v' C^{-1} v \rightarrow \min$ zu schätzen sind. Wir wollen untersuchen, welche Information über eine einzelne Beobachtung l_i in dem Vektor

$$\underline{w} = B' \underline{l} - c \quad (342-2)$$

der Widersprüche \underline{w} enthalten ist, da er als einziges Informationenthält, mit dem sich die Beobachtungen etwa auf grobe Fehler prüfen lassen. Mit $C_{ww} = B' C B$ erhalten wir

$$H(\underline{l}_i; \underline{w}) = \frac{1}{2} \log \frac{\sigma_{l_i}^2}{\sigma_{l_i}^2 - c_{l_i w} C_{ww}^{-1} c_{w l_i}} \quad (342-3)$$

und mit $l_i = e_i' l$ und daher $C_{l_i w} = e_i' C B$

$$H(\underline{l}_i; \underline{w}) = \frac{1}{2} \log \frac{\sigma_{l_i}^2}{\sigma_{l_i}^2 - e_i' C B (B' C B)^{-1} B' C e_i} \quad (342-4)$$

Nun ist aber der Nenner gerade gleich der Varianz $\sigma_{l_i}^2$ der i -ten ausgeglichenen Beobachtung $\underline{l}_i = \underline{l}_i + \underline{v}_i$ (vgl. GOTTHARDT 1968), so daß wir mit dem Redundanzanteil $r_i = \sigma_{v_i}^2 / \sigma_{l_i}^2$ (vgl. FÖRSTNER 1979a,b) und dem Unbekanntenteil $u_i = 1 - r_i$ schließlich

$$H(\underline{l}_i; \underline{w}) = \frac{1}{2} \log \frac{1}{u_i} = \frac{1}{2} \log \frac{1}{1 - r_i} \quad (342-5)$$

erhalten.

Dies bedeutet zunächst folgendes: Falls der Widerspruchsvektor \underline{w} und damit die Verbesserungen \underline{v}_i wenig Information über die Beobachtung \underline{l}_i enthalten, kann \underline{l}_i nicht geprüft werden. Diese Aussage ist identisch mit der aus der Zuverlässigkeitstheorie (BAARDA 1967, 1968; vgl FÖRSTNER 1987). Dort ist die untere Grenze für einen gerade noch erkennbaren (groben) Fehler in der Beobachtung \underline{l}_i proportional zu $1/\sqrt{r_i}$. Eine Beobachtung, die nicht oder schlecht kontrollierbar ist, liegt aber geometrisch isoliert zu den anderen, da die Bedingungen keine oder nur eine schwache Verbindung herstellen. Daher können wir geringe gegenseitige Information auch als Isoliertheit eines Attributs gegenüber den anderen deuten.

Ein Vergleich mit Gl. 341-3 legt nahe, daß es sich bei r_i um das Quadrat eines Korrelationskoeffizienten handelt. Tatsächlich gilt $\cos^2(e_i \cdot l_i, v) = r_i$ (vgl. EEG 1986, Gl. 12) und $\cos^2(e_i \cdot l_i, l+v) = u_i$ für die i -te Komponente $e_i \cdot l_i$ des Beobachtungsvektors l und, da l und v orthogonal sind $\cos^2(e_i \cdot l_i, v) + \cos^2(e_i \cdot l_i, l+v) = 1$. Da aber der Korrelationskoeffizient ρ_{uv} zweier Funktionen $\underline{u} = a' \underline{x}$ und $\underline{v} = b' \underline{x}$ identisch mit dem Cosinus des Winkels zwischen a und b ist, sind $\sqrt{r_i}$ bzw. $\sqrt{u_i}$ die Korrelationskoeffizienten zwischen $e_i \cdot l_i$ und \underline{v} bzw. $e_i \cdot l_i$ und \underline{l} . Damit ist völlige Analogie zwischen Gl. 342-5 und 341-3 hergestellt.

Bem.: Die Verbesserungen v sind linear von w abhängig: $v = -CB \cdot (B'CB)^{-1} w$ und enthalten die gleiche Information, da $H(\underline{v} | \underline{w}) = 0$ gilt. Wir haben aber das Modell 342-1 wegen der einfacheren Herleitung verwendet.

Insgesamt hat die Diskussion der gegenseitigen Information von Attributen zum Zweck der Bewertung für die Zuordnung reichhaltige Querverbindungen zu den klassischen Verfahren der Parameterschätzung und der Zuverlässigkeitstheorie hergestellt. Damit lassen sich die Maße, wenn sie für symbolische Attribute verwendet werden, mit denen numerischer Attribute vergleichen. Wie

oben erwähnt, steht die Normierung der einzelnen Größen noch aus, die eine vollständige Verschmelzung der Maße erlaubt. Die folgenden Beispiele werden aber zeigen, daß man auch bei symbolischen Attributen, sobald ein Abstandsmaß - welcher Art auch immer - definiert ist, auf die Gleichungen für numerische Attribute übergehen kann und so die entwickelten Konzepte verknüpfen kann.

3.4.3 Beispiele zur unüberwachten Bewertung von Attributen

Die unüberwachte Beurteilung von Attributen wollen wir an einer Reihe von Beispielen diskutieren. Sie sollen zeigen, daß das Konzept sehr allgemein ist, um isolierte Punkte, auffällige Symbolfolgen oder seltene Merkmale zu finden.

Allen Fällen liegt das gleiche Schema der Berechnung zugrunde:

1. Es wird für eine Anzahl von n Zufallsvariablen a_i eine Stichprobe (a_{ik}) ($k=1, \dots, m_k$) vom Umfang m_k vorgegeben.

Bei gleichem Stichprobenumfang m kann man die einander zugeordneten Werte (gleiches k) als Attribute eines Primitivs verstehen. Die Aufgabe besteht darin, das auffälligste Attribut zu finden, bzw. die Attribute nach ihrer Seltenheit zu ordnen.

Anderenfalls ist eine komplementäre Interpretation möglich: Die n Stichproben repräsentieren Primitive, die nach ihrer Auffälligkeit zu ordnen sind.

2. Es wird ein Distanzmaß für die Zuordnung der Stichproben vorgegeben.

Bei gleichem Stichprobenumfang und reellen Variablen kann dies ein Euklidischer Abstand sein, den wir auch im folgenden verwendet haben.

Bei ungleichem Stichprobenumfang haben wir die Stichprobe als Realisierung eines stochastischen Prozesses und damit als Primitiv interpretiert, so daß die Bestimmung des Distanzmaßes auf eine Zuordnung der Primitive hinausläuft. Als Distanzmaß haben wir bei Symbolfolgen die ungewogene Levenshtein-Distanz als Maß verwendet.

3. Aus der Distanzmatrix $D = (d_{ij})$ wird mit Hilfe einer positiv definiten Korrelationsfunktion $r(d)$ eine Korrelationsmatrix $C = (c_{ij}) = (r(d_{ij}))$ abgeleitet. Wir haben als Funktionen $r(d)$ die auf $r(0) = 1$ normierten Dichten der Gauß-, der Laplace- und der Cauchyverteilung untersucht. Von geringfügigen Einflüssen der Korrelationsfunktion auf die Rangfolge der Attribute bzw. Primitive abgesehen, stimmten die Ergebnisse überein. Die normierte Dichtefunktion

$$r_{ij}(d) = \frac{1}{1 + (d_{ij}/d_0)^2}$$

der Cauchy-Verteilung stellte sich als numerisch am stabilsten heraus und wurde daher in allen folgenden Beispielen verwendet. Als Referenzdistanz d_0 wählten wir, ebenfalls gleich in allen Beispielen, $d_{\max}/4$. Der Referenzabstand d_0 hat auch einen vernachlässigbaren Einfluß auf die Rangfolge, kann aber u. U. die relativen Unterschiede der Maße stark verändern.

Als Seltenheitsmaß verwendeten wir

$$S_i = 1 / H(\underline{x}_i; \underline{y})$$

Hierin ist x_i die betrachtete Stichprobe, y alle anderen Stichproben. $H(x_i; y)$ wurde nach Gl. 341-3 mit dem totalen Korrelationskoeffizienten nach Gl. 341-7 berechnet.

3.4.3.1 Isolierte und seltene Punkte

a. Das erste Beispiel soll die Ähnlichkeit der Beurteilung der Isoliertheit von Punkten mit der Beurteilung der Kontrollierbarkeit der Koordinaten bei einer Transformation demonstrieren.

Gegeben ist die Punktanzordnung der Fig. 3-1

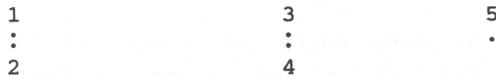


Fig. 3-2 Punktgruppe mit einem isolierten Punkt

Offensichtlich ist Punkt 5 isoliert. In der Tabelle ist die Seltenheit S_i und das Kontrollierbarkeitsmaß $1/r_i$ angegeben. Hierin ist r_i der Redundanzanteil, der sich bei einer Ähnlichkeitstransformation der Punkte ergeben würde (vgl. FÖRSTNER 1978)

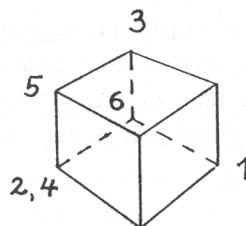
i	S_i	$1/r_i$
1	.364	1.75
2	.364	1.75
3	.363	1.29
4	.363	1.29
5	40.126	3.20

In beiden Fällen ergibt sich für Punkt 5 ein Maximum in der Trennbarkeit. Allerdings ist der Kontrast der Werte bei dem informationstheoretischen Maß deutlich größer. Dies ist stark durch die Referenzdistanz beeinflussbar. Der wesentliche Unterschied liegt in der Differenz des Kontrollierbarkeitsmaßes für die ersten beiden Punktpaare, die durch die Seltenheit kaum unterschieden werden. Dies läßt sich aus der unterschiedlichen Geometrie der Modelle erklären: Auf das Kontrollierbarkeitsmaß hat in diesem Fall der Abstand der Punkte zum Schwerpunkt einen bedeutenden Einfluß, während bei der Seltenheit nur die relative Dichte der Punkte im Sinne der Euklidischen Distanz eine Rolle spielt.

b. Das zweite Beispiel soll uns zeigen, daß identische Punkte zum Seltenheitsmaß 0 führen.

Die Tabelle enthält die dreidimensionalen Koordinaten und das Seltenheitsmaß von 6 Punkten

i	x y z	S
1	1 1 0	32.1
2	0 0 0	0
3	0 1 1	20
4	0 0 0	0
5	0 0 1	20.7
6	0 1 0	15.2



Die beiden identischen Punkte erhalten erwartungsgemäß das Seltenheitsmaß 0.

Wenn wir die Zeilen als Attribute interpretieren, so sind die Attribute 2 und 4 nicht unterscheidbar und Attribut 1 wäre dasjenige, das zuerst bei einer Zuordnung geprüft werden sollte.

Wir werden die Koordinaten später als Symbolfolge interpretieren und die Auffälligkeitsmaße miteinander vergleichen.

c. Als Anwendung wollen wir die Ecken eines Schachbretts daraufhin untersuchen, inwieweit sie für die Zuordnung geeignet sind. Es sind dies die Punkte, die mit einem Interestoperator gefunden würden.

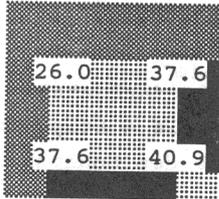


Fig 3-3 Teil eines Schachbretts
4 Eckpunkte mit Seltenheitsmaß

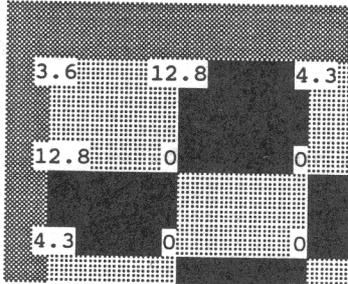


Fig 3-4 Teil eines Schachbretts
9 Eckpunkte mit Seltenheitsmaß

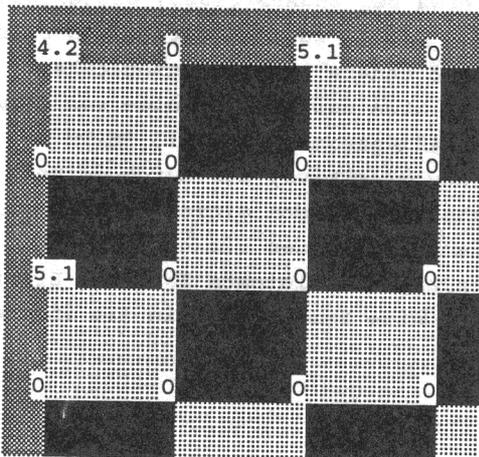


Fig 3-5 Teil eines Schachbretts
16 Eckpunkte mit Seltenheitsmaß

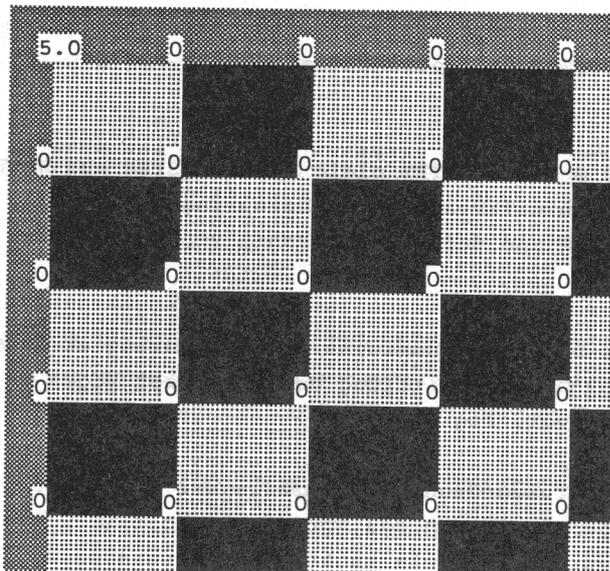


Fig 3-6 Teil eines Schachbretts
25 Eckpunkte mit Seltenheitsmaß

Als Attribute verwenden wir die Helligkeitswerte in den vier Quadranten. So kann jeder Eckpunkt des Schachbretts als Punkt im R^4 repräsentiert werden. Der Hintergrund habe den Helligkeitswert 2, die Felder des Schachbretts die Werte 1 und 3, so daß der Kontrast am Rand für alle Felder der gleiche ist (vgl. Fig. 2-2). Z. B. hat der linke obere Eckpunkt die Repräsentation (2,2,2,1).

In den Fig. 3-3 bis 3-6 sind die Seltenheitswerte für vier verschiedene Fälle dargestellt, die simulieren sollen, daß mehr und mehr von dem Schachbrett im Bild erscheint. Die 4 Eckpunkte in Fig. 3-3 zeigen kaum Unterschiede in der Auffälligkeit. Die linke obere Ecke des Schachbretts ist am wenigsten auffällig, da sie zu ihren Nachbarn große Ähnlichkeit hat. Von den 9 Punkten in der Fig. 3-4 sind nur die Randpunkte für die Zuordnung geeignet, da die Punkte in der Mitte doppelt vorkommen. Entsprechend sind in der Fig. 3-5 neben dem Eckpunkt des Schachbretts nur noch zwei Randpunkte auffällig. In Fig. 3-6 bleibt schließlich der Eckpunkt als einziger für die eine eindeutige Zuordnung erhalten.

Das Beispiel veranschaulicht die Abhängigkeit des Seltenheitsmaßes vom Kontext.

3.4.3.2 Auffällige Symbolfolgen

Symbolfolgen spielen in der Sprachanalyse eine wesentliche Rolle. In der Bildanalyse treten sie bei der Zuordnung von Bildern auf. Dabei werden, wie etwa bei BENARD (1983) diejenigen Kantenelemente, die Epipolarlinien schneiden, in geordneten Listen zusammengefaßt und repräsentieren dann die Grauwertfunktion der Epipolarlinien. Eine andere Anwendungsmöglichkeit besteht in der Codierung von Linien oder Kanten in Bildern, etwa durch eine Folge von geraden Kantenstücken, Kreissegmenten und nicht klassifizierbaren Liniensegmenten (vgl. SMITH und WOLF 1984). In beiden Fällen enthält die symbolische Beschreibung Primitive, die durch weitere Attribute gekennzeichnet sind. Die folgenden Beispiele behandeln Folgen einfacher Symbole, wie Buchstaben-, Ziffern- oder Zeichenketten. Als Ähnlichkeitsmaß verwenden wir die einfache Levenshtein-Distanz (vgl. KOHONEN 1984). Sie ist definiert als die kleinste Zahl von Eliminations- bzw. Austauschoperationen, die benötigt wird, um die eine Symbolfolge in die andere zu transformieren:

$$D(s,t) = \min (n_1 + n_2 + n_{12})$$

Hierin sind

- n_1 : die Zahl der in der ersten Symbolfolge s zu überspringenden Symbole
- n_2 : die Zahl der in der zweiten Symbolfolge t zu überspringenden Symbole
- n_{12} : die Zahl der auszutauschenden Symbole

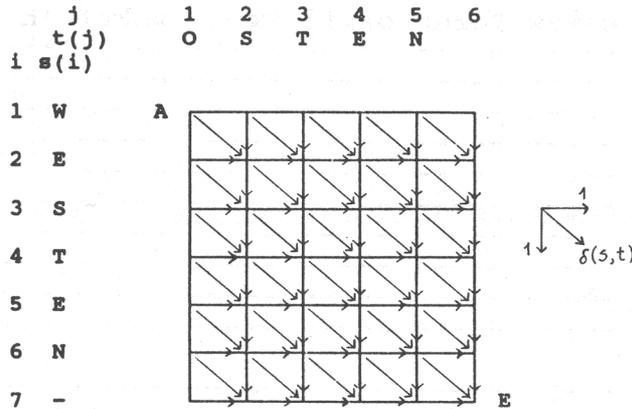
Z. B. ist die Levenshtein-Distanz von s = WESTEN und t = OSTEN gleich 2, da OSTEN aus WESTEN entweder durch Austauschen von W in WESTEN und O in OSTEN und durch Überspringen von E in WESTEN oder aber durch Überspringen von W in WESTEN und durch Austauschen von E in WESTEN und O in OSTEN entstehen kann.

Für allgemeine Symbolfolgen kann man die Distanz D durch dynamische Programmierung ermitteln. Insbesondere läßt sich die Bestimmung der Distanz D als Ermittlung des kürzesten Weges in einem gerichteten, bewerteten Graphen darstellen (vgl. Fig. 3-7)

Der Weg von A nach E über die Kanten entspricht der Zuordnung der Symbole der Symbolfolgen s(i) und t(j). Eine vertikale Kante entspricht dem Auslassen des vorigen Symbols in s(i), eine horizontale Kante entspricht dem Auslassen des vorigen Symbols in t(j), während eine diagonale Kante keiner Auslassung entspricht. Die Kosten vertikaler und horizontaler Kanten sind eins.

Dagegen hängt der Wert einer diagonalen Kante von $(i-1, j-1)$ nach (i, j) davon ab, ob die Symbole $s(i-1)$ und $t(j-1)$ identisch sind oder nicht, bei Identität ist der Wert der Kante 0, sonst 1. Damit läßt sich der Algorithmus zur Bestimmung der kürzesten Distanz wie folgt schreiben:

Fig. 3-7 Bestimmung der Levenshtein-Distanz als kürzestem Weg in einem gerichteten, bewerteten Graphen



1. $d(1,1) = 0$
2. for $i = 2$ to $|s|+1$
 $d(i,1) = d(i-1,1) + 1$
3. for $j = 2$ to $|t|+1$
 $d(1,j) = d(1,j-1) + 1$
4. for $i = 2$ to $|s|+1$
 for $j = 2$ to $|t|+1$
 $d1(i,j) = d(i-1,j) + 1$
 $d2(i,j) = d(i,j-1) + 1$
 $d12(i,j) = d(i-1,j-1) + \delta(s(i-1),t(j-1))$
 $d(i,j) = \min(d1(i,j), d2(i,j), d12(i,j))$
5. $D(s,t) = d(|s|+1, |t|+1)$

Die Lösung muß nicht eindeutig sein. Durch Rückwärtsrechnung kann man auch die Zuordnung bestimmen. Für das Paar $s = \text{WESTEN}$ und $t = \text{OSTEN}$ erhalten wir die Zuordnung der Fig. 3-8, in der der optimale Weg über die mit 0 und 1 gekennzeichneten Knoten verläuft. Dabei bedeutet eine 0 eine richtige Zuordnung, eine 1 eine falsche Zuordnung.

	O	S	T	E	N	
W	1	-	-	-	-	Fig. 3-8 Zuordnung der Symbolfolgen WESTEN und OSTEN
E	1	1	-	-	-	
S	-	0	-	-	-	
T	-	-	0	-	-	
E	-	-	-	0	-	
N	-	-	-	-	0	

Fig. 3-9 zeigt die Zuordnung von zwei längeren Symbolfolgen, die man sich als Texturcodierung zweier Epipolarlinien vorstellen kann. Aus der Zuordnung läßt sich das Parallaxenprofil ableiten, das offensichtlich nicht eindeutig ist.

Eine vollständige Beschreibung des verwendeten Algorithmus findet man bei HEIN (1977, vgl. auch OOMMEN 1987).

Wir wollen an drei Beispielen die Nutzung dieses Maßes für die Bestimmung der Auffälligkeit einer Symbolfolge innerhalb einer Liste von Symbolfolgen diskutieren.

Fig. 3-9 Zuordnung zweier Symbolfolgen
 Optimaler Weg über die mit 0 oder 1 gekennzeichneten
 Knoten
 0 = Identische Symbole
 1 = unterschiedliche Symbole

Match	54323	31221	asdfq	wetzw	#eruw	oryfl	vgöcj	ghkkf	+h.12	341
3	11010	-----	-----	-----	-----	-----	-----	-----	-----	---
4	-0-11	1----	-----	-----	-----	-----	-----	-----	-----	---
1	--111	10---	-----	-----	-----	-----	-----	-----	-----	---
2	-----	--0--	-----	-----	-----	-----	-----	-----	-----	---
2	-----	---0-	-----	-----	-----	-----	-----	-----	-----	---
1	-----	----0	-----	-----	-----	-----	-----	-----	-----	---
q	-----	-----	11110	-----	-----	-----	-----	-----	-----	---
w	-----	-----	-----	0--	-----	-----	-----	-----	-----	---
e	-----	-----	-----	-0-	-----	-----	-----	-----	-----	---
r	-----	-----	-----	--1-	-----	-----	-----	-----	-----	---
z	-----	-----	-----	--0-	-----	-----	-----	-----	-----	---
w	-----	-----	-----	---0	-----	-----	-----	-----	-----	---
e	-----	-----	-----	-----	10-	-----	-----	-----	-----	---
r	-----	-----	-----	-----	-0-	-----	-----	-----	-----	---
h	-----	-----	-----	-----	-1-	-----	-----	-----	-----	---
w	-----	-----	-----	-----	-0-	-----	-----	-----	-----	---
o	-----	-----	-----	-----	-----	0-	-----	-----	-----	---
r	-----	-----	-----	-----	-----	-0-	-----	-----	-----	---
y	-----	-----	-----	-----	-----	-0-	-----	-----	-----	---
f	-----	-----	-----	-----	-----	--0-	-----	-----	-----	---
d	-----	-----	-----	-----	-----	--11	-----	-----	-----	---
f	-----	-----	-----	-----	-----	--01	-----	-----	-----	---
d	-----	-----	-----	-----	-----	--11	-----	-----	-----	---
f	-----	-----	-----	-----	-----	--01	-----	-----	-----	---
d	-----	-----	-----	-----	-----	--11	-----	-----	-----	---
f	-----	-----	-----	-----	-----	--01	-----	-----	-----	---
d	-----	-----	-----	-----	-----	--11	-----	-----	-----	---
f	-----	-----	-----	-----	-----	--01	-----	-----	-----	---
ö	-----	-----	-----	-----	-----	--11	-----	-----	-----	---
g	-----	-----	-----	-----	-----	--11	-----	-----	-----	---
i	-----	-----	-----	-----	-----	--11	-----	-----	-----	---
f	-----	-----	-----	-----	-----	--01	-----	-----	-----	---
l	-----	-----	-----	-----	-----	--0-	-----	-----	-----	---
v	-----	-----	-----	-----	-----	-----	0-	-----	-----	---
g	-----	-----	-----	-----	-----	-----	-0-	-----	-----	---
ö	-----	-----	-----	-----	-----	-----	-0-	-----	-----	---
c	-----	-----	-----	-----	-----	-----	-0-	-----	-----	---
j	-----	-----	-----	-----	-----	-----	-0-	-----	-----	---
g	-----	-----	-----	-----	-----	-----	0-	-----	-----	---
x	-----	-----	-----	-----	-----	-----	-----	-1111	1----	---
c	-----	-----	-----	-----	-----	-----	-----	--111	11----	---
v	-----	-----	-----	-----	-----	-----	-----	--11	111--	---
1	-----	-----	-----	-----	-----	-----	-----	---1	1110-	---
2	-----	-----	-----	-----	-----	-----	-----	-----	---0	---
3	-----	-----	-----	-----	-----	-----	-----	-----	-----	0-
1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-10

Distance = 32.0

a: Gegeben seien die Symbolfolgen (diesmal auch Kleinbuchstaben)

Westen
Osten
Süden
Arktis

Die Distanzmatrix ergibt sich zu

$$D = \begin{bmatrix} 0 & 2 & 4 & 5 \\ 2 & 0 & 3 & 5 \\ 4 & 3 & 0 & 6 \\ 5 & 5 & 6 & 0 \end{bmatrix}$$

Daraus errechnen sich die Seltenheits- oder Auffälligkeitswerte

i	Folge	S
1	Arktis	115.1
2	Süden	27.1
3	Westen	8.9
4	Osten	7.5

Die Rangfolge der Auffälligkeit stimmt mit der Intuition überein. Insbesondere führt die große Zahl gemeinsamer Buchstaben von 'Westen' und 'Osten' zu einer kleinen Auffälligkeit, während s das Wort 'Arktis' aus den übrigen durch wenigstens 5 Operationen ergibt und daher am auffälligsten ist.

b. Im zweiten Beispiel wollen wir die 6 Punkte von Beispiel b. des vorigen Abschnitts aufgreifen. Wir interpretieren aber jetzt das Koordinatentripel als Folge dreier Symbole. Die Seltenheitswerte für die Interpretation als Koordinatentripel und als Symbolfolge sind .

i	Attribut	S(Koord.)	S(Symbolf.)
1	110	31.10	18.23
2	000	0	0
3	011	19.95	9.00
4	000	0	0
5	001	20.66	8.98
6	010	15.24	6.28

Die Rangfolge ist bis auf eine unerhebliche Verwechslung die gleiche. Der Grund ist der in diesem Fall geringe Unterschied der Distanzmatrizen.

c. Das dritte Beispiel soll die Abhängigkeit des Auffälligkeitsmaßes vom verwendeten Distanzmaß diskutieren.

Gegeben seien die 9 Worte:

Amsel
Baum
Braten
Fell
Haus
Purzelbaum
Samson
Simson
Veronica

Mit der Levenshtein-Distanz erhalten wir folgende Rangfolge der Auffälligkeit

	S
1. Purzelbaum	50.81
2. Veronica	36.15
3. Braten	10.07
4. Fell	7.95
5. Amsel	7.48
6. Baum	2.52
7. Haus	2.50
8. Simson	0.90
9. Samson	0.88

Es ist eine gewisse Tendenz zu erkennen, daß längere Worte gegenüber kürzeren ein höheres Auffälligkeitsmaß erhalten. Man ist daher geneigt, normierte Distanzen zu verwenden, etwa

$$D1(s,t) = D(s,t) / \sqrt{|s| \cdot |t|}$$

oder

$$D2(s,t) = D(s,t) / \max(|s|, |t|).$$

Damit erhalten wir die folgenden Rangordnungen

$D/\sqrt{ s t } \rightarrow$	S	$D/\max(s , t) \rightarrow$	S
1. Purzelbaum	19.11	1. Fell	19.02
2. Fell	17.60	2. Veronika	18.96
3. Veronika	17.24	3. Braten	13.63
4. Bräten	11.07	4. Amsel	12.84
5. Amsel	10.34	5. Purzelbaum	12.64
6. Baum	5.33	6. Haus	7.68
7. Haus	5.27	7. Baum	7.33
8. Samson	0.86	8. Simson	1.15
9. Simson	0.84	9. Samson	1.12

Es scheint ein Nachteil zu sein, daß die Rangfolge so stark vom verwendeten Distanzmaß abhängt. Unser Ausgangspunkt war es aber, eine Reihenfolge von Primitiven, hier Symbolfolgen, zu finden, die eine sichere Zuordnung verspricht. Dafür sind längere Symbolfolgen besser geeignet als kurze. Außerdem sind in allen drei Fällen die vier Worte mit den niedrigsten Auffälligkeitsswerten identisch. Sie sollten wegen Verwechslungsgefahr nicht als erste bei der Zuordnung verwendet werden. Unter diesen Gesichtspunkten leistet das vorgeschlagene Auffälligkeitsmaß genau den gewünschten Zweck.

4. ZUSAMMENFASSUNG UND AUSBLICK

1. Die Arbeit will einen Beitrag zu den theoretischen Grundlagen der Bildanalyse leisten. Die Qualitätsbeurteilung innerhalb der Verfahrensschritte steht dabei im Vordergrund. Sie verlangt eine durchgängige und konsistente Repräsentation der Unsicherheit aller Zwischenergebnisse, die sich daher auf ein einfaches Bildmodell zurückführen lassen und in ein brauchbares, d. h. interpretierbares Gütemaß für das Analyseergebnis überführen lassen sollte. Unter den technischen Hilfsmitteln zur Beschreibung von Unsicherheit nimmt die Wahrscheinlichkeitstheorie und die Statistik wegen ihrer vielfältigen Querbezüge eine dominante Rolle ein und durchzieht daher alle Teile der vorliegenden Arbeit.

Unter den zahlreichen Ursachen für die Schwierigkeiten bei der Bildanalyse, wie sie im 1. Kapitel zusammengestellt sind, haben wir drei Problemkreise ausgewählt, mit dem Ziel, dafür eine Vereinheitlichung der theoretischen Grundlagen zu versuchen. Dazu haben wir elementare Bildanalyseprozesse aus dem Bereich der Bildzuordnung und der Objektlokalisierung herangezogen und an ihnen exemplarisch aufgezeigt, wie sich ein solcher Vereinheitlichungsprozeß gestalten kann. Im einzelnen haben wir folgende Ergebnisse erzielt:

a. Der erste Schritt besteht im Versuch der Vereinheitlichung von Verfahren. Wir konnten zeigen, daß sich der vorgeschlagene Interestoperator zur Extraktion markanter Punkte, der sich auf ein funktionales Modell der diskreten Bildfunktion stützt, als Interpretation von Texturmerkmalskarten deuten läßt, die auf ein stochastisches Modell für die Bildfunktion zurückgreifen. Die Interpretation läßt sich auf die Extraktion von Kantenelementen übertragen und erlaubt damit eine - zumindest aus theoretischer Sicht einheitliche Darstellung des Übergangs von der ikonischen zu einer symbolischen Repräsentation des Bildinhalts.

Das Texturmodell führte parallel dazu auf ein informationserhaltendes Filter. Es läßt sich als adaptives Wiener-Filter formulieren und ermöglicht daher, wie bei den klassischen Restaurierungsmethoden, bei gegebenem Bildmodell eine Abschätzung der Genauigkeitseigenschaften des Ergebnisses. Es ist damit konzeptionell und, wie die Beispiele zeigen, auch praktisch anderen Kanten und Ecken erhaltenden Filtern überlegen. Damit verbunden war die Entwicklung eines robusten Schätzverfahrens für die Rauschvarianz.

Schließlich stellte sich das bei der Bildzuordnung von BOYER und KAK (1986, 1988) verwendete informationstheoretisch begründete Konzept für die Bestimmung der Ähnlichkeit symbolischer Attribute als geeignet heraus, eine Verbindung zwischen der Bewertung numerischer und symbolischer Attribute herzustellen. Es stellt eine Verallgemeinerung der Korrelationsverfahren und so auch der Kleinsten Quadrate Methoden zur Bildzuordnung dar. Damit wurde nicht nur eine hohe Flexibilität der Repräsentation gewonnen, sondern auch der Anschluß der statistischen Methoden der Meßtechnik an die heuristischen Suchverfahren der Künstlichen Intelligenz, wie sie bei der merkmalsgestützten Bildzuordnung und der Objekterkennung gebräuchlich sind, erreicht.

b. Der zweite Problemkreis betrifft die statistische Durchdringung des Übergangs von der ikonischen zur symbolischen Repräsentation des Bildinhalts. Ihr kommt eine Schlüsselrolle innerhalb der Bildanalyse zu. Denn von der damit ermöglichten Qualitätsabschätzung hängt die Beurteilung innerhalb der Folgeschritte unmittelbar ab.

Wir haben uns im Abschnitt 2.2 ausführlich mit den Genauigkeitseigenschaften der Extraktion markanter Punkte auseinandergesetzt. Obwohl das verwendete Rechenmodell gegenüber dem bei der Texturanalyse verwendeten, realistischeren Modell eine Näherung darstellt, erlaubt das Extraktionsverfahren, insbesondere aus den Widersprüchen zwischen Rechenmodell und Daten und der Geometrie der Bildfunktion, die Kovarianzmatrix des extrahierten Punktes abzuleiten. Das

Verfahren läßt sich auf Kantenelemente übertragen und stellt auch dort eine Qualitätsangabe zur Verfügung.

Der sich nach der Extraktion von Kantenelementen anschließende Aggregierungsprozeß ist ein typischer Weiterverarbeitungsschritt, bei dem die Genauigkeitsschätzungen der vorhergehenden Schritte verwendet werden und damit - indirekt - die Genauigkeit der diskreten Bildfunktionswerte für die Schätzung der Präzision, hier der geraden Kantenstücke, genutzt wird.

Uns stehen damit aus digitalen Bildern abgeleitete geometrische Größen wie in der Meßtechnik zur Verfügung, zusammen mit einer internen Qualitätsabschätzung, und zwar in einer für die Weiterverarbeitung idealen Repräsentation, der Kovarianzmatrix der Meßgrößen.

c. Für die Nutzung der extrahierten geometrischen Grundelemente, Primitive innerhalb eines Zuordnungsprozesses für die dreidimensionale Objektrekonstruktion oder die Objekterkennung - oder -lokalisierung, ist schließlich ihre Auffälligkeit bzw. Seltenheit und die Invarianz von zentraler Bedeutung; denn auch ohne quantitative Untersuchung der Wirkung dieser Vorinformation auf die durchschnittlichen Suchzeiten bei der Zuordnung ist ihr positiver Einfluß evident.

In unserem Zusammenhang ist zunächst auch weniger die Komplexität des Zuordnungsprozesses als die Definition der Begriffe Auffälligkeit und Invarianz von Interesse. Auf der Basis des informationstheoretischen Ähnlichkeitsmaßes gelang eine einheitliche Darstellung beider Maße, die sowohl für numerische als auch für symbolische Primitive, Merkmale oder Attribute geeignet ist.

Dabei stellten sich enge Verbindungen einerseits mit den Beurteilungstechniken bei der überwachten und unüberwachten Klassifikation als auch mit der Trennbarkeitstheorie in linearen Modellen heraus. Insbesondere konnte auch die "Markantheit" der nach Abschn. 2.2 extrahierten Punkte als lokale Auffälligkeit oder lokale Trennbarkeit (von Nachbarfenstern) interpretiert werden. Damit sind diese in der Bildanalyse häufig verwendeten Begriffe nicht nur verwandt, sondern - sofern man das geeignete Modell heranzieht - mathematisch identisch (Seltenheit, Markantheit, Auffälligkeit, Trennbarkeit) bzw. komplementär (Invarianz).

Inzwischen stellten CHEN und MULGAONKAR (1990) ein Verfahren vor, das - ebenfalls auf der Basis von Wahrscheinlichkeiten - die Nützlichkeit von Objektmerkmalen im Hinblick auf die optimale Rechenzeit des Suchverfahrens bestimmt. Die Bewertung stützt sich auf die Detektierbarkeit, die Fehlerrate bei der Merkmalsextraktion und die Wahrscheinlichkeit für die Richtigkeit der während des Suchprozesses postulierten Hypothesen.

2. Neben diesen theoretischen Ergebnissen lassen sich die in der Arbeit entwickelten Verfahren direkt anwenden bzw. haben unmittelbare praktische Bedeutung:

a. Wichtigster Gesichtspunkt ist die Ausschöpfung des in den Bildern enthaltenen Genauigkeitspotentials. Erste kontrollierte Versuche zur Ableitung digitaler Geländemodelle (HAHN und FÖRSTNER 1988) zeigen, daß mit dem vorgeschlagenen Operator zur Extraktion markanter Punkte Genauigkeiten erreichbar sind, die nur um einen Faktor kleiner als 2 ungünstiger sind als die vom Operateur erzielbaren. Dabei wird eine hohe Zuverlässigkeit, d. h. niedrige Ausfallquoten erreicht, wenn man sich an die von der Konzeption her bekannten Randbedingungen des Verfahrens (glatte Oberfläche) hält, also etwa Gebiete mit lichtem Wald oder Siedlungen ausschließt. Die Verwendung eines Kleinsten Quadratezuordnungsverfahrens an den vom Interestoperator ausgewählten Stellen führt dann noch zu einer Steigerung der Genauigkeit von einem Faktor 1.5 oder mehr und liefert damit Ergebnisse, die in der Qualität manuellen Messungen nicht nachstehen.

b. Die vorgeschlagenen Verfahren sind, abgesehen von der Verknüpfung der Kantenelemente zu geraden Kantenstücken, parallelisierbar und unter Nutzung geeigneter Rechnerarchitekturen realtimefähig. Etwa läßt sich die automatische Auswahl der Fenster von markanten Punkten (s. Abschn. 2.2.4) auf einem Pipelineprozessor so realisieren, daß die Mitten der selektierten Fenster bei zeilenweisem Einlesen der Bilder mit einer Verzögerung von 11 Bildzeilen zur Verfügung stehen, wenn mit 5 x 5-Bildfenstern für die Faltung und die Nichtmaximum-Unterdrückung gearbeitet wird (FÖRSTNER/GÜLCH 1987). Die Zeit für die Klassifikation der Fenster und die präzise Lokalisierung der markanten Punkte kann man bei entsprechender Parallelisierung demgegenüber vernachlässigen. Die auf diese Weise extrahierten Punkte kann man etwa für die Verfolgung von Objekten oder die Bestimmung der Trajektorie der Kamera aus Bildfolgen, z. B. bei der Fahrzeugnavigation oder für eine schnelle Vermessung der Form texturierter Oberflächen verwenden.

c. Die Repräsentation von Bildern durch gerade Kantenstücke hat wohl die breiteste Anwendung vor allem für die Bildinterpretation (vgl. NEVATIA/BABU 1980, HERMAN/KANADE 1986, BURNS et. al. 1986, VORHEES/POGGIO 1987, BOYER/KAK 1988). Eine wichtige Anwendung im Bereich der Photogrammetrie und der Fernerkundung ist die Lokalisierung natürlicher Paßpunkte in Luft- und Satellitenbildern. Gegenüber der ikonischen Repräsentation hat die Darstellung natürlicher Paßpunkte durch Kanten den Vorteil gegenüber Beleuchtungsänderungen weitgehend invariant zu sein. Darüberhinaus kann man die Unsicherheit von Strichvorlagen, etwa von Skizzen oder Karten auf die gleiche Weise, also durch die Kovarianzmatrix der Kantenendpunkte, repräsentieren wie die aus Bildern extrahierten geometrischen Elemente und so bei der Zuordnung auf die klassischen Konzepte insbesondere die Hypothesentests und die Zuverlässigkeitsmaße aus der Analyse geodätischer Messungen zurückgreifen (vgl. FÖRSTNER 1988). Diese Konzepte lassen sich auch unmittelbar auf die Zuordnung von Bildern zum Zweck der gegenseitigen Entzerrung oder der Objektvermessung übertragen und eröffnen damit vielfältige Anwendungsmöglichkeiten.

3. Die in der Arbeit entwickelten Einzellösungen lassen eine Reihe wesentlicher Fragen offen, die einerseits eine Detailanalyse verlangen, andererseits grundsätzlicher Natur sind und daher hier nur angerissen werden können.

a. Das informationserhaltende Filter ist auf seine Qualität hin, insbesondere im Vergleich zu anderen Verfahren zu untersuchen. Von speziellem Interesse ist die Abhängigkeit von der Art des Rauschen und der Fensterwahl. Da die hier gezeigten Beispiele sich auf eine Näherungslösung beziehen, ist zu untersuchen, inwieweit dadurch die Qualität der gefilterten Bildfunktion beeinflußt wird und inwieweit strengere Lösungen erreichbar sind, die etwa durch Verwendung von Mehr-Gitterverfahren, wie bei TERZOPOULOS (1986a), effizient realisiert werden können. Diese sollten aber auch in der Lage sein das Problem der Schätzung der Rauschvarianz, etwa mit Hilfe von Verfahren der Varianzkomponentenschätzung, zu lösen.

b. Der Interestoperator zur Selektion markanter Punkte basiert auf einer Näherung für das Bildmodell gegenüber dem des informationserhaltenden Filters. Die Auswirkung dieser Näherung ist unbekannt, kann aber ohne weiteres dazu führen, daß gegenüber einer strengen Lösung die Genauigkeit um den Faktor 2 geringer ausfällt. Erfahrungen mit dem Punktefinder in Multispektralbildern, etwa zur Satellitenbildentzerrung, fehlen bisher völlig.

c. Bei Verwendung kleiner Bildfenster ist der Operator zur Extraktion von Kantenelementen nicht notwendigerweise optimal, nicht nur weil dasselbe, nur genähert gültige Rechenmodell verwendet wird. Hier wäre nach der Güte, insbesondere auch dem Bias der tatsächlich extrahierten Kantenelemente in einem kontrollierten Versuch zu fragen.

Wichtiger ist allerdings die Entwicklung eines Verfahrens zur Extraktion von geraden oder auch anders geformten Kantenstücken, das parallelisierbar ist, wie etwa das von BURNS et. al.

(1986). Hierbei stellt sich ebenfalls die Frage nach dem Modell der Bildfunktion, das dem Verfahren zugrunde liegt, um zuverlässige Qualitätsangaben zu erhalten. Insbesondere fehlt bisher ein leistungsfähiges, parallelisierbares Verfahren zur Extraktion von geometrischen Grundelementen, wie Kreisen oder Ellipsen, das sowohl unter ungünstigen Rauschverhältnissen arbeitet als auch Aussagen über die Güte des Ergebnisses liefert.

d. Relationen zwischen den geometrischen Grundelementen haben wir in der Arbeit nicht behandelt. Hier sind, wie bei den Attributen, Relationen zweier Typen zu unterscheiden: numerische und symbolische. Die Zuordnung von Relationen zu diesen Typen ist nicht immer eindeutig. Numerische Relationen haben den Vorteil sich in natürlicher Weise in die statistische Modellierung einbinden zu lassen. Etwa ist die Relation "parallel(a,b)" bezüglich der Kanten a und b symbolisch und entzieht sich einer Koppelung mit geometrischen Größen stochastischer Eigenschaften, während die Relation "Winkel(a,b) = 0" numerisch ist und sich unmittelbar innerhalb eines geometrischen Modells verwenden läßt, wobei sogar der Grad der Parallelität durch die Angabe der Unsicherheit des Winkels etwa mit einer Standardabweichung beschrieben werden kann, eine selbstverständliche Repräsentationsform in der statistischen Analysetechnik. Sie wurde vor kurzem erst im Bereich der Künstlichen Intelligenz aufgegriffen zur Darstellung der Unsicherheit von Positionen eines Roboters in Bezug auf seine frühere oder eine andere Referenz, wobei die aus der Geodäsie bekannten S-Transformationen (BAARDA 1973) zur Transformation von Kovarianzmatrizen neu entwickelt wurden (SMITH et. al. 1987a,b).

Die hier angesprochene Querverbindung zwischen symbolischen und numerischen Relationen, die im Bereich geometrisch-physikalischer Modelle fast immer eine reellwertige Repräsentation mit den zugehörigen statistischen Eigenschaften erlaubt, läßt vermuten, daß bei meßtechnischen Bildanalyseproblemen, wie der Bildzuordnung und der Objektrekonstruktion oder -lokalisierung in allen Fällen und bei der Objekterkennung wenigstens zum Teil, rein symbolische Attribute bzw. Relationen vermieden werden können und so eine einheitliche Modellbildung in den allermeisten Fällen erreicht werden kann.

e. Damit ist die gemeinsame Grundlage symbolischer und numerischer Attribute von Abschnitt 3 angesprochen. Obwohl für Ähnlichkeit und, daraus abgeleitet für Seltenheit und Invarianz, ein einheitliches Konzept auf der Basis der Informationstheorie verwendet wurde, wurde kein gemeinsames Maß für numerische und symbolische Attribute gefunden. Dies liegt begründet in der unterschiedlichen Dimension der Information $-\log P(\underline{a} = a)$ und der differentiellen Information $-\log p(\underline{x} = x)$. Die behandelten Fragestellungen ließen eine Trennung von symbolischen und numerischen Attributen zu, so daß die Transinformation, als Differenz zweier Entropien ein für die Ähnlichkeit geeignetes Maß ist, das insbesondere bei numerischen Attributen, nach einmal gewähltem Maßstab für die Zufallsvariablen, maßstabsinvariant ist. Eine Koppelung der Distanzen symbolischer und numerischer Attribute, etwa durch Addition der Transinformationswerte, ist aber nicht unmittelbar möglich, wie etwa der Übergang vom symbolischen Prädikat Kantengerichtung zu einem numerischen Prädikat verdeutlicht.

Von besonderem Interesse ist aber in diesem Zusammenhang der Vorschlag von GEORGEFF und WALLACE (1984). Sie messen die "Erklärbarkeit" eines Datensatzes durch ein Modell mit der Zahl der bits, die wenigstens notwendig ist, um die Daten mit einer verlangten Genauigkeit zu beschreiben. Dadurch werden Modelle verschiedener Komplexität als Erklärungsgrundlage für vorgelegte Daten vergleichbar, ein Verfahren das den Kriterien zur Bestimmung der Ordnung autoregressiver Prozesse verwandt ist. Dabei werden offensichtlich symbolische und numerische Attribute, etwa in der Form der Prozeßordnung und des Prädiktionsfehlers, gemeinsam beurteilt. Das Beurteilungsmaß wurde unabhängig von RISSANEN (1983, 1987) unter dem Namen "Minimum Description Length" vorgeschlagen. Es wurde inzwischen im Bereich der Bildanalyse für verschiedenartige Aufgaben eingesetzt, etwa zur Bildzuordnung (SMITH/WOLF 1984), zur Bildrestaurierung

(LECLERC 1988) oder zur Identifizierung von Objekten, die durch generische Modelle vorgegeben sind, wie z. B. Gebäude oder Vegetationsbereiche (FUA/HANSON 1987a,b, 1988).

Da die Objekterkennung aus Bildern eine besondere Art der Erklärung von Daten darstellt, und dabei die Form der Objekte eine wesentliche Rolle spielt, wie Skizzen eindrücklich belegen, sei hier der Vorschlag von v. WEIZSÄCKER (1974) zum Problem der Information eines Gegenstandes genannt:

"Man kann nun als ein quantitatives Maß der Menge an Form eines Gegenstandes die Zahl von einfachen Alternativen¹ bezeichnen, die entschieden werden müssen, um seine Form zu beschreiben. In diesem Sinn mißt dann die Information, die in dem Gegenstand enthalten ist, genau die Menge seiner Form." (1974, S. 347f, ¹bits)

Gemeinsam mit den Konzepten von GEORGEFF/WALLACE und RISSANEN ist die Zahl der bits als Maß für den Informationsgehalt eines Gegenstands und die gleichartige Abhängigkeit dieses Maßes sowohl vom Vorwissen als auch von dem Grad der Auflösung. Im Zusammenhang mit der Objekterkennung stellt sich hier die Frage, ob die Konzepte sich als Grundlage einer Theorie der Form eignen. Ansätze dazu finden sich in den genannten Arbeiten von FUA und HANSON. Die Anwendbarkeit des Konzepts in der Bildanalyse wird durch die Arbeit von SMITH und WOLF (1984) überzeugend belegt. Die Bedeutung einer Theorie der Form, die gleichzeitig eine Theorie der Textur wäre, kann nicht hoch genug eingeschätzt werden und läßt den Versuch ihrer Entwicklung auf der Basis der Konzepte von GEORGEFF/WALLACE und RISSANEN lohnend erscheinen.

Verwendete Literatur

Abkürzungen:

CGIP	Computer Graphics and Image Processing,
CVGIP	Computer Vision, Graphics and Image Processing,
DGK	Deutsche Geodätische Kommission, München
IAP	International Archives of Photogrammetry and Remote Sensing
ICCV	First International Conference on Computer Vision, London 1987
IEEE T-PAMI	IEEE Transaction on Pattern Analysis and Machine Intelligence
IJCAI	International Joint Conference on Artificial Intelligence
IUW	Proceedings: Image Understanding Workshop, Los Angeles, 1987
PRL	Pattern Recognition Letters
SIPUS	Schriftenreihe des Instituts für Photogrammetrie der Universität Stuttgart
ZfV	Zeitschrift für Vermessungswesen

- Aloimonos J. Y., Weiss I. and Bandyopadhyay A. (1987): Active Vision, ICCV, London 1987, p. 35-54
- Akey M. L. and Mitchell R. (1984): Detection and Subpixel Location of Objects in Digitized Aerial Images, 7th Conf. on Pattern Recognition, Montreal, 1984
- Baarda W. (1967): Statistical Concepts in Geodesy, Netherlands Geodetic Commission, New Series, Vol. 2, No. 4, Delft 1967
- Baarda W. (1968): A Testing Procedure for Use in Geodetic Networks, Netherlands Geodetic Commission, New Series, Vol. 2, No. 5, Delft 1968
- Baarda W. (1973): S-Transformations and Criterion Matrices Netherlands Geodetic Commission, New Series, Vol. 5, No. 1, Delft 1973
- Bajcsy R. und Solina F. (1987): Three Dimensional Object Representation Revisited, ICCV, London 1987, p. 231-240
- Ballard D. H. and Brown C. M. (1983): Computer Vision, Prentice-Hall, N. J., 1983
- Barnard S. T and Thompson W. B. (1980): Disparity Analysis of Images, IEEE T-PAMI, Vol. 2, No. 5, 1980, p. 333-340
- Benard M. (1983): Automatic Stereophotogrammetry: A Method Based on Feature Detection and Dynamic Programming, Specialist Workshop on Pattern Recognition in Photogrammetry, Graz 1983
- Berger T. (1971): Rate Distortion Theory, Prentice-Hall, N. J., 1971
- Bigün J. and Granlund G. H. (1987): Optimal Orientation Detection of Linear Symmetry, ICCV, London 1987, p. 433-438
- Bigün J. (1990): A Structure Feature for Some Image Processing Applications Based on Spiral Functions, CVGIP 51, 2, 1990, pp. 166-194
- Blanz W. E. (1986): Feature Selection and Polynomial Classifiers for Industrial Decision Analysis Research Report, IBM Almaden Research Center, 1986
- Blostein S. T. and Huang T. S. (1987): Quantization Errors in Stereo Triangulation, ICCV, London 1987, p. 325-334
- Bock H. H. (1973): Automatische Klassifikation, Vandenhoeck & Ruprecht, Göttingen, 1973
- Bouthemy P. and Benveniste A. (1985): Modeling of Atmospheric Disturbances in Meteorological Pictures, IEEE T-PAMI, Vol. 6, No. 5, 1984, p. 587-600
- Boyer K. L. und Kak A. C. (1986): Symbolic Stereo from Structural Descriptions School of Electrical Engineering, Purdue University, West Lafayette, TR-EE 86-12
- Boyer K. L. und Kak A. C. (1988): Structural Stereopsis for 3-D Vision, IEEE T-PAMI, Vol. 10, No. 2, 1988, p. 144-166

- Burns J. B., Hanson A. R. and Riseman E. M. (1986): Extracting Straight Lines, IEEE T-PAMI, Vol. 8, No. 4, 1986, p. 425-455
- Canny J. (1986): A Computational Approach to Edge Detection, IEEE T-PAMI, Vol. 8, No. 6, 1986, p. 679-698
- Castleman K. R. (1979): Digital Image Processing, Prentice Hall, N. J., 1979
- Cheeseman P. (1985): In Defense of Probability, IJCAI IX, Los Angeles, 1985
- Chen C., Mulgaonkar P. G. (1990): Robust Vision-Programs Based on Statistical Feature Measurements, Int. Workshop on Robust Computer Vision, Seattle, WA, USA, 1-3 Oct. 1990, pp. 39-56
- Chin R. T. and Yeh C. L. (1983): Quantitative Evaluation of some Edge-Preserving Noise-Smoothing Techniques, CVGIP 23, 1983, pp. 67-91
- Cyganski D., Orr J. A., Cott T. A. and Dodson R. J. (1987): Development, Implementation, Testing and Application of an Affine Transform Invariant Curvature Function, ICCV, London 1987, p. 496-500
- Dreschler L. (1981): Ermittlung markanter Punkte auf den Bildern bewegter Objekte und Berechnung einer 3D-Beschreibung auf dieser Grundlage, Dissertation Fachbereich Informatik, Universität Hamburg, 1981
- Dorst L. (1986): Discrete Straight Line Segments: Parameters, Primitives and Properties, Dissertation, Delft, 1986
- Dorst L. and Duin R. P. W. (1984): Spirograph Theory: A Framework for Calculations on Digitized Straight Lines, IEEE T-PAMI, Vol. 6, No. 5, 1984, p. 632-639
- Ebner H. (1983): Berücksichtigung der lokalen Geländeform bei der Höheninterpolation mit finiten Elementen, Bildmessung und Luftbildwesen 51, 1983, S. 3-9
- Eeg J. (1986): On the Adjustment of Observations in the Presence of Blunders, Geodetic Institut, Technical Report No. 1, Kopenhagen 1986
- Enomoto H., Yonezaki N. and Watanabe Y. (1982): Application of Structure Lines to Surface Construction and 3-Dimensional Analysis, in Fu and Kunii 1982, p. 106-137
- Fang J. and Huang T. S. (1984): Some Experiments on Estimating the 3-D Motion Parameters of a Rigid Body from Two Consecutive Frames, IEEE T-PAMI, Vol. 6, No. 5, 1984, p. 545-554
- Förstner W. (1978): Die Suche nach groben Fehlern in photogrammetrischen Lageblöcken, DGK C 240, München 1978
- Förstner W. (1979a): Das Programm TRINA zur Ausgleichung und Gütebeurteilung geodätischer Lagenetze, ZfV 104, 1979, S. 61-72
- Förstner W. (1979b): Ein Verfahren zur Schätzung von Varianz- und Kovarianzkomponenten, Allg. Vermessungsnachrichten, 11-12, 1979, p. 446-453
- Förstner W. (1979c): On Internal and External Reliability of Networks, ASP-ACSM Convention, Washington, 1979
- Förstner W. (1982): On the Geometric Precision of Digital Correlation, IAP Vol. 24-III, Helsinki 1982, p. 176-189
- Förstner W. (1983): Reliability and Discernability of Extended Gauß-Markov Models, DGK A 98, München, 1983, p. 79-103
- Förstner W. (1984): Quality Assessment of Object Location and Pointtransfer Using Digital Image Correlation Techniques, IAP Vol. 25-A3a, Rio de Janeiro, 1984, pp. 197-219
- Förstner W. (1985): Determination of the Additive Noise Variance in Observed Autoregressive Processes Using Variance Component Estimation Techniques, Statistics & Decisions, Suppl. Issue No. 2, 1985, pp. 263-274
- Förstner W. (1986a): A Feature Based Correspondence Algorithm for Image Matching, IAP, 26-3/3, Rovaniemi 1986, pp. 150-166
- Förstner W. (1986b): Prinzip und Leistungsfähigkeit der Korrelation und Zuordnung digitaler Bilder, SIPUS 11, Stuttgart 1986, S. 69-90
- Förstner W. (1986c): On Automatic Measurement of Digital Surface Models, SIPUS 11, Stuttgart 1986, S. 69-90
- Förstner W. (1987): Reliability Analysis of Parameter Estimation in Linear Models with Applications to Mensuration Problems in Computer Vision, CVGIP, 40, 1987
- Förstner W. (1988): Model Based Detection and Location of Houses as Topographic Control Points in Aerial Images, IAP 27-3, Kyoto, 1988
- Förstner W. (1989): Image Analysis Techniques for Digital Photogrammetry, SIPUS 13, 1989
- Förstner W. and Gülch E. (1987): A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features, Proc. of Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken 1987, pp. 281-305

- Fraser C. S. und Brown D. C. (1986): Industrial Photogrammetry: New Developments and Recent Applications, Photogrammetric Record, Vol. 12, No. 68, 1986, pp. 197-217
- Fritsch D. and Düsedau G. (1987): On the Use of Curvature Measures in Digital Terrain Modelling, Proc. of Workshop "Progress in Digital Terrain Modelling" of WG III/3 of the Int. Soc. of Photogr. and Rem. Sensing, Lyngby, 1987, pp. 227-238
- Fu K. S. (1980): Digital Pattern Recognition, Communications and Cybernetics 10, Springer, 1980
- Fu K. S. and Kunii T. L. (1982): Picture Engineering, Springer, Berlin, 1982
- Fu K. S., Rosenfeld A. and Wolf J. J. (1980): Recent Developments in Digital Pattern Recognition, in Fu 1980, pp. 205-229
- Fua P. and Hanson A. J. (1987a): Using Generic Geometric Models for Intelligent Shape Extraction, IUW, Los Angeles 1987, pp. 227-233
- Fua P. and Hanson A. J. (1987b): Resegmentation Using Generic Shape: Locating Cultural Objects, PRL 5, 1987, pp. 243-252
- Fua P. and Hanson A. J. (1988): Extracting Generic Shape Using Model-Driven Optimization, IUW, 1988, pp. 994-1004
- Gale W. A. (Ed.) (1986): Artificial Intelligence and Statistics, Addison-Wesley, 1986
- Gallagher N. C. (1982): Median Filters: Properties and Applications, 16th Asilomar Conf. on Circuits, Systems and Computers, 1982
- Georgeff M. P. and Wallace C. S.: A General Selection Criterion for Inductive Inference, SRI Report
- Giri N. C. (1977): Multivariate Statistical Inference, Academic Press, 1977
- Gotthardt E. (1968): Einführung in die Ausgleichsrechnung, Wichmann, Karlsruhe, 1968
- Grafarend E. W. (1971): Mittlerer Punktfehler und Vorwärtseinschneiden, ZfV 96, 1971, pp. 41-54
- Gude Th., Märgner V. und Paulus E. (1985): Eine Systematik formaler Beschreibungen von Graphiken, Informatik-Fachberichte 107, Springer, 1985, S. 128-132
- Haggrén H. (1985): Photogrammetric Prototype System for Real-Time Engineering Applications, Optics in Engineering Measurement, SPIE Proc. Vol. 599, 1985, pp. 330-335
- Hahn M., Förstner W. (1989): The Applicability of a Feature Based and a Least Squares Matching Algorithm for DEM-Acquisition, IAP, 27, B9, Kyoto, pp. III137-150
- Haralick R. M. (1979): Statistical and Structural Approach to Texture, Proc. IEEE, Vol. 67, 1979, pp. 786-804
- Haralick R. M. (1984): Digital Step Edges from Zero Crossing of Second Directional Derivatives, IEEE T-PAMI, Vol. 6, No. 1, 1986, pp. 58-68
- Haralick R. M. (1985): Computer Vision Theory: The Lack Thereof, 3rd IEEE Workshop on Computer Vision, Bellaire, 1985
- Haralick R. M., Watson L. T. and Laffey T. J. (1983): The Topographic Primal Sketch, The Int. J. of Robotics Research, 2-1, 1983, pp. 50-72
- Harwood D., Subbaro M., Hakalhti H. and Davis L. S. (1987): A New Class of Edge Preserving Smoothing Filters, PRL 6, 1987, pp. 155-162
- Hein O. (1977): Graphentheorie für Anwender, BI-Hochschultaschenbuch, Bd. 83,
- Herman M. and Kanade T. (1986): The 3D Mosaic Scene Understanding System, in Pentland 1986, pp. 322-358
- Hölzler E. und Holzwarth (1975): Pulstechnik, Springer 1975
- Kanade T. (1987): Image Understanding Research at CMU, IUW, Los Angeles 1987, pp. 32-40
- Kass M., Witkin A. and Terzopoulos D. (1987): Snakes: Active Contour Models, ICCV, London 1987, pp. 259-268
- Kass M. and Witkin A. (1987): Analysing Oriented Patterns, CVGIP 37, 1987, pp. 362-385
- Kashyap R. L. and Lapsa P. M. (1984): Synthesis and Estimation of Random Fields Using Long Correlation Models, IEEE T-PAMI, Vol. 6, No. 2, 1984, pp. 800-809
- Koch K. R. (1987): Parameterschätzung und Hypothesentests, Dümmler, 1987
- Kohonen T. (1984): Self-Organization and Associative Memory, Springer, Berlin 1984
- Kories R. R. (1985): Bildzuordnungsverfahren für die Auswertung von Bildfolgen, SIPUS 11, Stuttgart 1986, S. 157-168
- Kuan D. T., Sawchuk A. A., Strand T. C. and Chavel P. (1985): Adaptive Noise Smoothing Filter for Images with Signal Dependent Noise, IEEE T-PAMI, Vol. 7, No. 2, 1985, pp. 165-177
- Leclerc Y. G. (1988): Image Partitioning for Constructing Stable Descriptions, IUW, 1988

- Lee J. L. (1981): Speckle Analysis and Smoothing of Synthetic Aperture Radar Images, CGIP 17, 1981, pp. 24-32
- Lee J. S. (1981): Refined Filtering of Image Noise Using Local Statistics, CGIP 15, 1981, pp. 380-389
- Lee J. S. (1983): Digital Image Smoothing and the Sigma Filter, CVGIP 24, 1983, pp. 255-269
- Li D. (1985): Theorie und Untersuchung der Trennbarkeit von groben Paßpunktfehlern und systematischen Fehlern bei der photogrammetrischen Punktbestimmung, DGK C 324, München, 1987
- Lindenberger J. (1987): Consideration of Observation Errors when Modelling Digital Terrain Profiles, Proc. of Workshop "Progress in Digital Terrain Modelling" of WG III/3 of the Int. Soc. of Photogr. and Rem. Sensing, Lyngby, 1987, pp. 227-238
- Longuet-Higgins M. S. (1957): The Statistical Analysis of a Random Moving Surface, Phil. Trans. Roy. Soc. London, Ser. A 249, Feb. 1957, pp. 321-387
- McClure D. E. (1980/1): Image Models in Pattern Theory, in Rosenfeld 1980/1, pp. 259-276
- Medioni G. and Nevatia R. (1984): Matching Images Using Linear Features, IEEE T-PAMI, Vol. 6, No. 6, 1984, pp. 675-685
- Mikhail E. M. and Ackermann F. (1976): Observations and Least Squares, Dun-Donnelly, N. Y., 1976
- Nagao M. and Matsuyama T. (1979): Edge Preserving Smoothing, CGIP, Vol. 9, 1979, pp. 394-407
- Nagel H.-H. and Enkelmann W. (1986): An Investigation of Smoothness Constraints for the Estimation of Displacement Vector Fields from Image Sequences, IEEE T-PAMI, Vol. 8, No. 5, 1986, pp. 565-593
- Nazif A. M. and Levine M. D. (1984): Low Level Image Segmentation: An Expert System, IEEE T-PAMI, Vol. 6, No. 5, 1984, pp. 555-577
- Nevatia R. and Babu K. R. (1980): Linear Feature Extraction and Description, CGIP, Vol. 13, 1980 pp. 257-269
- Nielsson N: (1971): Problem Solving Methods in Artificial Intelligence, McGraw-Hill, N. Y., 1971
- Nielsson N: (1982): Principles of Artificial Intelligence, Springer, Berlin, 1982
- Niemann H. (1981): Pattern Analysis, Springer, Berlin 1981
- Niemann H., Bunke H., Hofmann I., Sagerer G., Wolf F. and Feistel H. (1985): A Knowledge Based System for Analysis of Gated Blood Pool Studies, IEEE T-PAMI, Vol. 7, No. 3, 1985, pp. 246- 259
- O'Neill B. (1966): Elementary Differential Geometry, Academic Press, 1966
- Ohta Y., Takano K und Ikeda K. (1987): A Highspeed Matching System based on Dynamic Programming, ICCV, London 1987, pp. 335-342
- Oommen B. J. (1987): Recognition of Noisy Subsequences Using Constrained Edit Distances, IEEE T-PAMI, Vol. 9, No. 5, 1987, pp. 676-685
- Paderes F. C., Mikhail E. M. and Förstner W. (1984): Rectification of Single and Multiple Frames of Satellite Scanner Imagery Using Points and Edges as Control, NASA-Symp. on Mathematical Pattern Recognition and Image Analysis, Houston 1984
- Pagano M. (1974): Estimation of Models of Autoregressive Signal Plus White Noise, The Annals of Statistics, 1974, Vol. 2. No. 1, pp. 99-108
- Papoulis A. (1965): Probability, Random Variables and Stochastic Processes, McGraw-Hill, 1965
- Peet P. G. and Sahota T. S. (1985): Surface Curvature as a Measure for Image Texture, IEEE T-PAMI, Vol. 7, No. 6, 1985, pp. 734-738
- Pentland A. P. (1984): Fractal-Based Description of Natural Scenes, IEEE T-PAMI, Vol. 6, No. 2, 1984, pp. 661-674
- Pentland A. P. (Ed.) (1986): From Pixels to Predicates, Ablex Publishing Corpp., N. J., 1986
- Pentland A. P. (1987): Recognition by Parts, ICCV, London 1987, pp. 612-620
- Ponce J., Chelberg D. and Wallace M. (1987): Analytical Properties of Generalized Cylinders and their Projections, IUW, Los Angeles 1987, pp. 340-350
- Pregibon D. (1986): A DIY Guide to Statistical Strategy, in Gale 1986, pp. 389-400
- Rissanen I. (1983): A Universal Prior for Integers and Estimation by Minimum Description Length, The Annals of Statistics 2, 1983, pp. 416-431
- Rissanen I. (1987): Minimum Description Length-Principle, In Encyclopedia of Statistical Sciences, 5, 1987, pp. 523-527
- Rosenfeld A. (Ed.) (1980/1): Image Modeling, Academic Press, 1980/1
- Sarjakoski T. (1986): Knowledge based Blunder Treatment in Bundle Block Adjustment, IAP 26-3/3, Rovaniemi, 1986

- Schachter B. (1980/1): Long Crested Wave Models, in Rosenfeld 1980/1, pp. 327-342
- Schewe H. und Förstner W. (1986): The Program PALM for the Automatic Line and Surface Measurement Using Image Matching Techniques, IAP 26-3/2, Rovaniemi, 1986, pp. 608-622
- Schulte S. (1987): Modellierung von Beobachtungsreihen durch ein erweitertes Autoregressives Modell, DGK C 327, München, 1987
- Shannon C. E. and Weaver W. (1949): The Mathematical Theory of Communication, The University of Illinois Press, Urbana 1949 (deutsche Ausgabe: Oldenbourg, München 1976)
- Shapiro L. G. and Haralick R. M. (1985): A Metric for Comparing Relational Descriptions, IEEE T-PAMI, Vol. 7, No. 1, 1985, pp. 90-94
- Smith R., Self M. and Cheeseman P. (1987a): Estimating Uncertain Spatial Relationships in Robotics, Uncertainty in Artificial Intelligence, Vol. 2, NorthHolland, 1987
- Smith R., Self M. and Cheeseman P. (1987b): A Stochastic Map for Uncertain Spatial Relationships, to appear in Proc. of the 4th Int. Symp. on Robotics Research, MIT Press
- Smith G. B. and Wolf H. C. (1984): Image-to-Image Correspondence: Linear-Structure Matching, NASA-Symp. on Mathematical Pattern Recognition and Image Analysis, Houston 1984
- Straub B. (1986): Lernendes Verfahren zur Segmentierung industrieller Szenen, Informatik-Fachberichte 125, Springer, 1986, S. 24-28
- Szeliski R. (1987): Regularization Using Fractal Priors Proc. AAAI 6th National Conf. on Artificial Intelligence, pp. 749-754
- Terzopoulos D. (1986a): Image Analysis Using Multigrid Relaxation Methods, IEEE T-PAMI, Vol. 8, No. 2, 1986, pp. 129-139
- Terzopoulos D. (1986b): Regularization of Inverse Visual Problems Involving Discontinuities, IEEE T-PAMI, Vol. 8, No. 2, 1986, pp. 129-139
- Terzopoulos D. (1988): The Computation of Visible-Surface Representations, IEEE T-PAMI, Vol. 10, to appear
- Torre V. and Poggio T. A. (1986): On Edge Detection, IEEE T-PAMI, Vol. 8, No. 2, 1986, pp. 147-163
- Tou J. T. and Gonzalez R. C. (1974): Pattern Recognition Principles, Addison-Wesley, 1974
- Tou J. T. (1980/1): Pictorial Feature Extraction and Recognition via Image Modeling, in Rosenfeld 1980/1, 391-422
- Triendl E. and Kriegman D. J. (1987): Vision and Visual Exploration for the Stanford Mobile Robot, IUW, Los Angeles 1987, pp. 407-416
- Tsotsos, J. K. (1987): A 'Complexity Level' Analysis of Vision, ICCV, London 1987, pp. 346-355
- Turney J. L., Mudge T. N., and Volz R. A. (1985): Recognizing Partially Occluded Parts, IEEE T-PAMI, Vol. 7, No. 4, 1985, pp. 410-421
- Vosselman G. (1989): Symbolic Image Descriptions for Relational Matching, in Linkwitz K./Hangleiter U.: High Precision Navigation, Springer 1989, pp. 378-394
- Voorhees H. and Poggio T. (1987): Detecting Textons and Texture Boundaries in Natural Images, ICCV, London 1987, pp. 250-258
- Voorhees H. and Poggio T. (1987): Detecting Blobs as Textons in Natural Images, IUW, Los Angeles 1987, pp. 892-899
- Wallace C. S. and Georgeff M. P.: A general Objective for Inductive Inference, Comp. Sci. Techn. Report, Monash University, Australia
- Wang D. C. C., Vagnucci A. H. and Li C. C. (1981): Gradient Inverse Weighted Smoothing Scheme and the Evaluation of its Performance, CGIP 15, 1981, 167-181
- v. Weizsäcker C. F. (1974): Die Einheit der Natur, DTV, Wissenschaftliche Reihe, 1974
- v. Weizsäcker C. F. (1985): Aufbau der Physik, Hanser, 1985
- Wild E. (1983): Die Prädiktion mit Gewichtsfunktionen und deren Anwendung zur Beschreibung von Geländeflächen bei topographischen Geländeaufnahmen, DGK C 277, München, 1983
- Witkin A., Terzopoulos D. and Kass M. (1987): Signal Matching through Scale Space, Int. Journal for Computer Vision, 1987, pp. 133-144
- Zimmer H.-G. (1986): Deklarative Merkmalsbestimmung in Pyramidenstrukturen, Informatik-Fachberichte, Bd. 127, Springer, 1986, pp. 84-88
- Zimmermann G. und Kories R. (1984): Eine Familie von Bildmerkmalen für die Bewegungsbestimmung in Bildfolgen, Informatik-Fachberichte, Bd. 87, Springer, 1984, pp. 147-153

Anhang 1

Beweis von Gl. 215-15 und 215-20

Seien $d(x;m)$ und $D(x;m)$ die Dichte- und die Verteilungsfunktion von \underline{s} mit Parameter $m = E(\underline{s})$. Die Schätzung von $\alpha = D(x;m)$ für gegebenes x und vorliegender Stichprobe läßt sich auf ein Bernoulli-Experiment zurückführen.

Die Zahl der Erfolge, d. i. die Zahl \hat{n}_x der Werte der Stichprobe, die kleiner als x sind, ist Bernoulli-verteilt gemäß

$$\hat{n}_x \sim B(n, \alpha)$$

und hat daher die Varianz

$$V(\hat{n}_x) = n \cdot \alpha \cdot (1 - \alpha)$$

Wegen

$$\hat{\alpha} = \hat{n}_x / n$$

hat daher die geschätzte Wahrscheinlichkeit $\hat{\alpha}$ die Varianz

$$V(\hat{\alpha}) = \alpha \cdot (1 - \alpha) / n.$$

Nun ist die Schätzung von α bei gegebenem x und die Schätzung von x bei gegebenem α wegen $\alpha = D(x;m)$ und $D'(x;m) = d(x;m)$ durch

$$\sigma_\alpha = d(x;m) \sigma_x$$

verknüpft. Daher erhalten wir für die Varianz der Schätzung von m aus Gl. 215-12

$$\sigma_m^2 = \frac{m_0^2}{x_0^2} \frac{1 - \alpha}{d^2(x)} \frac{\alpha (1 - \alpha)}{n} \quad \text{q. e. d.}$$

Im Fall der X_2^2 -Verteilung können wir für kleine α und so auch kleine x folgende Näherungen verwenden:

$$d(x;m) = 1/m$$

$$D(x;m) = x/m$$

Damit ergibt sich mit $m_0 = 1$ und $\alpha = \alpha_0 = x_0$ und daher $V(\hat{\alpha}) \approx \alpha / n$

$$\sigma_m^2 = \frac{1}{\alpha^2} \frac{1}{1/m^2} \frac{\alpha}{n} = \frac{m^2}{n_x}$$

Anhang 2

Ableitung der Verteilung von s nach Gl. 215-16

Das Betragsquadrat s des Gradienten ergibt sich in signalfreiem Bild zu

$$s = \left[\frac{1}{2} (n_{11} - n_{01} + n_{10} - n_{00}) \right]^2 + \left[\frac{1}{2} (n_{11} - n_{10} + n_{01} - n_{00}) \right]^2$$

worin n_{ij} die Werte des Rauschens in einem 2x2 Fenster darstellen.

Mit $n_u = n_{11} - n_{00}$ und $n_v = n_{10} - n_{01}$ erhalten wir

$$\begin{aligned} s &= \left[\frac{1}{2} (n_u + n_v) \right]^2 + \left[\frac{1}{2} (n_u - n_v) \right]^2 \\ &= \frac{1}{2} (n_u^2 + n_v^2) \end{aligned}$$

Nun sind n_u und n_v unabhängig und normalverteilt mit Varianz $2 \sigma_n^2$.

Daher ist

$$E(s) = 2 \sigma_n^2$$

und $s / (2 \sigma_n^2)$ ist als Summe der Quadrate zweier normiert normal verteilter Zufallsvariablen χ^2 -verteilt. (q. e. d.)

Das Ergebnis läßt sich auf Betragsquadrate von Gradienten übertragen, die aus größeren Fenstern abgeleitet sind, also etwa auf Gradienten, die aus einem gefilterten Bild berechnet werden. Der Grund ist, daß ein normalverteiltes Bild durch ein lineares Filter in ein normalverteiltes Bild transformiert wird und Gradienten wegen der Orthogonalität der Filterkerne unabhängig sind.

Der Betrag des Gradienten ist daher immer Rayleigh-verteilt. Dies nutzen VORRHEES und POGGIO (1987) für ihr Verfahren.

Anhang 3

Figuren zu Teil 2