

Untersuchungen zur Detektion von Kanten in digitalen Farbbildern

Christian Drewniok

FBI-HH-B-138/88

9. Oktober 1988

Universität Hamburg
Fachbereich Informatik
Bodenstedtstraße 16
D-2000 Hamburg 50

Zusammenfassung

Bis heute wurden zahlreiche Verfahren zur Detektion von Kanten in Intensitätsbildern entwickelt. Hingegen finden sich nur sehr wenige Beiträge in der Fachliteratur, die die Möglichkeiten zur Kantendetektion in Farbbildern untersuchen. Die meisten dieser Beiträge stammen aus den siebziger Jahren und können dem heutigen Stand der Kunst nicht entsprechen.

In den Veröffentlichungen der letzten zwei Jahre finden sich Hinweise auf einen formal fundierten Farbansatz, der die Aspekte moderner Detektionsansätze berücksichtigt [Di Zenzo 86, Novak + Shafer 87]. Nach einer Einführung in die Thematik der Kantendetektion und einer Vorstellung der bisher bekannten Farbverfahren werden wir diesen Lösungsansatz ausführlich darstellen und auf dieser Basis verschiedene Detektionsverfahren für Farbbilder entwickeln. Die Verfahren werden sowohl theoretisch als auch experimentell analysiert und verglichen. Die Ergebnisse zeigen deutlich die Überlegenheit des vorgestellten Farbansatzes gegenüber vergleichbaren Intensitätsverfahren.

Abstract

Until recently numerous edge detection techniques for intensity images have been developed. However, there are few publications on color edge detection. Most were published in the early 1970s and thus, do not fulfill the current state-of-the-art requirements. Recently, *Di Zenzo 86* and *Novak + Shafer 87* presented a theoretical solution to the color edge detection problem which takes into account aspects of modern techniques. After an introduction to edge detection and a survey on color techniques we discuss this new solution in detail. Based on this, we develop a variety of edge detection schemes for color images. All schemes are compared theoretically and by experiment. Our results show clearly that color-based techniques are superior to intensity-based approaches.

Inhaltsverzeichnis

1	Einleitung	1
2	Diskontinuitäten in der Szene und im Bild	4
2.1	Kanten als physikalisches Phänomen	4
2.2	Kanten als Bildstrukturen	6
3	Die Deutung von Bildstrukturen	12
3.1	Das dichromatische Reflexionsmodell	14
3.2	Änderungen in der spektralen Intensitätsverteilung	17
4	Verfahren zur Detektion von Grauwertkanten	19
4.1	Hochpaßfilterung	20
4.2	Schablonenvergleich	20
4.3	Anpassung eines parametrischen Modells	21
4.4	Einfache Differenzoperatoren	23
5	Ableitungsorientierte Ansätze zur Kantendetektion	25
5.1	Ableitungsorientierte Detektion von Kanten	25
5.2	Verwendung verschiedener Operatorgrößen	32
5.3	Differenzieren als "ill-posed problem"	37
5.4	Herleitung einer Operatorfunktion als Optimierungsproblem	38
5.5	Faltungseigenschaften der Gaußfunktion	44
6	Verfahren zur Kantendetektion in Farbbildern	46
6.1	Nevatia 77	47
6.2	Robinson 76	49
6.3	Sethi 85	51
6.4	Shiozaki 86	52
6.5	Yakimovsky 76, Schröder 87	53
7	Ein ableitungsorientierter Farbansatz	56
7.1	Die formale Lösung	56
7.2	Veranschaulichung	59

7.3	Verallgemeinerung des Ansatzes	62
7.4	Wahl anderer Normen	64
7.5	Verschiedene Integrationsebenen	65
7.6	Zusammenfassung	68
8	Implementation und Experimente	71
8.1	Implementation eines Experimentalsystems zur Kantendetektion	71
8.1.1	Richtungsableitungen und Detektion lokaler Extrema	72
8.1.2	Repräsentation und Selektion von Kantenzügen	74
8.2	Experimente	78
8.2.1	Vergleich der verschiedenen Detektionsansätze	79
8.2.2	Vergleich von Farbvektoransatz und Intensitätsansatz	80
8.2.3	Bewertung der Ergebnisse des Farbvektoransatzes	86
8.2.4	Zusammenfassung	100
9	Zusammenfassung und Ausblick	104
A	Originalbilder	106
	Literaturverzeichnis	113

Verzeichnis der Abbildungen

2.1	Kanten verschiedenen physikalischen Ursprungs	5
2.2	Profil A: Rauschen	8
2.3	Profil B: Verdeckungskante	9
2.4	Profil C: Tangentiale Verdeckungskante	9
2.5	Profil D: Konvexe Orientierungskante	10
2.6	Profil E: Reflektanzkante	10
2.7	Profil F: Schlagschattenkante	11
2.8	Profil G: Selbstverschattung	11
3.1	Oberflächenreflexion (spiegelnd, diffus, gemischt)	14
3.2	Oberflächen- und Körperreflexion	15
3.3	Beleuchtungsgeometrie	16
3.4	Oberflächen- und Körperreflexion im Farbraum	17
4.1	Masken des Kirsch-Operators	21
4.2	Masken des Prewitt- und des Sobel-Operators	23
5.1	Lokalisation einer Stufenkante	28
5.2	Detektion von Extrema durch lokalen Vergleich	31
5.3	Difference-of-boxes-Ableitungsoperator	33
5.4	Gaußfunktion mit erster und zweiter Ableitung	35
5.5	Unterschied zwischen den Ableitungsoperatoren	43
7.1	Projektion der Einzelgradienten auf eine Gerade	60
7.2	Ergebnisse des Farbvektoransatzes	61
7.3	Veranschaulichung der Maximumsnorm	66
7.4	Möglicher Fehler bei Auswahl der maximalen Komponenten	69
8.1	Vergleich verschiedener Nachbarschaftsdefinitionen	76
8.2	Verwendete Selektionsparameter	77
8.3	Profile für die im Apfel-Bild gekennzeichnete Bildzeile	91
8.4	Mittelwerte der Bildbereiche für das Bild einer Ecke	96
8.5	Mittelwerte der Bildbereiche für das Bild eines Kreuzungspunktes	97

Verzeichnis der Bilder

1	Ausgewählte Bildbereiche für die Intensitätsprofile A bis G	8
2	Ergebnisse aller Ansätze im Vergleich	81
3	Intensitätsansatz und Farbvektoransatz im Vergleich, $\sigma = 1 \dots 3$	83
4	Intensitätsansatz und Farbvektoransatz im Vergleich, $\sigma = 4 \dots 6$	84
5	Differenz von Farbänderung und Intensitätsänderung für das Apfel-Bild	85
6	Ergebnis des Intensitätsansatzes für das Block-Bild ($\sigma = 2$)	87
7	Ergebnis des Farbvektoransatzes für das Block-Bild ($\sigma = 2$)	87
8	Ergebnis des Intensitätsansatzes für das Auto-Bild ($\sigma = 1$)	88
9	Ergebnis des Farbvektoransatzes für das Auto-Bild ($\sigma = 1$)	88
10	Differenz von Farbänderung und Intensitätsänderung für das Block-Bild	89
11	Differenz von Farbänderung und Intensitätsänderung für das Auto-Bild	89
12	Ausschnitt des Apfel-Bildes mit markierter Bildzeile (Farbvektoransatz)	90
13	Ausschnitt des Apfel-Bildes mit markierter Bildzeile (Intensitätsansatz)	90
14	Ergebnis des Farbvektoransatzes, dem Apfel-Bild überlagert	93
15	Ergebnis des Farbvektoransatzes für das Apfel-Bild, nach der Selektion .	93
16	Ergebnis des Farbvektoransatzes, dem Block-Bild überlagert	94
17	Ergebnis des Farbvektoransatzes für das Block-Bild, nach der Selektion	94
18	Ergebnis des Farbvektoransatzes, dem Auto-Bild überlagert	95
19	Ergebnis des Farbvektoransatzes für das Auto-Bild, nach der Selektion .	95
20	Änderungsbeträge für eine verrauschte Ecke	98
21	Änderungsrichtungen für eine verrauschte Ecke	98
22	Änderungsbeträge für einen verrauschten Kreuzungspunkt	99
23	Änderungsrichtungen für einen verrauschten Kreuzungspunkt	99
24	Ergebnis des Farbvektoransatzes für das Block-Bild (Stellung 0°)	101
25	Ergebnis des Farbvektoransatzes für das Block-Bild (Stellung 10°)	101
26	Ergebnis des Farbvektoransatzes für das Block-Bild (Stellung 20°)	102
27	Ergebnis des Farbvektoransatzes für das Block-Bild (Stellung 30°)	102
28	Apfelbild, Intensität	107
29	Apfelbild, Rotauszug	107
30	Apfelbild, Grünauszug	108
31	Apfelbild, Blauauszug	108
32	Blockbild, Intensität	109
33	Blockbild, Rotauszug	109

34	Blockbild, Grünauszug	110
35	Blockbild, Blauauszug	110
36	Autobild, Intensität	111
37	Autobild, Rotauszug	111
38	Autobild, Grünauszug	112
39	Autobild, Blauauszug	112

Kapitel 1

Einleitung

You digitize what's out there; the objects are delineated by their borders, which are the places where things change suddenly, so you look for the places of rapid change, figure out what the objects are, and voilà! all done. ... Unfortunately, life is not so easy.

[Blicher 84, S. 3]

Das Ziel eines *bildverstehenden Systems* ist es, durch die Verarbeitung von Bildern die Bedeutung einer abgebildeten Szene sowie der Veränderungen innerhalb der Szene zu erkennen. Grundlage für die Deutung einer Szene bilden die *geometrische Szenenrekonstruktion* und die *Objekterkennung* [Neumann 82]. Die geometrische Szenenrekonstruktion stellt eine Beschreibung der räumlichen Struktur, Anordnung, Beschaffenheit, Bewegung und der zeitlichen Veränderungen der Objekte in der Szene dar. Die Objekterkennung umfaßt die Identifikation bzw. die Zuordnung der so beschriebenen Objekte zu einer Objektklasse durch Vergleich mit entsprechenden Objektmodellen.

Bei der geometrischen Szenenrekonstruktion stellt sich das Problem, daß die explizite Information über die räumliche Tiefe bei der Aufnahme eines Bildes durch eine Kamera verloren geht. Das Wissen über die Entfernungen der Objekte relativ zueinander ist jedoch notwendig zur Ermittlung des Szenenaufbaus und der Objektbewegungen. Abhilfe kann die Verwendung von Stereobildpaaren und zeitlichen Bildfolgen schaffen: Aus den mehrfachen Ansichten lassen sich bei Kenntnis der Kameraparameter (Brennweite, Ort des optischen Zentrums, Richtung der optischen Achse usw.) Form, Lage und Bewegung von Objekten im Raum bestimmen. Voraussetzung dafür ist, daß die in den einzelnen Ansichten sichtbaren Strukturen, die Abbilder derselben physikalischen Struktur in der Szene sind, einander zugeordnet werden können. Dieses Zuordnungsproblem wird als *Korrespondenzproblem* bezeichnet. Die Entwicklung eines wissensbasierten Ansatzes zur Lösung des Korrespondenzproblems in Stereofarbbildfolgen bildet den Schwerpunkt des Projektes SISSY¹, in dessen Rahmen der vorliegende Bericht entstanden ist [Dreschler-Fischer 87].

Verfahren zur Bilddeutung setzen häufig Beschreibungen der Eingangsbilder durch

¹Stereo Image Sequences SYstem

elementare Bildstrukturen, sogenannte *Merkmale*, voraus, die zuvor mit Hilfe geeigneter Verfahren aus den Bildern extrahiert werden müssen. Bildbeschreibungen durch Merkmale bilden eine erste Abstraktionsebene, auf der von den Intensitäts- oder Farbwerten eines Bildes übergegangen wird zu einfachen lokalen Bildeigenschaften, die die im Bild enthaltene Information komprimiert wiedergeben. Die Zuordnung kann dann zwischen den in den einzelnen Bildern detektierten Merkmalen stattfinden. Typische Merkmale zur Bildbeschreibung auf dieser Ebene sind:

Homogene Flächen: Bildbereiche, die in sich eine gewisse Homogenität aufweisen und sich gleichzeitig von allen benachbarten Bildbereichen unterscheiden.

Markante Punkte: Bezüglich einer bestimmten Bildeigenschaft auffällige Punkte im Bild. Dies können isolierte Punkte, Eckpunkte von Flächen oder starke Krümmungen in linienhaften Strukturen sein.

Kanten: Orte im Bild, die starke Änderungen in der Intensität, Farbe, Textur o.ä. aufweisen.

Die genannten Merkmale sind nicht voneinander unabhängig. So würde man erwarten, daß Bereichsgrenzen mit Kanten zusammenfallen oder daß an Schnittpunkten von Kanten markante Punkte detektiert werden. Trotzdem führt die unterschiedliche Spezialisierung der Ansätze zu teilweise voneinander abweichenden Ergebnissen. Die Redundanz der Ergebnisse kann die Lösung des Korrespondenzproblems erleichtern.

Im Rahmen dieses Berichtes werden wir uns mit Aspekten der Detektion von Kanten auseinandersetzen. Bei der Korrespondenzanalyse können Kanten eine wichtige Rolle spielen: Durch die Überprüfung der Konsistenz der Zuordnungen entlang einer Kante lassen sich Aussagen über die Korrektheit von Zuordnungen gewinnen. Darüber hinaus enthalten Kanten weitere wichtige Informationen über die Objekte einer Szene, die sich unabhängig von der Korrespondenzanalyse zur geometrischen Szenenrekonstruktion nutzen lassen, wie wir in den Kapiteln 2 und 3 sehen werden.

Einen Schwerpunkt der Untersuchungen im Sissy-Projekt bildet die Frage, welche Rolle die Verwendung von Farbinformation bei der geometrischen Szenenrekonstruktion, insbesondere bei der Korrespondenzanalyse, spielen kann. Wir wollen deshalb untersuchen, wie sich die Farbinformation eines Bildes bei der Detektion von Kanten einsetzen läßt. Dazu werden wir zunächst darstellen, welche Phänomene in der Szene wir als Kante bezeichnen und welche Art von Strukturen in einem Bild durch Kanten in der Szene hervorgerufen werden. Anschließend werden wir einen Überblick über klassische Ansätze zur Detektion von Kanten in Intensitätsbildern geben und den Stand der Kunst darstellen. In Kapitel 6 werden wir uns dann Verfahren zur Detektion von Kanten in Farbbildern zuwenden. Den Schwerpunkt des Berichtes bildet Kapitel 7. Hier werden wir untersuchen, ob es möglich ist, ein Detektionsverfahren für Farbbilder zu entwickeln, das die Aspekte aktueller Intensitätsverfahren berücksichtigt. Aus der theoretischen Analyse werden wir Aussagen über verschieden Detektionsansätze ableiten, deren Gültigkeit

schließlich durch Experimente überprüft werden soll. Sowohl bei der Auswahl der Experimente als auch bei der Bewertung der Ergebnisse werden wir Kriterien zugrundelegen, die die besonderen Anforderungen innerhalb des SISSY-Projektes berücksichtigen.

Kapitel 2

Diskontinuitäten in der Szene und im Bild

The ultimate goal of edge detection is the characterization of intensity changes in the image in terms of the physical processes that originated them.

[Torre + Poggio 86, S. 147]

Obwohl der Begriff *Kante* sowohl für Kanten in einer Szene als auch für Kanten in einem Bild verwendet wird, müssen die beiden Bedeutungen dieses Begriffes streng unterschieden werden: Innerhalb einer Szene stellt eine Kante zunächst eine räumlich eingegrenzte, besondere Konstellation physikalischer Eigenschaften der realen Welt dar. Von einer Kante sprechen wir dann, wenn diese Konstellation geprägt ist durch die signifikante Änderung einer Eigenschaft.

Bei der Abbildung einer Szene durch eine Kamera lassen sich physikalische Eigenschaften in der Szene, wie räumliche Tiefe, Textur, Oberflächenorientierung oder Materialbeschaffenheit, nicht direkt messen. Den einzigen Zugriff eines Bildverarbeitungssystems auf Informationen dieser Art stellt das aufgenommene Bild dar. Ein *Bild* ist ein zweidimensionales diskretes Raster von Bildpunkten, denen die am jeweiligen Ort gemessenen diskretisierten Werte für Intensität oder Farbe zugeordnet sind. Im folgenden werden wir auch häufig von einer *Bildfunktion* sprechen. Der Definitionsbereich dieser Funktion umfaßt die Bildkoordinaten; der Wertebereich stellt die Menge der möglichen Meßwerte dar. Geht es nun um Kanten in Bildern, so verstehen wir unter einer Kante eine durch ein Modell beschreibbare Struktur in der Bildfunktion.

2.1 Kanten als physikalisches Phänomen

Anhand des schematischen Abbildes einer Szene in Abbildung 2.1 wollen wir versuchen, die wichtigsten in einer Szene zu beobachtenden Kanteneffekte nach ihrem physikalischen Ursprung zu unterscheiden und zu charakterisieren.

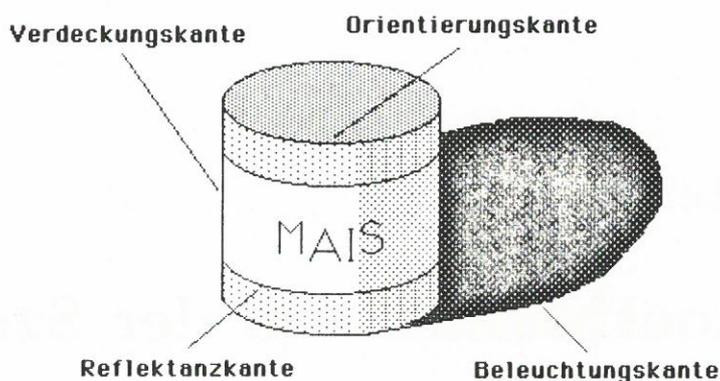


Abbildung 2.1: Kanten verschiedenen physikalischen Ursprungs¹

Treffen zwei Objektflächen unterschiedlicher Neigung aufeinander, so kann man die Linie, entlang der sich die Flächen berühren, als **Orientierungskante** bezeichnen. Eine solche Kante zeichnet sich also durch eine Diskontinuität in der Oberflächenorientierung aus. Wir setzen dabei voraus, daß sich die Oberflächenorientierung sprunghaft ändert. Verläuft die Änderung eher stetig, so wird die Lokalisation, d.h. die räumliche Eingrenzung der Kante, schwieriger. Entspricht die Änderung einem glatten Übergang, kann kaum noch von einer Kante gesprochen werden.

Trennlinien zwischen Flächen lassen sich — unabhängig von ihrer Orientierung — auch aufgrund von Diskontinuitäten der Helligkeit, Farbe, Textur o.ä. feststellen. Wir sprechen grob von **Reflektanzkanten**, wenn diese Diskontinuitäten durch Veränderungen in den Materialeigenschaften der betrachteten Flächen verursacht werden. Als Beispiel diene der Rand des Dosenetikettes.

Wird eine Diskontinuität ausschließlich durch besondere Verhältnisse der Beleuchtung verursacht, sprechen wir von einer **Beleuchtungskante**. Ein Beispiel dafür ist der Rand des von der Dose geworfenen Schattens. Auch Kanten, die durch Spiegelungen oder Glanzeffekte verursacht werden, stellen Beleuchtungskanten dar. Hier spielt jedoch zusätzlich zu den Beleuchtungsverhältnissen auch die Betrachtungsrichtung — und somit der Beobachtungsprozeß an sich — eine Rolle.

Dies gilt in ganz besonderer Weise für **Verdeckungskanten**, an denen der vom Betrachtungsstandort sichtbare Teil eines Objektes endet. Sie bilden einen abrupten Übergang zwischen einem Objekt und dem vom Objekt verdeckten Teil der Szene. Häufig stellen Verdeckungskanten gleichzeitig Orientierungskanten dar. Für die in unserer Abbildung gekennzeichnete Verdeckungskante ist dies jedoch nicht der Fall. Wir spre-

¹Abbildung 2.1 wurde uns freundlicherweise von Prof. Dr. Leonie Dreschler-Fischer zur Verfügung gestellt.

chen hier von einer *tangentialen* Verdeckungskante, da Blickrichtung und Oberflächennormale senkrecht zueinander stehen.

2.2 Kanten als Bildstrukturen

Wir haben gesehen, wie sich verschiedene Arten von Kanten in der Szene unterscheiden und als physikalische Objektmerkmale oder durch die Beleuchtungsverhältnisse und die Betrachtungsgeometrie erklären und beschreiben lassen. Welches Erscheinen erwarten wir nun für die Strukturen im Bild, die mit einem der beschriebenen physikalischen Phänomene in der Szene korrespondieren? *Horn 77* untersucht den Zusammenhang zwischen den Eigenschaften einer Szene und den gemessenen Intensitäten im Bild und liefert damit eine Grundlage zur physikalischen Deutung von Bildstrukturen. Der Zusammenhang zwischen Szene und Bild läßt sich für einen betrachteten Objektpunkt beschreiben durch eine Reflektanzfunktion, die den Einfluß der Mikrostruktur des Objektes sowie die Eigenschaften der Lichtquelle berücksichtigt. Sie hängt in starkem Maße ab von den Winkeln zwischen einfallendem Licht, Oberflächenorientierung und Betrachtungsrichtung (siehe hierzu auch Abb. 3.3).

Geht man von ebenen Objektoberflächen mit homogenen Materialeigenschaften und von einer weit entfernten Lichtquelle aus, so verändert sich die für verschiedene Objektpunkte gemessene Intensität in erster Linie in Abhängigkeit von der jeweiligen Oberflächenorientierung. Unter diesen Voraussetzungen formuliert *Horn* verschiedene Erwartungen bezüglich der Intensitäten entlang einer Geraden im Bild, die den Verlauf einer Diskontinuität senkrecht schneidet. Man kann sich bei der Analyse auf dieses sogenannte *Intensitätsprofil* beschränken, da Kanten eine gerichtete, nichtflächenhafte Struktur besitzen und häufig lokal als stückweise geradlinig angenommen werden können. Um die Aussagen *Horns* zu illustrieren, haben wir für ein aufgenommenes Bild die Intensitätsprofile entlang verschiedener Geraden im Bild ermittelt. Die ausgewählten Geraden sind in Bild 1 gekennzeichnet. Für jede Gerade sind dann der vergrößerte Bildausschnitt sowie das Intensitätsprofil dargestellt.²

Betrachtet man die Intensitäten in einem Bildbereich, der einer zusammenhängenden, ebenen Objektoberfläche entspricht, so ist für das Intensitätsprofil ein konstanter Verlauf zu erwarten. Die auftretenden leichten Variationen entstammen den bei der Bildaufnahme unvermeidbaren Rauscheinflüssen (Abb. 2.2). Senkrecht zu einer Verdeckungskante besitzt das Profil die Form einer Stufe — wir sprechen von einer *Stufenkante*. Die Höhen der beiden Niveaus resultieren aus den unterschiedlichen Orientierungen bzw. Reflektanzeigenschaften der sich verdeckenden Flächen (Abb. 2.3). Stimmen die Höhen der Niveaus überein, so läßt sich in der Bildfunktion keine signifikante Struktur mehr erkennen. Eine besondere Situation ergibt sich für *tangentiale* Verdeckungskanten: Durch die Krümmung der Objektoberfläche variiert die Intensität der in

²Bei der Beurteilung der Profile ist zu beachten, daß das verwendete Bild nicht unter idealen Laborbedingungen aufgenommen wurde. Die Annahmen bezüglich der Beleuchtung sowie bezüglich der Reflektanzeigenschaften der Objekte werden deshalb nicht ideal erfüllt sein.

Betrachtungsrichtung reflektierten Strahlung, was sich durch eine entsprechende Variation im Profil bemerkbar machen sollte. In Abbildung 2.4 wird dieser Effekt jedoch nicht deutlich. Ein ebenfalls stufenförmiges Profil ergibt sich senkrecht zu einer konvexen Oberflächendiskontinuität (Abb. 2.5). Bei einer konkaven Oberflächendiskontinuität muß unter Umständen die gegenseitige Bestrahlung der Flächen durch Reflexion berücksichtigt werden. Horn erwartet für diesen Fall eine sogenannte *Dachkante*. In dem von uns verwendeten Bild läßt sich ein solcher Effekt nicht beobachten.

Auch Reflektanzdiskontinuitäten verursachen Stufenkanten (Abb. 2.6). Während das Profil einer Schlagschattenkante ebenfalls eine evtl. durch Streuung des Lichtes unscharfe Stufe aufweist, muß man bei Intensitätsveränderungen durch Selbstverschattung schon eher von einer *Rampe* sprechen (Abb. 2.7 und 2.8).

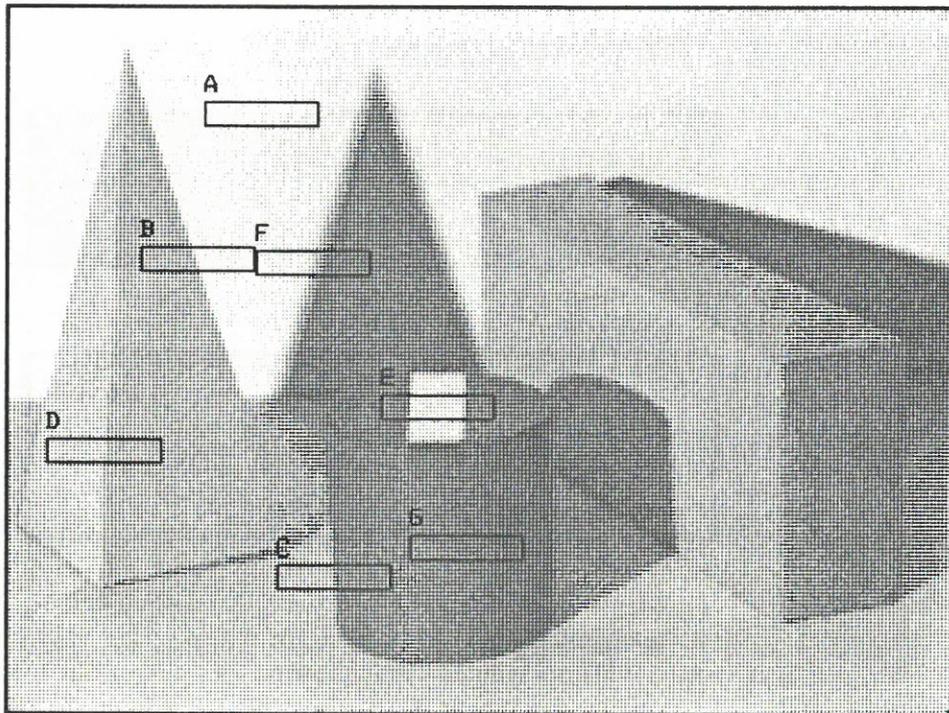


Bild 1: Ausgewählte Bildbereiche für die Intensitätsprofile A bis G

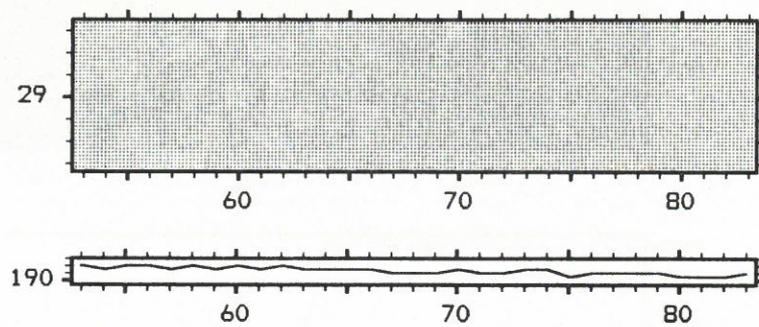


Abbildung 2.2: Profil A: Rauschen

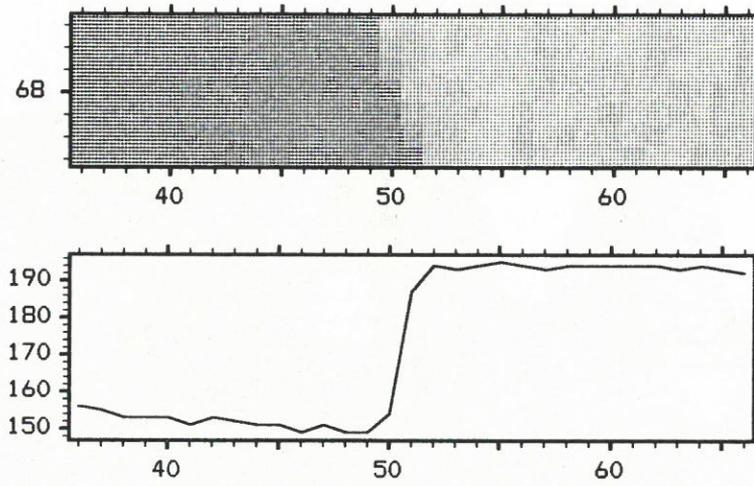


Abbildung 2.3: Profil B: Verdeckungskante

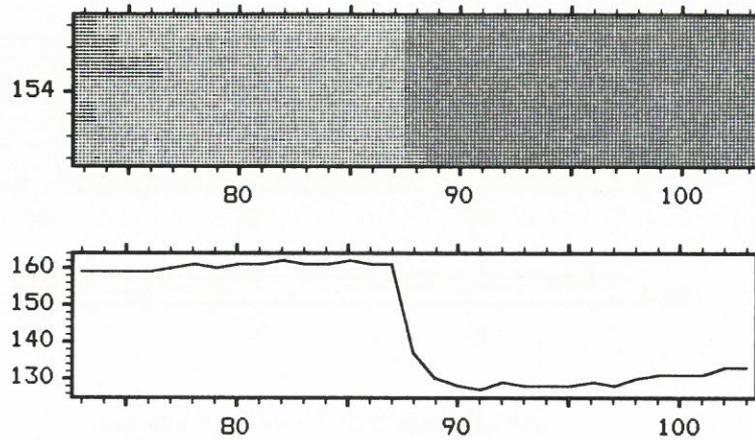


Abbildung 2.4: Profil C: Tangentiale Verdeckungskante

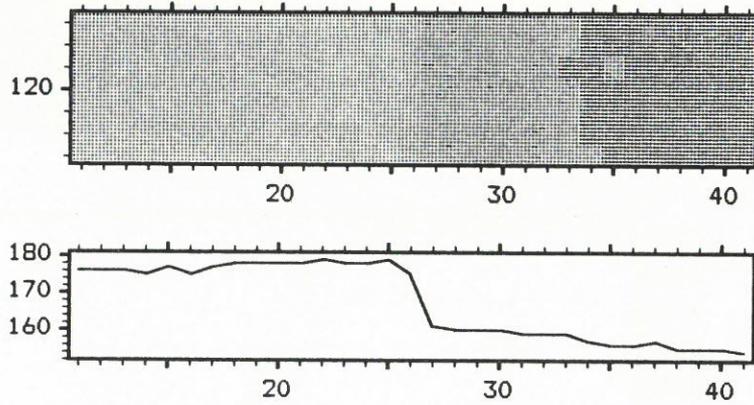


Abbildung 2.5: Profil D: Konvexe Orientierungskante

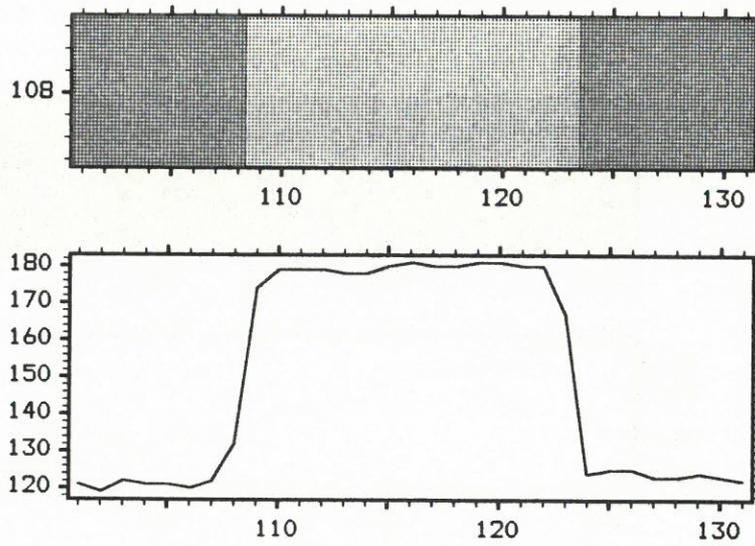


Abbildung 2.6: Profil E: Reflektanzkante

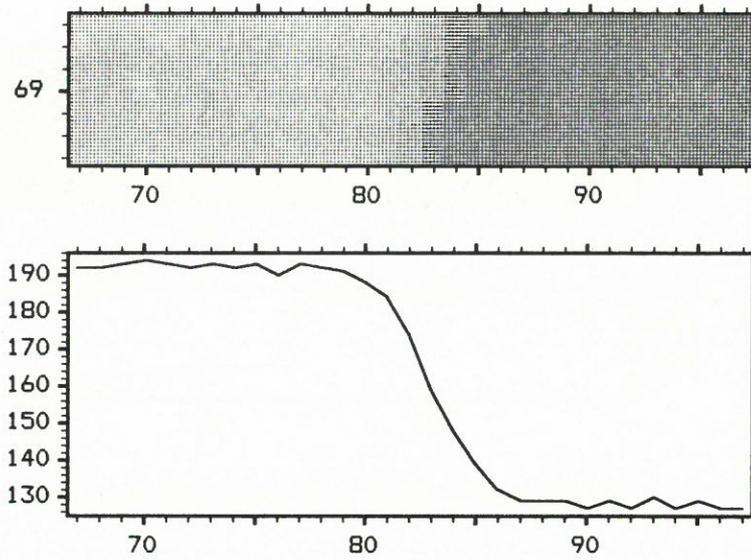


Abbildung 2.7: Profil F: Schlagschattenkante

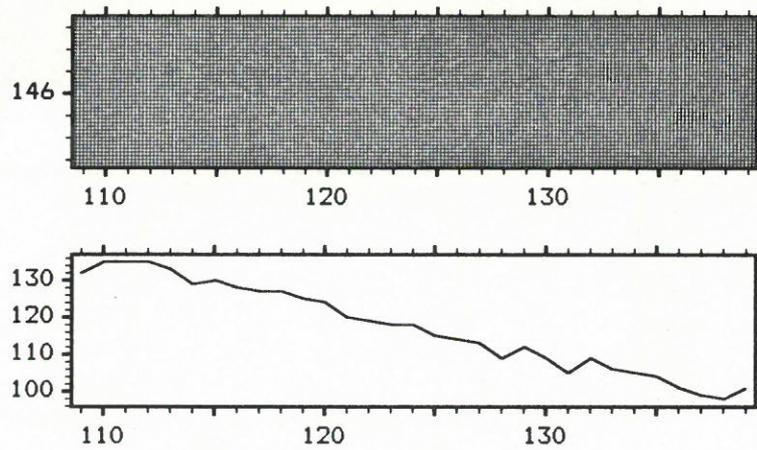


Abbildung 2.8: Profil G: Selbstverschattung

Kapitel 3

Die Deutung von Bildstrukturen

Das Ziel eines Kantendetektionsverfahrens besteht darin, Strukturen in Bildern zu detektieren, die sich einer Kante in der abgebildeten Szene zuordnen lassen. Zwischen den Kanten in der Szene und den im Bild als Kante detektierten Strukturen gibt es jedoch keine Eins-zu-Eins-Beziehung. Jedem Detektionsalgorithmus liegt ein Modell zugrunde, das die komplexen physikalischen Verhältnisse in der Szene, den Abbildungsprozeß sowie die interessierenden Bildstrukturen beschreibt. Ein Modell kann der komplexen Realität aber nicht vollständig gerecht werden. Dies führt dazu, daß vielen im Bild detektierten Strukturen keine physikalische Bedeutung zugeschrieben werden kann und viele physikalisch relevante Kanten in der Szene vom Algorithmus nicht als Bildstruktur detektiert werden.

Ist das Ziel der Verarbeitung von Bildern die geometrische Szenenrekonstruktion, so spielt neben der Entscheidung über die Relevanz detektierter Merkmale besonders die explizite Zuordnung einer physikalischen Bedeutung oder Ursache zu jeder detektierten Struktur im Bild eine wichtige Rolle. Wir bezeichnen diese Zuordnung als *Klassifikation*. Abschnitt 2.1 sollte gezeigt haben, daß Kanten in der Szene wichtige Informationen über die Objekte der Szene enthalten: Eine Orientierungskante läßt sich beispielsweise direkt zur geometrischen Rekonstruktion eines Objektes verwenden; Reflektanzkanten deuten auf eine Änderung in den Materialeigenschaften hin; Schattenkanten erlauben Aussagen über die Beleuchtung sowie über die Form und die Position von Objekten usw.

Läßt sich den im Bild detektierten Kanten eine physikalische Bedeutung zuordnen, so wird die mit ihnen verbundene Information zugänglich und für eine Interpretation des Bildes verwertbar. Man muß allerdings davon ausgehen, daß die physikalische Klassifikation von Bildstrukturen häufig erst durch Einsatz von Wissen aus höheren Deutungsebenen sowie im Zusammenspiel mit anderen Bildverarbeitungsmodulen (Punktefinder, Segmentationsverfahren, Schattierungsanalyse, Texturanalyse usw.) möglich wird.

Bei der Auswertung räumlicher und zeitlicher Korrespondenzen zur geometrischen Rekonstruktion der Szene kommt der physikalischen Klassifikation von Kanten eine besonders große Bedeutung zu. Wir haben gesehen, daß es Kanten im Bild gibt, deren Existenz und Lokalisation im Bild direkt mit dem Vorgang der Beobachtung verbunden ist. Beispiele dafür sind tangential Verdeckungskanten oder durch Spiegelungen

und Glanz hervorgerufene Kanten. Solchen Kanten läßt sich kein körpereigenes physikalisches Merkmal zuordnen. Verfahren zur merkmalsbasierten Korrespondenzanalyse setzen jedoch voraus, daß es sich bei den in Korrespondenz zu setzenden Strukturen um körperfeste Merkmale handelt. Unter dieser Voraussetzung lassen sich aus den Zuordnungen zwischen Merkmalen aus den Bildern eines Stereobildpaares oder einer zeitlichen Bildfolge Aussagen über die Position und Bewegung von Strukturen in der Szene gewinnen. Ist diese Voraussetzung jedoch nicht erfüllt, muß es zu Fehlern bei der Deutung kommen. Als Beispiel diene wieder Abbildung 2.1: Die gekennzeichnete tangentielle Verdeckungskante wird in beiden Bildern eines Stereobildpaares zu erkennen sein. Die in den beiden Bildern detektierten Kanten stellen jedoch wegen des veränderten Betrachterstandortes Abbilder unterschiedlicher Strukturen in der Szene dar! Die geometrische Auswertung der detektierten Kanten unter der Annahme, es handle sich in beiden Bildern um Abbilder desselben körperfesten Merkmals, wird also eine falsche Aussage über die Szene liefern. Durch Wissen über die physikalische Bedeutung jeder Kante im Bild ließen sich solche Fehler vermeiden.

Wir wollen in diesem Bericht das Problem der Kantenklassifikation nicht näher untersuchen. Da wir uns im folgenden jedoch insbesondere mit der Verarbeitung von Farbbildern auseinandersetzen werden, wollen wir kurz darstellen, welche Möglichkeiten es gibt, die Farbinformation eines Bildes zur Deutung von Bildstrukturen zu nutzen.

Bis vor wenigen Jahren wurde Farbinformation fast ausschließlich in Verfahren zur Segmentierung von Bildern in homogene Bereiche genutzt: Die Bildpunkte werden nach ihrer Farbe mit Hilfe von Vorwissen klassifiziert oder aufgrund ihrer Ähnlichkeit mit benachbarten Bildpunkten zu Bereichen zusammengefaßt. Die Ähnlichkeit wird dabei durch lokale Differenzmessungen oder durch Ballungsanalysen im Farbraum bestimmt. Einen guten Überblick über verschiedene Segmentierungstechniken gibt [Schröder 87]. Die Farbe wird in den genannten Verfahren als lokales Bildmerkmal betrachtet, dessen räumliche Konstanz zurückgeführt wird auf räumlich homogene Eigenschaften in der Szene. Ein tiefgehendes Modell über die physikalischen Ursachen von Farbe wird nicht zugrundegelegt.

Wie zuvor in der Grauwertbildverarbeitung zeigt die Entwicklung in der Farbbildverarbeitung seit Beginn der achtziger Jahre, daß die sorgfältige Analyse und Modellierung der zugrundeliegenden physikalischen Prozesse ein besseres Verständnis der zu beobachtenden Strukturen in Bildern erlaubt und die Entwicklung neuer Algorithmen zur Bilddeutung ermöglicht.

Einen guten Einblick in die physikalischen Grundlagen der spektralen Reflektanzeigenschaften von Körpern geben u.a. [Healey + Binford 87], [Shafer 85], [Klinker et al. 87] oder auch [Klinker et al. 88]. Shafer 85 stellt das sogenannte dichromatische Reflexionsmodell vor, das zugleich einfach und allgemein ist und hier wegen seiner großen Bedeutung kurz beschrieben werden soll. Anschließend beschreiben wir kurz einen ebenfalls häufig zitierten Ansatz von Rubin + Richards 82. Auch sie modellieren die spektralen Reflektanzeigenschaften von Objektoberflächen und nutzen dieses Modell zur Klassifikation von Kantenstrukturen.

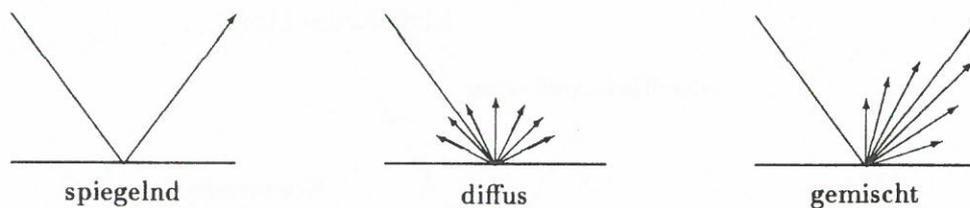


Abbildung 3.1: Oberflächenreflexion (spiegelnd, diffus, gemischt)

3.1 Das dichromatische Reflexionsmodell

Das visuell wahrnehmbare Erscheinungsbild eines Körpers resultiert aus der Wechselwirkung des einstrahlenden Lichtes mit der Materie des Körpers. Wahrgenommen wird das vom Körper reflektierte Licht. Dabei kann man zwei Arten von Reflexion unterscheiden.¹

Oberflächenreflexion

Ein Teil des einfallenden Lichtes wird direkt an der Oberfläche des Körpers reflektiert. Ist die Oberfläche glatt², dann wird das Licht so reflektiert, daß Einfallswinkel und Ausfallswinkel gleich sind und beide Winkel mit der Oberflächennormalen in einer Ebene liegen (*spiegelnde Reflexion*). Die Rauheit der Oberfläche führt zu einer mehr oder weniger diffusen Streuung des reflektierten Lichtes um die Richtung der ideal spiegelnden Reflexion (siehe Abb. 3.1). Die Stärke des reflektierten Lichtes im Vergleich zum einfallenden Licht ist abhängig von der Geometrie (Einfallswinkel, Ausfallswinkel, Oberflächenorientierung) sowie von der Wellenlänge des Lichtes. Im Bereich des sichtbaren Lichtes kann man jedoch für fast alle Materialien die Wellenlängenabhängigkeit der Oberflächenreflexion vernachlässigen, so daß das Spektrum des an der Objektoberfläche reflektierten Lichtes mit dem Spektrum des einfallenden Lichtes übereinstimmt.

Körperreflexion

Der nicht an der Oberfläche reflektierte Teil des Lichtes dringt in den Körper ein und wird hier durch die Pigmente gestreut und selektiv absorbiert. Dabei gelangt ein Teil der Strahlung zurück zur Oberfläche und tritt wieder aus dem Körper aus. Die Körperreflexion wird bestimmt durch die Streuungs- und Absorptionseigenschaften sowie durch die Verteilung der Pigmente. Im allgemeinen geht man von einer zufälligen Verteilung der Pigmente aus und erhält dann eine Gleichverteilung der Richtungen des reflektierten Lichtes. Das Spektrum des auf diese Weise reflektierten Lichtes wird geprägt durch

¹Das Modell beschränkt sich auf inhomogene Materialien, also auf solche, die aus einer transparenten Trägersubstanz und darin eingebetteten, mit dem Licht wechselwirkenden, Pigmenten besteht. Beispiele sind Kunststoff, Papier, Keramik sowie die meisten Anstrichfarben — nicht jedoch Metall. Die Selbststrahlung des Materials muß vernachlässigbar sein.

²Wir bezeichnen eine Oberfläche als glatt, wenn die Oberflächenstruktur klein ist gegen die Wellenlänge des Lichtes.

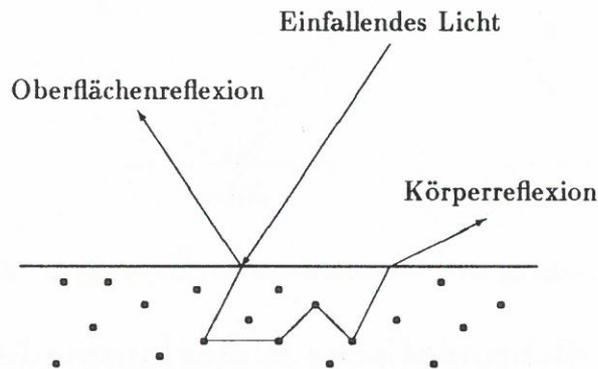


Abbildung 3.2: Oberflächen- und Körperreflexion

die selektiven Absorptionseigenschaften des Materials. Die zwei Arten der Reflexion veranschaulicht Abbildung 3.2.

Es sei hier darauf hingewiesen, daß in der englischsprachigen Literatur statt der Begriffe *surface reflection* und *body reflection* häufig die Bezeichnungen *specular reflection* und *diffuse reflection* gebraucht werden. Da die Unterscheidung specular/diffuse jedoch lediglich eine geometrische Charakterisierung darstellt, ist die Verwendung dieser Bezeichnungen irreführend, da beispielsweise auch an der Oberfläche reflektierte Strahlung diffus gestreut sein kann.

Das von Shafer vorgestellte Modell sagt aus, daß sich Oberflächen- und Körperreflexion an jedem Punkt der Oberfläche additiv überlagern, so daß sich für eine bestimmte Wellenlänge λ , Einfallswinkel i , Ausfallswinkel e und Phasenwinkel g (siehe Abb. 3.3) das reflektierte Licht in die Komponenten L_s (s steht für *surface*) und L_b (b steht für *body*) zerlegen läßt:

$$L(\lambda, i, e, g) = L_s(\lambda, i, e, g) + L_b(\lambda, i, e, g)$$

Da die spektralen Eigenschaften, wie oben erläutert, von der Geometrie sowie vom Ort auf der Oberfläche unabhängig sind, lassen sich L_s und L_b jeweils in einen geometrischen Term m und einen spektralen Term c zerlegen:

$$L(\lambda, i, e, g) = m_s(i, e, g) \cdot c_s(\lambda) + m_b(i, e, g) \cdot c_b(\lambda)$$

m_s und m_b bestimmen das Gewicht der jeweiligen Reflexionsart in Abhängigkeit von der Geometrie. c_s und c_b spiegeln die spektralen Eigenschaften des einfallenden Lichtes bzw. des Materials wieder. Sie geben an, wie stark jede Wellenlänge in der reflektierten Strahlung vertreten ist. Betrachtet man den von allen Wellenlängen gebildeten unendlichdimensionalen Wellenlängenraum, so kann man c_s und c_b als Punkte oder Vektoren in diesem Raum auffassen. Bei der Aufnahme eines Farbbildes mit Hilfe einer Kamera wird die wellenlängenabhängige Information des Bildes repräsentiert durch drei Farbwerte (rot, grün, blau). Der unendlichdimensionale Wellenlängenraum reduziert sich dann auf einen diskreten dreidimensionalen Farbraum (RGB-Raum). Shafer zeigt, daß

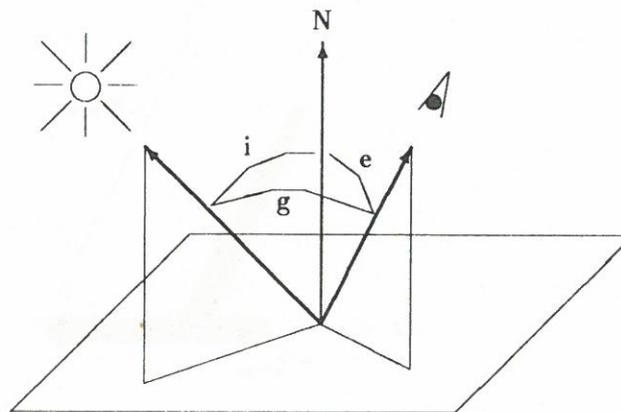


Abbildung 3.3: Beleuchtungsgeometrie. Einfallswinkel i , Ausfallswinkel e , Phasenwinkel g , Oberflächennormale N

das dichromatische Reflexionsmodell auch im RGB-Raum seine Gültigkeit behält. c_s und c_b können deshalb auch als RGB-Vektoren betrachtet werden. Die Richtung eines Farbvektors bestimmt die Farbe, während die Länge des Vektors ein Maß für die Intensität³ ist.

Betrachtet man die Farbwerte auf einer zusammenhängenden homogenen Objekt-oberfläche, so fallen laut Reflexionsmodell alle Werte in das durch die Farbvektoren c_s und c_b im Farbraum aufgespannte Parallelogramm. Abbildung 3.4 zeigt schematisch die Farbwerte für eine gleichmäßig gekrümmte Oberfläche. Aufgrund der wechselnden Oberflächenorientierung variiert die Intensität der reflektierten Strahlung. Man beobachtet deshalb eine lineare Ballung von Farbwerten entlang des c_b -Vektors. Die c_s -Komponente ist für die meisten Farbwerte gering. Nur in einem engbegrenzten Bereich auf der Oberfläche stimmt die Betrachtungsrichtung gut mit der Richtung idealer Oberflächenreflexion überein. Die in diesem Bereich zu beobachtenden Farbwerte besitzen eine starke c_s -Komponente. Bereiche, in denen die Oberflächenreflexion aufgrund der geometrischen Bedingungen die Körperreflexion überwiegt, bezeichnet man als *Glanzlichter*. Wie aus dem Modell ersichtlich ist, stellt man für Glanzlichter eine Verschiebung des gemessenen Spektrums von der körpereigenen Farbe hin zur Farbe der Beleuchtung fest.

Beobachtet man für eine Oberfläche tatsächlich zwei lineare Ballungen im Farbraum, die sich durch das Modell erklären lassen, kann man die Farbvektoren c_s und c_b bestimmen und durch Projektion der Farbwerte auf diese Vektoren zwei neue Bilder gewinnen, die zum einen die Oberflächenreflexion und zum anderen die Körperreflexion wiedergeben. Algorithmen zur Erkennung von Glanzlichtern auf dieser Basis beschreiben [Klinker et al. 88] und [Gershon et al. 87]. Jedes der beiden erhaltenen Bilder beinhaltet Informationen über die Szene, die separat einfacher zu interpretieren sind als in ihrer Überlagerung. So sollten beispielsweise *shape-from-shading*-Verfahren,

³Die Intensität entspricht dem arithmetischen Mittel der drei Farbwerte.

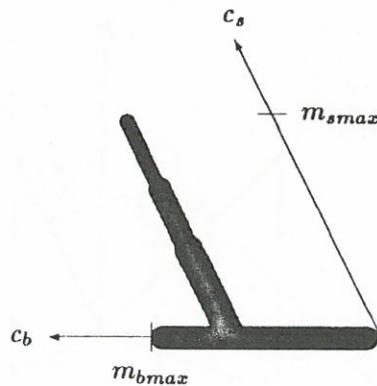


Abbildung 3.4: Oberflächen- und Körperreflexion im Farbraum

die aus der Schattierung die Form der Oberfläche erschließen, an Sicherheit gewinnen, wenn störende Effekte durch die Oberflächenreflexion ausgeschlossen sind. Andererseits ist es möglich, von den Glanzlichtern Rückschlüsse auf die Form der Objektoberflächen zu ziehen. Einen solchen Ansatz könnte man als *shape-from-gloss* bezeichnen.

Ebenso ließen sich aufgrund von Glanzlichtern detektierte Diskontinuitäten als Beleuchtungskanten klassifizieren. Wir haben bereits gesehen, daß Kanten, die nur aufgrund besonderer geometrischer Bedingungen beobachtet werden und keinem körperfesten Merkmal in der Szene zugeordnet werden können, Verfahren zur Bilddeutung fehlleiten können. Durch Erkennung von Glanzlichteffekten ließen sich derartige Fehler teilweise vermeiden.

3.2 Änderungen in der spektralen Intensitätsverteilung

Einen anderen Ansatz zur Klassifikation von Kanten mit Hilfe von Farbinformation beschreiben Rubin + Richards 82. Ausgehend von einem einfachen Reflektanzmodell, untersuchen sie die Veränderungen in der spektralen Intensitätsverteilung für benachbarte Bildbereiche, die durch eine Kante voneinander getrennt sind. Ihr Ziel ist die Erkennung solcher Kanten, die auf eine Materialänderung hinweisen. Die Strategie, die sie verfolgen, bezeichnen sie als *negative view*: Kann man ausschließen, daß es sich bei der untersuchten Kante um eine Beleuchtungskante, eine Orientierungskante oder eine Kante aufgrund einer Veränderung der Pigmentdichte handelt, so muß es sich um eine Reflektanzkante aufgrund einer Materialänderung handeln, wenn man annimmt, daß es keine weiteren Arten von Kanten in der Szene gibt.

Je nach der physikalischen Ursache der Kante sind unterschiedliche Veränderungen in der spektralen Intensitätsverteilung zwischen den beiden Bereichen zu beobachten. Rubin und Richards zeigen, unter Annahme eines Reflektanzmodells für matte Flächen, daß die beiden Verteilungskurven eine über alle Wellenlängen konstante Relation (*größer* oder *kleiner*) zueinander besitzen. Nur wenn es sich bei der Kante um eine Reflektanz-

kante handelt, läßt sich bezüglich der Veränderungen in der spektralen Verteilung keine Annahme machen, da sich die beiden aufeinandertreffenden Materialien in beliebiger Weise voneinander unterscheiden können. Beobachtet man, daß sich die zwei Verteilungskurven schneiden, so daß keine konstante Relation über alle Wellenlängen gegeben ist, kann man alle anderen Kanteneffekte ausschließen und annehmen, daß es sich um eine Reflektanzkante handelt.

Die diesem Ansatz zugrundeliegenden Annahmen sind sehr einschränkend und lassen z.B. Effekte aufgrund von Oberflächenreflexion oder indirekter Beleuchtung, die bei Schattenkanten eine Rolle spielen kann, unberücksichtigt. Auch die Unterscheidung zwischen Reflektanzkanten und Verdeckungskanten wäre nach diesem Kriterium nicht möglich. *Gershon et al. 86* versuchen in ihrem Ansatz den Einfluß indirekter Beleuchtung zu berücksichtigen, um Fehlklassifikationen von Schattenkanten zu vermeiden. Dies setzt allerdings Vorwissen über die indirekte Beleuchtung voraus.

Kapitel 4

Verfahren zur Detektion von Grauwertkanten

In den ersten Kapiteln dieses Berichtes haben wir untersucht, welche Phänomene in der Szene wir als Kante bezeichnen und welche Strukturen dadurch in Bildern hervorgerufen werden. Wir wenden uns nun Ansätzen zur Detektion von Bildmerkmalen zu, die idealerweise mit einer Kante in der Szene korrespondieren sollten. Dabei wollen wir uns zunächst auf die Detektion von Grauwertkanten, also von Diskontinuitäten in der Bildintensität, beschränken.

Kanten in der Szene können, wie wir gesehen haben, im Bild unterschiedliche Strukturen hervorrufen. Man kann nicht erwarten, daß es ein Detektionsschema gibt, das die Detektion aller denkbaren Kantenstrukturen erlaubt. Es ist aber bereits deutlich geworden, welche besondere Rolle Stufenkanten in Bildern spielen. Sie korrespondieren mit Verdeckungskanten, Reflektanzkanten, Orientierungskanten oder Schlagschattenkanten und bilden somit den häufigsten und wichtigsten Kantentyp. Den meisten der hier beschriebenen Verfahren liegt deshalb ein *Stufenkantenmodell* zugrunde. Das Modell sagt aus, daß eine Stufenkante senkrecht zur Kantenrichtung eine stufenförmige Diskontinuität der Intensität aufweist, während die Bildfunktion entlang der Kantenrichtung als konstant angenommen wird. Implizit wird damit angenommen, daß die Kante lokal geradlinig verläuft. Die Annahmen des Stufenkantenmodells sind sowohl bei der Entwicklung eines Ansatzes als auch bei Aussagen über die Güte von Ergebnissen zu berücksichtigen. Die Detektion anderer Kantenarten, z.B. dachförmiger Kanten, verlangt meist die Entwicklung anderer Verfahren.

Wir geben einen Überblick über verschiedene klassische Ansätze zur Kantendetektion, meist unter Nennung eines typischen Vertreters. Anschließend werden wir uns den ableitungsorientierten Kantendetektionsverfahren zuwenden. Diese Verfahren stellen nach unserer Ansicht den aktuellen Stand der Kunst dar und sollen deshalb einen Schwerpunkt dieses Berichtes bilden. Überblicke über Kantendetektionsverfahren finden sich in der Literatur u.a. in [Rosenfeld + Kak 82], [Ballard + Brown 82], [Davis 75] oder [Blicher 84]. Aus der Konzeption des SISSY-Projektes ergeben sich einige Anforderungen, die die Auswahl der von uns betrachtenden Verfahren grundsätzlich einschränken:

- Wir werden nur **eigenständige Verfahren** zur Kantendetektion betrachten. Da die Redundanz der Ergebnisse der verschiedenen Merkmalsfinder (Punkte, Bereiche, Kanten) im SISSY-Projekt zur Unterstützung einer sicheren Bilddeutung dienen soll, scheidet Ansätze aus, die beispielsweise Kanten als Bereichsgrenzen aus dem Ergebnis eines Segmentierungsverfahrens extrahieren.
- Die Art der zu verarbeitenden Bilder ist in keiner Weise eingeschränkt. Deshalb soll ein Verfahren auf der untersten Stufe der Verarbeitung **kein szenenspezifisches Wissen** einsetzen. Ebenso sind bei der Festlegung der Parameter des Verfahrens (z.B. Schwellwerten) Spezialisierungen für bestimmte Bildinhalte zu vermeiden.
- In Anbetracht der sehr großen Datenmenge einer Stereo-Farbbildfolge ist ein zeitlich effizientes Verfahren wünschenswert. Wir werden uns deshalb auf **parallele Verfahren**¹ beschränken, da diese sich gegebenenfalls durch spezielle, parallele Hardware unterstützen lassen. Sequentielle Ansätze wie *Heuristische Suche* oder *Dynamische Programmierung* werfen zudem Probleme durch die Startpunkt- und Reihenfolgeabhängigkeit auf.

4.1 Hochpaßfilterung

Stufenkanten stellen starke Änderungen der Intensität im Bild dar. Betrachtet man die Bildfunktion im Frequenzraum, so zeigt sich, daß solche Änderungen besonders große Anteile hoher Frequenzen besitzen. Die Verstärkung der hochfrequenten und die Dämpfung der niederfrequenten Anteile der Bildfunktion sollte demnach Kanten im Bild verstärken. Da sich aber auch Rauschen meist in den hohen Bildfrequenzen bemerkbar macht, werden mit den Kanten gleichzeitig die rauschbedingten Störungen verstärkt. Filtertechniken zur Kantenverstärkung können in der Bildaufbereitung dienlich sein; sie werden den Anforderungen einer robusten Kantendetektion jedoch nicht gerecht.

4.2 Schablonenvergleich

Lokale Operationen zur Verarbeitung von Bildern lassen sich häufig durch *Faltungen* der Bildfunktion mit speziellen *Operatormasken* realisieren. Die diskrete, zweidimensionale Faltungsoperation des Bildes I mit einer Operatormaske f der Größe $N \times M$ an der Bildposition (x_0, y_0) ist definiert durch:

$$(I * f)(x_0, y_0) = \sum_{i=-N/2}^{+N/2} \sum_{j=-M/2}^{+M/2} I(x_0 - i, y_0 - j) \cdot f(i, j)$$

¹Bei parallelen Verfahren ist die Entscheidung über ein lokales Kantenelement unabhängig von vorhergehenden Entscheidungen über andere Kantenelemente. Deshalb ist prinzipiell die parallele Ausführung des Verfahrens für alle Bildpunkte gleichzeitig möglich.

-1	0	1
-1	0	1
-1	0	1

1	1	1
0	0	0
-1	-1	-1

0	1	1
-1	0	1
-1	-1	0

1	1	0
1	0	-1
0	-1	-1

Abbildung 4.1: Masken des Kirsch-Operators

Die Werte der diskreten Bildfunktion werden im von der Maske abgedeckten Bildausschnitt mit den entsprechenden Koeffizienten der Maske gewichtet und summiert. Eine Faltungsoption stellt somit eine *lineare Operation* auf dem Bild dar.

Ein Beispiel für die Anwendung von Faltungsoptionen stellen Ansätze zur Kantendetektion durch Schablonenvergleich dar. *Schablonen* (engl. templates) stellen Idealisierungen der zu detektierenden Bildstrukturen in Form von Operatormasken dar. Die Güte der Übereinstimmung eines Bildausschnitts mit einer Schablone läßt sich durch Faltung der Bildfunktion mit der entsprechenden Operatormaske bewerten. Hohe Beträge im Faltungsergebnis deuten auf gute, niedrige Werte auf schlechte Übereinstimmung hin. Hinweise auf Kanten lassen sich anhand hoher Faltungswerte detektieren, wenn die Schablonen ideale Stufenkanten repräsentieren. Jede Schablone repräsentiert dabei Kanten einer bestimmten Richtung. Somit läßt sich auch eine Aussage über die Richtung der detektierten Kante gewinnen. Ansätze zur Kantendetektion durch Schablonenvergleich unterscheiden sich nach Größe der verwendeten Masken und Anzahl der verschiedenen durch eine Maske repräsentierten Kantenrichtungen.

Ein typisches Beispiel ist der *Kirsch-Operator* (siehe dazu [Ballard + Brown 82, S. 79f]). Kirsch definiert Masken für vier verschiedene Richtungen (siehe Abb. 4.1). Für jeden Bildpunkt wird die Bildfunktion mit den vier Masken gefaltet. Der Index der Maske, die das größte Faltungsergebnis liefert, wird als Hinweis auf die Kantenrichtung registriert. Das Faltungsergebnis selbst wird als Maß für den Kantenkontrast verwendet. Kanten werden schließlich dort markiert, wo der Kontrastwert einen bestimmten Schwellwert übersteigt.

4.3 Anpassung eines parametrischen Modells

Verfahren dieser Klasse versuchen, die Parameter eines zugrundeliegenden Kantenmodells so festzulegen, daß lokal eine möglichst gute Übereinstimmung zwischen der Bildfunktion und der durch die aktuellen Parameterwerte bestimmten Approximation erreicht wird. Die Güte der Übereinstimmung wird durch eine Fehlerfunktion bewertet und gibt Auskunft darüber, wie wahrscheinlich das Vorhanden- oder Nichtvorhandensein einer durch das Modell beschreibbaren Kante im betrachteten Bildausschnitt ist. Durch Minimierung der Fehlerfunktion erhält man die Parameterwerte, die zu der *besten* Modellanpassung führen. Anhand der ermittelten Werte lassen sich dann Aussagen über Lage, Richtung und Kontrast der vermuteten Kante machen.

Einen klassischen Ansatz dieser Art beschreibt Hueckel 71. Die Bildfunktion I in einem kreisförmigen Bildausschnitt D wird beschrieben durch eine Modellfunktion F mit den Parametern p_1 bis p_5 , die die Lage und Richtung einer idealen Stufenkante sowie die konstanten Grauwerte rechts und links der Kante festlegen. Als Fehler ε der Modellanpassung an den Bildausschnitt wird der quadratische Fehler verwendet:

$$\varepsilon(p_1, \dots, p_5) = \int_D [I(x, y) - F(x, y, p_1, \dots, p_5)]^2 dx dy$$

Um das Minimierungsproblem zu vereinfachen, werden Bild- und Modellfunktion in einem System von Basisfunktionen repräsentiert. Die Angabe einer vollständigen und orthogonalen Familie von Funktionen $\{H_i | i = 0 \dots \infty\}$ über D erlaubt die Darstellung jeder Funktion I über D als Vektor von Koeffizienten a_i , so daß

$$a_i = \int_D H_i(x, y) \cdot I(x, y) dx dy \quad (i = 0 \dots \infty)$$

Die Basisfunktionen entsprechen periodischen Funktionen, deren Raumfrequenzen monoton ansteigen.² Beschreiben wir I und F durch entsprechende Vektoren (a_0, a_1, \dots) und (f_0, f_1, \dots) , ergibt sich als Fehlerfunktion:

$$\varepsilon(p_1, \dots, p_5) = \sum_{i=0}^{\infty} [a_i - f_i]^2$$

Die Minimierung der Fehlerfunktion läßt sich effizient durchführen, wenn man sich bei der Repräsentation der Funktionen auf die ersten Koeffizienten des Spektrums beschränkt. Dies entspricht der Unterdrückung hochfrequenter Anteile in der Bildfunktion, die in erster Linie Rauschen enthalten sollten. Hueckel schlägt vor, die Repräsentation von I und F auf die ersten acht Basisfunktionen der von ihm gewählten Familie zu beschränken. Die Minimierung der nun endlichen Fehlerreihe läßt sich so auf die Berechnung der Extremwerte einer Gleichung reduzieren, die außer von den Koeffizienten $a_0 \dots a_7$ nur noch von der Kantenrichtung abhängt. Ist auf diese Weise die Kantenrichtung bestimmt, lassen sich auch die übrigen Parameter der Approximation angeben. Der Kosinus des Winkels zwischen den Vektoren (a_0, \dots, a_7) und (f_0, \dots, f_7) wird als Gütemaß der Anpassung verwendet. Der Algorithmus umfaßt demnach folgende Schritte:

1. Bestimmung der Koeffizienten $a_0 \dots a_7$ durch Faltung der Bildfunktion mit Masken, die die Basisfunktionen $H_0 \dots H_7$ diskret approximieren.
2. Bestimmen der Kantenrichtung durch Extremwertberechnung.
3. Ermittlung der übrigen Parameter sowie des Gütewertes.

Durch Festlegung globaler Schwellwerte für die Kontrast- und Gütewerte der einzelnen Bildpunkte läßt sich das Eingabebild schließlich in ein binäres Kantenbild überführen.

²Verwendet man als Basisfunktionen komplexe Exponentialfunktionen, deren Raumfrequenzen ansteigende Vielfache einer Grundfrequenz sind, so entspricht der Koeffizientenvektor dem Fourierspektrum von I .

-1	0	1	1	1	1	-1	0	1	1	2	1
-1	0	1	0	0	0	-2	0	2	0	0	0
-1	0	1	-1	-1	-1	-1	0	1	-1	-2	-1

(a) Prewitt
(b) Sobel

Abbildung 4.2: Masken des Prewitt- und des Sobel-Operators

4.4 Einfache Differenzoperatoren

Bereits früh geboren wurde die Idee, Kanten anhand großer Grauwertdifferenzen zwischen benachbarten Bildpunkten zu detektieren. Der älteste Vertreter dieser Klasse von Verfahren ist der *Roberts-Operator* (siehe auch [Ballard + Brown 82, S. 76f]). Hier werden die Grauwertdifferenzen zwischen horizontal und vertikal benachbarten Bildpunkten berechnet:³

$$\Delta_x = I(x+1, y) - I(x, y) \quad \Delta_y = I(x, y+1) - I(x, y)$$

Mit Hilfe der Differenzen Δ_x und Δ_y lassen sich Richtung (φ) und Kontrast (l) einer Kante wie folgt schätzen:

$$\varphi = \arctan\left(\frac{\Delta_y}{\Delta_x}\right) \quad \text{und} \quad l = \sqrt{\Delta_x^2 + \Delta_y^2}$$

Dieses Vorgehen stellt, wie wir in Kapitel 5 sehen werden, die einfachste Approximation des *Gradientenoperators* dar. In die Berechnung der Differenzen gehen nur sehr wenige Grauwerte ein, so daß sich bereits ein einziger durch Rauschen beeinflusster Wert sehr störend bemerkbar machen kann. Dies gilt besonders für die Schätzung der Kantenrichtung. Um den Einfluß von Rauschen zu verringern, kann man Differenzen zwischen den Mittelwerten benachbarter Bildbereiche statt zwischen einzelnen Grauwerten berechnen, wie es Prewitt und Sobel vorschlagen (siehe [Ballard + Brown 82, S. 77]). Die Mittelung der Grauwerte besitzt eine rauschunterdrückende Wirkung und verbessert so die Schätzung. In Abbildung 4.2 sehen wir die Masken des Sobel- und des Prewitt-Operators. Der Sobel-Operator berücksichtigt zusätzlich die unterschiedliche Entfernung der benachbarten Bildpunkte zum Zentrum durch stärkere Wichtung der horizontalen und vertikalen Nachbarn.

Interessant ist an dieser Stelle die Beziehung zwischen den Differenzoperatoren und dem Ansatz zur lokalen Flächenapproximation: Geht man davon aus, daß sich die Bildfunktion lokal im betrachteten Maskenfenster durch eine ebene Fläche beschreiben läßt,

³Die Berechnung der Differenzen ist auch zwischen diagonal benachbarten Bildpunkten möglich.

kann man die Funktion wie folgt approximieren:

$$I(x, y) = a \cdot x + b \cdot y + c$$

Die Flächenparameter a , b und c lassen sich durch Minimierung des quadratischen Fehlers

$$\sum_{x,y} [I(x, y) - (ax + by + c)]^2$$

bestimmen (siehe z.B. [Hille 82]). Auf einfache Weise läßt sich zeigen, daß die mit Hilfe der Prewitt-Masken berechneten Werte (bis auf einen konstanten Faktor) mit den Schätzwerten für die Parameter a und b übereinstimmen. a und b geben aber gerade die Neigung der approximierten Fläche in x - und y -Richtung an.

Durch Minimierung des quadratischen Fehlers lassen sich auch Polynome höherer Ordnung lokal an die Bildfunktion anpassen. *Beudet 78* stellt zahlreiche Masken unterschiedlicher Größe zur Berechnung verschiedener Differenzen durch Approximation von Polynomen unterschiedlicher Ordnung vor.

Kapitel 5

Ableitungsorientierte Ansätze zur Kantendetektion

Einen besonderen Schwerpunkt bei unserem Überblick über Kantendetektionsverfahren wollen wir auf die ableitungsorientierten Ansätze legen. Diese Ansätze bilden einen Großteil der bisher bekannten Verfahren. Zudem gewannen sie in den achtziger Jahren — besonders durch die hier vorgestellten Ansätze — stark an Bedeutung und bilden heute sozusagen den Stand der Kunst.

Wir werden zunächst die Verwendung von Ableitungsoperationen zur Detektion von Kanten motivieren und verschiedene Möglichkeiten zur praktischen Umsetzung ableitungsorientierter Ansätze vorstellen. Das Ergebnis dieser Betrachtungen wird ein Detektionsschema sein, das die Grundlage für die von uns durchgeführten Experimente bildet. Die nachfolgenden Abschnitte dieses Kapitels behandeln dann einige wichtige, die Realisierung von diskreten Ableitungsoperationen betreffende Aspekte.

5.1 Ableitungsorientierte Detektion von Kanten

Die Grundidee der Kantendetektion mit Hilfe von Ableitungsoperationen wurde bereits angedeutet: Hohe Werte in den Differenzen benachbarter Bildpunkte weisen auf das Vorhandensein einer Kante hin. Die Berechnung solcher Differenzen stellt eine einfache diskrete Approximation räumlicher Ableitungen der Bildfunktion dar. Die vorgestellten Operatoren von Roberts, Sobel und Prewitt dienen also der Schätzung der direktionalen Ableitungen der Bildfunktion nach x bzw. y .

Betrachtet man eine ideale Stufenkante, ließe sich folgende Frage stellen: Wie lassen sich Stufenkanten mit Hilfe direktionaler Ableitungen detektieren, wenn gerade am Ort der Stufe die Bildfunktion einen Sprung unendlicher Steigung besitzt und somit nicht differenzierbar ist? Die Antwort auf diese Frage liegt im Abbildungsprozeß: Die Aufnahme eines Bildes über ein optisches System und einen elektronischen Sensor führt zu einer Bandbeschränkung des Frequenzbereiches der Bildfunktion; somit kann die Bildfunktion keine Frequenzen aufweisen, die über einer bestimmten oberen Grenzfrequenz

liegen. Ein Sprung unendlicher Steigung ist nicht möglich.

Welche Möglichkeiten bieten sich nun an, Kanten mit Hilfe von Ableitungsoperationen zu detektieren? Denkbar wäre es, die Bildfunktion in beliebig viele verschiedene Richtungen abzuleiten, um so Kanten jeder Orientierung detektieren zu können. Dieser hohe Aufwand läßt sich durch Einsatz des *Gradientenoperators* vermeiden: Werden die Richtungsableitungen in zwei zueinander orthogonale Richtungen (z.B. x - und y -Richtung) berechnet, erhält man die Komponenten des Gradientenoperators ∇ (I sei die Bildfunktion):¹

$$\nabla I = \begin{pmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{pmatrix} = \begin{pmatrix} I_x \\ I_y \end{pmatrix}$$

Die Richtung des Gradientenvektors ∇I entspricht der Richtung der stärksten Änderung der Bildfunktion. Der Gradient sollte sich somit senkrecht zu einer Stufenkante ausrichten, wenn wir davon ausgehen, daß die Bildfunktion entlang einer Stufenkante lokal als konstant angenommen werden kann. Die Länge des Gradientenvektors ist ein Maß für die Stärke der Änderung und somit für den Kontrast einer Stufenkante. Richtung und Länge des Vektors ergeben sich aus den Komponenten wie folgt (siehe auch Abschnitt 4.4):

$$\arctan\left(\frac{I_y}{I_x}\right) \quad \text{und} \quad \sqrt{I_x^2 + I_y^2}$$

Ein hoher Gradientenbetrag kann daher als Hinweis auf eine Kante gedeutet werden. Allerdings ergeben sich hohe Werte für den Gradienten auch, wenn die Bildintensität in einer Richtung stark linear variiert.

Sind die beiden Komponenten des Gradienten bekannt, so läßt sich der Wert der Ableitung der Bildfunktion in eine Richtung \vec{n} bestimmen durch Projektion des Gradientenvektors auf die Richtung \vec{n} :

$$\frac{\partial I}{\partial \vec{n}} = \nabla I \cdot \vec{n}$$

Erforderlich sind demnach nur zwei Ableitungsoperationen in zueinander orthogonalen Richtungen, um den Wert der Ableitung in einer beliebigen Richtung zu bestimmen. An dieser Stelle müssen zwei Anmerkungen gemacht werden:

1. Die Beziehung zwischen Gradient und Richtungsableitung gilt im kontinuierlichen Fall. Wir haben es aber stets mit diskreten und verrauschten Bildfunktionen zu tun. In diesem Fall kann die Berechnung vieler verschiedener Richtungsableitungen sinnvoll sein, um eine größere Robustheit gegen Rauschen zu erzielen.
2. Die Beziehung zwischen Gradientenrichtung und Kantenrichtung gilt nur, solange im betrachteten Bildausschnitt nur eine Kante sichtbar ist. Im Bereich von Schnittpunkten von Kanten kommt es zu einer Überlagerung verschiedener Kanteneffekte,

¹Durch die tiefgesetzten Indizes x und y seien — auch im folgenden — Richtungsableitungen nach x bzw. y notiert.

so daß das Ergebnis des Gradientenoperators eine Art Mittelung darstellt. Der Gradient liefert unter diesen Umständen keine exakte Information über jede einzelne Kante mehr. Im Bereich von Schnittpunkten werden deshalb gradientenbasierte Detektionsverfahren Schwächen zeigen.

Der Weg der gradientenbasierten Kantendetektion ist vorgezeichnet: Man bestimme die Komponenten des Gradienten für jeden Bildpunkt und suche nach hohen Gradientenbeträgen. Wie läßt sich aber eine Kante exakt lokalisieren? Bei den einfachen Verfahren von Roberts, Prewitt, Sobel o.ä. wird als Kriterium eine globale Schwelle für den Gradientenbetrag festgelegt. Jeder Bildpunkt, der einen ausreichend hohen Gradientenbetrag aufweist, wird als Kantenpunkt markiert. Wenn man davon ausgeht, daß die Komponenten des Gradienten mit Hilfe von Differenzen der Mittelwerte lokaler Umgebungen des Bildpunktes bestimmt werden, ist zu erwarten, daß im Bereich einer Kante ein mehr oder weniger breites Band hoher Gradientenbeträge gefunden wird. Um dünne, linienhafte Kantenzüge zu erhalten, wird eine anschließende Verdünnungsoperation erforderlich. Solche Operationen nutzen die in der Bildfunktion enthaltene Information nicht vollständig, so daß im allgemeinen keine exakte Lokalisation der Kanten zu erwarten ist.

Eine natürlichere Definition für den Ort einer Kante ergibt sich, wenn man das (eindimensionale) Kantenprofil, sowie die entsprechenden Werte der ersten und zweiten direktionalen Ableitung senkrecht zur Kante betrachtet (siehe Abbildung 5.1). Es erscheint naheliegend, als Kantenposition den Ort stärkster Änderung der Bildfunktion in Gradientenrichtung, also senkrecht zur Kante, zu definieren. Wir erhalten dann kein breites Kantenband mehr, sondern exakt einen Kantenpunkt. Der Ort stärkster Änderung der Bildfunktion sollte mit einem Extremum in der ersten und einem Nulldurchgang in der zweiten direktionalen Ableitung in Gradientenrichtung korrespondieren. Als formale Definition für den Ort einer Kante erhalten wir:

$$\frac{\partial^2 I}{\partial \vec{n}^2} = 0$$

\vec{n} sei hier der Einheitsvektor in Gradientenrichtung:

$$\vec{n} = \frac{\nabla I}{|\nabla I|}$$

Hier spielt der mit der Ableitungsoperation verbundene Mittelungseffekt eine wichtige Rolle: Würden mit Hilfe eines Ableitungsoperators die exakten Steigungen der Bildfunktion für jeden Bildpunkt bestimmt werden, so ließen sich anhand der Ableitungswerte niedrige aber steile Flanken in der Bildfunktion, die durch das Rauschen verursacht werden, nicht von echten Stufen unterscheiden. Nur wenn in die Schätzung der Ableitung eine größere Umgebung von Bildpunkten eingeht, läßt sich in verrauschten Bildern eine sichere Detektion und Lokalisation von Kanten erzielen. Die Berechnung von Differenzen der Mittelwerte von Umgebungen setzt zu einer exakten Lokalisation einer Kante eine gewisse Symmetrie der Bildfunktion rechts und links der Kante voraus. Weist die Bildfunktion in Gradientenrichtung beispielsweise beiderseits der Kante eine leichte lineare

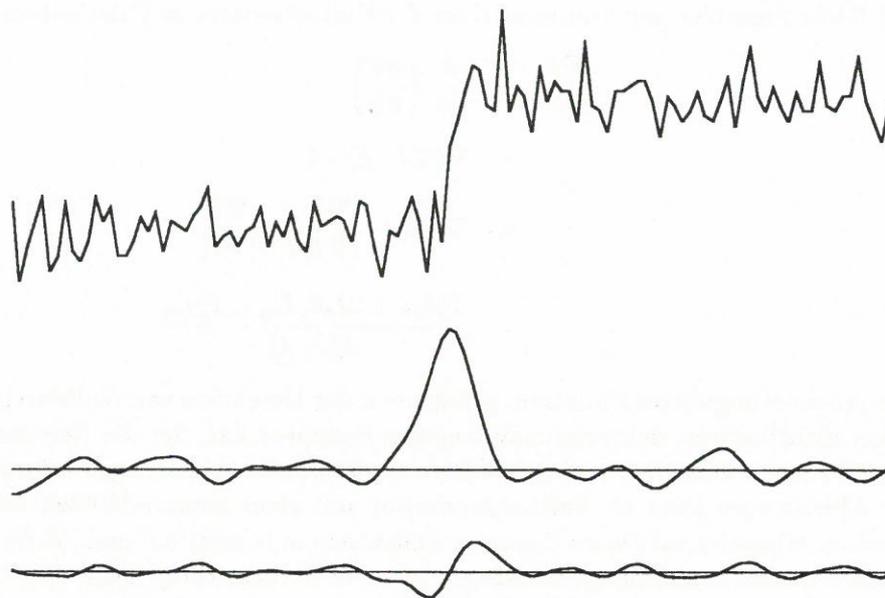


Abbildung 5.1: Lokalisation einer Stufenkante: Der Ort stärkster Änderung der Bildfunktion korrespondiert mit einem Extremum in der ersten und mit einem Nulldurchgang in der zweiten Ableitung der geglätteten Bildfunktion.

Variation unterschiedlicher Steigung auf, so kann diese Unsymmetrie des Profils durch die Mittelung zu einer Verschiebung des Nulldurchganges bzw. Extremums führen.

Auch die Definition von Kanten durch Nulldurchgänge in der zweiten Ableitung in Gradientenrichtung läßt verschiedene Möglichkeiten der praktischen Umsetzung zu. Die direkteste, wenn auch nicht eleganteste, wäre die, mit Hilfe eines Gradientenoperators die Gradientenrichtung zu bestimmen, um anschließend mit einer entsprechend spezialisierten Operatormaske das Bild in diese Richtung zweimal abzuleiten und dann nach Nulldurchgängen zu suchen. Auch hier ergibt sich das Problem, für viele verschiedene Orientierungen Ableitungsoperatoren zu realisieren, was wegen des diskreten Bildrasters für Zwischenrichtungen nur auf wenig effiziente Weise möglich ist.

Eine elegantere Lösung ergibt sich, wenn wir uns an die Beziehung zwischen Gradient und Richtungsableitung erinnern (\vec{n} sei der Einheitsvektor in Gradientenrichtung):²

$$\begin{aligned} \frac{\partial^2 I}{\partial \vec{n}^2} &= \frac{\partial}{\partial \vec{n}} \left(\frac{\partial I}{\partial \vec{n}} \right) \\ &= \nabla (\nabla I \cdot \vec{n}) \cdot \vec{n} \\ &= \nabla \left(\nabla I \cdot \frac{\nabla I}{|\nabla I|} \right) \cdot \frac{\nabla I}{|\nabla I|} \\ &= \frac{I_x^2 I_{xx} + 2I_x I_y I_{xy} + I_y^2 I_{yy}}{I_x^2 + I_y^2} \end{aligned}$$

Die Auswertung dieser Funktion, gefolgt von der Detektion von Nulldurchgängen, stellt einen nichtlinearen richtungsunabhängigen Operator dar, der die Berechnung von fünf verschiedenen einfachen und gemischten direktionalen Ableitungen erfordert. Jede dieser Ableitungen kann als Faltungsoperation mit einer entsprechenden Maske realisiert werden. Hinweise auf diesen Operator finden sich in [Canny 83] und [Torre + Poggio 86]. Da der Operator richtungsunabhängig ist — er berücksichtigt lokal die Gradientenrichtung — lassen sich durch Suche nach Nulldurchgängen im Operatorergebnisbild Kanten beliebiger Orientierung detektieren. Zu beachten ist hier allerdings, daß wir es mit diskreten Bildfunktionen zu tun haben. Nulldurchgänge lassen sich deshalb nicht nur dort detektieren, wo die zweite Ableitung der Bildfunktion den Wert Null aufweist, sondern auch zwischen den Abtaststellen des Bildes, wenn die zweifach abgeleitete Bildfunktion zwischen zwei Abtaststellen ihr Vorzeichen ändert. Die genaue Umsetzung des vorgeschlagenen Operators erfordert außerdem die Berücksichtigung der Gradientenrichtung bei der Suche nach Vorzeichenwechseln.

Marr + Hildreth 80 schlagen als Vereinfachung des nichtlinearen Operators die Verwendung des Laplaceoperators ∇^2 vor:

$$\nabla^2 I = I_{xx} + I_{yy}$$

Die mit Hilfe des Laplaceoperators im Bereich einer Kante gefundenen Nulldurchgänge stimmen mit den Ergebnissen des nichtlinearen Operators überein, solange die Bildfunk-

²Die tiefgestellten Indizes kennzeichnen auch hier die verschiedenen Ableitungen nach x und y .

tion entlang der Kante höchstens linear variiert. Wir stellen uns dazu eine vertikale³ Kante vor. Ist die Bildfunktion in Kantenrichtung konstant, so werden alle Ableitungen nach y Null. Der oben hergeleitete nichtlineare Operator liefert in diesem Fall:

$$\frac{\partial^2 I}{\partial \bar{n}^2} = I_{xx}$$

Dieses Ergebnis liefert auch der Laplaceoperator. Variiert die Bildfunktion im Bereich der Kante in y -Richtung linear, liefern I_{yy} und I_{xy} Null. Auch in diesem Fall bestimmt der Nulldurchgang von I_{xx} die Kantenposition, da der verbleibende Term im Bereich der Kante nur dort einen Nulldurchgang besitzt, wo I_{xx} Null wird:

$$\frac{\partial^2 I}{\partial \bar{n}^2} = \frac{I_x^2 I_{xx}}{I_x^2 + I_y^2}$$

Der Laplaceoperator besitzt die schöne Eigenschaft, der einfachste lineare und rotationsinvariante Ableitungsoperator zu sein (siehe [Rosenfeld + Kak 82, Bd. 1, S. 238f]). Seine Realisierung ist mit Hilfe einer einzigen Faltungsoption möglich. Marr und Hildreth schlagen die Kombination des Laplaceoperators mit einem gaußförmigen Mittelungsoperator G vor. Der entstehende Operator $\nabla^2 G$ wird wegen seiner Form häufig als *mexican hat operator* bezeichnet.

Die Schwachstellen des Laplaceoperators gegenüber dem nichtlinearen Operator sind deutlich erkennbar: Die Verwendung des Laplaceoperators setzt die strikte Einhaltung des Stufenkantenmodells voraus, wobei lineare Variationen entlang einer Kante noch akzeptiert werden. Auf systematische oder rauschbedingte Variationen der Bildfunktion in Kantenrichtung, die von diesem Modell abweichen, reagiert der Operator empfindlich, was eine Verschlechterung von Detektion und Lokalisation zur Folge haben muß. Dies ist der Preis, der für die Nichteinbeziehung der Gradientenrichtung gezahlt werden muß. Auch die großen Fehler, die durch den Laplaceoperator gefundene Nulldurchgangslinien an Eck- und Kreuzungspunkten von Kanten aufweisen, haben hier ihre Ursache.

Canny 83 vergleicht beide Ansätze unter Berücksichtigung des Einflusses von Rauschen auf formale Weise. Seine Analyse besagt, daß sich durch die Verwendung des nichtlinearen Operators, der die Gradientenrichtung lokal berücksichtigt, gegenüber dem Marr-Hildreth-Operator eine um den Faktor 2 verbesserte Detektion und eine um den Faktor 1.6 verbesserte Lokalisation erzielen läßt. Ein sehr ausführlicher Vergleich der Operatoren, sowie eine genaue Analyse der geometrischen Eigenschaften von Nulldurchgängen findet sich in [Torre + Poggio 86].

Canny 83 schlägt einen Ansatz vor, der zum einen den hohen Rechenaufwand des nichtlinearen Operators verringert, zum anderen aber die lokale Richtungsspezifität beibehält. Kanten werden nicht durch Nulldurchgänge der zweiten, sondern durch Extrema der ersten Ableitung in Gradientenrichtung definiert. Der Wert der Ableitung in Gradientenrichtung ist für jeden Bildpunkt gegeben durch den zuvor bestimmten

³Die vertikale Richtung entspreche im folgenden stets der y -Richtung, die horizontale Richtung der x -Richtung.

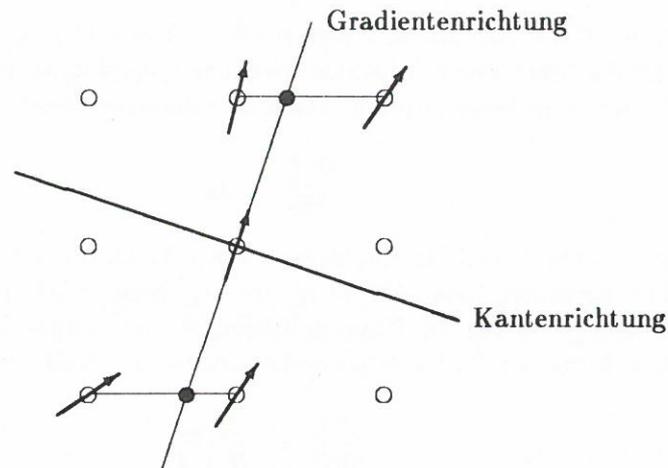


Abbildung 5.2: Detektion von Extrema durch lokalen Vergleich

Gradientenbetrag. Ob ein Bildpunkt einen Gradientenbetrag aufweist, der in Gradientenrichtung ein Extremum darstellt, läßt sich durch einen einfachen lokalen Vergleich feststellen. Zum Vergleich stehen die Gradientenbeträge der horizontal, vertikal und diagonal benachbarten Bildpunkte zu Verfügung. Weist die Gradientenrichtung einen Zwischenwert auf, muß zwischen jeweils zwei Nachbarn interpoliert⁴ werden (siehe Abb. 5.2). Die Idee einer solchen, als *non maximum suppression* bezeichneten, Operation zur Lokalisation von Kanten findet sich bereits in [Rosenfeld + Thurston 71].

Auch an dieser Stelle müssen zwei Anmerkungen gemacht werden: Der Vergleich der Gradientenbeträge zur Prüfung der Extremalbedingung setzt voraus, daß die Richtungen aller in den Vergleich eingehenden Gradienten mit der Gradientenrichtung des betrachteten, zentralen Bildpunktes übereinstimmt. In der Praxis wird diese Annahme sicher oft ausreichend gut erfüllt sein. An Eck- und Kreuzungspunkten wird diese Annahme jedoch verletzt. Um hier korrekt vorzugehen, müßten alle zu vergleichenden Gradientenvektoren auf die maßgebliche Gradientenrichtung projiziert werden, um dann die Beträge der Projektionen miteinander zu vergleichen.⁵ Dies bedeutet allerdings einen wesentlich höheren Rechenaufwand. Die zweite Anmerkung betrifft den räumlichen Abstand zweier zu vergleichender Werte. Wie wir in Abschnitt 5.2 sehen werden, ist es sinnvoll, die Größe der lokalen Umgebungen, die in die Mittelung und Differenzbildung eingehen, zu variieren. Wie der Zusammenhang aussieht zwischen der Größe der Umgebungen und dem Abstand, in dem die Gradientenbeträge miteinander verglichen werden und ob dies in einem praktischen Algorithmus berücksichtigt werden muß, wird von Canny nicht untersucht.

Das von Canny vorgeschlagene und unseren eigenen Experimenten zugrundeliegende ableitungsorientierte Detektionsschema umfaßt somit folgende Schritte:

⁴Canny schlägt eine lineare Interpolation vor.

⁵Entsprechendes gilt bei der richtungsspezifischen Suche nach Vorzeichenwechseln bei Verwendung des nichtlinearen Operators.

1. Berechnung der Richtungsableitungen I_x und I_y durch zwei Faltungsoperationen.
2. Bestimmung von Gradientenrichtung und -betrag aus den Komponenten I_x und I_y .
3. Detektion lokaler Extrema durch Vergleich der, gegebenenfalls interpolierten, Gradientenbeträge in Gradientenrichtung.

Es sei noch darauf hingewiesen, daß als lokale Extrema in diesem Zusammenhang nur positive Maxima und negative Minima betrachtet werden. Negative Maxima und positive Minima korrespondieren mit stellenweise flacheren Passagen in einer ansonsten steileren Flanke und stellen somit kein Merkmal dar, das mit der gegebenen Definition einer Kante übereinstimmt.

Nachdem wir das Grundgerüst eines ableitungsorientierten Detektionsschemas ausführlich dargestellt haben, werden wir nun auf zwei Fragen eingehen, die bisher nur angedeutet wurden:

- Welche Form sollte die in die Berechnung der Richtungsableitungen eingehende Mittelungsfunktion besitzen?
- Wie groß sollten die in die Mittelung eingehenden lokalen Umgebungen sein?

Verschiedene Untersuchungen in den letzten Jahren, die von unterschiedlichen Fragestellungen ausgehen, deuten auf die besondere Eignung der Gaußfunktion als Mittelungsfunktion in Verbindung mit Ableitungsoperationen hin. Die Argumente seien hier kurz zusammengestellt:

- Die Gaußmittelung erlaubt die Verarbeitung eines Bildes in unterschiedlichen Auflösungssebenen.
- Sie bildet den optimalen Kompromiß zwischen sicherer Detektion und Lokalisation.
- Sie dient als Regularisierungsfiler.
- Sie besitzt Eigenschaften, die eine einfache und effiziente Implementation von Detektionsverfahren ermöglichen.

In den nun folgenden Abschnitten wollen wir diese Aussagen kurz erläutern.

5.2 Verwendung verschiedener Operatorgrößen

Um die Verwendung von Operatoren verschiedener Größe zu motivieren, wollen wir uns vorstellen, wir würden durch das Objektiv einer Kamera hindurch einen Baum betrachten. Von der Entfernungseinstellung am Objektiv hängt es ab, wie scharf, d.h. mit welcher räumlichen *Auflösung*, wir den Baum betrachten. Variieren wir die Entfernungseinstellung, so läßt sich beobachten, daß in verschiedenen Bildauflösungen verschiedene

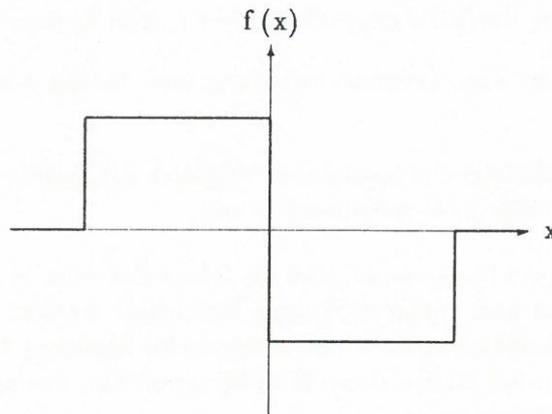


Abbildung 5.3: Difference-of-boxes-Ableitungsoperator

Strukturen erkennbar werden. Bei unscharfer Einstellung treten als Kanten in erster Linie die äußeren Konturen des Baumes hervor, wenn er sich deutlich gegen den Hintergrund absetzt. In einer etwas schärferen Einstellung erkennt man die Gliederung des Baumes in Stamm, Äste und Laub. Bei der schärfsten Einstellung sind möglicherweise sogar die feinen Strukturen auf den Blättern und in der Rinde erkennbar. So besitzt jede Struktur im Bild eine eigene Größenordnung, die einen ihr angepassten Betrachtungsmaßstab verlangt. Obwohl eine bestimmte Kantenstruktur in einem weiten Bereich räumlicher Auflösungen erkennbar sein kann, ist doch einsichtig, daß es eine Auflösung gibt, die ihrer physikalischen Natur und somit dem zu ihrer Erkennung geeigneten Betrachtungsmaßstab am besten entspricht. Ein Verfahren zur Detektion von Kanten in unterschiedlichsten Bildern muß dieser Eigenschaft von Kantenstrukturen Rechnung tragen.

Detektionsverfahren, die auf der Berechnung räumlicher Ableitungen der Bildfunktion basieren, bilden Differenzen zwischen den gemittelten Intensitäten lokaler Umgebungen eines betrachteten Bildpunktes. Die Veränderung der Größen der in die Mittelung einbezogenen Umgebungen erlaubt eine Anpassung des Betrachtungsmaßstabes an die lokale Struktur des Bildes.

Eines der ersten Verfahren, das aus dieser Motivation heraus Operatoren unterschiedlicher Größe verwendet, ist das von *Rosenfeld + Thurston 71* vorgeschlagene. Mit Hilfe von sogenannten *Difference-of-boxes-Funktionen* (siehe Abb. 5.3) wird die erste Ableitung des Bildes berechnet. Die verwendete *Operatorweite* wird für eine bestimmte Position im Bild solange vergrößert, bis kein nennenswerter Zuwachs im Ableitungsbetrag mehr feststellbar ist. Anschließend werden Kanten als lokale Extrema des Ableitungsbetrages detektiert.

Bei der Veränderung der Operatorweite ist folgendes zu beobachten: Mit zunehmender Größe des Operators wird der Mittelungseffekt stärker; das Rauschen wird zunehmend unterdrückt, was eine sicherere Detektion erlaubt. Gleichzeitig tritt ein immer stärker werdender Verschmierungseffekt auf; die Lokalisation einer detektierten Kante

wird immer ungenauer. Detektion und Lokalisation stehen also in Konflikt zueinander: Eine Verbesserung der Detektion ist nur auf Kosten der Lokalisation möglich und umgekehrt. Canny 83 bezeichnet dies als *Unsicherheitsprinzip*. Im Gegensatz zu Rosenfeld und Thurston, die wegen des rauschunterdrückenden Effekts möglichst große Operatoren bevorzugen, schlägt Canny vor, lokal den kleinsten Operator zu verwenden, der eine ausreichende Robustheit gegen das Rauschen besitzt, um eine gute Lokalisation der Kante zu erhalten. Eine Aussage über die Robustheit des Operators gegen Rauschen erlaubt das *Signal-zu-Rausch-Verhältnis*, das das Verhältnis zwischen der Stärke des Kantensignals und dem Signal aufgrund von Rauschen mißt. Wir werden für diesen Begriff im weiteren die Abkürzung SNR (*Signal to Noise Ratio*) verwenden.

Die noch zu klärende Frage ist die nach der geeigneten Form der Mittelungsfunktion. Gefordert haben wir sowohl die Möglichkeit, Kanten unterschiedlicher Auflösung zu detektieren, als auch eine detektierte Kante möglichst gut zu lokalisieren. Marr + Hildreth 80 weisen ausgehend von diesen Forderungen darauf hin, daß die Gaußfunktion als Mittelungsfunktion das Produkt der Bandbreiten im Orts- und Frequenzraum minimiert. Eine kleine Bandbreite im Ortsraum bedeutet eine gute Lokalisation; eine kleine Bandbreite im Frequenzraum bedeutet eine enge Eingrenzung der räumlichen Auflösung der zu detektierenden Strukturen. Ein minimales Produkt beider Bandbreiten entspricht somit dem optimalen Kompromiß zwischen beiden Forderungen. Die Gaußfunktion (siehe Abb. 5.4)

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

ist parametrisierbar durch die Standardabweichung σ , die die Operatorweite und somit den Betrachtungsmaßstab bei der Analyse des Bildes festlegt. Durch Veränderung von σ läßt sich für eine gegebene Kante das Gewicht entweder auf eine robustere Detektion oder eine bessere Lokalisation verschieben.

Seit Beginn der achtziger Jahre wurden viele Detektionsverfahren entwickelt, die in Kombination mit Ableitungsoperatoren eine gaußgewichtete Mittelungsfunktion verwenden. Häufig wird die Detektion von Kanten für unterschiedliche Auflösungen — also σ -Werte — durchgeführt. Betrachtet man die Nulldurchgänge der zweiten Ableitung einer eindimensionalen Bildfunktion über eine kontinuierliche Folge von σ -Werten hinweg, so lassen sich die detektierten Strukturen der verschiedenen Auflösungsebenen einander zuordnen und zu einer vollständigen Beschreibung der Bildinformation verwenden, wie Witkin 83 zeigt. Yuille + Poggio 86 führen diese Untersuchungen für zweidimensionale Bildfunktionen fort. Ein wichtiges Resultat der beiden Arbeiten ist die Aussage über das sogenannte *nice scaling behavior* von Nulldurchgängen der zweiten Ableitung der gaußgeglätteten Bildfunktion:

Bei Vergrößerung der Operatorweite (σ) können Nulldurchgänge verschwinden, miteinander verschmelzen, jedoch nicht neu entstehen.

Diese Eigenschaft gilt nur bei Verwendung der Gaußfunktion als Mittelungsfilter. Gewöhnlich werden wir Ableitungen der Bildfunktion nur für einige diskrete σ -Werte berechnen. Damit ergibt sich das nichttriviale und bis heute — trotz vieler Vorschläge —

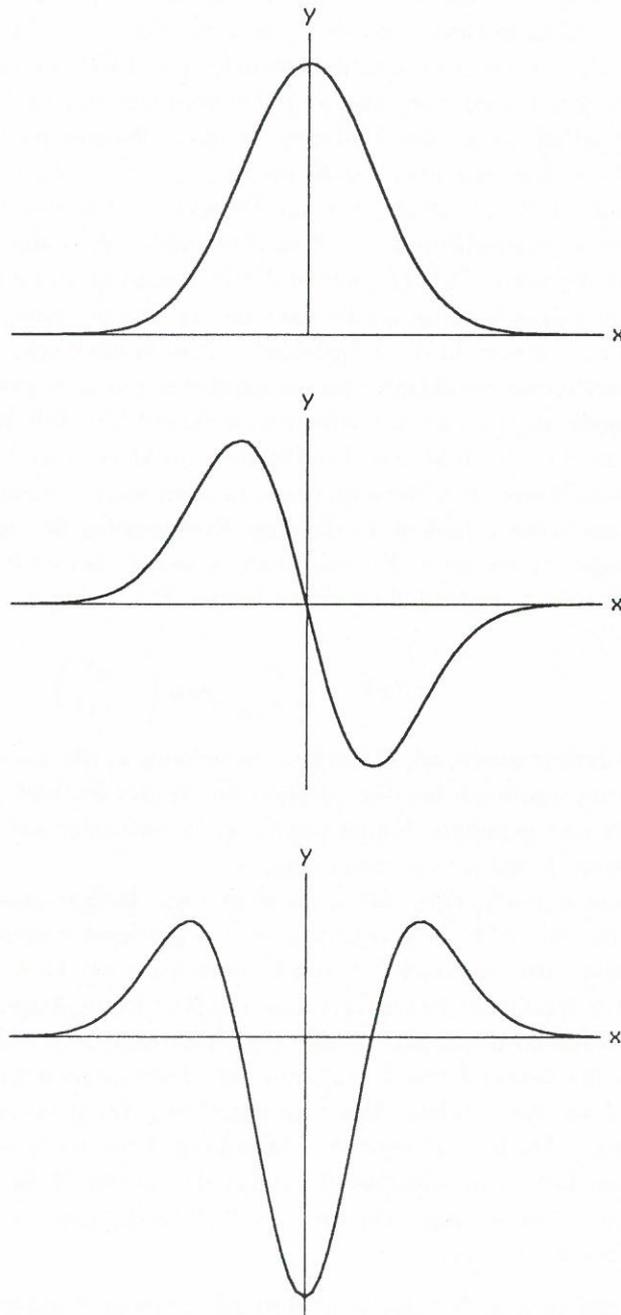


Abbildung 5.4: Gaußfunktion mit erster und zweiter Ableitung. Zur besseren Darstellung wurden die drei Funktionen unterschiedlich skaliert.

wohl nicht befriedigend gelöste Problem der Integration der Ergebnisse aus den verschiedenen Auflösungsebenen. Durch die unvorhersehbaren Veränderungen in der Lokalisation der detektierten Strukturen ist eine eindeutige Zuordnung zwischen den Ebenen, wie im kontinuierlichen Fall, nicht mehr möglich. Aussagen darüber, welche Strukturen durch dasselbe physikalische Merkmal verursacht wurden, welche durch räumliche Überlappung mehrerer Effekte neu entstanden oder verschwunden sind oder welche Auflösung die zur Detektion oder Lokalisation einer bestimmten Struktur geeignetste ist, sind kaum noch möglich.

Marr + Hildreth 80 betrachten die übereinstimmende Lokalisation von Nulldurchgängen in mehreren in der Auflösungshierarchie aufeinanderfolgenden Auflösungsebenen als Hinweis darauf, daß diese Nulldurchgänge durch dasselbe physikalische Phänomen verursacht wurden. Canny 83 stellt einen Ansatz zur *Merkmals-synthese* vor: Ausgehend von den Kanten, die mit Hilfe des kleinsten Operators gefunden wurden, wird aus dem Operatorergebnisbild durch Glättung das Ergebnisbild der nächst größeren Operatorweite synthetisiert. Dann wird das tatsächliche Operatorergebnisbild für diese Größe berechnet. Neue Kanten werden nur akzeptiert, wenn ihr berechneter Operatorwert signifikant von dem synthetisierten abweicht. Andernfalls wird davon ausgegangen, daß es sich um eine bereits detektierte Kante handelt, die bei einer kleineren Operatorweite besser lokalisiert werden kann. Dieses Vorgehen läßt sich für die nachfolgenden Operatorweiten wiederholen. Der Ansatz hat den Vorteil, daß zwischen den detektierten Kanten benachbarter Auflösungsebenen keine exakt übereinstimmende Lokalisation erforderlich ist, um sie einander zuzuordnen und daß die Detektion zweier unabhängiger Effekte, die sich in zwei deutlich verschiedenen Auflösungsbereichen bemerkbar machen, jedoch in ihrer Position übereinstimmen, möglich wird (z.B. Überlagerung einer *weichen* Schattenkante und einer *scharfen* Reflektanzkante).

Basierend auf einer speziellen Normierungsbedingung für die verwendete Gaußfunktion, schlagen Korn + Erdtel 85 vor, die Operatorweite ausgehend vom kleinsten Operator zu vergrößern, bis ein maximales Operatorergebnis erreicht wird. Dieses Vorgehen erinnert an den Ansatz von Rosenfeld + Thurston 71.

Bei der Auswahl einer geeigneten Operatorgröße stellt sich noch ein weiteres Problem, das bereits angedeutet wurde. Die meisten Verfahren setzen voraus, daß im betrachteten Bildausschnitt nur ein Kanteneffekt zu beobachten ist. Kommt es an Kreuzungs- oder Eckpunkten von Kanten oder durch dicht benachbarte Kanten zu einer gegenseitigen Beeinflussung unterschiedlicher Kanteneffekte, wird das zugrundeliegende Kantenmodell verletzt und es kommt zu unvorhersehbaren Störungen der Detektion und Lokalisation. Die Wahrscheinlichkeit einer Überlagerung mehrerer Effekte wächst mit zunehmender Operatorweite, so daß die gewonnene Information möglicherweise immer unsicherer wird.

Zusammenfassend läßt sich feststellen, daß die Integration der Ergebnisse von verschiedenen Auflösungsebenen auch heute noch ein offenes Problem darstellt.

5.3 Differenzieren als “ill-posed problem”

Die Ausführungen in diesem Abschnitt stützen sich in starkem Maße auf [Torre + Poggio 86], [Poggio et al. 85] und [Poggio + Torre 84]. Für weitergehende Details sei auf diese Arbeiten verwiesen.

Ein mathematisches Problem ist *well-posed*, wenn seine Lösung

- existiert,
- eindeutig ist,
- und stetig von den Eingangsdaten abhängt.

Die meisten Probleme der klassischen Physik sind “well-posed”. *Inverse* Probleme genügen den genannten Bedingungen jedoch gewöhnlich nicht. Ein inverses Problem ergibt sich, wenn man versucht, von den Ergebnisdaten eines Abbildungsprozesses auf die Ursprungsdaten zurückzuschließen, die Abbildung somit umzukehren. Im konkreten Fall der Bildverarbeitung betrachten wir die Abbildung der Bildfunktion z über einen rauschbehafteten Abbildungsoperator A auf die gemessenen Intensitätswerte y :

$$y = Az$$

Ein inverses und zugleich *ill-posed* Problem wäre die Rekonstruktion der exakten Bildfunktion z aus der — vom Rauschen beeinflussten — Funktion y . Ein typisches “ill-posed problem” ist auch das Differenzieren der Bildfunktion. Das Ziel ist hierbei, durch Differenzieren von y eine gute Approximation der Ableitung von z zu erhalten. Das Rauschen kann jedoch dazu führen, daß sich zwischen den Ableitungen von y und z erhebliche Unterschiede ergeben, selbst wenn y und z selbst nur geringe Unterschiede voneinander aufweisen (siehe [Torre + Poggio 86]). Dies gilt in zunehmendem Maße für höhere Ableitungen.

Zur Lösung von “ill-posed problems” wurden in der Vergangenheit verschiedene *Regularisierungstechniken* entwickelt. Die zugrundeliegende Idee der Regularisierung ist dabei, das betrachtete Problem als Minimierungsproblem zu formulieren und den Lösungsraum so einzuschränken, daß ein eindeutiges Minimum gefunden wird. Für das beschriebene Rekonstruktionsproblem ergibt sich folgende Aufgabe: Gesucht ist die Funktion \hat{z} , die den Ausdruck

$$\|A\hat{z} - y\|^2 + \lambda \cdot \|P\hat{z}\|^2$$

minimiert. Festzulegen ist dazu ein geeignetes Abstandsmaß, das die Abweichung zu den tatsächlich gemessenen Daten bewertet ($\|A\hat{z} - y\|^2$), ein Regularisierungsfunktional ($\|P\hat{z}\|^2$), das die physikalischen Randbedingungen für eine Lösung ausdrückt, sowie der Regularisierungsfaktor λ , der als Gewichtungsfaktor zwischen Datennähe und Regularität dient. Für den konkreten Fall des Differenzierens schlagen Torre + Poggio 86 folgenden Ausdruck vor:

$$\sum_i (y_i - f(x_i))^2 + \lambda \cdot \int (f''(x))^2 dx$$

Die Werte $y_i = f(x_i) + \varepsilon_i$ entsprechen den diskreten, gemessenen Intensitäten, die sich unter Einfluß der Meßfehler ε_i aus der exakten Bildfunktion $f(x_i)$ ergeben. Als Regularisierungsfunktional dient das Integral über die Krümmungen von f . Gefordert wird damit eine Glattheit der ursprünglichen Funktion f , was in diesem Fall wegen der Bandbegrenzung des Bildsignals gerechtfertigt ist.

Poggio et al. 85 zeigen, daß die Lösung dieses Minimierungsproblems gefunden werden kann durch Faltung der Daten y_i mit einem Regularisierungsfiler R . Die Form dieses Filters stimmt gut mit einer Gaußfunktion überein. Wie sie zeigen, übernimmt die Standardabweichung σ die Rolle des Regularisierungsfaktors λ . Der Konflikt zwischen Datennähe und Regularität steht also im Zusammenhang mit dem bereits beschriebenen Konflikt zwischen Lokalisation und Detektion. Die Glättung des Bildes mit einem Gaußfilter vor oder in Verbindung mit einer Ableitungsoperation erfüllt also dieser Argumentation folgend die wünschenswerte Funktion einer Regularisierung des Differenzierens.

5.4 Herleitung einer Operatorfunktion als Optimierungsproblem

Einen weiteren Hinweis auf die geeignete Form einer Operatorfunktion zur Kantendetektion liefert die Arbeit von *Canny 83* (siehe hierzu auch [*Canny 83b*] oder [*Canny 86*]). *Canny* betrachtet das Problem, Kanten in verrauschten Bildern anhand von Extrema im Ergebnis der Faltung des Bildes mit einer Operatorfunktion zu detektieren, und stellt die Frage nach der Form des optimalen Operators.

Ausgangspunkt bilden drei Kriterien, die *Canny* an ein Kantendetektionsverfahren stellt. Durch mathematische Formulierung der Kriterien und Maximierung der erhaltenen Gleichungen gewinnt er eine Lösung für den Fall einer eindimensionalen kontinuierlichen Bildfunktion. Gefordert werden:

Sichere Detektion: Alle physikalisch relevanten Kanten sollen detektiert werden. Alle detektierten Kanten sollen mit einer Kante in der Szene korrespondieren.

Gute Lokalisation: Der Ort einer detektierten Kante soll möglichst gut mit der tatsächlichen Kantenposition übereinstimmen.

Eindeutige Detektion: Jede Kante soll nur einmal detektiert werden und nicht zur Detektion mehrerer Kanten in unmittelbarer Nachbarschaft führen.

Das dritte Kriterium erscheint auf den ersten Blick überflüssig, denn weitere Markierungen einer bereits detektierten Kante lassen sich nicht mit den Forderungen nach guter Detektion und guter Lokalisation vereinbaren. Im Rahmen der mathematischen Behandlung wird sich aber zeigen, daß dieses Kriterium explizit gemacht werden muß.

Bei der mathematischen Behandlung beschränkt sich *Canny* auf die Betrachtung von Stufenkanten. In [*Canny 83*] wird gezeigt, wie sich der Ansatz auch auf andere

Bildmerkmale übertragen läßt. Als eindimensionale Bildfunktion einer Stufenkante an der Position $x = 0$ unter Einfluß einer Rauschgröße $n(x)$ sei definiert:

$$I(x) = A \cdot u(x) + n(x)$$

wobei A die Höhe der Stufe angibt und für u folgendes gelte:

$$u(x) = \begin{cases} 0 & \text{wenn } x < 0 \\ 1 & \text{sonst} \end{cases}$$

Bei dem Rauschen soll es sich um *weißes, normalverteiltes* Rauschen handeln:

- Die Größe der Abweichung der Bildfunktion ist normalverteilt mit Mittelwert Null.
- Die Störungen an zwei verschiedenen Orten im Bild sind unabhängig voneinander.

Das Detektionskriterium

Die Güte der Detektion läßt sich durch das Signal-zu-Rausch-Verhältnis (SNR) bewerten. Zur Bestimmung des lokalen SNR wird das Ergebnis der Faltung der Bildfunktion mit der Operatorfunktion in eine Signal- und eine Rauschkomponente zerlegt. Die Faltung der kontinuierlichen Bildfunktion I mit der Operatorfunktion $f(x)$ an der Stelle x_0 ergibt:

$$(I * f)(x_0) = \int_{-\infty}^{+\infty} I(x_0 - x) \cdot f(x) dx$$

Für die Position der Kante ($x_0 = 0$) gilt:

$$\begin{aligned} (I * f)(0) &= \int_{-\infty}^{+\infty} I(-x) \cdot f(x) dx \\ &= \int_{-\infty}^{+\infty} (A \cdot u(-x) + n(-x)) \cdot f(x) dx \\ &= \int_{-\infty}^{+\infty} A \cdot u(-x) \cdot f(x) dx + \int_{-\infty}^{+\infty} n(-x) \cdot f(x) dx \end{aligned}$$

Das erste Integral liefert die Signalkomponente:

$$A \cdot \int_{-\infty}^0 f(x) dx$$

Da der Rauschanteil eine zufällige Größe darstellt, betrachtet man hier das mittlere Quadrat ($E[\]$ bedeutet *Erwartungswert von*):

$$E \left[\left(\int_{-\infty}^{+\infty} n(-x) \cdot f(x) dx \right)^2 \right]$$

Da es sich um weißes, also räumlich unkorreliertes Rauschen mit Mittelwert Null handelt, heben sich alle gemischten Terme des Quadrates in der Summe zu Null auf, so daß nur die quadratischen Terme zurückbleiben:

$$E \left[\int_{-\infty}^{+\infty} n^2(-x) \cdot f^2(x) dx \right]$$

Legt man folgendes fest:

$$n_0^2 = E [n^2(x)]$$

so ergibt sich für die (quadrierte) Rauschkomponente:

$$n_0^2 \cdot \int_{-\infty}^{+\infty} f^2(x) dx$$

Als Verhältnis beider Komponenten erhält man schließlich:

$$SNR = \frac{A \cdot \int_{-\infty}^0 f(x) dx}{n_0 \cdot \sqrt{\int_{-\infty}^{+\infty} f^2(x) dx}} = \frac{A}{n_0} \cdot \Sigma$$

wobei das so definierte Σ nur noch von der zu bestimmenden Operatorfunktion abhängt.

Das Lokalisationskriterium

Als Bewertung der Lokalisation dient die mittlere Abweichung des Maximums im Faltungsergebnis von der wirklichen Position der Kante (hier $x = 0$) aufgrund des Rauschens. Die Lage des Maximums wird durch den Nulldurchgang der ersten Ableitung des Faltungsergebnisses bestimmt:

$$\frac{d}{dx_0} \int_{-\infty}^{+\infty} I(x_0 - x) \cdot f(x) dx = 0$$

Da Integral und Differentiation vertauschbar sind und im Faltungsintegral wahlweise nur eine der beiden Größen von x_0 abhängt, entspricht dies:

$$\int_{-\infty}^{+\infty} I(x_0 - x) \cdot f'(x) dx = 0$$

Auch hier wird eine Zerlegung in Signal- und Rauschanteil vorgenommen:

$$\int_{-\infty}^{+\infty} A \cdot u(x_0 - x) \cdot f'(x) dx + \int_{-\infty}^{+\infty} n(x_0 - x) \cdot f'(x) dx = 0$$

Nach Umstellung ergibt sich:

$$\int_{-\infty}^{x_0} f'(x) dx = -\frac{1}{A} \cdot \int_{-\infty}^{+\infty} n(x_0 - x) \cdot f'(x) dx$$

Für $f(x)$ werden grundsätzlich zwei Eigenschaften vorausgesetzt:

1. $f(x)$ ist antisymmetrisch. Jede Funktion läßt sich in eine symmetrische und eine antisymmetrische Komponente zerlegen. Es ist einsichtig, daß die symmetrische Komponente nichts zur Detektion beitragen kann, wenn wir nach Maxima im Faltungsergebnis suchen. Sie erhöht lediglich den Rauscheinfluß.
2. $f(x)$ geht im Unendlichen gegen Null. Eine Kante stellt eine lokale Struktur im Bild dar, so daß weit entfernte Bildbereiche als unabhängig von dieser Struktur gelten und keinen Einfluß auf die Detektion oder Lokalisation der Kante haben sollten.

Damit ergibt sich für die linke Seite der Gleichung:

$$\int_{-\infty}^{x_0} f'(x) dx = f(x_0)$$

und durch Annäherung mit Hilfe einer Taylor-Reihe :

$$\approx x_0 \cdot f'(0)$$

Für den Ort des Maximums ergibt sich demnach:

$$x_0 = -\frac{1}{A \cdot f'(0)} \cdot \int_{-\infty}^{+\infty} n(x_0 - x) \cdot f'(x) dx$$

Als mittlere quadratische Abweichung von der tatsächlichen Position erhält man:

$$E|x_0^2| = \frac{n_0^2}{A^2 \cdot f'(0)^2} \cdot \int_{-\infty}^{+\infty} f'(x)^2 dx$$

Die zu maximierende Größe läßt sich wie folgt formulieren:

$$\frac{1}{\sqrt{E|x_0^2|}} = \frac{A \cdot |f'(0)|}{n_0 \cdot \sqrt{\int_{-\infty}^{+\infty} f'(x)^2 dx}} = \frac{A}{n_0} \cdot \Lambda$$

wobei Λ nur noch von der Operatorfunktion abhängt.

Canny weist an dieser Stelle auf den bereits dargestellten Konflikt zwischen Lokalisation und Detektion hin. Um diesem *Unsicherheitsprinzip* Rechnung zu tragen, wird das Produkt von Σ und Λ maximiert. Canny zeigt, daß dieses Produkt invariant ist gegen Skalierung der Funktion f . Die Maximierung liefert also eine *Klasse von Funktionen* als Lösung, deren festzulegender Parameter die Operatorweite ist. Eine Skalierung von f verbessert die Lokalisation unter Verschlechterung der Detektion bzw. umgekehrt. Das Produkt $\Sigma \cdot \Lambda$ aber bleibt konstant.

Die Durchführung der Maximierung soll hier nicht erläutert werden. Sie liefert als Klasse von optimalen Operatorfunktionen die Difference-of-boxes-Funktionen, die wir im Zusammenhang mit der Verwendung unterschiedlicher Operatorweiten kennengelernt haben.

Das Eindeutigkeitskriterium

In verrauschten Bildern ergibt sich bei Verwendung der Difference-of-boxes-Funktionen das Problem, daß in der Nähe einer Stufenkante mehrere Maxima detektiert werden, so daß nicht zu entscheiden ist, welches Maximum der wirklichen Kantenposition entspricht. Hier wird deutlich, warum die explizite Formulierung des dritten Kriteriums notwendig ist: Detektions- und Lokalisationskriterium betrachten nur das Operatorergebnis an einem einzigen Bildpunkt — die Auswirkungen auf benachbarte Punkte bleiben unberücksichtigt.

Um die Anzahl der Maxima in der Nähe einer Stufenkante zu begrenzen, formuliert Canny ein Kriterium, das den mittleren Abstand zweier Rauschmaxima in Beziehung zur Operatorweite setzt. Wir benötigen also einen Ausdruck, der die funktionale Abhängigkeit zwischen dem Abstand benachbarter Rauschmaxima und der Operatorfunktion f beschreibt. Hier läßt sich ein Ergebnis von Rice 45 verwenden, wenn man statt der Abstände der Maxima die Abstände der Nulldurchgänge in der Ableitung des Operatorergebnisses betrachtet. Bei Anwendung einer Operatorfunktion g auf normalverteiltes Rauschen ergibt sich als mittlerer Abstand der Nulldurchgänge:

$$x_{ave} = \pi \cdot \sqrt{\frac{-R(0)}{R''(0)}}$$

wobei $R(\tau)$ die Autokorrelationsfunktion von g ist:

$$R(\tau) = \int_{-\infty}^{+\infty} g(x) \cdot g(x + \tau) dx$$

Für die zweite Ableitung ergibt sich durch partielle Integration und unter der Annahme, daß g im Unendlichen gegen Null geht:

$$\begin{aligned} R''(\tau) &= \int_{-\infty}^{+\infty} g(x) \cdot g''(x + \tau) dx \\ &= - \int_{-\infty}^{+\infty} g'(x) \cdot g'(x + \tau) dx \end{aligned}$$

Für $\tau = 0$ gilt also:

$$R(0) = \int_{-\infty}^{+\infty} g^2(x) dx \quad \text{und} \quad R''(0) = - \int_{-\infty}^{+\infty} g'^2(x) dx.$$

Setzt man in unserem Fall $g = f'$, so ergibt sich:

$$x_{ave} = \pi \cdot \sqrt{\frac{\int_{-\infty}^{+\infty} f'^2(x) dx}{\int_{-\infty}^{+\infty} f''^2(x) dx}}$$

Mit der Forderung $x_{ave} = k \cdot W$ wird nun der mittlere Abstand der Nulldurchgänge auf den k -ten Teil der Operatorweite W festgelegt. Damit erhalten wir ein drittes Kriterium, das den Lösungsraum für das Optimierungsproblem weiter einschränkt.

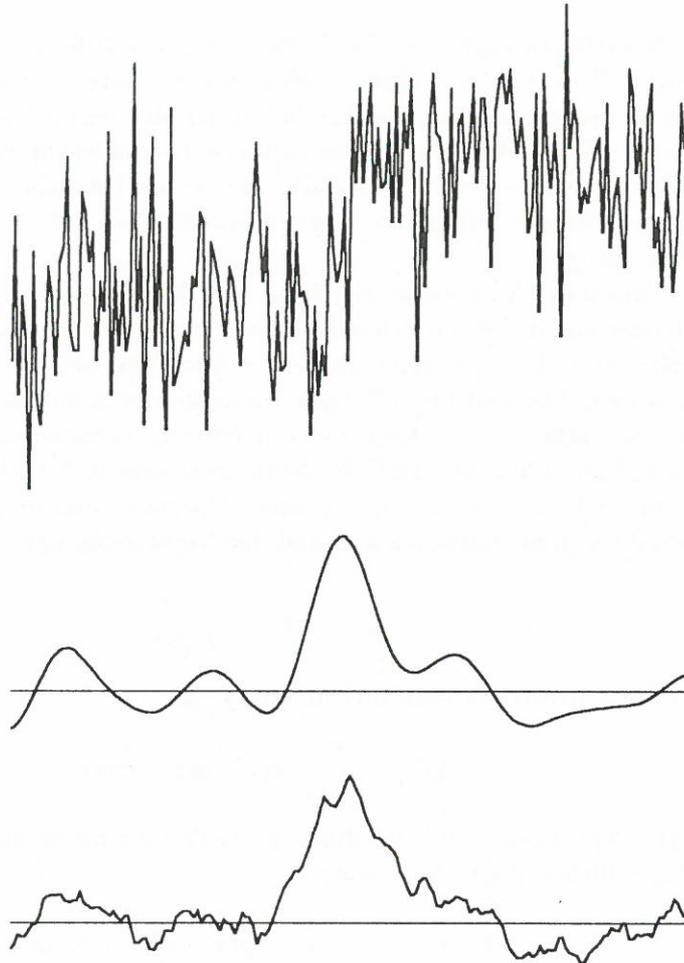


Abbildung 5.5: Unterschied bei Anwendung der ersten Ableitung der Gaußfunktion und eines Difference-of-boxes-Operators auf eine stark verrauschte Stufenkante

Nach Hinzunahme des Eindeutigkeitskriteriums liefert die Optimierung als Lösung eine Klasse von Funktionen, die sehr gut durch die erste Ableitung einer Gaußfunktion approximiert werden können (siehe Abb. 5.4).

Den Unterschied zwischen den Ergebnissen des gaußförmigen Ableitungsoperators und des Difference-of-boxes-Operators macht Abbildung 5.5 deutlich: Für eine stark verrauschte bzw. kontrastarme Stufenkante können bei Anwendung des Difference-of-boxes-Operators im Bereich der Kante mehrere starke Maxima detektiert werden, während sich bei Verwendung der ersten Ableitung der Gaußfunktion ein einziges klares Maximum ausbildet.

Canny gewinnt eine optimale Operatorfunktion ausgehend von sehr allgemeinen

Gütekriterien durch ein numerisches Optimierungsverfahren. Damit unterscheidet sich seine Entwicklung eines Operators grundsätzlich von vielen anderen Ansätzen. Trotzdem erhalten wir auch hier einen weiteren Hinweis auf die besondere Eignung der Gaußfunktion als Mittelungsoperator sowie eine Rechtfertigung des ableitungsorientierten Vorgehens zur Kantendetektion.

5.5 Faltungseigenschaften der Gaußfunktion

Viele der vorgenannten Argumentationen liefern Ergebnisse, die die Verwendung der Gaußfunktion nicht zwingend fordern, die aber zeigen, daß die Gaußfunktion die jeweils optimale Lösung gut approximiert. Der Grund für die so beliebte Verwendung der Gaußfunktion als annähernd optimale Lösung liegt in ihren besonderen mathematischen Eigenschaften, die eine einfache und sehr effiziente Realisation von Detektionsverfahren erlauben. Da diese Eigenschaften immer wieder in praktischen Verfahren ausgenutzt werden, wollen wir sie hier darstellen.

Die wichtigste Eigenschaft der zweidimensionalen Gaußfunktion sowie ihrer räumlichen Ableitungen ist die *Separierbarkeit*. Wir betrachten dazu die eindimensionale Gaußfunktion $g(z) = \exp(-\frac{z^2}{2\sigma^2})$. Für die zweidimensionale Gaußfunktion gilt:

$$\begin{aligned} G(x, y) &= \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \\ &= \exp\left(-\frac{x^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{y^2}{2\sigma^2}\right) \\ &= g(x) \cdot g(y) \end{aligned}$$

Für die zweidimensionale Faltung des Bildes mit einer Gaußfunktion an der Stelle (a, b) erhalten wir damit:

$$\begin{aligned} (G * I)(a, b) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G(x, y) \cdot I(a - x, b - y) dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x) \cdot g(y) \cdot I(a - x, b - y) dx dy \\ &= \int_{-\infty}^{+\infty} g(y) \cdot \left(\int_{-\infty}^{+\infty} g(x) \cdot I(a - x, b - y) dx \right) dy \end{aligned}$$

Die Faltung des Bildes mit einer zweidimensionalen Gaußfunktion läßt sich also realisieren durch die Hintereinanderausführung von zwei eindimensionalen Faltungen in x - und y -Richtung. Man spricht auch von der *Dekomposition* der Gaußfunktion. Wird die Gaußfunktion in einem praktischen Algorithmus durch eine diskrete Operatormaske approximiert, benötigt die zweidimensionale Faltung des Bildes mit einer $N \times N$ Bildpunkte großen Maske für jede Bildposition N^2 Multiplikationen. Nutzt man jedoch die

Separierbarkeit und faltet zunächst die Bildzeilen und anschließend die resultierenden Zeilen spaltenweise mit einer $1 \times N$ Maske, die die eindimensionale Gaußfunktion approximiert, so sind nur $2N$ Multiplikationen für eine Faltung erforderlich. Bei Verwendung großer Operatoren bedeutet dies eine beträchtliche Ersparnis an Rechenzeit. Bezeichnet man mit $(I * g)_x$ die eindimensionale Faltung in x -Richtung und mit $(I * g)_y$ entsprechend in y -Richtung, läßt sich das Resultat wie folgt formulieren:

$$I * G = ((I * g)_x * g)_y = ((I * g)_y * g)_x$$

Die zweite Eigenschaft, die häufig ausgenutzt wird, ist die Differentiationsregel der Faltungsoperation:

$$\frac{\partial(I * G)}{\partial x} = I * \frac{\partial G}{\partial x}$$

Sie gilt, weil Differentiation und Faltungsintegral miteinander vertauschbar sind und innerhalb des Faltungsintegrals nur eine der beiden Funktionen von der speziellen x -Koordinate abhängt. Glättung und Ableitung können also in einem Schritt realisiert werden als Faltung der Bildfunktion mit entsprechenden direktionalen Ableitungen der Gaußfunktion. Durch Kombination beider Eigenschaften erhalten wir:

$$\frac{\partial(I * G)}{\partial x} = ((I * g)_y * \frac{dg}{dx})_x = ((I * g)_y * g')_x = ((I * g')_x * g)_y$$

und entsprechend

$$\frac{\partial(I * G)}{\partial y} = ((I * g)_x * g')_y$$

Dies gilt entsprechend für höhere oder gemischte Ableitungen. Umgekehrt entspricht die Faltung des Bildes mit einer Ableitung der Gaußfunktion dem Differenzieren des zuvor gaußgeglätteten Bildes.

Kapitel 6

Verfahren zur Kantendetektion in Farbbildern

Nachdem wir die wichtigsten Aspekte der Detektion von Kanten dargestellt und einen Überblick über verschiedene praktische Ansätze zur Kantendetektion in Intensitätsbildern gegeben haben, wollen wir nun untersuchen, welche Möglichkeiten sich zur Detektion von Kanten in Farbbildern anbieten.

Die in den letzten zwanzig Jahren veröffentlichte Fachliteratur weist nur sehr wenige Beiträge auf, die sich mit dieser Problematik auseinandersetzen. In den folgenden Abschnitten wollen wir die Beiträge, die uns zugänglich waren, darstellen und kurz diskutieren. Der Schwerpunkt dieser Betrachtungen soll auf der Frage liegen, in welcher Form die vorgestellten Ansätze die Farbinformation des Bildes auswerten. Ansätzen zur ableitungsorientierten Kantendetektion in Farbbildern werden wir, wie schon beim Überblick über die Intensitätsverfahren, ein eigenes Kapitel widmen.

Einige weitere Beiträge, die uns vorlagen, fanden keine Aufnahme in diese Übersicht. Der Ansatz von *Huntsberger + Descalzi 85* dient der Detektion von Bereichsgrenzen für ein zuvor segmentiertes Bild. *Machuca + Phillips 83* untersuchen Methoden der Analyse von Vektorfeldern und ihre Anwendbarkeit in der Bildverarbeitung. Dem von ihnen vorgeschlagenen Kantendetektionsverfahren liegt allerdings kein Stufenkantenmodell zugrunde. Sie betrachten Kanten mit dachförmigem Profil. *Sankar 78*, *Pietikainen + Harwood 86* und *Solinsky 85* ersetzen die Grauwertdifferenzen direkt benachbarter Bildpunkte auf die eine oder andere Weise durch Vektordifferenzen.

Bevor wir mit der Übersicht beginnen, sei an dieser Stelle die Frage nach der Wahl eines geeigneten *Farbkoordinatensystems* aufgegriffen. Das von einer herkömmlichen Farbkamera gelieferte Bildsignal besitzt zunächst drei Komponenten. Man spricht vom Rot-, Grün- und Blaukanal. Jeder Kanal reagiert auf verschiedene spektrale Bereiche des in die Kamera einfallenden Lichtes. Die größte Sensitivität der Kanäle liegt im Rot-, Grün- bzw. Blaubereich des Spektrums. Das Farbsystem, in dem die so gemessenen Farbvektoren dargestellt werden, wird als RGB-Raum bezeichnet. Da jeder Kanal für einen ganzen Spektralbereich sensitiv ist und sich diese Bereiche überschneiden, sind die Informationen in den drei Farbkanälen im allgemeinen nicht voneinander unabhängig.

Es stellt sich die Frage, ob die Farbinformation eines Bildes durch die Darstellung im RGB-Raum in geeigneter Weise repräsentiert wird oder ob das Bild vor einer Verarbeitung durch eine entsprechende Transformation in eine andere Darstellung überführt werden sollte. In der Literatur finden sich viele Vorschläge für Farbtransformationen. Die Motivationen für die verschiedenen Transformationen liegen häufig in technischen Anforderungen (z.B. Fernsehnormen) oder in Beobachtungen, die am Sehsystem des Menschen gemacht wurden.

Wenn im Zusammenhang mit Verarbeitungsprozessen, wie der Kantendetektion, die Rede von Farbtransformationen ist, so beschränkt sich dies meist auf die Darstellung der verschiedenen Transformationen oder auf den visuellen Vergleich der in den verschiedenen Repräsentationen gewonnenen Resultate. Kriterien, die Aussagen darüber ermöglichen, unter welchen Umständen welche Transformationen für welche Aufgabstellungen geeignet sind, fehlen bisher. Solche Kriterien ließen sich nur durch sorgfältige Analyse des Bildentstehungsprozesses und der in die Verarbeitungsprozesse eingehenden Modelle gewinnen.

Die Transformation des Farbsystems stellt eine *Umgewichtung* der spektralen Bildinformation dar. Gänzlich neue Informationen über das Bild sind auf diese Weise nicht zu erzielen. Der gezielte Einsatz einer Transformation setzt Wissen über die zu verarbeitenden Bilder und den Verarbeitungsprozeß voraus. Solange die Art der zu verarbeitenden Bilder in keiner Weise eingeschränkt wird, bleibt die Frage nach dem geeigneten Farbsystem — wegen fehlender Kriterien — offen.

Ein besonderer Aspekt bei der Untersuchung verschiedener Farbsysteme ist die Abhängigkeit bzw. Unabhängigkeit der Informationen in den verschiedenen Kanälen. Im allgemeinen muß in allen bekannten Farbsystemen von einer Abhängigkeit der Kanäle ausgegangen werden. Eine Ausnahme bilden Systeme, die aus der Transformation eines RGB-Bildes mit Hilfe der *Karhunen-Loève-Transformation* hervorgegangen sind. Diese Transformation sorgt für eine Dekorrelation der Kanäle. Die Transformationsparameter müssen allerdings für jedes Bild neu bestimmt werden. Die erreichte Unabhängigkeit der Kanäle gilt außerdem nur global. Lokal können weiterhin Abhängigkeiten auftreten.

Im folgenden wird — wenn nicht anders angegeben — stets davon ausgegangen, daß Farbbilder durch Farbvektoren im RGB-Raum repräsentiert sind. Wenn wichtige Wechselwirkungen zwischen den Eigenschaften eines Verfahrens und dem gewählten Farbsystem bestehen, werden wir darauf hinweisen.

6.1 Nevatia 77

Einen der ersten Ansätze zur Übertragung eines Kantendetektionsverfahrens von Intensitätsbildern auf Farbbilder beschreibt *Nevatia 77*. Hier wird der bereits beschriebene Ansatz von *Hueckel 71* zur Approximation eines Kantenmodells an die Bildfunktion auf Farb-, oder allgemeiner, Multispektralbilder erweitert.

Ausgehend von Beobachtungen am Menschen sieht *Nevatia* Farbe als eine die Helligkeit ergänzende Information an, die eine reichere Wahrnehmung erlaubt. Deshalb betrachtet er bei der Definition einer Farbkante nicht den RGB-Raum, sondern einen

Farbraum, der durch die Intensität sowie zwei Farbkomponenten aufgespannt wird.¹ Die verwendeten Komponenten sind:

$$Y = c_1 \cdot R + c_2 \cdot G + c_3 \cdot B \quad \text{mit} \quad c_1 + c_2 + c_3 = 1$$

$$T_1 = \frac{R}{R + G + B} \quad T_2 = \frac{G}{R + G + B}$$

Die Konstanten c_i sind so gewählt, daß Y der Intensität entspricht. Während eine Kante in einem Intensitätsbild in Hueckels Ansatz definiert wird durch die Parameterwerte einer an das Bild approximierten Modellfunktion F , besteht eine Farbkante nach Nevatia aus drei Teilen F_1, F_2, F_3 , die den an die Bildkomponenten Y, T_1 und T_2 angepaßten Modellfunktionen entsprechen. Betrachtet man die F_i als unabhängig voneinander, dann ergibt sich der zu minimierende Gesamtfehler ε der Approximation als Summe der Fehler der drei Einzelapproximationen:

$$\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$$

Nevatia verlangt nun als einzige Wechselbeziehung zwischen den drei Modellfunktionen die Übereinstimmung der Kantenrichtungen. Dabei läßt sich — wie im Falle von Grauwertbildern — ausnutzen, daß sich die Kantenrichtung unabhängig von den übrigen Kantenparametern bestimmen läßt. Die Minimierung des Gesamtfehlers ε in Abhängigkeit von der Kantenrichtung verlangt das Lösen einer Gleichung achter Ordnung. Um diesen Aufwand zu vermeiden, führt Nevatia die Minimierung für jede Bildkomponente separat durch und nähert die gemeinsame Kantenrichtung durch das Mittel der mit den jeweiligen Kontrastwerten gewichteten Einzelrichtungen an. Ist die gemeinsame Richtung bestimmt, müssen die übrigen Parameter der drei Einzelapproximationen neu berechnet werden. Der Algorithmus besteht somit aus folgenden Teilen:

1. Separate Approximation einer Modellfunktion an die drei Komponenten.
2. Mittelung der Richtungsparameter zur Festlegung einer gemeinsamen Kantenrichtung.
3. Neuberechnung der übrigen Parameter.

Eine weitere Integration der ermittelten Parameterwerte der Einzelapproximationen wird nicht vorgenommen. Eine Farbkante kann sich also in nur einer, zwei oder allen Bildkomponenten bemerkbar machen. Kanten, die in den Komponenten T_1 oder T_2 auftreten, werden *chromatische* Kanten genannt. Nevatia interessiert besonders der Vergleich zwischen den chromatischen Kanten und denen, die in der Intensitätskomponente detektiert werden. Er macht folgende Beobachtungen:

- Fast alle chromatischen Kanten sind auch Intensitätskanten.

¹Die Wahl des Farbkoordinatensystems ist jedoch durch den Ansatz nicht prinzipiell eingeschränkt.

- Es gibt weniger chromatische Kanten als Intensitätskanten. Chromatische Kanten sind aber meist aussagekräftiger (weniger Rauscheffekte, physikalische Relevanz der detektierten Kanten).

Nevatia zieht daraus den Schluß, daß die Bildintensität den größten Teil der Kanteninformation beinhaltet, während es auch Situationen gibt, in denen Kanten nur aufgrund ihres Farbkontrastes detektiert werden. Der Hauptwert der Farbinformation liegt in der größeren Robustheit und Zuverlässigkeit. Von einer Beantwortung der Frage, ob sich der Mehraufwand der Farbbildverarbeitung lohnt, sieht Nevatia wegen der geringen Anzahl durchgeführter Experimente ab.

Anmerkungen

Der beschriebene Ansatz behandelt die drei Komponenten eines Farbbildes im Grunde unabhängig voneinander. Wir erhalten zu einem Bildpunkt nicht die Information über *eine* sondern über *drei* Kanten, die mehr oder weniger stark korreliert sein können. Den einzigen Schritt hin zu einer Integration der Informationen aus den Einzelkanälen stellt die Forderung nach einer eindeutigen Kantenrichtung dar. Die Mittelung der separat bestimmten Richtungen wird in den meisten Fällen einen rauschunterdrückenden Effekt besitzen. Sie setzt allerdings die Korrelation der drei Farbkanäle voraus. Fraglich ist, welchen Nutzen die Information über eine gemeinsame Kantenrichtung besitzt, wenn die Position der Kante in den einzelnen Bildkomponenten variieren kann.

6.2 Robinson 76

Die Beiträge von Robinson und Nevatia gehören zu den Veröffentlichungen zum Thema Kantendetektion in Farbbildern, die immer wieder zitiert werden. Es sind zudem die frühesten Beiträge. In [Robinson 76] geht es zum einen um die Übertragung eines Verfahrens zur Kantendetektion durch Schablonenvergleich auf Farbbilder, zum anderen um die Wahl eines für die Kantendetektion geeigneten Farbkoordinatensystems.

Robinson schlägt folgende Möglichkeiten zur Detektion von Kanten in Farbbildern vor:

- Beschränkung auf eine Komponente und Anwendung eines Intensitätsverfahrens. Es sollte die Komponente mit der größten Energie gewählt werden.
- Dekorrelation der Komponenten durch eine Karhunen-Loève-Transformation und separate Behandlung der entstehenden Komponenten.
- Separate Behandlung der Komponenten und Kombination der Ergebnisse. Bei der Kombination sollte die Korrelation der Bildkanäle berücksichtigt werden.

Betont wird besonders die große Bedeutung der Wahl eines Farbkoordinatensystems. Es werden verschiedene Farbsysteme vorgestellt, die sich als lineare oder nichtlineare Transformationen aus dem RGB-Raum ergeben. Anhand eines Beispielbildes werden

die Kovarianzen sowie der Energiegehalt der einzelnen Kanäle für die verschiedenen Systeme in Tabellen gegenübergestellt.

Die Detektion von Kanten in Farbbildern geschieht durch Anwendung des Verfahrens von Kirsch auf jede Bildkomponente. Registriert wird für jeden Bildpunkt des Farbbildes als Kantenkontrast das Maximum der drei in den einzelnen Komponenten bestimmten Kontrastwerte sowie der Index der entsprechenden Maske als Hinweis auf die Kantenrichtung. Es wird also die deutlichste der drei möglichen Kanten registriert.

Um einen besseren quantitativen Vergleich zwischen den verschiedenen Farbsystemen in Bezug auf das Kantendetektionsproblem zu ermöglichen, wird ein *edge activity index* (EAI) eingeführt. Der EAI wird für jeden Bildpunkt und jede Bildkomponente definiert als das Verhältnis des maximalen Faltungsergebnisses zum mittleren Faltungsergebnis bei Faltung mit den verschiedenen Kirsch-Masken. Um eine Aussage für das gesamte Bild zu erhalten, wird der EAI für jede Komponente über alle Kantenpunkte gemittelt.

Die Ergebnisse des Kantendetektionsalgorithmus sowie ein Vergleich der EAI-Werte werden für das oben erwähnte Beispielbild und die verschiedenen Farbsysteme präsentiert. Der Vergleich zeigt, daß ein niedriger Energiegehalt einer Komponente nicht gleichbedeutend ist mit einem niedrigen EAI-Wert. Aus dem Vergleich der Ergebnisbilder und der Tabellen zieht Robinson das Fazit:

A general conclusion that can be drawn is that edge extraction should be performed not in the RGB-space but in one of the other color coordinate systems studied in this paper.

Anmerkungen

Bei der Beurteilung dieses Beitrages ist zu berücksichtigen, daß er aus einer Zeit stammt, in der auch die Detektion von Kanten in Intensitätsbildern noch nicht zu befriedigenden Ergebnissen führte. Dies zeigt sich besonders an den präsentierten Kantenbildern.

Die Farbbildverarbeitung beschränkt sich auf die separate Behandlung der einzelnen Bildkomponenten und die einfache Überlagerung der Ergebnisse unter Berücksichtigung des jeweils stärksten Hinweises auf eine Kante.

Etwas ausführlicher fällt der Vergleich der verschiedenen Farbsysteme aus. Es werden jedoch keine Kriterien angegeben, nach denen die Eignung eines Farbkoordinatensystems als Grundlage zur Kantendetektion zu beurteilen wäre. Alle betrachteten Transformationen verlagern den größten Teil der Bildenergie (80-90%) in die erste Bildkomponente, die meist die Bildintensität darstellt. Mit Ausnahme der Karhunen-Loève-Transformation zeigen die Kovarianzen der Komponenten keine wesentlichen Unterschiede gegenüber dem RGB-Raum. Dies gilt auch für die EAI-Werte. Die gezeigten Ergebnisbilder lassen kaum Unterschiede erkennen und werden nicht eingehender diskutiert. Somit ist die Schlußfolgerung Robinsons — die Ablehnung des RGB-Raums als geeignete Repräsentation — weder im Text noch durch die Tabellen oder durch die resultierenden Kantenbilder motiviert.

6.3 Sethi 85

Das Problem der Detektion von Diskontinuitäten findet sich nicht nur bei der Analyse von Grauwert- oder Farbbildern, sondern auch bei der Verarbeitung von verschiedenen anderen bildhaften Datenstrukturen, wie z.B. Verschiebungsvektorfeldern, Vektorfeldern, die die Oberflächenorientierung wiedergeben, u.ä. Dies ist der Anlaß für Sethi 85, einen einfachen gradientenbasierten Ansatz (Prewitt) für die Anwendung auf vektorwertige Bildfunktionen zu verallgemeinern. Man betrachte dazu die n -dimensionale Bildfunktion²

$$\vec{C}(x, y) = (C_1(x, y), C_2(x, y), \dots, C_n(x, y))^T$$

In einer lokalen Umgebung des Bildvektors $\vec{C}(x_0, y_0)$ wird die Ähnlichkeit jedes benachbarten Bildvektors zum Vektor $\vec{C}(x_0, y_0)$ durch einen Skalar bewertet. Mit Hilfe der Prewitt-Masken läßt sich dann der Gradient der Ähnlichkeitswerte bestimmen. Kanten werden definiert als Orte im Bild, die einen starken Ähnlichkeitsgradienten aufweisen.

Sethi verwendet zur Bewertung der Ähnlichkeit $s(x, y)$ zwischen den Vektoren $\vec{C}(x, y)$ und $\vec{C}(x_0, y_0)$ das normierte innere Produkt der Vektoren:

$$\begin{aligned} s(x, y) &= \frac{\vec{C}(x, y) \cdot \vec{C}(x_0, y_0)}{\|\vec{C}(x, y)\| \cdot \|\vec{C}(x_0, y_0)\|} \\ &= \frac{\sum_{i=1}^n C_i(x, y) \cdot C_i(x_0, y_0)}{\|\vec{C}(x, y)\| \cdot \|\vec{C}(x_0, y_0)\|} \end{aligned}$$

Im folgenden stehe $\vec{U}(x, y)$ für den auf Eins normierten Bildvektor:

$$\vec{U}(x, y) = \frac{\vec{C}(x, y)}{\|\vec{C}(x, y)\|}$$

Wir betrachten nun am Beispiel der x -Komponente Δ_x , wie sich die Gradientenkomponenten durch Faltung mit den Prewitt-Masken (siehe Abb. 4.2) aus den Ähnlichkeitswerten ergeben:

$$\begin{aligned} \Delta_x &= \sum_{k=-1}^1 \left(\sum_{i=1}^n U_i(x_0 + 1, y_0 + k) \cdot U_i(x_0, y_0) \right) - \\ &\quad \left(\sum_{i=1}^n U_i(x_0 - 1, y_0 + k) \cdot U_i(x_0, y_0) \right) \\ &= \sum_{k=-1}^1 \sum_{i=1}^n U_i(x_0, y_0) \cdot (U_i(x_0 + 1, y_0 + k) - U_i(x_0 - 1, y_0 + k)) \\ &= \sum_{i=1}^n \left(U_i(x_0, y_0) \cdot \sum_{k=-1}^1 (U_i(x_0 + 1, y_0 + k) - U_i(x_0 - 1, y_0 + k)) \right) \end{aligned}$$

²Das hochgesetzte T steht für die Transposition eines Vektors oder einer Matrix.

Die x -Komponente des Gradienten ergibt sich aus der Summe der mit der jeweiligen Vektorkomponente gewichteten x -Komponenten der Einzelgradienten. Entsprechendes gilt für Δ_y . Für den Ähnlichkeitsgradienten ∇s erhält man somit:

$$\nabla s(x_0, y_0) = \sum_{i=1}^n U_i(x_0, y_0) \cdot \nabla U_i(x_0, y_0)$$

Für den speziellen Fall einer Farbbildfunktion mit den Bildkomponenten R,G,B ergibt sich:

$$\nabla s = R \cdot \nabla R + G \cdot \nabla G + B \cdot \nabla B$$

Es sei daran erinnert, daß die Farbvektoren des Bildes zuvor auf die Länge Eins normiert werden! Die Gradienten der normierten Bildkomponenten können deutlich verschieden sein von den Gradienten der unnormierten Komponenten, da jeder Bildvektor mit einem eigenen Faktor skaliert wird.

Anmerkungen

Die Verwendung des inneren Produktes der Bildvektoren als Ähnlichkeitsmaß zur Detektion von Diskontinuitäten kann nur in besonderen Fällen sinnvoll sein. Das unnormierte innere Produkt ist untauglich, da hier das Produkt der Längen der betrachteten Vektoren, nicht aber die Differenz der Längen, einen starken Einfluß auf die Ähnlichkeit zweier Vektoren hat. Dieser störende Einfluß läßt sich durch die Normierung vermeiden. Damit reduziert sich das Maß der Ähnlichkeit aber auch auf die Größe des Winkels zwischen beiden Vektoren. Ihre Längen und auch der Unterschied ihrer Längen spielen keine Rolle.

Für Farbbildvektoren im RGB-Raum würde dies bedeuten: Kanten, die sich nur durch eine Differenz in ihrer Länge — also in der Intensität —, nicht aber durch einen Farbunterschied bemerkbar machen, ließen sich nicht detektieren, wie groß der Intensitätsunterschied auch sein mag.

Man muß im allgemeinen davon ausgehen, daß sowohl die Richtungs- als auch die Längendifferenz der Vektoren bei der Suche nach Diskontinuitäten im Bild zu berücksichtigen sind, obwohl Fälle denkbar sind, in denen die ausschließliche Betrachtung der Richtungsunterschiede zur Erkennung der interessierenden Merkmale ausreicht.

6.4 Shiozaki 86

Shiozaki 86 stellt einen nichtlinearen, lokalen Operator zur Detektion von Kanten in Intensitätsbildern vor, sowie einen Ansatz zur Übertragung dieses Verfahrens auf Farbbilder.

Der Operator mißt die *Entropie* der Intensität in einem Bildausschnitt. Die Intensitäts-Entropie H wird wie folgt definiert:

$$H = -\frac{1}{\log(n+1)} \cdot \sum_{i=0}^n p_i \cdot \log p_i \quad \text{mit} \quad p_i = \frac{I_i}{\sum_{j=0}^n I_j}$$

wobei die I_i ($i = 0 \dots n$) den Intensitätswerten der $n + 1$ Bildpunkte im betrachteten Bildausschnitt entsprechen. Die so definierte Entropie steht im Zusammenhang mit der Varianz der Grauwerte im Bildausschnitt:

- Große Entropie bedeutet große Homogenität der Grauwerte.
- Kleine Entropie deutet auf eine große Varianz der Grauwerte hin.

Kanten im Bild sollten sich also, aufgrund des starken Grauwertkontrastes, durch niedrige Entropiewerte auszeichnen. Im Gegensatz zu anderen Kantendetektoren ist der hier vorgestellte Operator jedoch nicht nur vom Kontrast, sondern auch von der mittleren Helligkeit in der lokalen Umgebung abhängig (der Operator reagiert empfindlicher im Bereich niedriger Intensität). Er ist rotationsinvariant und — laut Autor — in mit dem Laplaceoperator vergleichbarer Weise rauschanfällig.

Die Farbentropie H läßt sich definieren als gewichtete Summe der Grauwertentropien H_R, H_G, H_B , die in den drei Farbkanälen separat gemessen werden:

$$H = q_R \cdot H_R + q_G \cdot H_G + q_B \cdot H_B$$

$$q_R = \frac{R}{R + G + B} \quad \text{entsprechend für } q_G, q_B$$

wobei R, G, B den drei Farbwerten des betrachteten Farbvektors \vec{C} entsprechen.

Die Ergebnisse des Verfahrens werden anhand verschiedener Bildvorlagen, im Vergleich zu anderen einfachen Kantendetektoren sowie im Vergleich zwischen Intensitäts- und Farbbildern präsentiert.

Anmerkungen

Die von Shiozaki definierte Entropie berücksichtigt nur die Varianz der Grauwerte in einer lokalen Umgebung. Die geometrischen Eigenschaften von Kanten bleiben ungenutzt: die räumliche Verteilung der Grauwerte im betrachteten Ausschnitt spielt für die Höhe der Entropie keine Rolle. Zudem ist die Berechnung der Entropie, verglichen mit anderen einfachen lokalen Operatoren rechenintensiv. Die Abhängigkeit der Operatorempfindlichkeit von der mittleren Bildintensität ist im allgemeinen nicht wünschenswert.

Auch bei Shiozaki findet die Verarbeitung von Farbinformation durch separate Behandlung der Farbkanäle und Kombination der Einzelmessungen statt. Dies gestaltet sich besonders einfach, da der Entropieoperator nur eine skalare Größe und keine Richtungsinformation liefert. Die Wichtung mit den normierten Farbwerten wird nicht begründet. Weder für den Intensitäts- noch für den Farbansatz wird die Wirkungsweise des Operators näher untersucht. Auch die Unterschiede zwischen den verschiedenen Ergebnisbildern werden nicht ausreichend diskutiert.

6.5 Yakimovsky 76, Schröder 87

Yakimovsky 76 nutzt Methoden der Statistik zur Detektion von Kanten in Intensitätsbildern. Dazu legt er folgendes Stufenkantenmodell zugrunde:

Zu einem möglichen Kantenelement K seien zwei disjunkte Umgebungen N_1 und N_2 auf beiden Seiten von K definiert. Die Bildfunktion über N_1 und entsprechend N_2 wird als konstant angenommen, überlagert von additivem, normalverteiltem Rauschen. Die Verteilung der Grauwerte in N_1 bzw. N_2 kann somit durch eine Normalverteilung $N(\mu_1, \sigma_1)$ bzw. $N(\mu_2, \sigma_2)$ beschrieben werden.

Stellt K tatsächlich eine Kante im Bild dar, so sollten sich die beiden Verteilungen voneinander unterscheiden. Stammen die Grauwerte beider Umgebungen jedoch aus einer homogenen Bildregion, so sollte sich ihre Verteilung gut durch eine gemeinsame Verteilung $N(\mu_0, \sigma_0)$ beschreiben lassen. Die Entscheidung, ob K als Kantenelement akzeptiert wird oder nicht, läßt sich treffen durch Abwägen der Hypothesen H_0 und H_1 :

H_0 : Die Grauwerte aus N_1 und N_2 stammen aus *einer* Normalverteilung $N(\mu_0, \sigma_0)$.

H_1 : Die Grauwerte aus N_1 und N_2 stammen aus *zwei* verschiedenen Normalverteilungen $N(\mu_1, \sigma_1)$ und $N(\mu_2, \sigma_2)$.

In der Statistik wird als Entscheidungsmaß zwischen den beiden Hypothesen das Verhältnis ihrer Wahrscheinlichkeiten, das sogenannte *Likelihood-Verhältnis*, verwendet. Ist dieses Verhältnis größer als ein Schwellwert t , so wird die Hypothese H_0 zurückgewiesen und K wird als Kantenelement akzeptiert.

Um das Likelihood-Verhältnis angeben zu können, müssen zunächst die Verteilungsparameter $\mu_0, \mu_1, \mu_2, \sigma_0, \sigma_1, \sigma_2$ geschätzt werden. Die Maximum-Likelihood-Schätzungen für den Mittelwert und die Varianz einer Normalverteilung sind (die x_i seien die zur Schätzung verwendeten Grauwerte):

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2$$

Die Wahrscheinlichkeitsdichte dafür, daß x_i einer Verteilung $N(\hat{\mu}, \hat{\sigma})$ entstammt, ist:

$$p_{\hat{\mu}, \hat{\sigma}}(x_i) = \frac{1}{\sqrt{2\pi\hat{\sigma}}} \cdot \exp\left(-\frac{(x_i - \hat{\mu})^2}{2\hat{\sigma}^2}\right)$$

Sind die Beobachtungen x_i unabhängig voneinander, so ist die Wahrscheinlichkeitsdichte für die Gesamtprobe $x_1 \dots x_n$ das Produkt der einzelnen Dichten:

$$\begin{aligned} p_{\hat{\mu}, \hat{\sigma}}(x_1 \dots x_n) &= \prod_{i=1}^n p_{\hat{\mu}, \hat{\sigma}}(x_i) \\ &= \left(\frac{1}{\sqrt{2\pi\hat{\sigma}}}\right)^n \cdot \exp\left(-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n (x_i - \hat{\mu})^2\right) \\ &= \left(\frac{1}{\sqrt{2\pi\hat{\sigma}}}\right)^n \cdot \exp\left(-\frac{n}{2}\right) \end{aligned}$$

Seien die $x_1 \dots x_n$ die Grauwerte aus N_1 und $y_1 \dots y_m$ die Grauwerte aus N_2 , so erhalten wir als Likelihood-Verhältnis:

$$\frac{p_{\hat{\mu}_1, \hat{\sigma}_1}(x_1 \dots x_n) \cdot p_{\hat{\mu}_2, \hat{\sigma}_2}(y_1 \dots y_m)}{p_{\hat{\mu}_0, \hat{\sigma}_0}(x_1 \dots x_n, y_1 \dots y_m)} = \frac{\hat{\sigma}_0^{n+m}}{\hat{\sigma}_1^n \cdot \hat{\sigma}_2^m}$$

Schröder 87 hat diesen Ansatz auf vektorwertige Bildfunktionen und somit für multivariate Verteilungen verallgemeinert. Aus den Mittelwerten werden dann Mittelwertvektoren, aus den Varianzen werden Kovarianzmatrizen. Als Likelihood-Verhältnis erhält Schröder das Verhältnis der Determinanten der Kovarianzmatrizen $\hat{\Sigma}_0, \hat{\Sigma}_1, \hat{\Sigma}_2$:

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

Obwohl Schröders Ziel die Entwicklung eines statistischen Ansatzes zur Bereichssegmentierung von Farbbildern war, wurden von ihm auch Experimente zur Detektion von Farbkanten mit Hilfe des von ihm erweiterten Likelihood-Verhältnis-Testes durchgeführt. Auf die Präsentation von Ergebnisbildern verzichten wir an dieser Stelle jedoch. Ein fairer Vergleich zu den später gezeigten Resultaten wäre nur bei zeitaufwendiger Implementation eines vollständigen Detektionsverfahrens möglich.

Anmerkungen

Der vorgestellte Ansatz von Yakimovsky zeichnet sich gegenüber vielen anderen durch formale Fundiertheit und explizite Beschreibung des zugrundeliegenden Modells aus. Wie Schröder zeigt, läßt sich der Ansatz direkt auf vektorwertige Bildfunktionen, also auch auf Farbbilder, erweitern. Das so erweiterte Verfahren wertet die Farbinformation des Bildes in direkter Weise aus. Es handelt sich hier also um ein echtes Farbbildverfahren.

Zusammenfassend läßt sich über die in diesem Kapitel vorgestellten Ansätze zur Kantendetektion in Farbbildern folgendes sagen:

- Die Integration der Informationen aus den einzelnen Farbkanälen geschieht, mit wenigen Ausnahmen, auf eine Weise, die einer formalen Fundiertheit und einer einsichtigen Motivation entbehrt.
- Die als Vorbild dienenden Grauwertverfahren stammen überwiegend aus den siebziger Jahren und können dem Stand der Kunst in der Kantendetektion nicht entsprechen.

Kapitel 7

Ein ableitungsorientierter Farbansatz

Angesichts der unbefriedigenden Resultate, die von den bisher vorgestellten Ansätzen zur Kantendetektion in Farbbildern zu erwarten sind, wollen wir uns nun fragen, ob es möglich ist, einen Farbansatz zu entwickeln, der die in Kapitel 5 dargestellten Aspekte aktueller ableitungsorientierter Grauwertverfahren berücksichtigt. Solche Verfahren basieren auf der Schätzung der Gradienteninformation. Der mathematische Gradientenoperator ist jedoch nur für skalare Funktionen definiert. Die zu lösende Frage lautet also: Gibt es ein multidimensionales Analogon zum Gradientenoperator?

Wie wir im folgenden sehen werden, läßt sich auch für vektorwertige Bildfunktionen ein Operator entwickeln, der die Richtung und Stärke der stärksten Änderung der Bildfunktion ermittelt. Präsentationen der formalen Lösung finden sich in [Di Zenzo 86] und [Novak + Shafer 87]. Während Novak und Shafer die Lösung in genialer Kürze angeben, weist der Bericht von Di Zenzo eine größere Ausführlichkeit, jedoch auch eine schwer nachvollziehbare Motivation und eine sehr verwirrende Notation auf. Wir wollen deshalb im nun folgenden Abschnitt den Lösungsweg ausführlich beschreiben.

7.1 Die formale Lösung

Unser Ziel ist es, mit Hilfe eines lokalen Operators die Richtung zu bestimmen, in der sich die Farbbildfunktion am stärksten ändert, sowie einen Wert anzugeben, der die Größe der Änderung beschreibt. Zur Herleitung des Operators betrachten wir die Farbbildfunktion \vec{C} sowie den durch den Winkel φ zur positiven x -Achse festgelegten Richtungsvektor \vec{n} :

$$\vec{C}(x, y) = \begin{pmatrix} R(x, y) \\ G(x, y) \\ B(x, y) \end{pmatrix} \quad \vec{n} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$$

Die direktionale Ableitung der vektorwertigen Funktion \vec{C} in Richtung \vec{n} liefert einen *Ableitungsvektor*, dessen Komponenten die Richtungsableitungen der ursprünglichen Komponenten sind:

$$\frac{\partial \vec{C}}{\partial \vec{n}} = \begin{pmatrix} \frac{\partial R}{\partial \vec{n}} \\ \frac{\partial G}{\partial \vec{n}} \\ \frac{\partial B}{\partial \vec{n}} \end{pmatrix}$$

Durch Anwenden der bereits bekannten Beziehung zwischen Gradient und Richtungsableitung und Formulierung des Ergebnisses als Matrixprodukt erhalten wir:

$$\frac{\partial \vec{C}}{\partial \vec{n}} = \begin{pmatrix} \nabla R \cdot \vec{n} \\ \nabla G \cdot \vec{n} \\ \nabla B \cdot \vec{n} \end{pmatrix} = \begin{pmatrix} R_x & R_y \\ G_x & G_y \\ B_x & B_y \end{pmatrix} \cdot \vec{n} = J \cdot \vec{n}$$

Die durch die Richtungsableitungen der Komponenten nach x und y definierte Matrix J wird als *Jacobimatrix* (oder auch *Funktionalmatrix*) bezeichnet.

Um einen skalaren Wert für die Stärke der Änderung in Richtung \vec{n} zu erhalten, verwenden wir die euklidische Norm und betrachten die Länge des Ableitungsvektors als Maß für die Stärke der Änderung (l^2 sei die quadrierte Länge des Vektors):

$$l^2 = \|J \cdot \vec{n}\|^2 = (J \cdot \vec{n})^T \cdot (J \cdot \vec{n}) = \vec{n}^T \cdot (J^T J) \cdot \vec{n}$$

Für die symmetrische 2×2 -Matrix $J^T J$ seien die folgenden Koeffizienten definiert:

$$J^T J = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{mit}$$

$$a_{11} = R_x^2 + G_x^2 + B_x^2$$

$$a_{22} = R_y^2 + G_y^2 + B_y^2$$

$$a_{12} = R_x R_y + G_x G_y + B_x B_y$$

$$a_{21} = a_{12}$$

Wir erhalten für die (quadrierte) Stärke der Änderung in Richtung \vec{n} :

$$l^2 = a_{11} \cdot \cos^2 \varphi + 2a_{12} \cdot \sin \varphi \cos \varphi + a_{22} \cdot \sin^2 \varphi$$

Die Richtung stärkster Änderung der Bildfunktion \vec{C} entspricht nun der Richtung, für die l^2 ein Maximum aufweist. Wir leiten also die Gleichung für l^2 nach φ ab und bestimmen das Maximum durch Nullsetzen:

$$(a_{22} - a_{11}) \cdot \sin \varphi \cos \varphi + a_{12} \cdot (\cos^2 \varphi - \sin^2 \varphi) = 0$$

Hier wird eine Fallunterscheidung erforderlich:

- Ist $a_{12} = 0$ und $a_{11} = a_{22}$, erhalten wir keine Lösung für φ . Es gibt keine ausgezeichnete Richtung; die Änderung der Bildfunktion ist für alle Richtungen gleich.
- Ist $a_{12} = 0$ und $a_{11} \neq a_{22}$, so muß entweder $\sin \varphi = 0$ oder $\cos \varphi = 0$ gelten. Als Lösungen erhalten wir $\varphi_1 = 0^\circ$ und $\varphi_2 = 90^\circ$.
- Ist umgekehrt $a_{12} \neq 0$ und $a_{11} = a_{22}$, so muß gelten $\cos^2 \varphi = \sin^2 \varphi$. Die Lösungen sind dann $\varphi_1 = 45^\circ$ und $\varphi_2 = 135^\circ$.
- Im Normalfall wird gelten $a_{12} \neq 0$ und $a_{11} \neq a_{22}$, so daß wir für die Lösungswinkel φ_1 und φ_2 folgende Bedingung erhalten:

$$\tan \varphi_{1,2} = \frac{1}{2a_{12}} \cdot \left((a_{22} - a_{11}) \pm \sqrt{(a_{22} - a_{11})^2 + 4a_{12}^2} \right)$$

Oder einfacher:

$$\tan 2\varphi_{1,2} = \frac{2 \tan \varphi_{1,2}}{1 - \tan^2 \varphi_{1,2}} = \frac{2a_{12}}{a_{11} - a_{22}}$$

Wir erhalten jeweils zwei Lösungen, von denen eine mit der Richtung maximaler Änderung und die zweite mit der Richtung minimaler Änderung korrespondiert. Um zu prüfen, welche Lösung dem Maximum entspricht, müßten beide Lösungen in die Betragsgleichung für l^2 eingesetzt werden.

Es ergibt sich ein zweiter eleganterer Lösungsweg, wenn wir unser Problem als *Eigenwertproblem* betrachten. Dem mathematisch versierten Betrachter fällt auf, daß der Term $\vec{n}^T \cdot (J^T J) \cdot \vec{n}$ dem *Rayleigh-Quotienten* der Matrix $J^T J$ entspricht. Der Rayleigh-Quotient $R(\vec{x})$ einer Matrix A ist definiert als:

$$R(\vec{x}) = \frac{\vec{x}^T \cdot A \cdot \vec{x}}{\vec{x}^T \cdot \vec{x}}$$

Wir sind nun an den Extremwerten von $R(\vec{x})$ interessiert. Wir leiten R deshalb nach \vec{x} ab und setzen gleich Null:¹

$$\begin{aligned} \frac{dR(\vec{x})}{d\vec{x}} &= 0 \\ \Leftrightarrow \frac{1}{\vec{x}^T \vec{x}} \cdot 2A\vec{x} - \frac{\vec{x}^T A \vec{x}}{\vec{x}^T \vec{x}} \cdot 2\vec{x} &= \frac{2}{\vec{x}^T \vec{x}} \cdot (A\vec{x} - R(\vec{x})\vec{x}) = 0 \\ \Leftrightarrow A\vec{x} &= R(\vec{x})\vec{x} \end{aligned}$$

Durch eine Gleichung der Form $A\vec{x} = \lambda\vec{x}$ sind aber gerade die Eigenvektoren der Matrix A definiert: Die Vektoren \vec{x} , zu denen es einen Wert λ gibt, so daß die Gleichung erfüllt ist, heißen *Eigenvektoren* der Matrix A . Die λ -Werte heißen *Eigenwerte*. Die Extremwerte des Rayleigh-Quotienten stimmen demnach mit den Eigenwerten der Matrix A überein.

¹Es handelt sich diesmal — trotz der Ähnlichkeit in der Notation — um die Ableitung nach einem Vektor und nicht um eine Richtungsableitung.

Somit sind die Eigenwerte der Matrix $J^T J$ zugleich die Extremwerte der Gleichung für l^2 . Die Richtung maximaler Änderung entspricht der Richtung des zum betragsmäßig größten Eigenwert gehörenden Eigenvektors.

Wir suchen nun nach den Lösungen der Eigenwertgleichung (A sei eine $n \times n$ -Matrix, \vec{x} ein n -dimensionaler Vektor und E die $n \times n$ -Einheitsmatrix):

$$A\vec{x} = \lambda\vec{x} \Leftrightarrow A\vec{x} - \lambda\vec{x} = 0 \Leftrightarrow (A - \lambda E) \cdot \vec{x} = 0$$

Eine nichttriviale Lösung für \vec{x} existiert, wenn

$$\det(A - \lambda E) = 0$$

Setzen wir für A die uns interessierende Matrix $J^T J$ ein, erhalten wir als Extremwerte:

$$\lambda_{1,2} = \frac{1}{2} \cdot \left((a_{11} + a_{22}) \pm \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2} \right)$$

Da a_{11} und a_{22} als Summen von Quadraten stets positiv sein müssen, entspricht die Lösung mit der positiven Wurzel dem Betragsmaximum. Den zu λ_{max} gehörenden Eigenvektor erhalten wir durch Einsetzen in die Eigenwertgleichung ($t \in \mathcal{R}$):

$$(A - \lambda_{max} E) \cdot \vec{x} = 0 \Leftrightarrow \vec{x}_{max} = t \cdot \begin{pmatrix} \frac{a_{12}}{\lambda_{max} - a_{11}} \\ 1 \end{pmatrix}$$

Als Richtung maximaler Änderung erhalten wir:

$$\varphi_{max} = \arctan \left(\frac{\lambda_{max} - a_{11}}{a_{12}} \right)$$

Basierend auf der Schätzung der Richtungsableitungen der einzelnen Farbkanäle nach x und y ($R_x, R_y, G_x, G_y, B_x, B_y$), lassen sich, wie die hergeleitete Lösung zeigt, Richtung und Betrag der stärksten Änderung der Farbbildfunktion bestimmen. Wir wollen diesen Ansatz im folgenden als *Farbvektoransatz* bezeichnen. Die Integration der Informationen aus den einzelnen Kanälen geschieht auf formal fundierte Weise unter Berücksichtigung der Vektorwertigkeit der Bildfunktion. Wir besitzen damit ein Farbanalogon zum Gradientenoperator.

Durch Detektion lokaler Extrema im Änderungsbetrag können wir wie im Grauwertverfahren Kantenelemente detektieren und ihnen Richtungs- und Kontrastwerte zuordnen. Kombinieren wir die Ableitungsoperationen zur Schätzung der Gradientenkomponenten mit einer Mittelungsoperation unter Verwendung einer Gaußfunktion, so finden im Farbvektoransatz auch die in Kapitel 5 dargestellten Aspekte Berücksichtigung.

7.2 Veranschaulichung

Welche Richtung von dem Farbvektoransatz als Richtung stärkster Änderung der Farbbildfunktion ermittelt wird, läßt sich an Vektordiagrammen veranschaulichen. Die (qua-

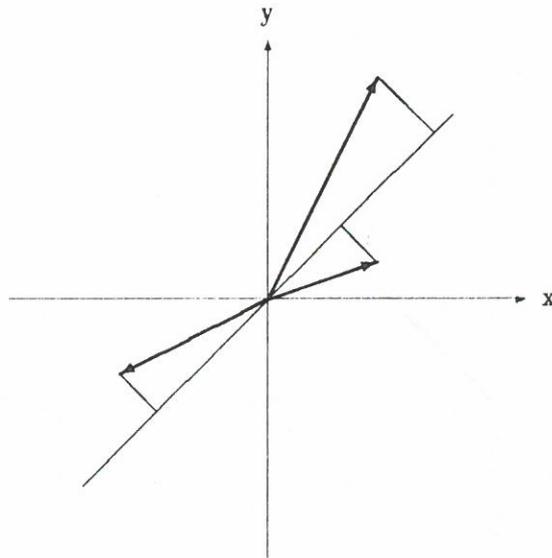


Abbildung 7.1: Projektion der Einzelgradienten auf eine Gerade

drierte) Größe der Änderung in eine bestimmte Richtung \vec{n} entspricht, gemäß der verwendeten euklidischen Norm, der Summe der quadrierten Komponenten des Ableitungsvektors:

$$l^2 = \|\mathbf{J} \cdot \vec{n}\|^2 = (\nabla R \cdot \vec{n})^2 + (\nabla G \cdot \vec{n})^2 + (\nabla B \cdot \vec{n})^2$$

Die Einzelkomponenten wiederum entsprechen den Projektionen der Einzelgradienten² auf die Richtung \vec{n} . Abbildung 7.1 macht dies deutlich. Wir sehen dort die drei Einzelgradienten sowie ihre Projektionen auf eine Gerade, deren Richtung durch \vec{n} definiert sei. Die Richtung stärkster Änderung entspricht — gemäß unserer Definition — der Richtung, für die die Summe der quadrierten Projektionen maximal ist. Die durch den Farbvektoransatz gefundene Lösung stellt somit eine Art Mittelung von Richtung und Betrag der Einzelgradienten dar. Abbildung 7.2 zeigt verschiedene Beispiele.³ Wir sehen die drei Einzelgradienten, sowie einen hervorgehobenen Vektor, der dem jeweils resultierenden Ergebnisvektor entspricht. Die Beispiele zeigen den Fall eines einzigen dominanten Einzelgradienten (a), den Fall von zwei gut antikorrelierten, starken Einzelgradienten (b), sowie den Fall zweier unkorrelierter, starker Einzelgradienten (c). Die Abbildungen machen folgendes deutlich:

- Die Mittelung der Einzelgradienten setzt voraus, daß die zu beobachtende Diskontinuität signifikanter wird, wenn sich die Richtungen der Einzelgradienten ähnlicher werden.

²Wir bezeichnen im folgenden als *Einzelgradienten* die Gradienten der einzelnen Bildkanäle.

³Um die Vergleichbarkeit zwischen Farb- und Intensitätsverfahren zu gewährleisten, weisen wir in unseren Experimenten den Ergebnisvektoren die Länge $\sqrt{\frac{1}{3}l^2}$ zu.

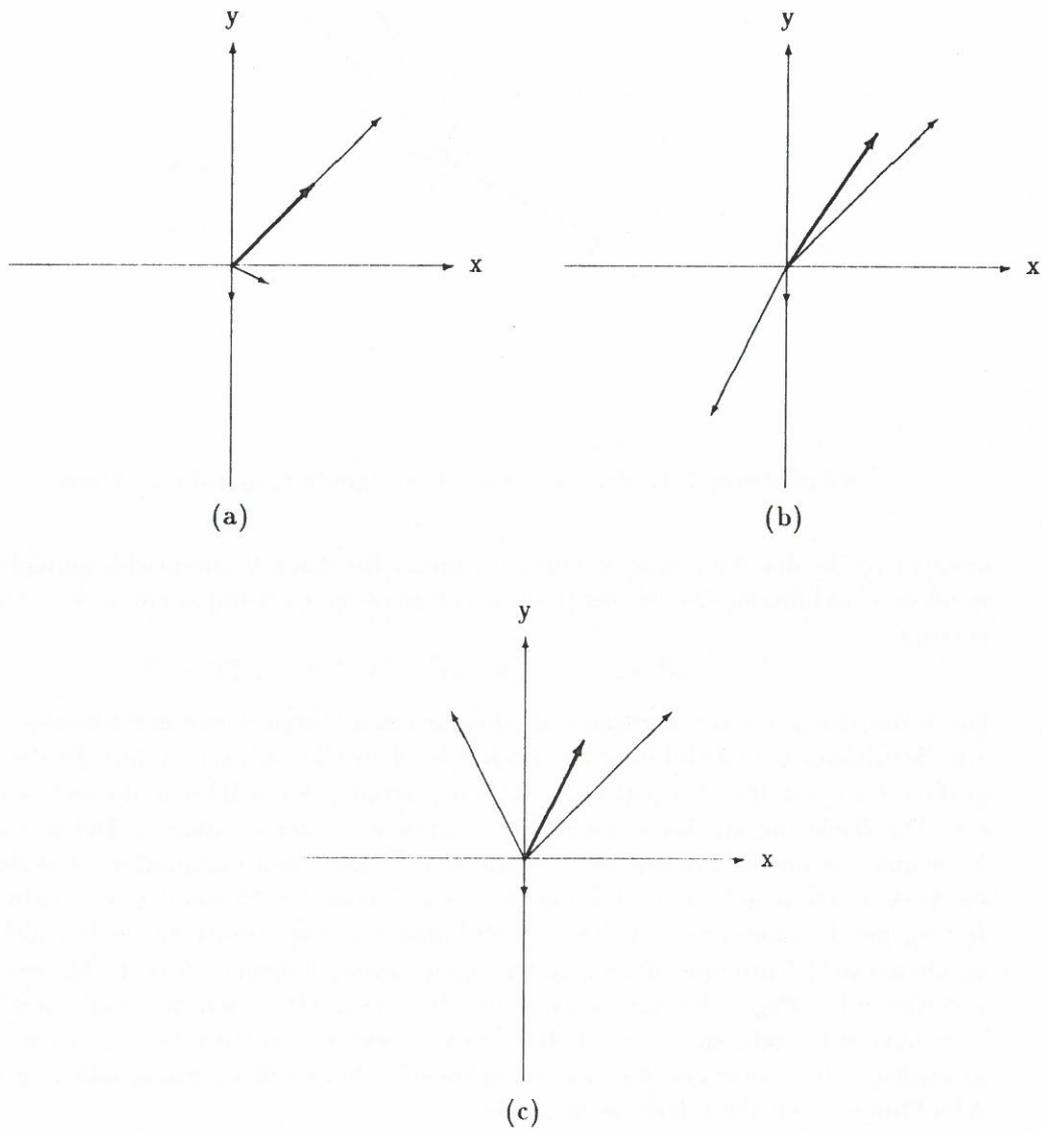


Abbildung 7.2: Ergebnisse des Farbvektoransatzes

- Winkeldifferenzen von $\pm 180^\circ$ spielen für das Ergebnis keine Rolle. Einzelgradienten unterstützen sich gegenseitig, wenn ihre Richtungen stark korreliert oder antikorreliert sind.

Beide Beobachtungen entsprechen unseren Erwartungen. Wenn wir nach Diskontinuitäten im Bild suchen, erwarten wir, daß stark korrelierte Effekte in den einzelnen Kanälen den Gesamteffekt verstärken. Lassen sich jedoch mehrere starke Einzelgradienten unterschiedlicher Richtungen beobachten, kann man wohl kaum eine Stufenkante im Bild als Ursache vermuten. Schwierigkeiten sind zu erwarten, wenn sich eine kontrastarme Kante in einem Kanal größtmäßig kaum abhebt von den rauschbedingten Effekten in den anderen Kanälen. Die Mittelung wird dann zu einer Verfälschung der Richtungsinformation führen.

Der gleich starke Einfluß korrelierter und antikorrelierter Einzelgradienten resultiert aus der Quadrierung der Projektionslängen. Sie bewirkt, daß die Vorzeichen der Projektionen auf den Gesamteffekt keinen Einfluß haben. Dies erlaubt gerade die Detektion von Kanten, die sich allein aufgrund der Intensitätsinformation nicht detektieren ließen. Betrachten wir beispielsweise eine vertikale Kante, die sich in horizontaler Richtung durch eine starke Zunahme im Rotkanal und eine ebenso starke Abnahme im Blaukanal auszeichnet. Die exakte Antikorrelation der Effekte führt dazu, daß keine Intensitätsänderung zu beobachten ist. Im Farbansatz verstärken sich die Einzeleffekte jedoch gegenseitig und erzeugen eine leicht zu detektierende Diskontinuität.

7.3 Verallgemeinerung des Ansatzes

Obwohl wir in unseren Betrachtungen stets von dem Problem ausgegangen sind, Kanten in Farbbildern zu detektieren, sieht man leicht, daß der hier vorgestellte Lösungsansatz sowohl unabhängig ist von der Dimension der Bildvektoren und des Definitionsbereiches der Bildfunktion als auch von der Wahl eines speziellen Systems von Bildkanälen.

Unabhängigkeit von der Dimension

Die Unabhängigkeit von der Dimension der Bildvektoren ist leicht formal zu zeigen. Unsere Bildfunktion \vec{C} sei nun m -dimensional und bestehe aus den Komponenten C_1 bis C_m . Für die Richtungsableitung nach \vec{n} erhalten wir:

$$\frac{\partial \vec{C}}{\partial \vec{n}} = \begin{pmatrix} \frac{\partial C_1}{\partial \vec{n}} \\ \vdots \\ \frac{\partial C_m}{\partial \vec{n}} \end{pmatrix} = \begin{pmatrix} \nabla C_1 \cdot \vec{n} \\ \vdots \\ \nabla C_m \cdot \vec{n} \end{pmatrix} = \begin{pmatrix} C_{1x} & C_{1y} \\ \vdots & \vdots \\ C_{mx} & C_{my} \end{pmatrix} \cdot \vec{n}$$

Die Jacobimatrix besitzt nun m Zeilen. Die Größe der Matrix $J^T J$ bleibt jedoch unverändert. Wir erhalten für die Koeffizienten der Matrix:

$$a_{11} = C_{1x}^2 + C_{2x}^2 + \dots + C_{mx}^2$$

$$\begin{aligned}
 a_{22} &= C_{1_y}^2 + C_{2_y}^2 + \dots + C_{m_y}^2 \\
 a_{12} &= C_{1_x} C_{1_y} + C_{2_x} C_{2_y} + \dots + C_{m_x} C_{m_y} \\
 a_{21} &= a_{12}
 \end{aligned}$$

An der Lösung für Richtung und Betrag in Abhängigkeit von den Koeffizienten ändert sich durch die neue Dimensionalität nichts! Interessant ist es auch, den Spezialfall $n = 1$ zu betrachten. Wir bezeichnen die einzige Komponente der Bildfunktion \vec{C} mit I :

$$\frac{\partial \vec{C}}{\partial \vec{n}} = \left(\frac{\partial I}{\partial \vec{n}} \right) = (\nabla I \cdot \vec{n}) = \begin{pmatrix} I_x & I_y \end{pmatrix} \cdot \vec{n}$$

Die Richtung maximaler Änderung entspricht der Richtung des Gradienten von I . Dies zeigen auch die resultierenden Ergebnisse für Richtung und Betrag:

$$\lambda_{max} = I_x^2 + I_y^2 \quad \text{und} \quad \tan \varphi = \frac{I_y}{I_x}$$

Wie die Wahl der Bezeichnung I bereits nahelegte, sehen wir sehr schön, daß der Gradientenansatz auf einem Intensitätsbild als Spezialisierung des beliebigdimensionalen Lösungsansatzes betrachtet werden kann.

Wir betrachteten bisher die Dimension der Bildvektoren, also des Wertebereiches der Bildfunktion. Auf ebenso einfache Weise läßt sich die Unabhängigkeit des beschriebenen Lösungsansatzes von der Dimension des Definitionsbereiches zeigen. Der Definitionsbereich sei im folgenden n -dimensional. Für die Matrix $J^T J$ erhalten wir somit eine symmetrische $n \times n$ -Matrix. Während sich für das Eigenwertproblem im Falle einer 2×2 -Matrix eine sehr einfache analytische Lösung angeben läßt, wird man bei größeren Matrizen im allgemeinen auf numerische Lösungsverfahren zur Eigenwertbestimmung zurückgreifen müssen.

Unabhängigkeit vom System

Obwohl der Farbvektoransatz nicht allein für Darstellungen von Bildern im RGB-System gültig ist, sind bei der Übertragung auf andere Bildsysteme zwei wichtige Voraussetzungen zu berücksichtigen:

- Gleichgroße Differenzen in den verschiedenen Bildkanälen müssen für den zu detektierenden Gesamteffekt eine vergleichbare Wirkung besitzen.
- Richtungen und Beträge der Einzelgradienten werden gemittelt. Dies setzt voraus, daß die (Anti-) Korrelation der Einzelgradienten den Gesamteffekt verstärkt.

Das Problem der Vergleichbarkeit ergibt sich immer, wenn mehrere gemessene Merkmale zu einem Merkmalsraum zusammengefaßt werden. Entsprechende Probleme ergeben sich z.B. bei Klassifikationsansätzen, die eine Ballungsanalyse in einem mehrdimensionalen Merkmalsraum durchführen. Die Lösung dieser Probleme besteht in einer entsprechenden Gewichtung und Normierung der einzelnen Kanäle.

Auf den Mittelungseffekt sind wir bereits eingegangen. In den meisten Fällen wird die Annahme, daß sich die Einzeleffekte gemäß ihrer Korrelation verstärken, gut erfüllt sein. Eine Dekorrelation der Bildkanäle, etwa durch eine Karhunen-Loève-Transformation, ist aber natürlich nicht sinnvoll. Wenn man davon ausgehen muß, daß die in den einzelnen Kanälen zu beobachtenden Effekte überwiegend voneinander unabhängig sind, stellt sich die Frage, ob eine gemeinsame Behandlung der Kanäle überhaupt sinnvoll ist.

Der vorgestellte Lösungsansatz dient, wie wir gesehen haben, nicht nur der Detektion von Kanten in Farbbildern, sondern kann — unter Beachtung der beschriebenen Voraussetzungen — zur Detektion von Diskontinuitäten in beliebigdimensionalen Bildfunktionen verwendet werden. Denkbare Anwendungen (die noch nicht näher von uns untersucht werden konnten) sind beispielsweise:

- Detektion von Kanten in mehrkanaligen medizinischen Bildern.
- Separierung bewegter Objekte in dichten Verschiebungsvektorfeldern.
- Gewinnung von Hinweisen auf markante Punkte durch Detektion von Diskontinuitäten in Gradientenvektorfeldern.

7.4 Wahl anderer Normen

Die Wahl der euklidischen Norm als Maß für die Stärke der Änderung ist zwar nahelegend, jedoch nicht zwingend. Sie führt zu einer Lösung, die gut mit unseren Erwartungen an einen Farbansatz übereinstimmt, die jedoch im praktischen Einsatz viel Rechenaufwand erfordert. Wir wollen deshalb prüfen, ob sich durch Wahl einer anderen Norm eine ebenso akzeptable, aber effizientere Lösung ergibt. Wir betrachten als Alternativen die Berechnung der Summe der Komponenten des Ableitungsvektors,⁴ sowie die Maximumsnorm, bei der die Größe der maximalen Komponente als Maß für die Stärke der Änderung dient. Denkbar wäre außerdem die Absolutnorm, bei der die Summe der Absolutbeträge der Komponenten gebildet wird. Diese Norm läßt sich jedoch schlecht analytisch behandeln. Zudem sollten sich zur euklidischen Norm nur geringe Unterschiede ergeben, da auch bei der Absolutnorm das Vorzeichen der Projektionen unterdrückt wird. Der Unterschied besteht im wesentlichen in der unterschiedlichen Wichtung kleiner und großer Differenzen durch die Quadrierung.

Summe der Komponenten

Wie leicht zu zeigen ist, führt die Verwendung der Summe der Komponenten des Ableitungsvektors zu einem Ansatz, der mit dem Intensitätsansatz übereinstimmt:

$$I = \nabla R \cdot \vec{n} + \nabla G \cdot \vec{n} + \nabla B \cdot \vec{n}$$

⁴Die Berechnung der Summe dient hier der Abbildung des Ableitungsvektors auf einen Skalar. Diese Abbildung erfüllt jedoch nicht die mathematischen Bedingungen einer Norm.

$$\begin{aligned}
 &= (\nabla R + \nabla G + \nabla B) \cdot \vec{n} \\
 &= \nabla(R + G + B) \cdot \vec{n}
 \end{aligned}$$

l wird maximal, wenn \vec{n} mit der Richtung des Intensitätsgradienten übereinstimmt, womit die Äquivalenz zum Intensitätsansatz hergestellt ist.⁵ Der Nachteil des Intensitätsansatzes bei der Verarbeitung von Farbbildern wurde bereits in Abschnitt 7.2 dargestellt.

Maximumsnorm

Die Maximumsnorm weist l die betragsmäßig maximale Komponente des Ableitungsvektors zu:

$$l = \max (|\nabla R \cdot \vec{n}|, |\nabla G \cdot \vec{n}|, |\nabla B \cdot \vec{n}|)$$

Betrachtet man jede Komponente für sich, so erhält man ein Maximum, wenn \vec{n} mit der Richtung des jeweiligen Einzelgradienten übereinstimmt. Als Kandidaten für das Gesamtmaximum kommen also ∇R , ∇G und ∇B in Frage. Von diesen drei Einzelgradienten wird durch die Maximumsnorm derjenige ausgewählt, der den stärksten Betrag aufweist. In einem praktischen Verfahren müssen also nur die drei Einzelgradienten bestimmt und derjenige mit dem größten Betrag ausgewählt werden.

Bei diesem Ansatz kommt es nicht mehr zu einer Mittelung der Einzelgradienten. Der dominante Einzelgradient wird jedoch in vielen Fällen ein guter Repräsentant für die Verhältnisse im betrachteten Bildausschnitt sein. Zu beachten ist, daß die Korrelation bzw. Nichtkorrelation der Einzelgradienten keinen Einfluß mehr auf das Resultat hat! Ob nur *ein* starker Einzelgradient vorhanden ist, oder aber mehrere starke Gradienten durch gute Korrelation auf einen gemeinsamen Effekt hindeuten, spielt keine Rolle. Abbildung 7.3 macht dies deutlich: Der dargestellte Gradient sei der durch die Maximumsnorm ausgewählte; die beiden anderen Einzelgradienten können nun beliebig innerhalb des Kreises liegen, ohne daß ihre Lage Einfluß auf das Ergebnis hat. Im Extremfall könnten die beiden anderen Einzelgradienten gerade senkrecht zum maximalen Gradienten stehen und dabei jeweils nur eine etwas geringere Länge als dieser aufweisen.

7.5 Verschiedene Integrationsebenen

Will man sich, aus praktischen Überlegungen, von dem Farbvektoransatz lösen, so ergibt sich eine Reihe von Möglichkeiten, die Informationen der einzelnen Bildkanäle miteinander zu kombinieren. Wir legen dabei folgendes Detektionsschema zugrunde:

1. Berechnung der Richtungsableitungen nach x und y .
2. Bestimmung von Richtung und Betrag des Gradienten aus den Ableitungswerten.

⁵Es sei daran erinnert, daß sich die Intensität als arithmetisches Mittel aus den Farbkomponenten R , G und B ergibt. Die Intensität stimmt somit bis auf den Skalierungsfaktor $\frac{1}{3}$ mit der Summe $R + G + B$ überein.

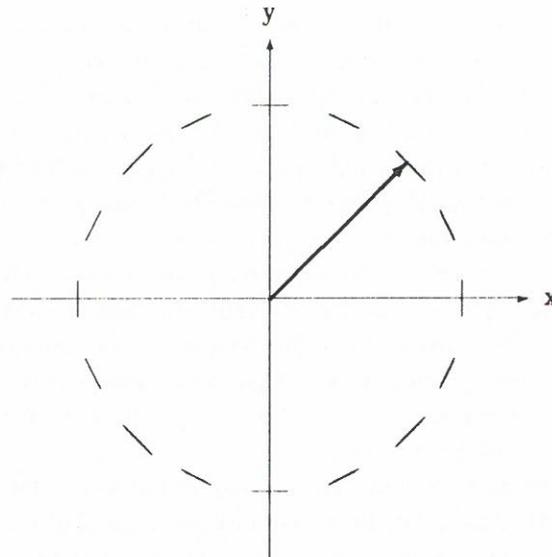


Abbildung 7.3: Veranschaulichung der Maximumsnorm

3. Detektion lokaler Extrema.

Je nachdem, auf welcher Ebene des Schemas die Integration der Informationen aus den einzelnen Kanälen stattfindet, lassen sich verschiedene mögliche Ansätze unterscheiden (siehe hierzu auch [Novak + Shafer 87]):

- Kombination der Farbkanäle zu einem einzigen Kanal vor einer Verarbeitung und Verwendung des Grauwertansatzes.
- Kombination der Richtungsableitungen nach x und y zu den x - und y -Komponenten eines Gesamtgradienten.
- Kombination der Einzelgradienten zu einem Gesamtgradienten.
- Separate Behandlung der Kanäle und Kombination der resultierenden Kantenbilder.

Wir wollen die sich anbietenden Ansätze kurz untersuchen, um damit die Grundlage für eine vergleichende Bewertung in den nachfolgenden Experimenten zu schaffen.

Die Verwendung des Grauwertansatzes auf einem Kanal, der zuvor durch Transformation aus allen Bildkanälen oder durch Auswahl eines einzigen Kanals gewonnen wird, stellt eine sehr effiziente Möglichkeit dar. Es wird hierbei vorausgesetzt, daß sich die vektorwertige Bildfunktion auf eine skalare Funktion reduzieren läßt, ohne daß dabei für die konkrete Anwendung wichtige Informationen verlorengehen. Ein Beispiel, das wir zum experimentellen Vergleich heranziehen werden, ist die Detektion von Kanten auf Intensitätsbildern. Weitere Betrachtungen über unterschiedliche Farbtransformationen werden wir aus den bereits genannten Gründen nicht vornehmen.

Die Kombination der separat erzeugten Kantenbilder zu einem Gesamtergebnis stellt hingegen die aufwendigste Alternative dar. Zudem wird die Mehrdimensionalität der Bildfunktion nicht zur Detektion ausgenutzt. Bei Verwendung eines Systems unabhängiger Bildkanäle ist dieser Ansatz sicherlich angebracht. Sind die Bildkanäle jedoch stark korreliert, können sich bei der Kombination der Kantenbilder Probleme durch die wahrscheinlich in den einzelnen Kanälen variierende Lokalisation der Kantenelemente ergeben. Wir wollen Ansätze dieser Art nicht weiter verfolgen.

Zur Kombination der Einzelgradienten bietet sich die Vektorsumme sowie die Auswahl des betragsmäßig stärksten Einzelgradienten an. Beide Möglichkeiten haben wir bereits diskutiert. Die Verwendung der Vektorsumme entspricht der Berechnung der Summe der Komponenten des Ableitungsvektors und somit dem Intensitätsverfahren. Die Auswahl des dominanten Gradienten entspricht dem formalen Lösungsansatz bei Verwendung der Maximumsnorm.

Als letzte Alternative wollen wir die Kombinationen der Richtungsableitungen zu einem Gesamtgradienten betrachten. Gemäß einer gewählten Norm ergibt sich aus den Ableitungen nach x die x -Komponente des Gesamtgradienten und aus den Ableitungen nach y entsprechend die y -Komponente. Derartige Ansätze werden z.B. von *Di Zenzo 86* erwähnt. Betrachtet werden die euklidische Norm, die Absolutnorm, die Maximumsnorm sowie die Berechnung der Summe der Richtungsableitungen.⁶

Summe der Richtungsableitungen

Für den Gesamtgradienten definieren wir:

$$\nabla \vec{C} \equiv \begin{pmatrix} R_x + G_x + B_x \\ R_y + G_y + B_y \end{pmatrix} = \nabla R + \nabla G + \nabla B = \nabla(R + G + B)$$

Somit ergibt sich auch hier eine Äquivalenz zum Intensitätsansatz.

Euklidische Norm und Absolutnorm

Die Verwendung der Absolutnorm und der euklidischen Norm führt durch die Unterdrückung der Vorzeichen zu der Projektion aller Einzelgradienten in den ersten Quadranten des Gradientenrichtungsraumes. Machen wir uns die Konsequenz an einem Beispiel klar: Es sei nur ein starker Einzelgradient im Bildausschnitt vorhanden. Dieser weise einen Winkel von 135° zur positiven x -Achse auf.⁷ Durch die Projektion dieses Gradienten in den ersten Quadranten erhalten wir ein Ergebnis, das einen Fehler von 90° gegenüber der korrekten Lösung aufweist. Statt anschließend senkrecht zur Kante nach einem lokalen Extremum zu suchen, wird das Detektionsschema die Gradientenwerte in Kantenrichtung miteinander vergleichen. Dieser gravierende Fehler wird sich für alle Gradientenvektoren bemerkbar machen, die im zweiten oder vierten Quadranten liegen, während die Unterscheidung zwischen erstem und drittem Quadranten für die

⁶ Auch hier können wir bei Berechnung der Summe nicht von einer Norm sprechen.

⁷ Der Winkel wird gegen den Uhrzeigersinn gemessen.

Detektion der Diskontinuität keine Rolle spielt. Am deutlichsten wird sich der Fehler auf Gradienten mit einem Winkel von 135° , also Kanten mit einer Orientierung von etwa 45° , auswirken. Liegen hingegen alle Einzelgradienten im ersten oder dritten Quadranten, so sollte mit diesen Ansätzen im Gegensatz zum Intensitätsansatz die Detektion reiner Farbkanten möglich sein, da sich die antikorrelierten Effekte wegen der Nichtbeachtung des Vorzeichens nicht gegenseitig aufheben.

Maximumsnorm

Würden wir bei Verwendung der Maximumsnorm die jeweils maximalen Absolutbeträge auswählen, entstünden erneut die gerade beschriebenen Fehler. Wir wollen deshalb die Maximumsnorm so definieren, daß als x -Komponente des Gesamtgradienten die Richtungsableitung nach x *inklusive ihres Vorzeichens* ausgewählt wird, die den größten Absolutbetrag aufweist. Entsprechend wird die y -Komponente bestimmt. Wenn es im Bildausschnitt einen dominanten Einzelgradienten gibt, werden wir mit diesem Ansatz häufig dasselbe Resultat erzielen wie bei Auswahl des maximalen Einzelgradienten. Eine Fehlerquelle ergibt sich jedoch, wenn mehrere Einzelgradienten vergleichbarer Größe zu beobachten sind. Gravierende Fehler können dann entstehen, wenn die resultierenden Komponenten des Gesamtgradienten von zwei verschiedenen Einzelgradienten stammen.

Abbildung 7.4 macht dies an einem Beispiel deutlich. Wir sehen zwei starke Einzelgradienten, die durch ihre gute Antikorrelation auf eine starke Diskontinuität hinweisen. Im ersten Fall besitzt ein Gradientenvektor sowohl die betragsmäßig größte x - als auch die größte y -Komponente. Der gemäß unserer Maximumsnorm resultierende Vektor entspricht deshalb diesem Einzelgradienten. Im zweiten Fall weist der zweite Einzelgradient eine etwas größere y -Komponente auf, was dazu führt, daß sich der resultierende Vektor aus der x -Komponente des einen und der y -Komponente des anderen Einzelgradienten zusammensetzt. Auch hier ergibt sich ein Fehler von fast 90° , der aus einer nur winzigen Variation der Vektorlängen resultiert. Die eingezeichneten Rechtecke deuten, wie der Kreis bei der Betrachtung der Maximumsnorm in Verbindung mit dem formalen Lösungsansatz, den Bereich an, in dem die Einzelgradienten liegen dürfen, um den resultierenden Vektor zu erzeugen.

Eine hohe Wahrscheinlichkeit für mehrere vergleichbar starke Einzelgradienten ergibt sich in Bereichen kontrastarmer Kanten und starken Rauschens. Obwohl die Ergebnisse dieses Ansatzes vergleichbar sein werden mit den Ergebnissen bei Auswahl des stärksten Einzelgradienten, sollte deshalb eine größere Rauschempfindlichkeit dieses Ansatzes besonders in kontrastarmen Bereichen zu beobachten sein.

7.6 Zusammenfassung

In diesem Kapitel haben wir verschiedene ableitungsorientierte Ansätze zur Detektion von Kanten in Farbbildern vorgestellt. Die Qualität der damit zu erzielenden Ergebnisse muß nun anhand von praktischen Experimenten beurteilt werden. Aus der theoretischen

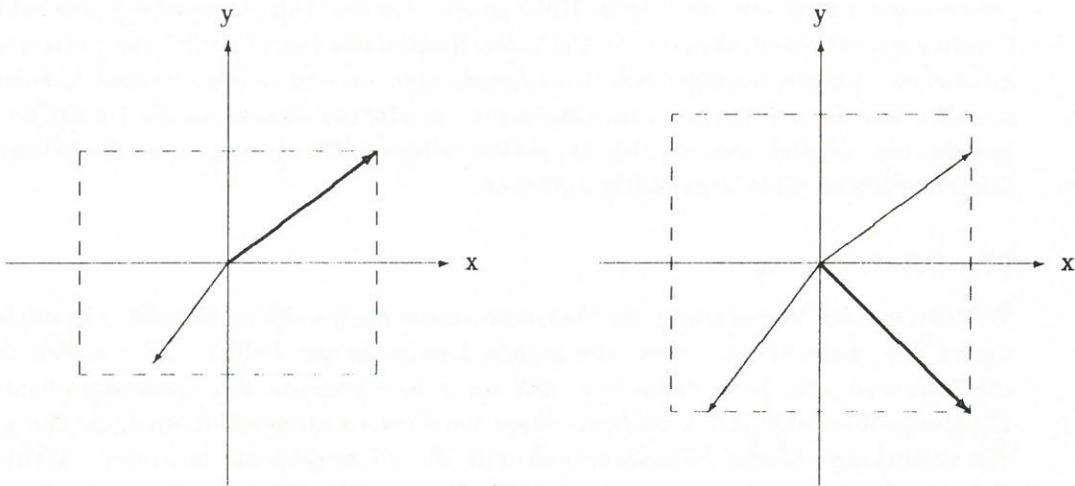


Abbildung 7.4: Möglicher Fehler bei Auswahl der maximalen Komponenten

Analyse heraus lassen sich über die zu vergleichenden Ansätze folgende Aussagen machen:

Farbvektoransatz:

Richtung und Betrag der Einzelgradienten werden gemittelt. Der Ansatz erlaubt die Detektion von Kanten, die allein aufgrund der Intensitätsinformation nicht oder nur unsicher detektiert werden können. Die Ergebnisse sollten deshalb den Ergebnissen des Intensitätsverfahrens überlegen sein.

Intensitätsansatz:

Der Intensitätsansatz entspricht dem formalen Lösungsansatz bei Verwendung der Summe der Komponenten des Ableitungsvektors. Er ergibt sich ebenso aus der Vektorsumme der Einzelgradienten sowie bei Verwendung der Summe der Richtungsableitungen. Die Detektion reiner Farbkanten ist nicht möglich.

Auswahl des maximalen Gradienten:

Dieses Verfahren ergibt sich aus dem formalen Lösungsansatz bei Verwendung der Maximumnorm. Die Ergebnisse sollten ähnlich gut sein, wie die des Farbvektoransatzes. Eine Mittelung findet jedoch nicht statt. Die Stärke der Korrelation der Einzelgradienten hat deshalb keinen Einfluß auf das Resultat. Dies müßte eine schlechtere Unterscheidbarkeit signifikanter und nichtsignifikanter Strukturen anhand des Änderungsbetrages zur Folge haben.

Kombination der Richtungsableitungen unter Verwendung der euklidischen Norm und der Absolutnorm:

Durch die Projektion aller Einzelgradienten in den ersten Quadranten des Gradientenrichtungsraumes sollten diese Ansätze Schwächen zeigen, wenn die zu detektierenden

Kanten einen Winkel von ungefähr 45° zur positiven x -Achse aufweisen.

Wahl der jeweils maximalen Richtungsableitung:

Die Ergebnisse sollten denen ähnlich sein, die bei Auswahl des maximalen Einzelgradienten erhalten werden. In kontrastarmen Bereichen sollte sich allerdings eine größere Rauschanfälligkeit zeigen.

Kapitel 8

Implementation und Experimente

In Kapitel 7 haben wir verschiedene Ansätze zur ableitungsorientierten Detektion von Kanten in Farbbildern vorgestellt. Wir wollen nun in verschiedenen Experimenten untersuchen, wie sich die praktischen Ergebnisse der einzelnen Verfahren voneinander unterscheiden und wie die Ergebnisse im Hinblick auf eine Anwendung im SISSY-Projekt zu bewerten sind. Die Beurteilung von Ergebnissen setzt die Kenntnis einiger Details des implementierten Experimentalsystems voraus. Wir müssen deshalb im folgenden Abschnitt zunächst auf Aspekte der Implementation eingehen, bevor wir die durchgeführten Experimente beschreiben können.

8.1 Implementation eines Experimentalsystems zur Kantendetektion

Um die Ergebnisse der verschiedenen Detektionsansätze praktisch vergleichen zu können, haben wir ein modulares Experimentalsystem implementiert, dessen Module den einzelnen Verarbeitungsschritten des verwendeten Detektionsschemas entsprechen:

1. Ein Modul zur Durchführung von Ableitungsoperationen auf den einzelnen Bildkanälen.
2. Verschiedene Module zur Ermittlung von Betrag und Richtung der lokal stärksten Änderung der Bildfunktion durch Kombination der Richtungsableitungen. Implementiert wurden Module für folgende Ansätze: Intensitätsansatz, Farbvektoransatz, Auswahl des maximalen Einzelgradienten, Auswahl der Richtungsableitungen nach euklidischer Norm, Absolutnorm und Maximumsnorm.
3. Ein Modul zur Detektion lokaler Extrema anhand der Betrags- und Richtungswerte.

4. Ein Modul zur Segmentierung der einzelnen Kantenelemente zu Kantenzügen und zur Auswahl der relevanten Kantenstrukturen.

Die folgenden Unterabschnitte beschreiben die zu einer Reimplementation sowie die zur Beurteilung der praktischen Ergebnisse erforderlichen Details dieser Implementation.

8.1.1 Richtungsableitungen und Detektion lokaler Extrema

Alle beschriebenen Detektionsansätze basieren auf der Schätzung der Richtungsableitungen der einzelnen Bildkanäle nach x und y . Die Ableitungsoperationen lassen sich realisieren durch Faltungen der Bildfunktion mit entsprechenden Operatormasken. Wir nutzen dabei die Separierbarkeit der Gaußfunktion aus und ersetzen die zweidimensionale Faltungsoperation durch zwei aufeinanderfolgende eindimensionale Faltungen. Die eindimensionalen Masken bilden diskrete Approximationen der eindimensionalen Gaußfunktion bzw. ihrer ersten Ableitung.

Die einfachste Art der Approximation einer kontinuierlichen Funktion ist die Abtastung der Funktion in festgelegten Intervallen. Die Gaußfunktion und ihre Ableitungen gehen relativ schnell gegen Null, so daß sich ihr Hauptwirkungsbereich bereits durch relativ wenige Abtastwerte erfassen läßt. Canny 83 schlägt vor, die Abtastung abubrechen, wenn alle Funktionswerte erfaßt sind, die betragsmäßig größer sind als ein Tausendstel des größten Funktionswertes. Betrachten wir die Werte einer zweidimensionalen Maske, die sich aus der Kombination der eindimensionalen Masken ergibt, so erhalten wir nach diesem Abbruchkriterium einen Abtastbereich von -4σ bis $+4\sigma$. Die zu einer Gaußfunktion der Operatorweite $\sigma = 2$ gehörende zweidimensionale Operatormaske hat demnach eine Ausdehnung von 17×17 Bildpunkten.

Zu prüfen wäre, ob sich die Werte der Operatormasken in geeigneter Weise ganzzahlig repräsentieren ließen, um so aufwendige reellwertige Rechenoperationen zu vermeiden. Bisher verwenden wir jedoch reelle Funktionswerte zur Darstellung der Operatorfunktionen.

Eine weitere Frage betrifft die Skalierung der Operatorfunktionen. Zur Detektion und Selektion von Kanten ist zunächst keine besondere Skalierung erforderlich, da hierfür nur der relative Vergleich der Faltungsergebnisse, nicht aber ihr absoluter Betrag eine Rolle spielt. Bei Verwendung unterschiedlicher Operatorweiten erweist es sich jedoch als zweckmäßig, die verwendeten Operatorfunktionen so zu skalieren, daß die Ergebnisse für verschiedene Operatorweiten miteinander vergleichbar sind. Durch die folgende Skalierungsbedingung wird dies gewährleistet (siehe auch [Korn + Erdtel 85]). Es seien

$$g(z) = \exp\left(-\frac{z^2}{2\sigma^2}\right) \quad \text{und} \quad g'(z) = -\frac{z}{\sigma^2} \cdot \exp\left(-\frac{z^2}{2\sigma^2}\right)$$

die unskalierte, eindimensionale Gaußfunktion und ihre erste Ableitung. Die zweidimensionale Operatorfunktion zur Schätzung der Ableitung nach x ergibt sich aus den, nun durch den Faktor $\frac{1}{k}$ skalierten, eindimensionalen Funktionen entsprechend der Dekomposition zu

$$G_x(x, y) = \frac{1}{k} g'(x) \cdot \frac{1}{k} g(y)$$

Als Skalierungsbedingung fordern wir:¹

$$\sum_{i=-N}^0 \sum_{j=-N}^N G_x(i, j) = 1$$

$$\Leftrightarrow k^2 = \sum_{i=-N}^0 g'(i) \cdot \sum_{j=-N}^N g(j)$$

Bildhaft gesprochen fordern wir, daß sich die Operatorwerte, die den positiven (bzw. negativen) Hügel der zweidimensionalen Operatorfunktion bilden, betragsmäßig zu eins summieren. Um diese Forderung zu motivieren, betrachten wir eine ideale vertikale Stufenkante unendlicher Ausdehnung. Mit Hilfe der wie beschrieben skalierten Operatorfunktionen erhalten wir — unabhängig von der gewählten Operatorweite — als Ergebnis der Ableitungsoperation die Höhe der Stufe, also den Betrag des Grauwertkontrastes der Kante. Alle berechneten Ergebnisse sollten daher in der Praxis im Bereich tatsächlich auftretender Bildkontraste liegen. Dies wiederum läßt sich bei allen nachfolgenden Verarbeitungsschritten ausnutzen.

Zur Darstellung der Faltungsergebnisse sowie zur Durchführung von Schwellwertoperationen werden die Ergebnisse der Ableitungsoperationen im Betragsbereich von 0.0 bis 80.0 auf den ganzzahligen Intensitätsbereich 0 bis 255 skaliert. Beträge, die größer als 80.0 sind, werden auf 255 abgebildet. Ein solches Vorgehen ist aufgrund der speziellen Skalierung unabhängig von der Operatorweite möglich. Werte, die über 80.0 hinausgehen, sollten in jedem Fall signifikante Strukturen darstellen. Nach unseren bisherigen Erfahrungen wird der Bereich von 0.0 bis 80.0 durch die Intensitätsskala ausreichend fein aufgelöst, um Schwellwerte zur Unterscheidung von Kantenstrukturen und Rauscheffekten festzulegen.

Obwohl die Operatorweite ein frei wählbarer Verarbeitungsparameter ist, wurde bisher, wegen der damit verbundenen Probleme, kein Ansatz zur Integration der Ergebnisse von verschiedenen Auflösungsebenen realisiert. Unter der Zielsetzung möglichst genauer Lokalisation werden wir in den Experimenten möglichst kleine Operatorweiten einsetzen. Die kleinste Operatorweite, die ohne Fehler durch Unterabtastung des Bildes einzusetzen ist, entspricht in unserer Anwendung dem Intervall, mit dem die Bildfunktion abgetastet wird ($\sigma = 1$ Bildpunkt, siehe hierzu [Canny 83, S. 71f]).

Aus den Faltungsergebnissen der Richtungsableitungen nach x und y , die bei der Verarbeitung von Farbbildern für alle Farbkanäle separat ermittelt werden, lassen sich je nach Art des Ansatzes für jeden Bildpunkt Betrag und Richtung der stärksten Änderung der Bildfunktion bestimmen.² Nach dem bereits beschriebenen Verfahren lassen sich nun durch lokalen Vergleich der Änderungsbeträge Extrema als Kanten markieren. Dazu wird gegebenenfalls eine Interpolation der zu vergleichenden Werte erforderlich.

¹Die zweidimensionale Operatormaske umfasse $(2N + 1) \times (2N + 1)$ Bildpunkte.

²Am Bildrand entsteht ein Bereich von Bildpunkten, für die keine Faltung ohne Fehler durchgeführt werden kann, da der Einzugsbereich der Operatorfunktion über den Bildbereich hinausgeht. Dieser Rand wird bei der Verarbeitung des Bildes nicht berücksichtigt.

Wir beziehen uns als Beispiel auf Abbildung 5.2. Sei $l(x, y)$ der Änderungsbetrag an den Koordinaten (x, y) . l_0 sei der Änderungsbetrag des zentralen Bildpunktes (x_0, y_0) . Die für diesen Bildpunkt bestimmte Richtung stärkster Änderung weise einen Winkel φ zur Vertikalen auf.³ Die für den Vergleich erforderlichen Werte l_1 und l_2 können durch lineare Interpolation ermittelt werden:

$$\begin{aligned} l_1 &= \tan \varphi \cdot l(x_0 + 1, y_0 + 1) + (1 - \tan \varphi) \cdot l(x_0, y_0 + 1) \\ l_2 &= \tan \varphi \cdot l(x_0 - 1, y_0 - 1) + (1 - \tan \varphi) \cdot l(x_0, y_0 - 1) \end{aligned}$$

An der Position (x_0, y_0) wird eine Kante markiert, wenn

$$l_0 > l_1 \quad \text{und} \quad l_0 > l_2.$$

Sollen rauschfreie, synthetische Bilder verarbeitet werden, muß eine der beiden *echt größer*-Bedingungen in eine *größer gleich*-Forderung abgeschwächt werden. Liegt eine ideale Stufenkante genau zwischen zwei Bildpunkten, so stimmen die zu vergleichenden Ableitungswerte aufgrund des fehlenden Rauschens exakt überein und es würde andernfalls kein Maximum detektiert werden. Das Auftreten dieses Effektes in realen Bildern ist extrem unwahrscheinlich.

8.1.2 Repräsentation und Selektion von Kantenzügen

Nach einer Durchführung der bisher dargestellten Verarbeitungsschritte besitzen wir ein binäres Kantenbild, in dem alle detektierten lokalen Extrema markiert sind. Zu jedem Bildpunkt stehen außerdem die Werte für Betrag und Richtung der stärksten Änderung der Bildfunktion zur Verfügung. Um die gewonnene Kanteninformation in geeigneter Weise für die weitere Analyse bereitzustellen, ist als letzter Schritt in unserem Verfahren die Segmentierung des Kantenbildes in Kantenzüge vorgesehen. Zudem ist eine Reduktion des Datenvolumens durch Unterdrückung nichtrelevanter und nichtverwertbarer Strukturen (in erster Linie Rauscheffekte) wünschenswert, soweit eine Entscheidung über die Relevanz der Strukturen bereits auf dieser Ebene getroffen werden kann.

Obwohl eine tiefergehende Analyse der mit dieser Aufgabenstellung verbundenen Probleme nicht Teil unserer Untersuchungen sein konnte, wurde ein einfacher Ansatz zur Segmentierung von Kantenpunkten und zur Selektion von Kantenzügen implementiert, um eine visuelle Beurteilung der mit den verschiedenen Detektionsansätzen gewonnenen Ergebnisse zu erleichtern. Wir weisen jedoch darauf hin, daß in diesen Verarbeitungsschritt Heuristiken eingehen, die das Endresultat beeinflussen. Dies ist bei einem Vergleich unterschiedlicher Detektionsansätze zu berücksichtigen. Für einen solchen Vergleich sollten immer auch die Kantenbilder herangezogen werden, die noch keiner Auswahloperation unterzogen wurden. Andererseits zeigen die Ergebnisse der Auswahloperationen, daß es in der Tat möglich sein sollte, aus der Vielzahl detektierter Strukturen die für die Interpretation wesentlichen zu extrahieren.

³ φ sei in unserem Beispiel kleiner als 45° .

Über die Relevanz detektierter Kanten wird meist durch die Festlegung eines *globalen Schwellwertes* entschieden: Ein Kantenelement wird als relevant akzeptiert, wenn der Änderungsbetrag einen global festgesetzten Mindestwert übersteigt. Dieses Vorgehen ist bei der Verarbeitung von Realweltszenen jedoch unbefriedigend, da solche Bilder stark wechselnde Kontraste aufweisen können. Kontrastarme Kanten liefern, ebenso wie Rauschen, nur kleine Änderungswerte, so daß mit der Festlegung eines globalen Schwellwertes neben Rauscheffekten auch kontrastarme Kanten unterdrückt werden. Dies führt nicht nur zu dem Verlust, sondern auch zu einer *Zerstückelung* von Kantenzügen, wenn nur einige der zugehörigen Kantenpunkte einen zu kleinen Betrag aufweisen. Eine perfekte Selektion durch Festlegung einer Schwelle ist unmöglich: Entweder wird der Schwellwert so hoch gewählt, daß garantiert keine Rauscheffekte mehr vorhanden sind, mit dem Risiko, Teile relevanter Kanten zu verlieren, oder aber der Schwellwert wird so niedrig gewählt, daß keine relevanten Kanten verlorengehen, mit dem Risiko, auch eine Vielzahl nichtverwertbarer Strukturen zu erhalten.

Einen geeigneten Kompromiß zwischen beiden Extremen stellt der Vorschlag von Canny 83 dar, zwei Schwellwerte zur Selektion zu verwenden.⁴ Dazu werden nicht mehr einzelne Kantenpunkte betrachtet, sondern zusammenhängende Kantenzüge: Ein Kantenzug wird selektiert, wenn die Änderungsbeträge aller Kantenpunkte über einem niedrigen Schwellwert und wenigstens einige davon über einem hohen Schwellwert liegen.

Dieses sogenannte *hysteresis thresholding* setzt die Segmentierung der Kantenpunkte in Kantenzüge voraus. Dazu werden in unserem Verfahren die einzelnen Kantenpunkte aufgrund ihrer Nachbarschaft miteinander verbunden. Die so erhaltenen Kantenzüge erlauben nicht nur Schwellwertoperationen bezüglich der in ihnen auftretenden Änderungsbeträge, sondern auch Selektionsoperationen, die beispielsweise die Länge der Kantenzüge berücksichtigen. Die verbleibenden Kantenzüge werden schließlich als Kettenkode⁵ repräsentiert und für die weitere Analyse bereitgestellt.

Bei der Segmentierung des Kantenbildes wirken sich zwei Eigenschaften des verwendeten Detektionsschemas positiv aus:

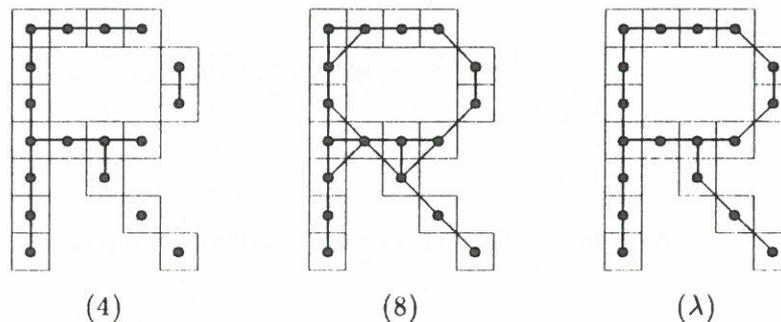
1. Die von den detektierten Kantenpunkten gebildeten Strukturen sind meist einen Bildpunkt breit. Eine Verdünnung oder Skeletierung zusammenhängender Strukturen ist nicht erforderlich.
2. Die idealerweise zu verbindenden Kantenpunkte bilden tatsächlich zusammenhängende, d.h. nicht lückenhafte, Strukturen.

Kanten im Bild bilden also im wesentlichen zusammenhängende, linienhafte Strukturen; die Verbindung der Kantenpunkte ist aufgrund einfacher Nachbarschaftsbeziehungen möglich. Denkbare Nachbarschaftsdefinitionen sind:

4-Nachbarschaft: Zwei Bildpunkte heißen 4-benachbart, wenn sie in horizontaler oder vertikaler Richtung direkt aneinander angrenzen.

⁴Die Idee der Verwendung von zwei Schwellwerten zur Verringerung von rauschbedingten Störungen findet sich auch in [Rosenfeld + Kak 82, Bd. 2, S. 68].

⁵Die Kettenkode-Repräsentation enthält die Bildkoordinate eines Startpunktes, sowie eine Folge von Richtungsangaben, die die Bestimmung des jeweils nächsten Kantenpunktes im Kantenzug erlaubt.

Abbildung 8.1: 4-, 8-, und λ -Nachbarschaft im Vergleich

8-Nachbarschaft: Zwei Bildpunkte heißen 8-benachbart, wenn sie in horizontaler, vertikaler oder diagonaler Richtung direkt aneinander angrenzen.

Beide Nachbarschaftsdefinitionen weisen Mängel auf, wie in Abb. 8.1 deutlich wird: Die 4-Nachbarschaft erlaubt keine Verbindung "über Eck", während die Verwendung der 8-Nachbarschaft zur Bildung zahlreicher zusätzlicher Kreuzungspunkte führen würde. Eine beide Nachteile vermeidende Nachbarschaftsdefinition ist die von *Speck 83* formal eingeführte λ -Nachbarschaft.

λ -Nachbarschaft: Zwei Bildpunkte heißen λ -benachbart, wenn sie 8-benachbart sind, aber nicht über einen dritten Bildpunkt im 4-Zusammenhang stehen.

Der Vorteil dieser Definition wird in Abb. 8.1 ersichtlich. Verbindet man benachbarte Kantenpunkte gemäß der λ -Nachbarschaft und zerlegt den so erhaltenen Verbindungsgraphen in kreuzungsfreie Segmente maximaler Länge, so erhält man eine erste Segmentierung der Kantenpunkte in Kantenzüge. Der Aufwand kann an dieser Stelle verringert werden, wenn der in das "hysteresis thresholding" eingehende niedrigere Schwellwert bereits vor der Segmentierung auf die Höhe der Änderungswerte der Kantenpunkte angesetzt wird. Der Schwellwert sollte so niedrig gewählt sein, daß ausschließlich nicht-relevante Strukturen unterdrückt werden. Aus dem Segmentationsergebnis können nun Kantenzüge selektiert werden, indem verschiedene Anforderungen an die Segmente gestellt werden. Die vorgesehenen Auswahlkriterien sind:

Längen-Kriterium: Kantenzüge müssen eine gewisse Mindestlänge besitzen.

Kontrast-Kriterium: Kantenzüge müssen mindestens eine vorgegebene Anzahl⁶ von Kantenpunkten enthalten, deren Änderungsbeträge über einer bestimmten Schwelle⁷ liegen.

⁶Die Festlegung einer *relativen* Anzahl erweist sich für kurze Kantenzüge als wenig geeignet.

⁷Dieser Schwellwert entspricht dem höheren der beiden Schwellwerte, die in das "hysteresis thresholding" eingehen.

	Mindestlänge	Mindestanzahl
Typ 0	10	6
Typ 1	6	6
Typ 2	3	3

Abbildung 8.2: Verwendete Selektionsparameter

Um die Unterdrückung von Rauscheffekten zu erleichtern, ist es sinnvoll, verschiedene Typen von Segmenten zu unterscheiden:

Typ 0: Das Segment ist isoliert, d.h., kein Endpunkt stellt einen Kreuzungspunkt mit anderen Segmenten dar.

Typ 1: Ein Endpunkt des Segments bildet einen Kreuzungspunkt.

Typ 2: Beide Endpunkte bilden Kreuzungspunkte.

Dies ermöglicht eine typspezifische Festlegung der Selektionsparameter (Mindestlänge, Schwellwert, Mindestanzahl). Auf diese Weise läßt sich beispielsweise für einen isolierten Kantenzug eine größere Mindestlänge fordern, als für einen Kantenzug, der über Kreuzungspunkte mit anderen Kantenzügen verbunden ist. Nach Festlegung der Parameter beginnt der iterative Selektionsprozeß:

1. Zu jedem Segment des Verbindungsgraphen wird der Segmenttyp bestimmt.
2. Alle Segmente, die den typspezifischen Anforderungen nicht genügen, werden aus dem Verbindungsgraphen gelöscht.
3. Das Entfernen von Segmenten kann dazu führen, daß Kreuzungspunkte gelöscht werden. Dies kann gegebenenfalls ein Verschmelzen von Segmenten erforderlich machen, um wieder eine Zerlegung in kreuzungsfreie Segmente maximaler Länge zu erhalten.
4. Wurde kein Segment entfernt, ist der Prozeß beendet; andernfalls wird der Prozeß bei Schritt 1 fortgesetzt.

Zur Zeit müssen alle Parameter des Selektionsverfahrens durch den Benutzer festgelegt werden. Um die Abhängigkeit vom jeweiligen Bild möglichst gering zu halten, werden wir bei unseren Experimenten die Parameter für Mindestlänge und Mindestanzahl für alle Bilder gleich wählen. In Abhängigkeit vom Bild geschieht nur die Festlegung eines niedrigen und eines hohen Schwellwertes. Der niedrige Schwellwert wird bereits beim Aufbau des Verbindungsgraphen berücksichtigt, der hohe dient für alle Segmenttypen als Selektionsparameter. Die in den Experimenten verwendeten Parameterwerte sind Abbildung 8.2 zu entnehmen. Grundsätzlichen Fragen, die sich hier stellen, sind:

- Gibt es ein allgemeines — bildunabhängiges — Selektionsschema?
- Lassen sich die in das Schema eingehenden Schwellwerte in Abhängigkeit vom Bild automatisch bestimmen?

Beide Fragen können wir im Rahmen unserer Untersuchungen aus Zeitgründen nicht weiter verfolgen.

8.2 Experimente

Im Rahmen unserer Experimente wollen wir folgende Fragestellungen untersuchen:

- Wie unterscheiden sich die Ergebnisse der verschiedenen Detektionsansätze? Lassen sich die aus der theoretischen Analyse gewonnenen Erkenntnisse bestätigen?
- Ergeben sich durch die Berücksichtigung von Farbinformation Vorteile gegenüber dem Intensitätsverfahren?
- Wie sind die Ergebnisse des Farbvektoransatzes im Hinblick auf eine Anwendung im SISSY-Projekt zu bewerten?

Bei der Bewertung der Ergebnisse wollen wir Kriterien zugrundelegen, die die besonderen Anforderungen an ein Kantendetektionsverfahren innerhalb des SISSY-Projektes berücksichtigen. Das vornehmliche Ziel des SISSY-Projektes ist die geometrische Auswertung räumlicher und zeitlicher Korrespondenzen zur Schätzung von Lage- und Bewegungsparametern. Aus dieser Zielsetzung ergeben sich folgende Forderungen an ein Kantendetektionsverfahren:

1. **Sichere Detektion**, d.h. Robustheit gegen Rauschen sowie physikalische Relevanz detektierter Merkmale.
2. Möglichst **gute Lokalisation** detektierter Merkmale zur präzisen Schätzung der Lage- und Bewegungsparameter.
3. **Gute Detektion und Lokalisation an Schnitt- und Eckpunkten** von Kanten, um die Redundanz der Ergebnisse von Kanten- und Punktefinder nutzen zu können.
4. **Stabilität** bei kleinen Veränderungen, die sich aufgrund sich bewegender Objekte, eines sich leicht verändernden Betrachterstandortes oder auch für eine statische Szene über die Bilder einer zeitlichen Bildfolge hinweg durch Rauschen ergeben können.
5. **Reichhaltige Beschreibung** detektierter Kanten zur Vereinfachung der Korrespondenzanalyse. Eine reichhaltige Beschreibung vereinfacht die Unterscheidung einander ähnlicher Merkmale.

Die Forderungen nach guter Detektion und Lokalisation können unabhängig von einer konkreten Anwendung an ein Kantendetektionsverfahren gestellt werden. Im Rahmen des Sissy-Projektes ist zu beachten, daß bei Realweltaufnahmen die Aufnahmebedingungen nur wenig kontrolliert werden können. Im Gegensatz zu industriellen Szenen oder Laborszenen läßt sich die Qualität der aufgenommenen Bilder nicht durch Manipulationen an der Szene verbessern.

Das Ziel der physikalischen Relevanz aller detektierten Strukturen wird vermutlich für Verfahren, die kein weiteres Wissen ausnutzen, auf dieser untersten Verarbeitungsstufe immer ein unerreichbares Ideal bleiben.

Bei der Lokalisation der Merkmale ist zu bedenken, daß eine in Bildpunkten gemessene geringe Ungenauigkeit bei der Tiefenmessung aus Stereo-Bildern durch Triangulation bereits bedeutende Fehler verursachen kann.

Schnitt- und Eckpunkte⁸ von Kanten stellen bei der Detektion und Lokalisation ein großes Problem dar, da die meisten Verfahren davon ausgehen, daß sich im Einzugsbereich eines Operators nur ein Kanteneffekt bemerkbar macht. Darüber hinaus wird oft vorausgesetzt, daß eine Kante im betrachteten Bildausschnitt geradlinig verläuft.

Die Stabilität der Merkmale ist erforderlich, um die Zuordnung detektierter Strukturen zwischen den Bildern eines Stereobildpaares oder einer Bildfolge zu erleichtern.

Die Reichhaltigkeit der Beschreibung der detektierten Kanten ist relativ unabhängig vom verwendeten Detektionsverfahren. Die meisten Ansätze erlauben Aussagen über Ort, Richtung und Kontrast eines Kantenelements. Zu Kantenzügen zusammengefaßte Kantenelemente bieten darüber hinaus zusätzliche Beschreibungsmöglichkeiten, z.B. bezüglich der Form. Auf die geeignete Darstellung dieser Informationen werden wir nicht weiter eingehen.

Wir beginnen die Darstellung unserer Experimente mit der Gegenüberstellung der Ergebnisse aller implementierten Ansätze. Anschließend werden wir den Vergleich zwischen Farbvektoransatz und Intensitätsverfahren vertiefen. Schließlich werden wir versuchen, die Ergebnisse des Farbvektoransatzes nach den oben aufgeführten Kriterien zu bewerten. Die in den Experimenten verwendeten Originalbilder finden sich in Anhang A. Wir bezeichnen die Bilder im folgenden als *Apfel-Bild*, *Block-Bild* und *Auto-Bild*.

8.2.1 Vergleich der verschiedenen Detektionsansätze

Wie sich in unseren Versuchen gezeigt hat, werden die Unterschiede zwischen den verschiedenen Detektionsansätzen besonders deutlich an den Ergebnissen für das Apfel-Bild. Wir beschränken uns deshalb beim Vergleich aller Ansätze auf die Präsentation der Ergebnisse für das Apfel-Bild. Die beschriebenen Unterschiede lassen sich aber auch in den Resultaten für die anderen Bilder beobachten.

Bild 2 zeigt die binären Kantenbilder, die sich bei Verwendung der verschiedenen Detektionsansätze für das Apfel-Bild ergeben. Die Kantenbilder enthalten alle lokalen Extrema ohne Rücksicht auf die Höhe der Änderungsbeträge. Zur Schätzung der

⁸Unter einem Eckpunkt verstehen wir in diesem Zusammenhang einen Punkt starker Krümmung oder Krümmungsänderung in einem Kantenzug.

Richtungsableitungen wurden Operatoren der Weite $\sigma = 2$ verwendet. Es sei daran erinnert, daß alle folgenden Ergebnisbilder einen mehr oder weniger breiten Rand aufweisen werden, in dem keine Verarbeitung des Bildes stattfand, da der Einzugsbereich des Operators über den Bildrand hinausgegangen wäre.

Wie in Bild 28 erkennbar ist, weist der Hintergrund des Apfel-Bildes eine starke Textur auf. Die in dieser Textur durch den **Intensitätsansatz** ermittelten Änderungsbeträge liegen zum Teil in derselben Größenordnung wie die Beträge an der Außenkontur der Äpfel. Die Außenkonturen werden deshalb bei Verwendung des Intensitätsansatzes nicht immer geschlossen detektiert. Dies zeigt sich besonders an den Konturen der beiden rechten Äpfel. Auch die kontrastarmen Konturen im Schattenbereich in der linken unteren Bildecke werden nicht detektiert.

Der **Farbvektoransatz** zeigt an den Außenkonturen der Äpfel dagegen sehr gute Ergebnisse. Der texturierte Hintergrund setzt sich aufgrund der blauen Färbung ausreichend deutlich von den Äpfeln ab, so daß gut geschlossene Konturen detektiert werden. Selbst im Schattenbereich kann die Außenkontur des Apfels detektiert werden.

Bei **Auswahl des maximalen Einzelgradienten** erhalten wir Ergebnisse, die durchaus vergleichbar sind mit den Ergebnissen des Farbvektoransatzes. An einigen Schnittpunkten der Außenkonturen ergeben sich Abweichungen durch eine unterschiedliche Verbindung der sich schneidenden Kanten. Lediglich im Schattenbereich kann man dem Farbvektoransatz eine gewisse Überlegenheit zusprechen.

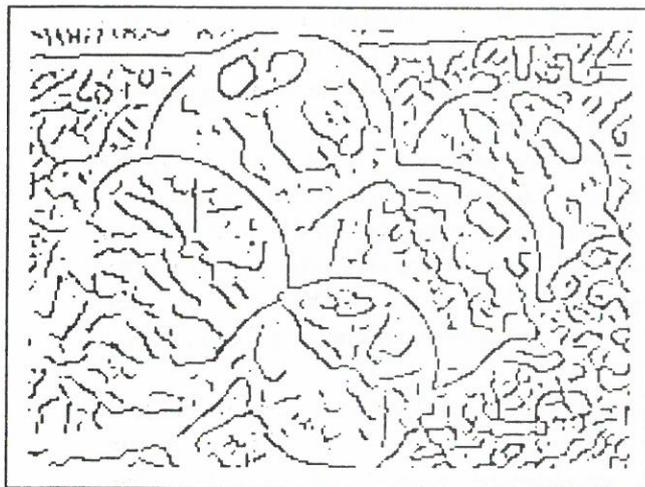
Das Ergebnis bei **Auswahl der maximalen Richtungsableitungen** bestätigt die theoretischen Erwartungen in kontrastarmen Bildbereichen: im Schattenbereich ist gegenüber allen anderen Farbansätzen eine deutlich stärkere Rauschempfindlichkeit festzustellen. Ansonsten zeigen sich zum Ergebnis bei Auswahl des maximalen Gradienten kaum Unterschiede.

Die Ergebnisse bei **Kombination der Richtungsableitungen nach euklidischer Norm und nach Absolutnorm** stimmen gut überein. Auch hier bestätigen sich die Erwartungen: Die Ergebnisse sind überwiegend zufriedenstellend; nur bei Kanten mit einer Richtung von ungefähr 45° zur Horizontalen zeigen beide Ansätze Schwächen.

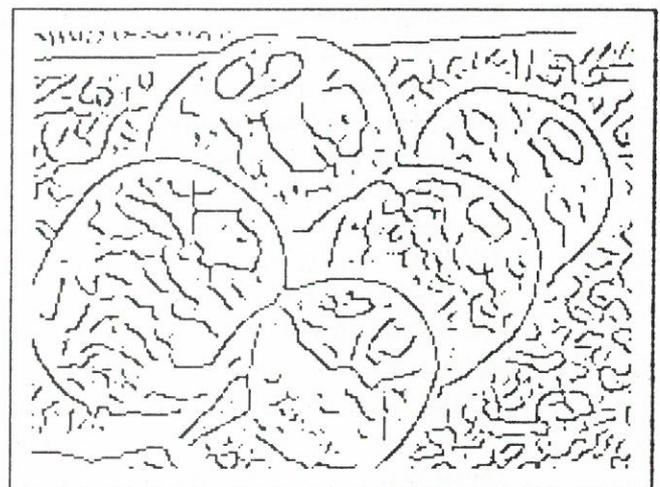
Wie die Diskussion der Ergebnisse zeigt, lassen sich die aus der theoretischen Analyse gewonnenen Erwartungen sehr gut praktisch bestätigen. Auch die Überlegenheit des Farbvektoransatzes gegenüber dem Intensitätsansatz und den alternativen Farbansätzen zeigt sich deutlich. Vergleichbare Ergebnisse lassen sich nur bei Auswahl des maximalen Einzelgradienten erzielen. Im folgenden werden wir uns auf den Vergleich zwischen Farbvektoransatz und Intensitätsansatz beschränken.

8.2.2 Vergleich von Farbvektoransatz und Intensitätsansatz

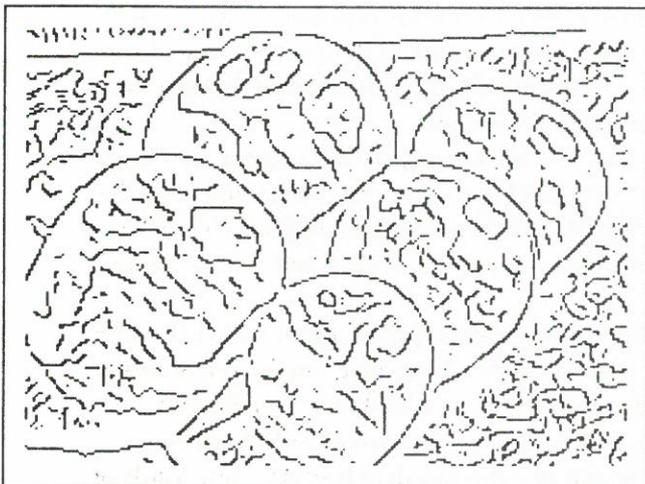
Wie wir gesehen haben, ergeben sich zwischen den Ergebnissen des Intensitätsansatzes und des Farbvektoransatzes recht deutliche Unterschiede. Um zu überprüfen, welche Rolle die gewählte Operatorweite bei diesem Vergleich spielt, haben wir das Apfel-Bild mit Operatoren der Weiten $\sigma = 1 \dots 6$ verarbeitet. Die Resultate für Intensitäts- und Farbvektoransatz zeigen die Bilder 3 und 4. Wie man erkennt, läßt sich die Überle-



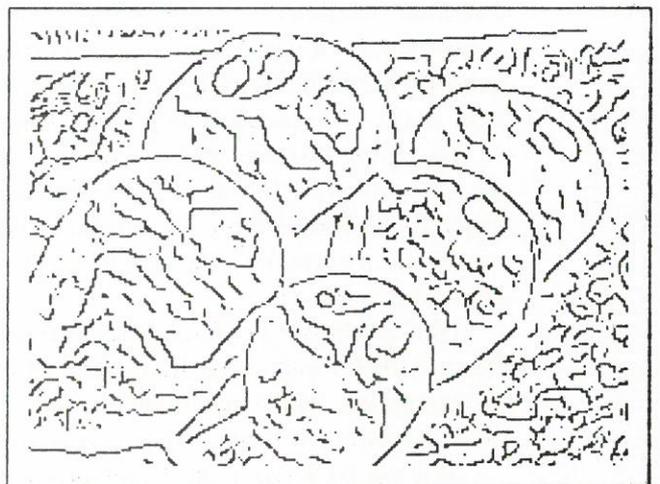
Intensitätsansatz



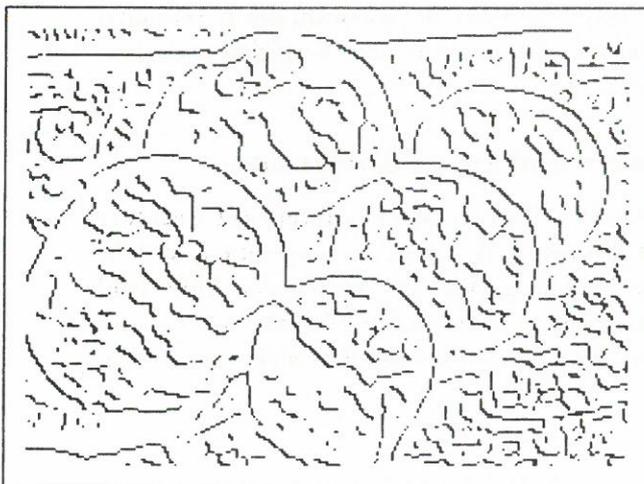
Farbvektoransatz



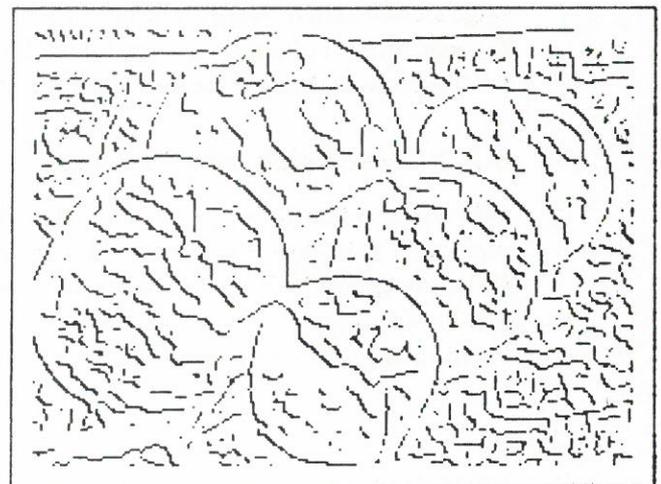
Auswahl des max. Einzelgradienten



Auswahl der max. Richtungsableitungen



Kombination nach euklidischer Norm



Kombination nach Absolutnorm

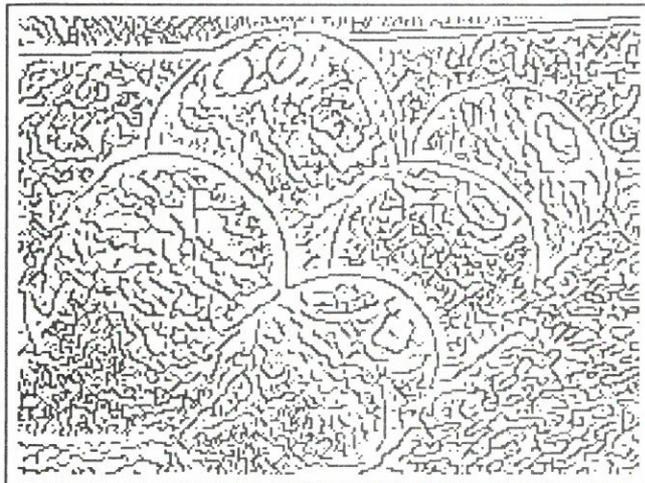
Bild 2: Ergebnisse aller Ansätze im Vergleich

genheit des Farbvektoransatzes gegenüber dem Intensitätsansatz für alle verwendeten Operatorweiten beobachten.

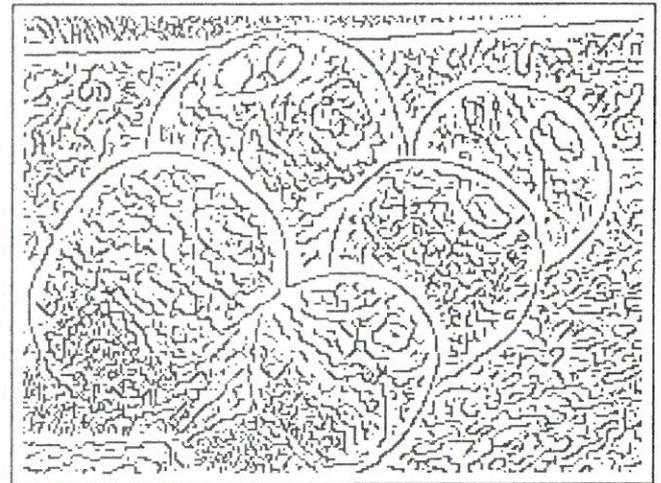
Für welche Positionen im Bild eine Überlegenheit des Farbvektoransatzes zu erwarten ist, zeigt Bild 5. Wir sehen hier die Differenz zwischen den Änderungsbeträgen, die durch den Farbvektoransatz ermittelt wurden und denen, die mit Hilfe des Intensitätsansatzes bestimmt wurden. Dunkle Bereiche im Bild besitzen wesentlich höhere Farbänderungsbeträge als Intensitätsänderungsbeträge. In weißen Bereichen stimmen beide Beträge überein. Da sich eine Intensitätsänderung im Normalfall auch im Ergebnis des Farbansatzes bemerkbar macht, gibt es keine Bereiche im Bild (von Rauscheffekten abgesehen), in denen der Betrag der Intensitätsänderung den Betrag der Farbänderung übersteigt. Im Vergleich mit den zuvor gezeigten Ergebnissen erkennt man, daß die dunklen Bildbereiche gerade mit den Bereichen übereinstimmen, in denen der Intensitätsansatz Schwächen zeigt.

Die Bilder 6 bis 11 zeigen die Ergebnisse des Intensitäts- und des Farbvektoransatzes für das Block- und das Auto-Bild. Das Block-Bild wurde mit der Operatorweite $\sigma = 2$, das Auto-Bild mit der Operatorweite $\sigma = 1$ verarbeitet. Das Block-Bild weist sehr klare Kantenstrukturen auf. Schwierigkeiten für Kantendetektionsverfahren ergeben sich durch die teilweise geringen Kontraste an den Orientierungskanten der einzelnen Blöcke (siehe Bild 32). Ein Beispiel dafür ist der rechte Rand der Oberseite des langen Quaders. Diese Kante wird zwar mit beiden Ansätzen detektiert, sie weist jedoch nur sehr geringe Änderungsbeträge auf. Auch der Kontrast zwischen Quaderoberseite (linker Rand) und Hintergrund ist sehr schwach. Dies zeigt sich im Ergebnis des Intensitätsansatzes: Die Verdeckungskante wird nur im oberen Teil detektiert. Im unteren Verlauf wird ein Schattierungseffekt im Hintergrund so dominant, daß die Kante in den Hintergrund hinein fortgesetzt wird. Dieser Effekt tritt bei Verwendung des Farbvektoransatzes nicht auf: Der Farbunterschied zwischen Quader und Hintergrund reicht aus, um die Verdeckungskante in ihrem vollen Verlauf zu detektieren. Ein weiterer deutlicher Unterschied zeigt sich für die Oberseite des Pyramidenstumpfes. Die Oberseite ist blau und setzt sich farblich stark von den beige Seiten der Pyramide ab. Der Rand der Oberseite wird deshalb durch den Farbvektoransatz ohne Probleme detektiert. In der Intensität weist dieser Rand nur einen sehr geringen Kontrast auf. Mit Hilfe des Intensitätsansatzes wird der Rand zwar teilweise detektiert, der Kontrast ist jedoch so gering, daß die detektierten Strukturen nicht anhand der Höhe der Änderungsbeträge von Rauscheffekten zu unterscheiden sein werden. Auch für das Block-Bild zeigen sich die genannten Unterschiede im Differenzbild von Farbänderungsbetrag und Intensitätsänderungsbetrag (Bild 10).

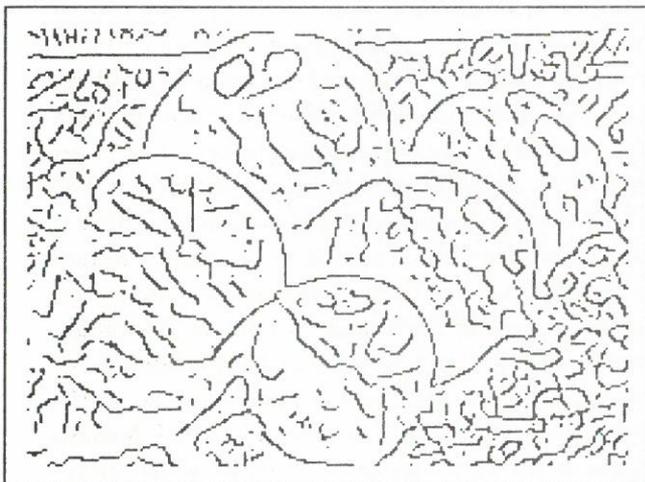
Im Auto-Bild sind nur sehr geringe Unterschiede zwischen den Ergebnissen beider Ansätze erkennbar. Insgesamt führt die kräftige Färbung der Fahrzeugteile dazu, daß die durch den Farbvektoransatz ermittelten Änderungsbeträge meist deutlich über den Beträgen der Intensitätsänderung liegen (siehe Bild 11). Ein interessanter Unterschied ist am Übergang zwischen Windschutzscheibe und Motorhaube zu beobachten: Mit Hilfe des Farbvektoransatzes wird sowohl der Rand der Windschutzscheibe als auch der rechte Rand des hellen Streifens auf der Motorhaube detektiert. Bei Verwendung des Intensitätsansatzes wird der Rand des Streifens so dominant, daß der Übergang zur



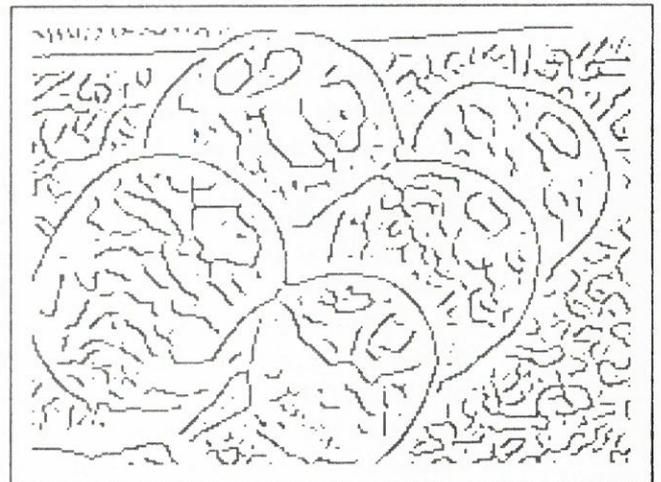
Intensitätsansatz, $\sigma = 1$



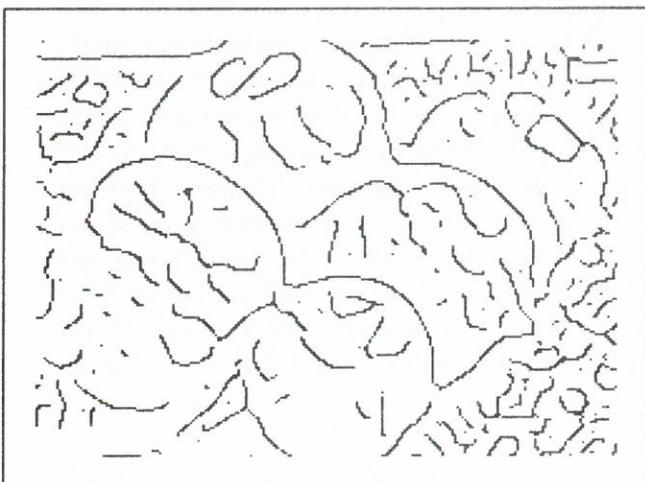
Farbvektoransatz, $\sigma = 1$



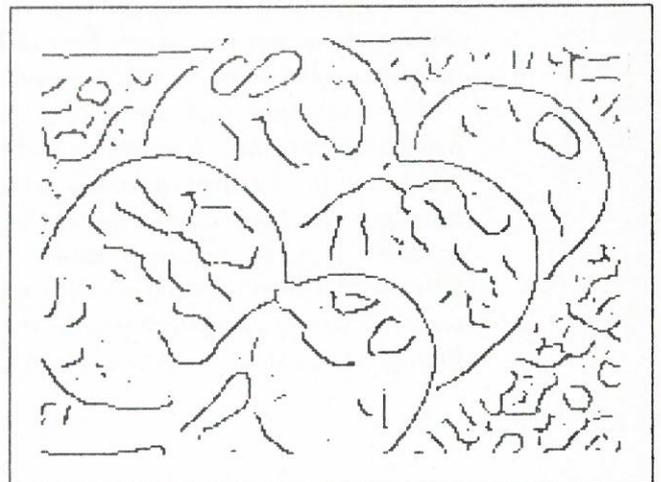
Intensitätsansatz, $\sigma = 2$



Farbvektoransatz, $\sigma = 2$

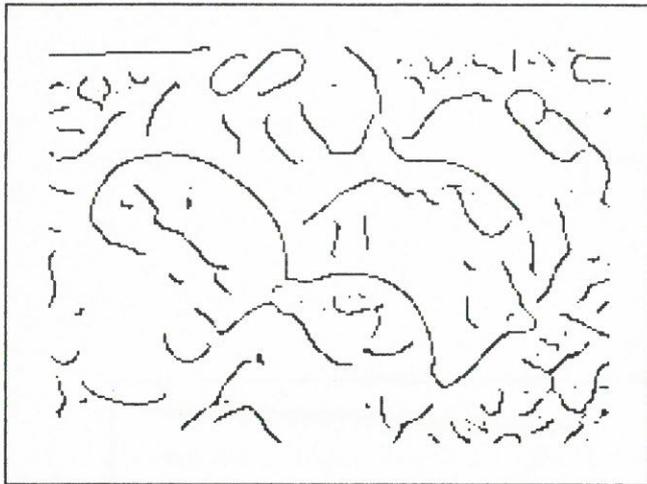


Intensitätsansatz, $\sigma = 3$

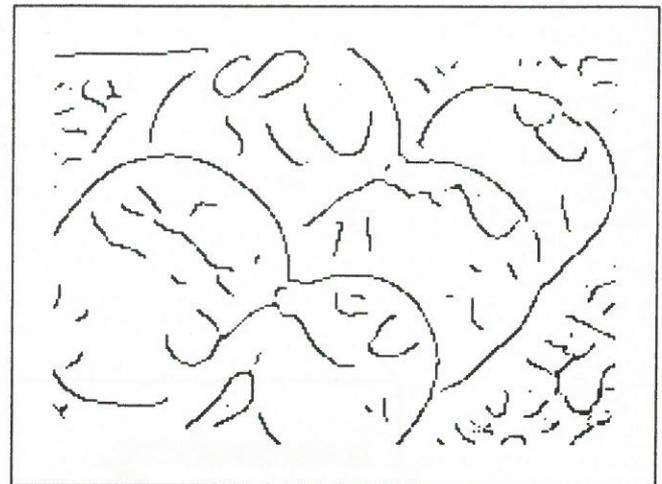


Farbvektoransatz, $\sigma = 3$

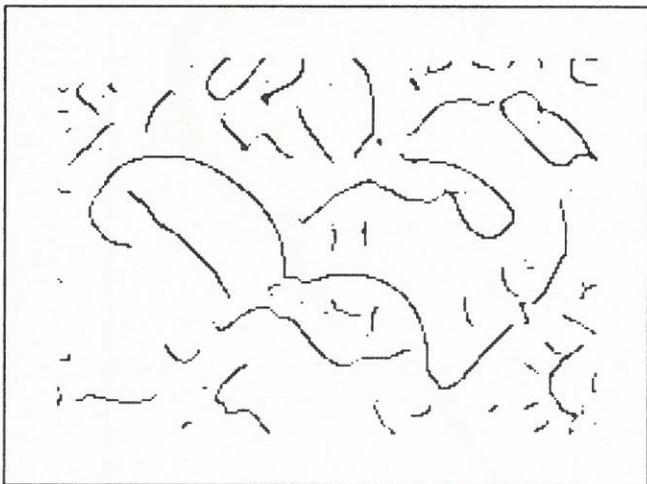
Bild 3: Intensitätsansatz und Farbvektoransatz im Vergleich, $\sigma = 1 \dots 3$



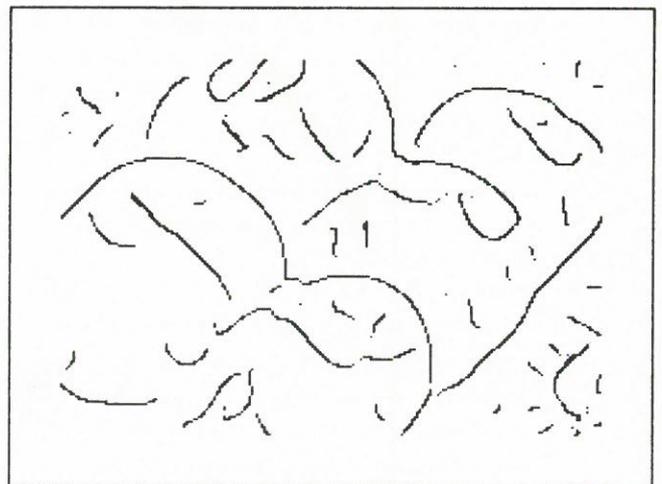
Intensitätsansatz, $\sigma = 4$



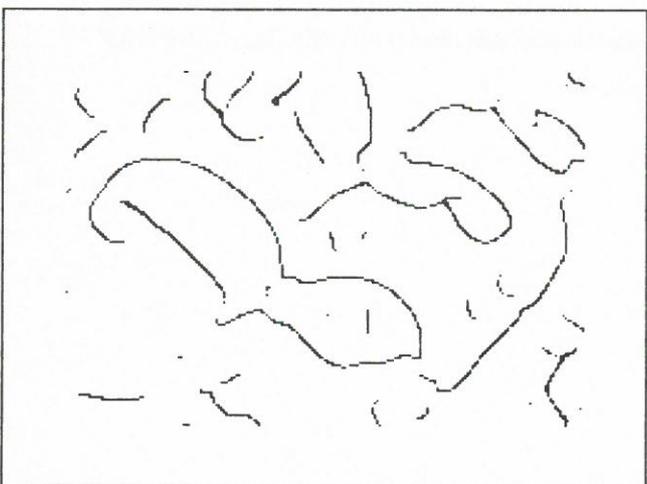
Farbvektoransatz, $\sigma = 4$



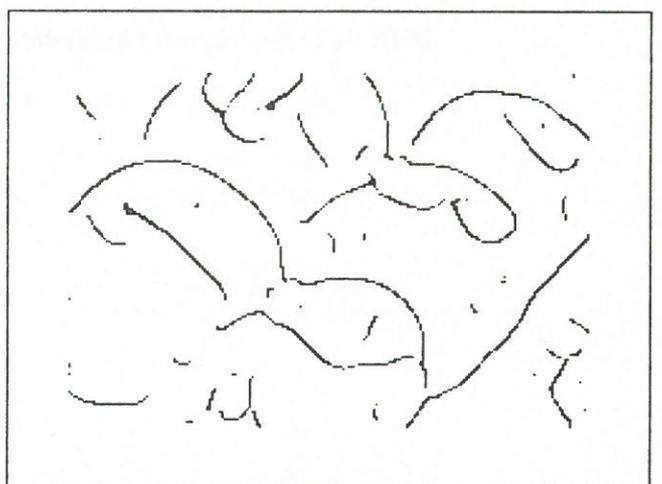
Intensitätsansatz, $\sigma = 5$



Farbvektoransatz, $\sigma = 5$



Intensitätsansatz, $\sigma = 6$



Farbvektoransatz, $\sigma = 6$

Bild 4: Intensitätsansatz und Farbvektoransatz im Vergleich, $\sigma = 4 \dots 6$



Bild 5: Differenz von Farbänderung und Intensitätsänderung für das Apfel-Bild

Windschutzscheibe im mittleren Verlauf nicht mehr detektiert wird. Stattdessen wird eine kontrastarme Kante im Wageninneren detektiert.

Einige der in den Ergebnissen erkennbaren Unterschiede lassen sich sehr gut anhand von Intensitäts- bzw. Farbprofilen analysieren. Wir wollen dies an einem besonders auffälligen Beispiel im Apfel-Bild tun. Bild 12 zeigt einen Ausschnitt aus dem Apfel-Bild. Der eingetragene Rahmen kennzeichnet den Teil einer Bildzeile mit den dort mit Hilfe des Farbvektoransatzes detektierten Kanten. Bild 13 zeigt den entsprechenden Ausschnitt mit dem Ergebnis des Intensitätsansatzes. Während die rechte der beiden detektierten Kanten von beiden Ansätzen gleich lokalisiert wird, unterscheiden sich die Positionen der linken Kante deutlich voneinander. Die mit Hilfe des Farbvektoransatzes detektierte Kante entspricht der Außenkontur des Apfels; die mit Hilfe des Intensitätsansatzes detektierte Kante ergibt sich scheinbar aufgrund der Schattierung des Apfels. Eine Erklärung liefern die in Abbildung 8.3 gezeigten Profile. Sie stellen den Verlauf der Bildfunktion in jedem Farbkanal bzw. im Intensitätsbild für die gekennzeichnete Bildzeile dar. Die vertikalen Linien markieren die Positionen der mit Hilfe des Farbvektoransatzes bzw. des Intensitätsansatzes detektierten Kanten. Die linke durch den Farbvektoransatz detektierte Kante zeichnet sich durch starke Änderungen im Rot- und im Blaukanal aus. Die rechte Kante ist in allen drei Kanälen zu beobachten. Im Intensitätsprofil prägt sich hingegen nur die rechte der beiden Kanten aus. Die Außenkontur des linken Apfels kann durch den Intensitätsansatz nicht detektiert werden, da sich die Änderungen im Rot- und Blaukanal aufgrund ihres entgegengesetzten Vorzeichens in der Intensität gegenseitig aufheben. Stattdessen wird eine Kante im Bereich der nach links rampenartig ansteigenden Intensität detektiert.

8.2.3 Bewertung der Ergebnisse des Farbvektoransatzes

Es folgt nun der Versuch einer Bewertung der mit Hilfe des Farbvektoransatzes erzielten Ergebnisse unter Berücksichtigung der oben formulierten Kriterien.

Detektion und Lokalisation

Genaue Aussagen über die Güte von Detektion und Lokalisation sind nur sehr schwer zu treffen. Quantitative Bewertungen würden die Aufnahme exakt vermessener Szenen verlangen. Zudem ist eine Bewertung der Ergebnisse im Hinblick auf eine Anwendung im SISSY-Projekt erst möglich, wenn die detektierten Merkmale tatsächlich in der Korrespondenzanalyse zur Schätzung von Lage- und Bewegungsparametern eingesetzt werden. Wir wollen uns deshalb an dieser Stelle auf die visuelle Beurteilung der Ergebnisse beschränken.

Wir betrachten dazu die Bilder 14, 16 und 18. Die Bilder zeigen die Ergebnisse des Farbvektoransatzes für Apfel-, Block- und Auto-Bild, jeweils dem Intensitätsbild überlagert. Zumindest für das Apfel- und das Block-Bild fällt die visuelle Beurteilung relativ leicht: Soweit dies durch Betrachtung zu beurteilen ist, scheinen die in den Szenen vorhandenen relevanten Kantenstrukturen durch den Farbvektoransatz detektiert zu

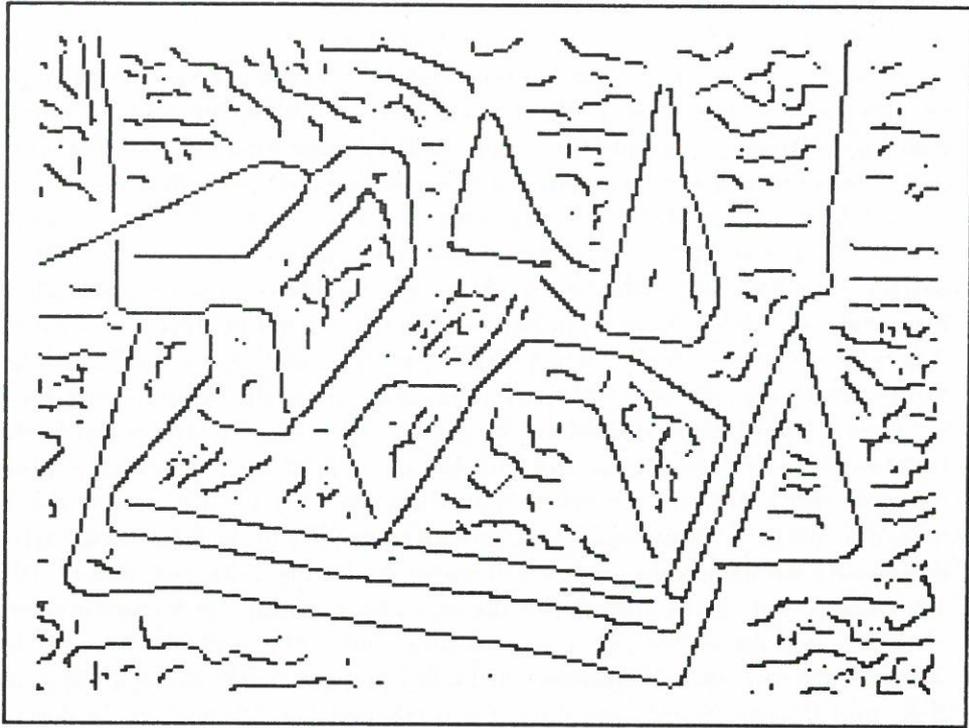


Bild 6: Ergebnis des Intensitätsansatzes für das Block-Bild ($\sigma = 2$)

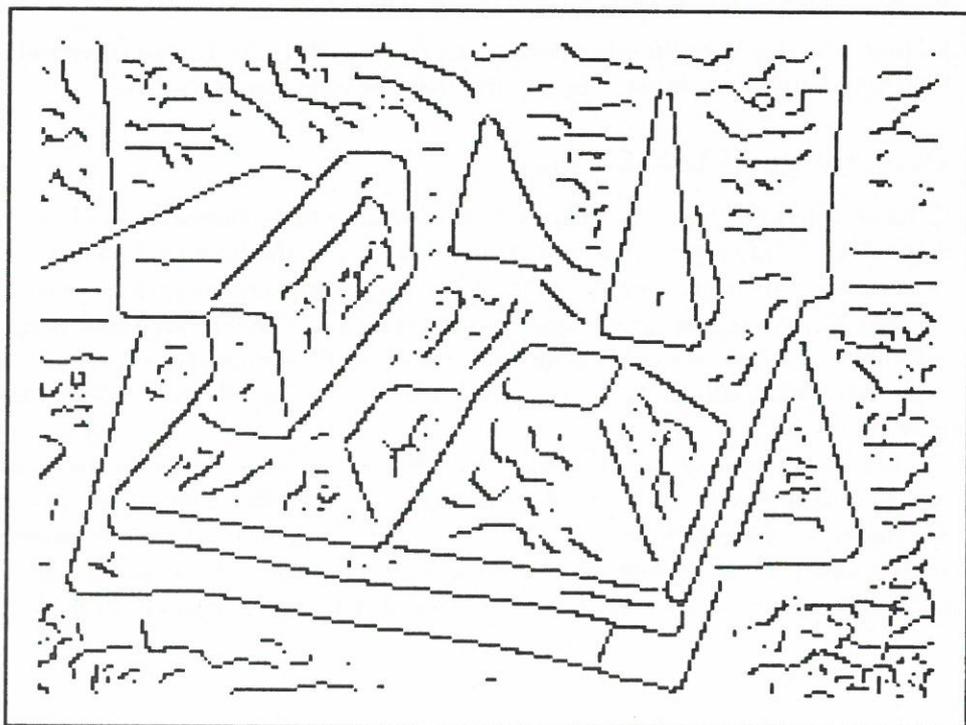


Bild 7: Ergebnis des Farbvektoransatzes für das Block-Bild ($\sigma = 2$)

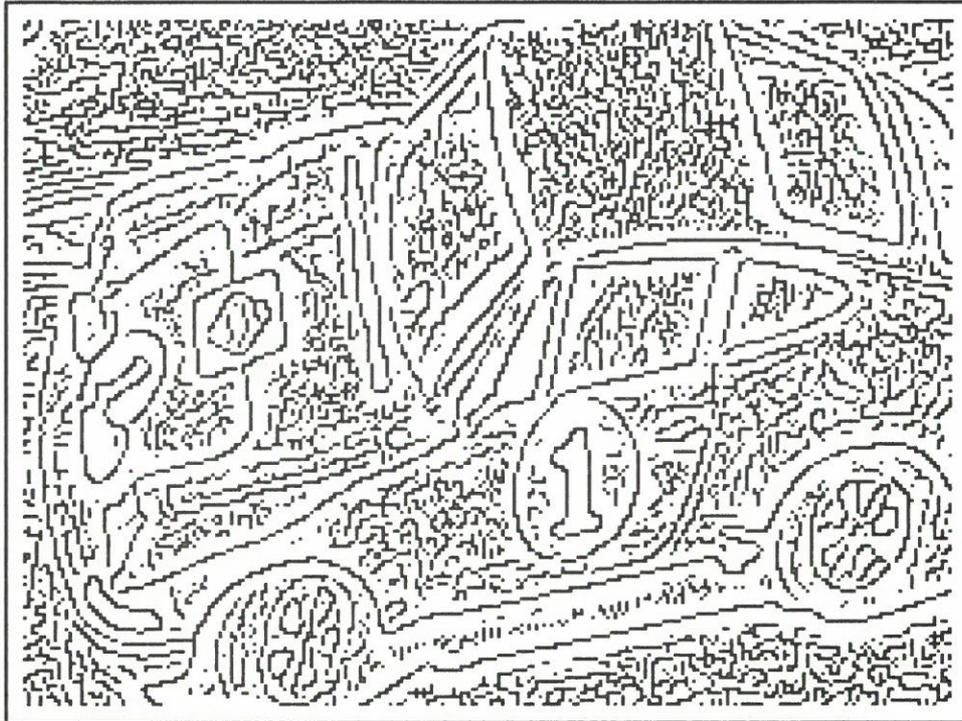


Bild 8: Ergebnis des Intensitätsansatzes für das Auto-Bild ($\sigma = 1$)

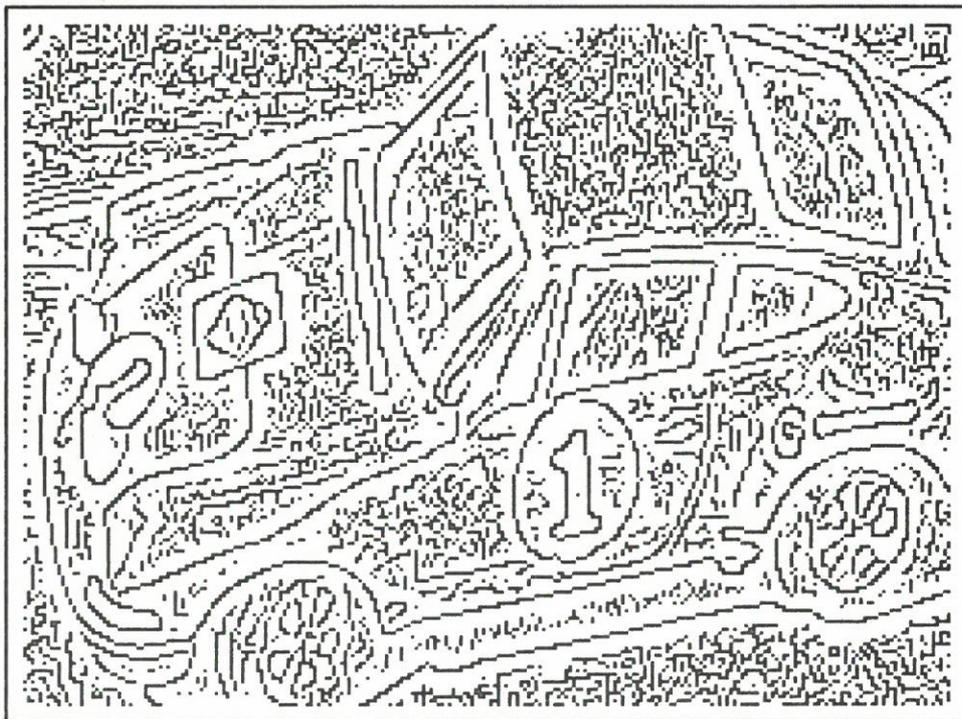


Bild 9: Ergebnis des Farbvektoransatzes für das Auto-Bild ($\sigma = 1$)

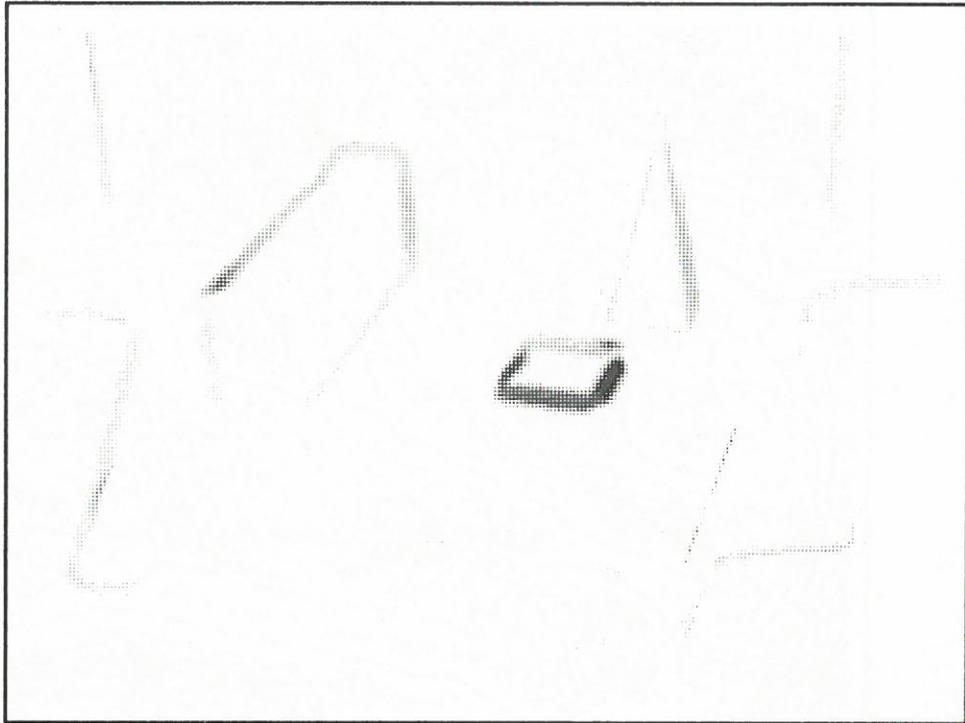


Bild 10: Differenz von Farbänderung und Intensitätsänderung für das Block-Bild

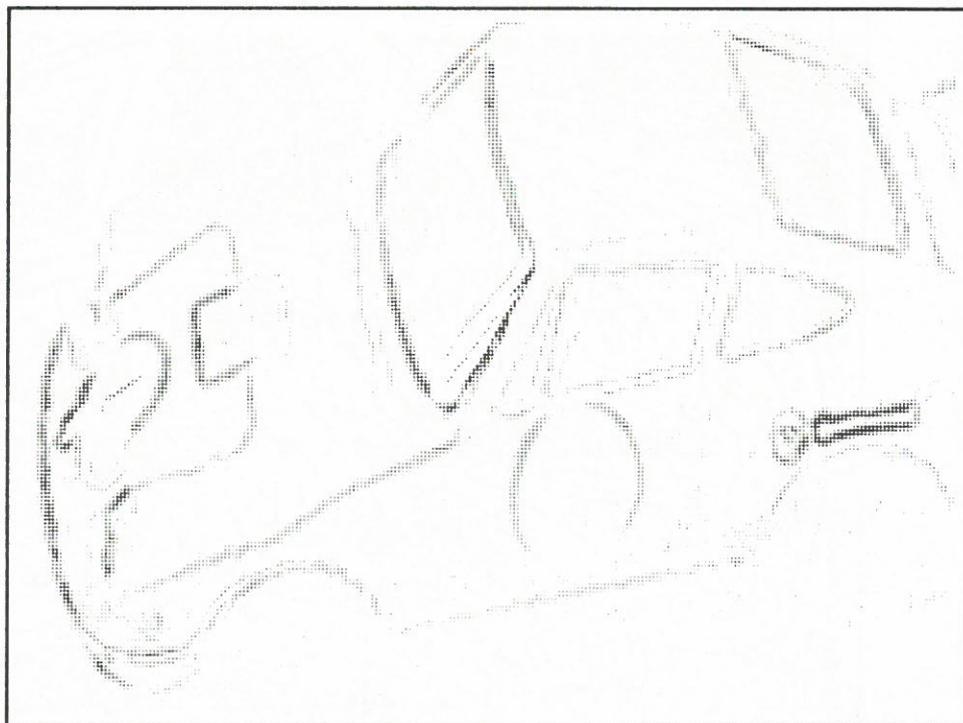


Bild 11: Differenz von Farbänderung und Intensitätsänderung für das Auto-Bild

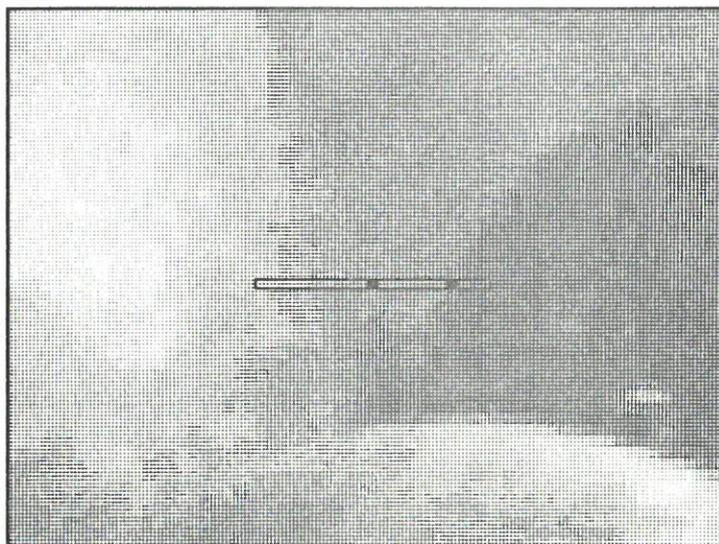


Bild 12: Ausschnitt des Apfel-Bildes mit gekennzeichneter Bildzeile und den mit Hilfe des Farbvektoransatzes detektierten Kanten



Bild 13: Ausschnitt des Apfel-Bildes mit gekennzeichneter Bildzeile und den mit Hilfe des Intensitätsansatzes detektierten Kanten

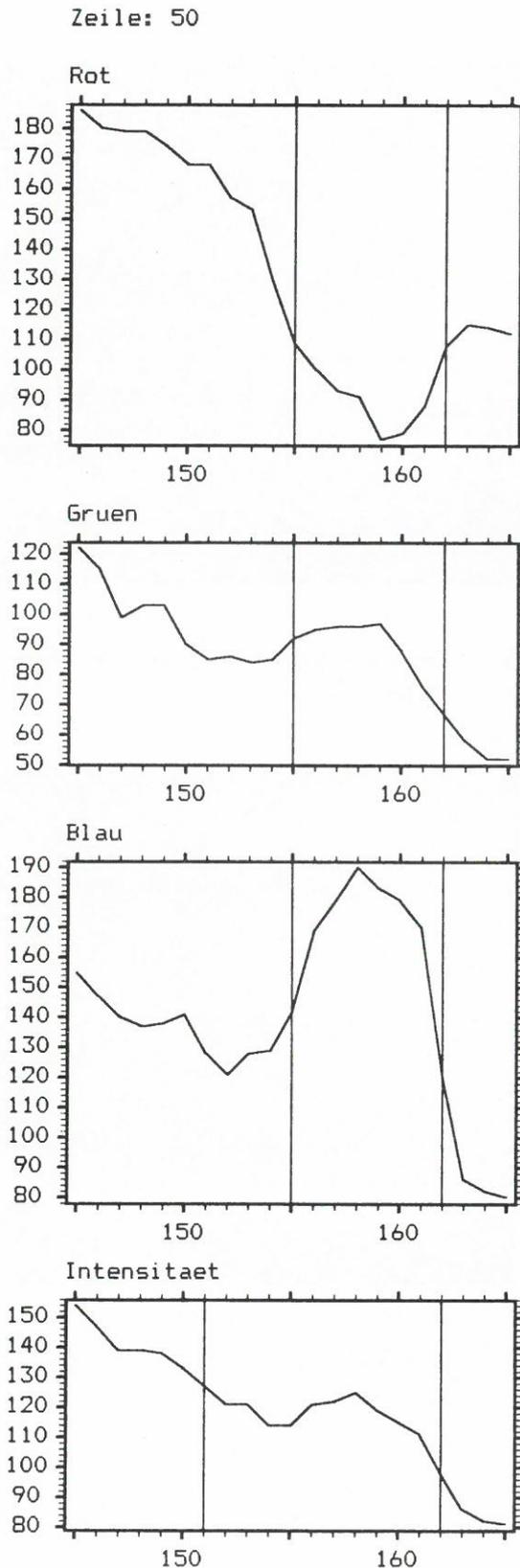


Abbildung 8.3: Profile für die gekennzeichnete Bildzeile (siehe Bild 12 und 13) mit den mit Hilfe des Farbvektoransatzes (rot, grün, blau) bzw. des Intensitätsansatzes (Intensität) detektierten Kanten

werden. Detektionslücken zeigen sich in erster Linie dort, wo sich mehrere Kanten im Bild schneiden oder sich sehr nahe kommen, und dort, wo auch für den menschlichen Betrachter ein Kontrast in Intensität oder Farbe kaum oder gar nicht wahrnehmbar ist (vergleiche hierzu auch die Originalbilder im Anhang). Auch die Lokalisation der Kanten scheint gut mit der eigenen visuellen Einschätzung übereinzustimmen.

Die gezeigten Bilder enthalten alle detektierten Kanten, ohne Rücksicht auf ihre Relevanz. Das Ziel einer sicheren Detektion ist aber auch, in erster Linie solche Strukturen im Bild zu detektieren, die einer für die Bildinterpretation relevanten Struktur in der Szene entsprechen. Wie wir schon öfter angedeutet haben, ist es fraglich, ob eine Entscheidung über die Relevanz eines detektierten Merkmals bereits auf dieser Verarbeitungsebene nur aufgrund lokaler Bildeigenschaften getroffen werden kann. Die Frage, in welcher Weise solche Entscheidungen innerhalb des im Rahmen des SISSY-Projektes zu entwickelnden Gesamtsystems getroffen werden sollen, ist noch ungeklärt. Wir wollen aber demonstrieren, daß es zumindest für die von uns verwendeten Bilder möglich ist, die wichtigsten Kantenstrukturen aus der Menge aller detektierten Kanten zu extrahieren. Die Bilder 15, 17 und 19 zeigen die Resultate des von uns entworfenen Selektionsalgorithmus. Die Selektionsparameter wurden für alle drei Bilder gleich gewählt, mit Ausnahme der zwei in das Verfahren eingehenden Schwellwerte, die von den Kontrastverhältnissen im jeweiligen Bild abhängen. Nähere Details zum Selektionsalgorithmus entnehme man Abschnitt 8.1.2. Obwohl die drei Bilder sehr unterschiedlich sind, werden mit Hilfe des Selektionsverfahrens, bei gleichgewählten Selektionsparametern, in allen Fällen sehr gute Ergebnisse erzielt. Dies hat sich in weiteren Versuchen mit anderen Bildern bestätigt.

Im Ergebnis für das Apfel-Bild fallen besonders die vielen Kanten auf den vorderen beiden Äpfeln auf. Die gekrümmte Oberfläche der Äpfel führt zu einer starken Variation in der Stärke der Reflexion und zu Glanzeffekten. Die auf der Oberfläche ermittelten Gradienten weisen deshalb hohe Beträge auf. In diesen Bereichen detektierte lokale Extrema setzen sich daher deutlich von Rausch- oder Textureffekten ab. Möglichkeiten zur Vermeidung derartiger Effekte haben wir bisher nicht untersucht (siehe beispielsweise [Binford 81] zum Stichwort *lateral inhibition*).

Im Ergebnis für das Block-Bild fällt auf, daß die Orientierungskante an der rechten Oberseite des langen Quaders nicht durch das Selektionsschema erhalten werden konnte. Auch die Berücksichtigung der Farbänderung reicht nicht aus, um diese Struktur von Rauscheffekten zu unterscheiden.

Detektion und Lokalisation an Schnitt- und Eckpunkten

Wir haben bereits an früherer Stelle darauf hingewiesen, daß gradientenbasierte Kantendetektionsansätze an Schnitt- und Eckpunkten von Kanten Schwächen zeigen. Das zugrundeliegende Stufenkantenmodell wird an solchen Orten im Bild verletzt. Es wird angenommen, daß im betrachteten Bildausschnitt nur eine einzige Kante vorhanden ist, die lokal geradlinig verläuft. Nur unter diesen Voraussetzungen kann man davon ausgehen, daß die Gradientenrichtung senkrecht zur Kante steht. Befinden sich im Bildaus-

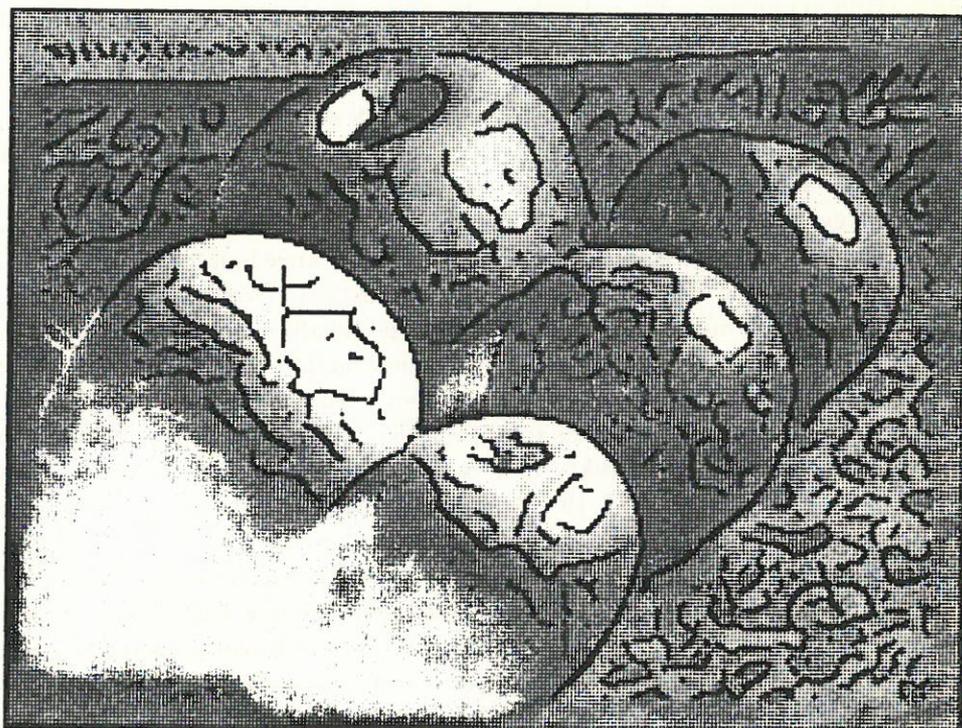


Bild 14: Ergebnis des Farbvektoransatzes, dem Apfel-Bild überlagert

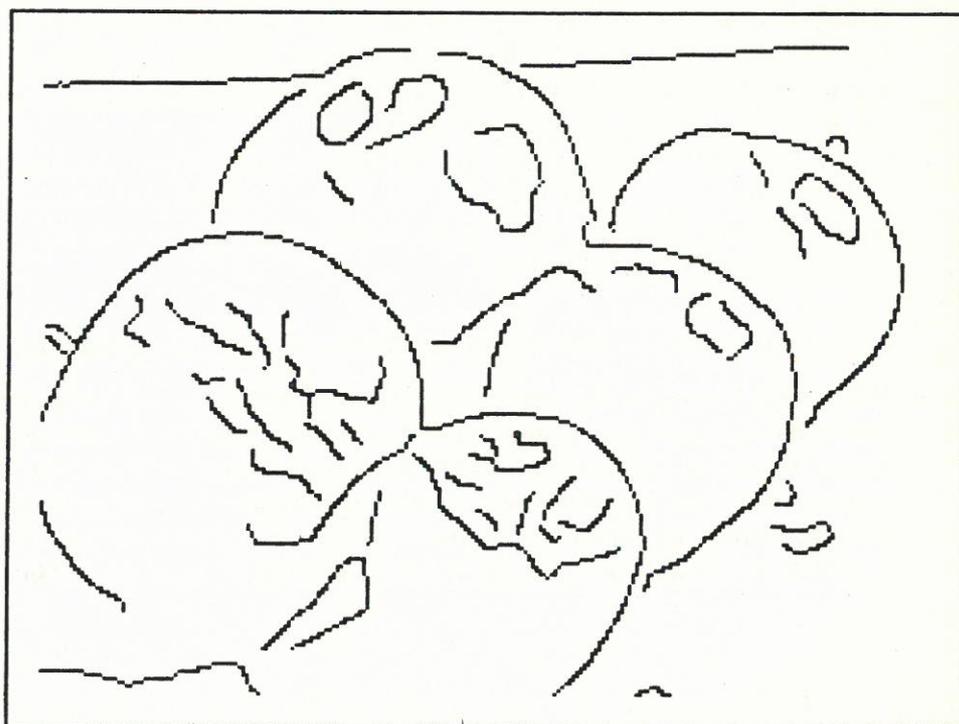


Bild 15: Ergebnis des Farbvektoransatzes für das Apfel-Bild, nach der Selektion

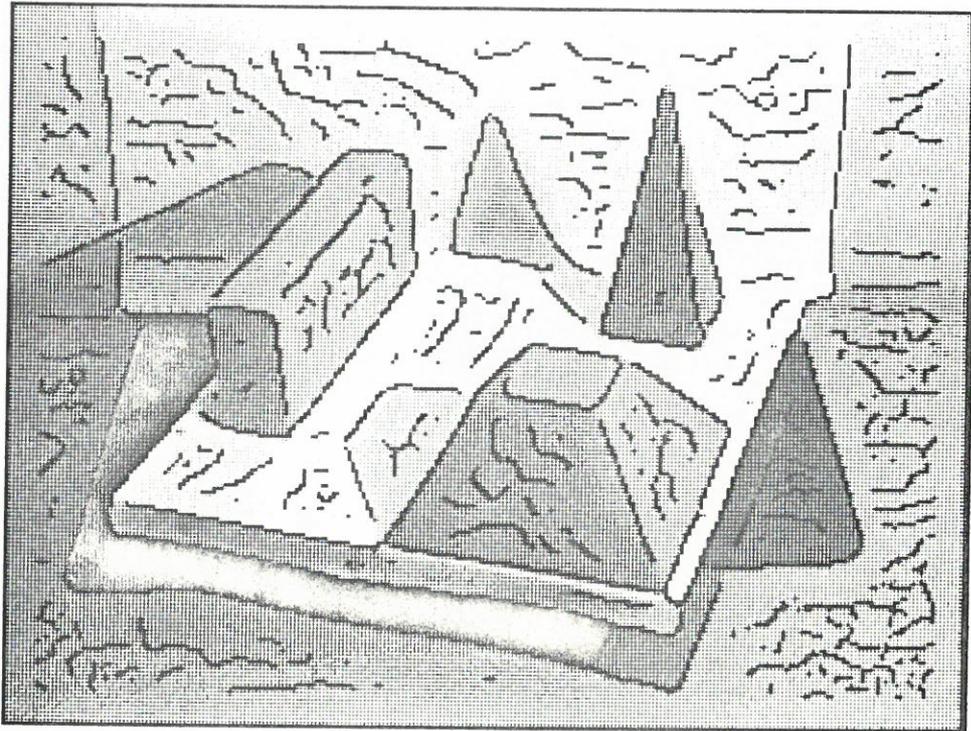


Bild 16: Ergebnis des Farbvektoransatzes, dem Block-Bild überlagert

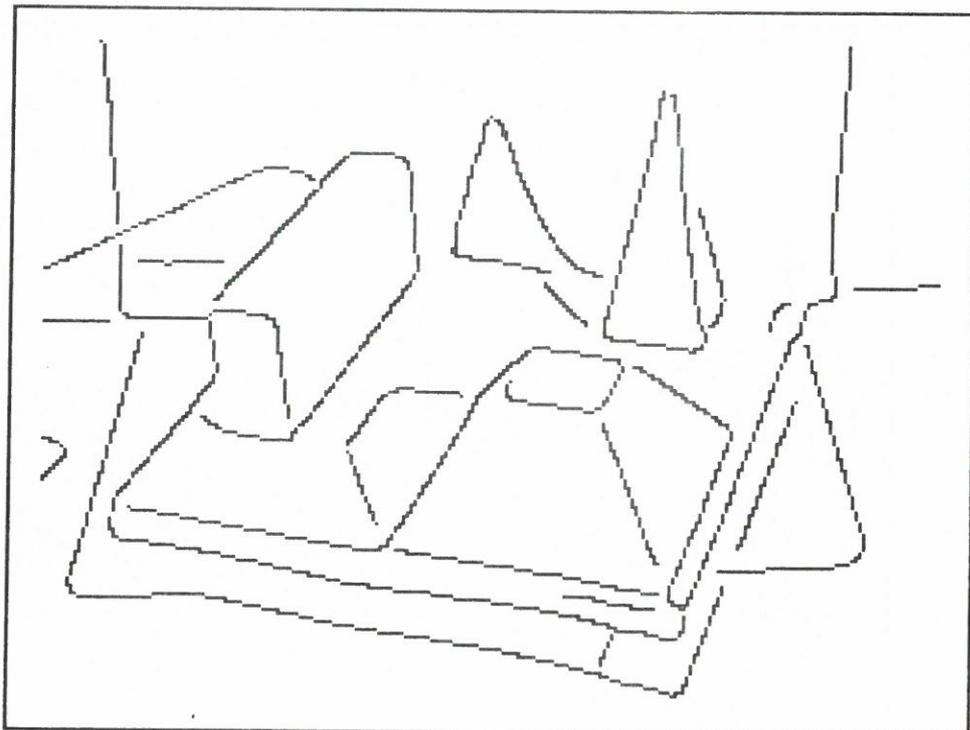


Bild 17: Ergebnis des Farbvektoransatzes für das Block-Bild, nach der Selektion

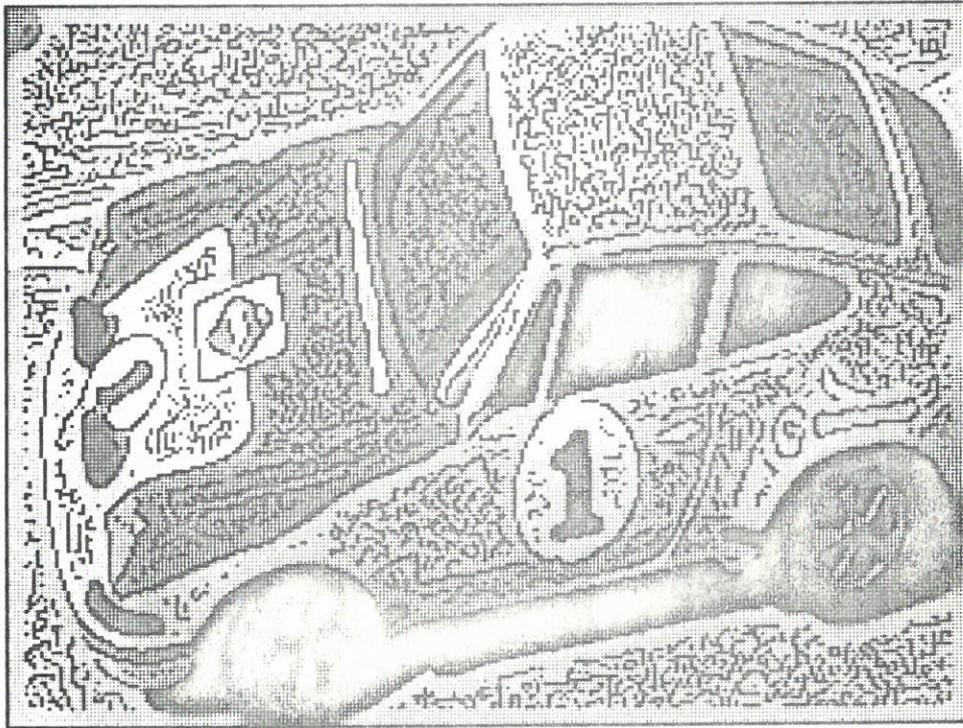


Bild 18: Ergebnis des Farbvektoransatzes, dem Auto-Bild überlagert

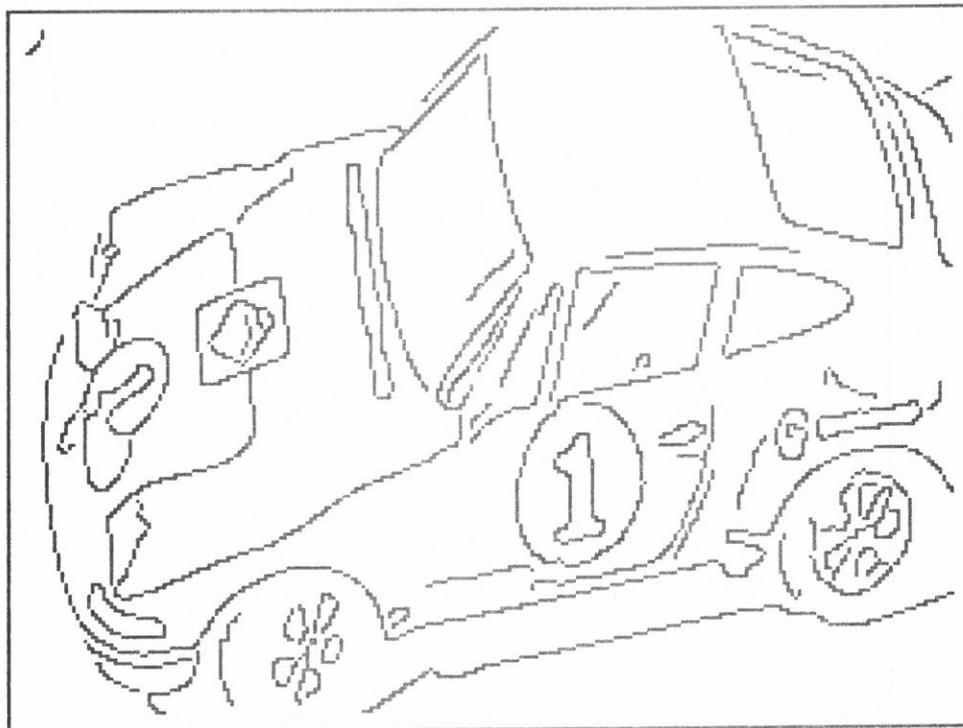


Bild 19: Ergebnis des Farbvektoransatzes für das Auto-Bild, nach der Selektion

	rot : 100
	grün: 100
	blau: 100
rot : 160	
grün: 140	
blau: 120	

Abbildung 8.4: Mittelwerte der Bildbereiche für das Bild einer Ecke. Varianz: $\sigma = 8.0$

schnitt mehrere Kanten, so entspricht die Schätzung des Gradienten einer Mittelung der einzelnen Kanteneffekte. Dabei wirken sich solche Kanten besonders stark aus, die einen hohen Kontrast aufweisen.

Bild 20 zeigt die mit Hilfe des Farbvektoransatzes bestimmten Änderungsbeträge für eine synthetisch erzeugte und verrauschte Ecke (schwarz entspricht starken Änderungen, weiß entspricht niedrigen Änderungen). Die Mittelwerte und Varianzen der verschiedenen Bereiche des synthetischen Bildes sind Abbildung 8.4 zu entnehmen. Das Bild wurde mit der Operatorweite $\sigma = 3$ verarbeitet. Bild 21 zeigt die ermittelten Gradientenrichtungen. Dort, wo im Bereich des Operators nur eine Kante vorhanden ist, ermöglicht der geschätzte Gradient die korrekte Detektion und Lokalisation einer Kante. Im Bereich der Ecke ergibt sich jedoch ein allmählicher Übergang der Gradientenrichtungen sowie eine gegenseitige Beeinflussung des horizontalen und des vertikalen Kontrastes. Das Resultat ist eine Abrundung der Ecke, die keinen ausgeprägten Eckpunkt aufweist. In Bild 22 sehen wir die Änderungsbeträge für einen ebenfalls synthetisch erzeugten Kreuzungspunkt. Die Mittelwerte und Varianzen der verschiedenen Bereiche sind Abbildung 8.5 zu entnehmen. Wieder wurde das verrauschte Bild mit der Operatorweite $\sigma = 3$ verarbeitet. Bild 23 zeigt die entsprechenden Gradientenrichtungen. Die Änderungsbeträge zeigen deutlich, daß die vertikal verlaufenden Kanten den stärksten Kontrast aufweisen. Sie gehen in die Mittelung im Bereich des Kreuzungspunktes am stärksten ein. Dies führt dazu, daß sich die Gradienten im Kreuzungsbereich auch noch rechts und links der vertikalen Kanten fast horizontal ausrichten. Bei der Suche nach lokalen Extrema wird der Vergleich der Änderungsbeträge in Gradientenrichtung durchgeführt. In horizontaler Richtung besitzt der Gradientenbetrag aber in der Bildmitte sein Maximum, so daß rechts und links davon keine Extrema detektiert werden. Ein Vergleich in vertikaler Richtung findet erst dort statt, wo der Einfluß der vertikalen Kanten so klein geworden ist, daß sich die Gradienten senkrecht zu den horizontalen Kanten ausrichten. Im Bereich eines Kreuzungspunktes werden also diejenigen Kanten durchgängig detektiert, die den stärksten Kontrast aufweisen. Die kontrastärmeren

rot : 90	rot : 140
grün: 100	grün: 30
blau: 100	blau: 80
rot : 120	rot : 100
grün: 150	grün: 50
blau: 80	blau: 30

Abbildung 8.5: Mittelwerte der Bildbereiche für das Bild eines Kreuzungspunktes. Varianz: $\sigma = 8.0$

Kanten brechen kurz vor dem eigentlichen Kreuzungspunkt ab.

Ein gradientenbasiertes Kantendetektionsverfahren kann im Bereich von Schnitt- und Eckpunkten weder eine sichere Detektion noch eine exakte Lokalisation gewährleisten. Die Redundanz zwischen den Ergebnissen des Kantendetektors und des Punktefinders ist deshalb nur bedingt nutzbar. Denkbar wäre eine Zuordnung zwischen beiden Merkmalsarten, wenn die Lokalisation der Merkmale gut übereinstimmt. Die exaktere Lokalisation sollten in diesen Fällen die markanten Punkte aufweisen. Grundsätzlich wäre es wünschenswert, wenn sich die Ergebnisse von Kantendetektor und Punktefinder ideal ergänzten. Dazu müßte ein Verfahren zur Detektion markanter Punkte gerade solche Fälle durch ein geeignetes Modell berücksichtigen, für die das Kantendetektionsverfahren Schwächen zeigt.

Stabilität von Merkmalen

Zur Beurteilung der Stabilität der detektierten Merkmale haben wir bisher keine Experimente durchgeführt. Darüberhinaus gilt auch hier, daß eine aussagekräftige Bewertung erst möglich ist, wenn die detektierten Kanten bei der Korrespondenzanalyse eingesetzt werden. Es müßte außerdem geklärt sein, in welcher Form die Kanteninformation in die Analyse eingeht.

Um zu veranschaulichen, welche Veränderungen in den Ergebnissen für die Bilder einer Bildfolge zu erwarten sind, haben wir die bereits bekannte Block-Szene bei verschiedenen Stellungen des in der Szene sichtbaren Drehtellers aufgenommen. Die Stellungen des Drehtellers unterscheiden sich von Aufnahme zu Aufnahme um einen Winkel von 10° . Die Bilder 24 bis 27 zeigen die bereits nachverarbeiteten Ergebnisse des Farbvektoransatzes den Intensitätsbildern überlagert. Zwischen den Ergebnissen der einzelnen Bilder zeigt sich eine gute Übereinstimmung: Die Menge der als Kante im Bild detektierten Kanten in der Szene stimmt für alle Bilder fast überein, mit Ausnahme solcher

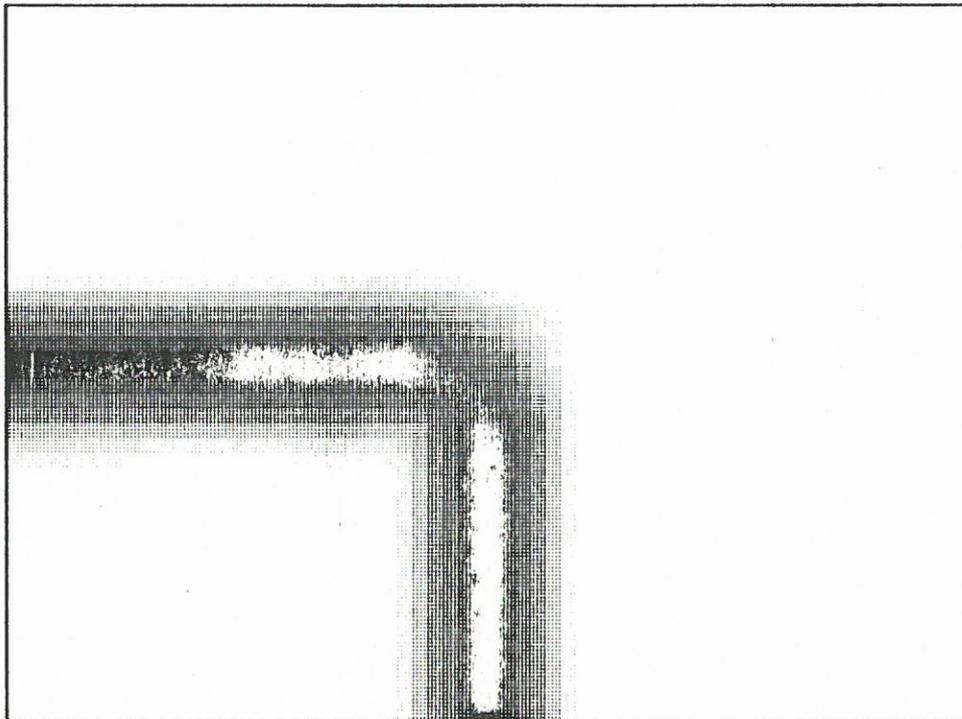


Bild 20: Mit Hilfe des Farbvektoransatzes ermittelte Änderungsbeträge für eine verrauschte Ecke (schwarz entspricht starken Änderungen). Siehe dazu Abbildung 8.4

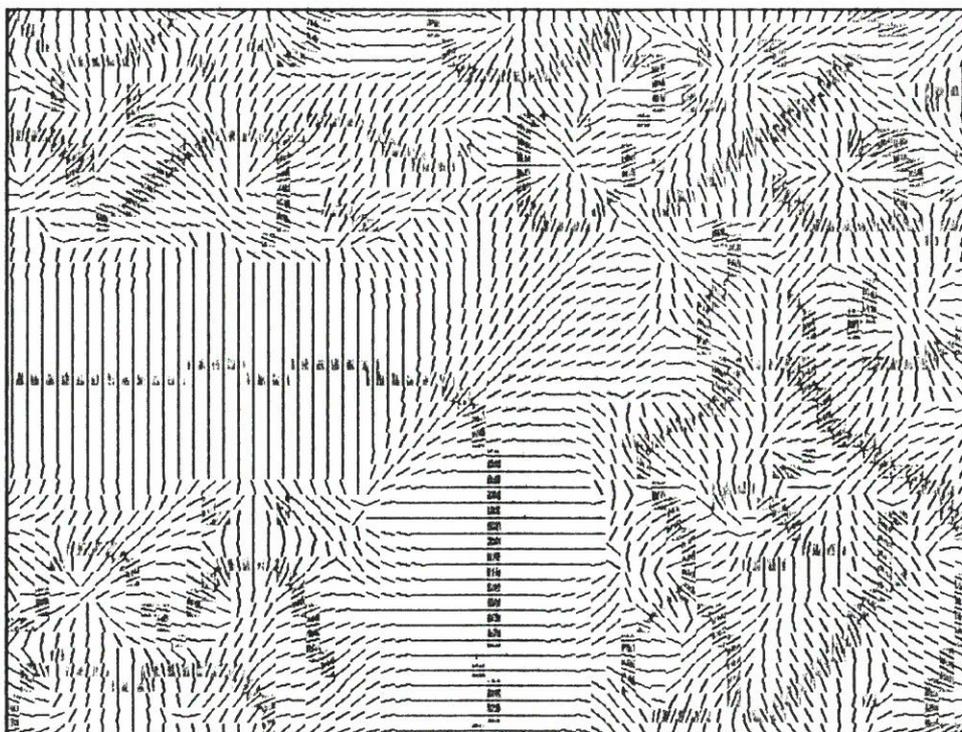


Bild 21: Mit Hilfe des Farbvektoransatzes ermittelte Richtungen stärkster Änderung für eine verrauschte Ecke. Lokale Extrema sind grau markiert. Siehe dazu Abbildung 8.4

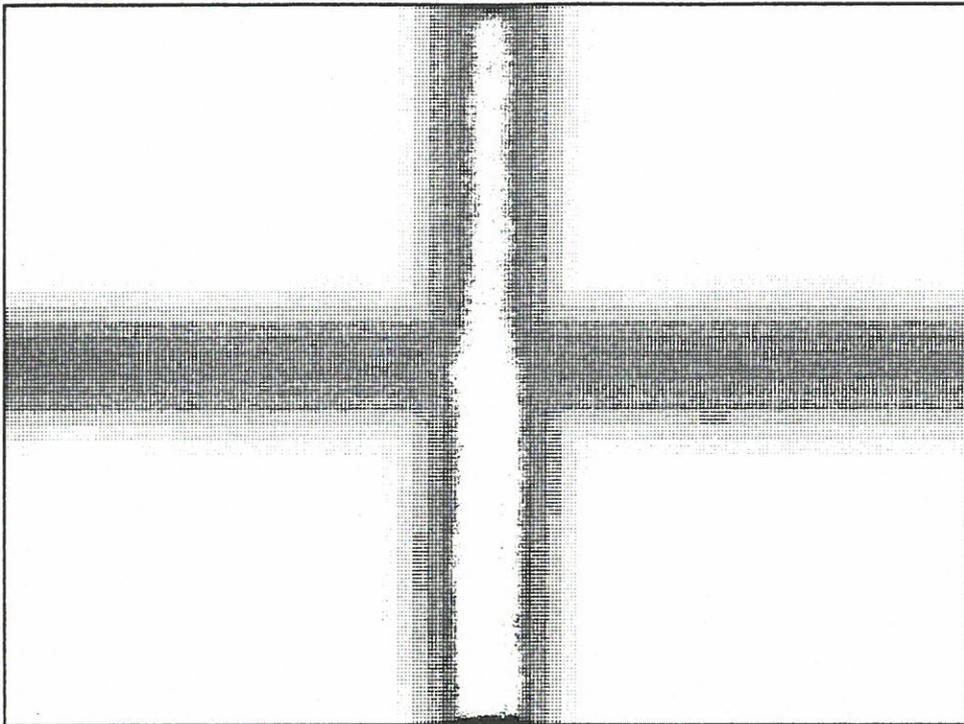


Bild 22: Mit Hilfe des Farbvektoransatzes ermittelte Änderungsbeträge für einen verrauschten Kreuzungspunkt (schwarz entspricht starken Änderungen). Siehe dazu Abbildung 8.5

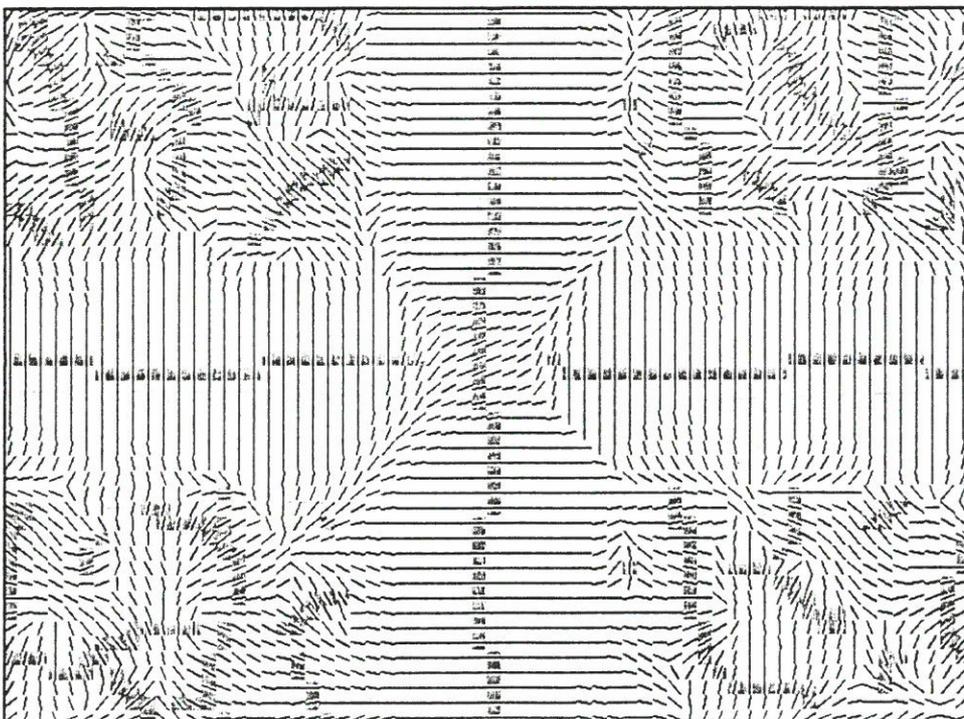


Bild 23: Mit Hilfe des Farbvektoransatzes ermittelte Richtungen stärkster Änderung für einen verrauschten Kreuzungspunkt. Lokale Extrema sind grau markiert. Siehe dazu Abbildung 8.5

Kanten, die durch die Rotation nur in einigen Bildern sichtbar sind. Ein auffälliges Gegenbeispiel stellt die rechte, vordere Orientierungskante des Pyramidenstumpfes dar: Durch die Rotation des Drehtellers verändern sich die Beleuchtungsverhältnisse derart, daß der an dieser Kante bestimmte Kontrast in den letzten beiden Bildern zu einer Detektion und Selektion ausreicht.

Erkennbar ist in den Ergebnissen auch, daß die selektierten Kantenzüge in der Regel sehr lang sind. Dies liegt daran, daß sich die Kantenzüge an zu erwartenden Kreuzungspunkten nicht tatsächlich schneiden. Kantenzüge, die unterschiedlichen Kanten in der Szene entsprechen können, werden stattdessen miteinander verschmolzen. Wie die Kanten im Bereich eines Kreuzungspunktes miteinander verbunden werden, hängt sehr von den lokalen Kontrastverhältnissen ab. Diese können aber von Bild zu Bild variieren. Ein gutes Beispiel stellt dafür der den langen Quader umfassende Kantenzug dar: In den Stellungen 0° bis 20° umfasst dieser Kantenzug den Quader fast vollständig. In der Stellung 30° haben sich die Kontrastverhältnisse derart verändert, daß der Kantenzug am oberen linken Ende des Quaders in den Schatten hinein fortgesetzt wird. Konsequenz der unsicheren Detektion und Lokalisation an Kreuzungspunkten ist also eine Instabilität im Ergebnis der Segmentierung. Um derartige Effekte zu vermeiden, müßten die Kantenzüge so zerlegt werden, daß die einzelnen Teilkanten tatsächlich einzelnen Kanten in der Szene entsprechen (siehe z.B. [Asada + Brady 86]). Dies würde außerdem die Verwendung der segmentierten Kantenzüge in der Korrespondenzanalyse erleichtern.

8.2.4 Zusammenfassung

Die in unseren Experimenten gewonnenen Erkenntnisse lassen sich wie folgt zusammenfassen:

- Die theoretischen Erwartungen an die Ergebnisse der verschiedenen Detektionsansätze lassen sich gut praktisch bestätigen.
- Insbesondere bestätigt sich die erwartete Überlegenheit des Farbvektoransatzes gegenüber dem Intensitätsansatz sowie gegenüber den alternativen Farbansätzen. Unter den alternativen Farbansätzen werden die besten Resultate bei Auswahl des maximalen Einzelgradienten erzielt.
- Im Vergleich zum Intensitätsansatz weisen die Ergebnisse des Farbvektoransatzes eine größere Zuverlässigkeit auf, da Kanten auch in solchen Bereichen noch sicher detektiert werden, in denen die Intensitätsinformation allein zu einer Detektion nicht ausreicht.
- Die überwiegende Zahl relevanter Kantenstrukturen in den verwendeten Bildern wird mit Hilfe des Farbvektoransatzes detektiert und, soweit bewertbar, gut lokalisiert.
- Das vorgestellte Selektionsschema demonstriert, daß die Selektion der relevanten Strukturen aus der Menge aller detektierten Kanten möglich sein sollte.

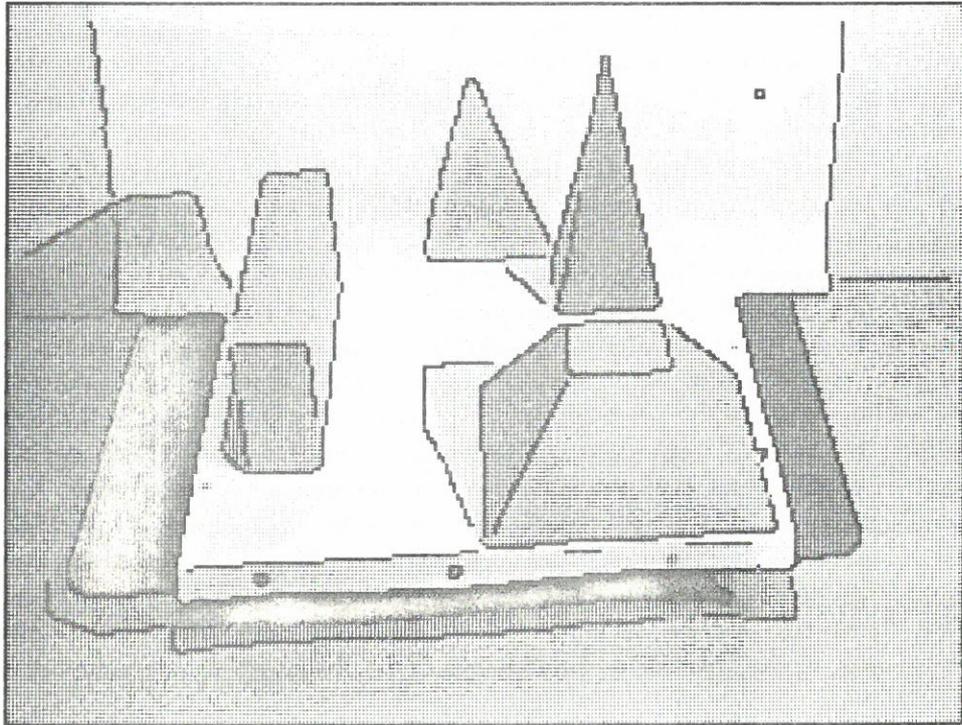


Bild 24: Ergebnis des Farbvektoransatzes für das Block-Bild (Stellung 0°)

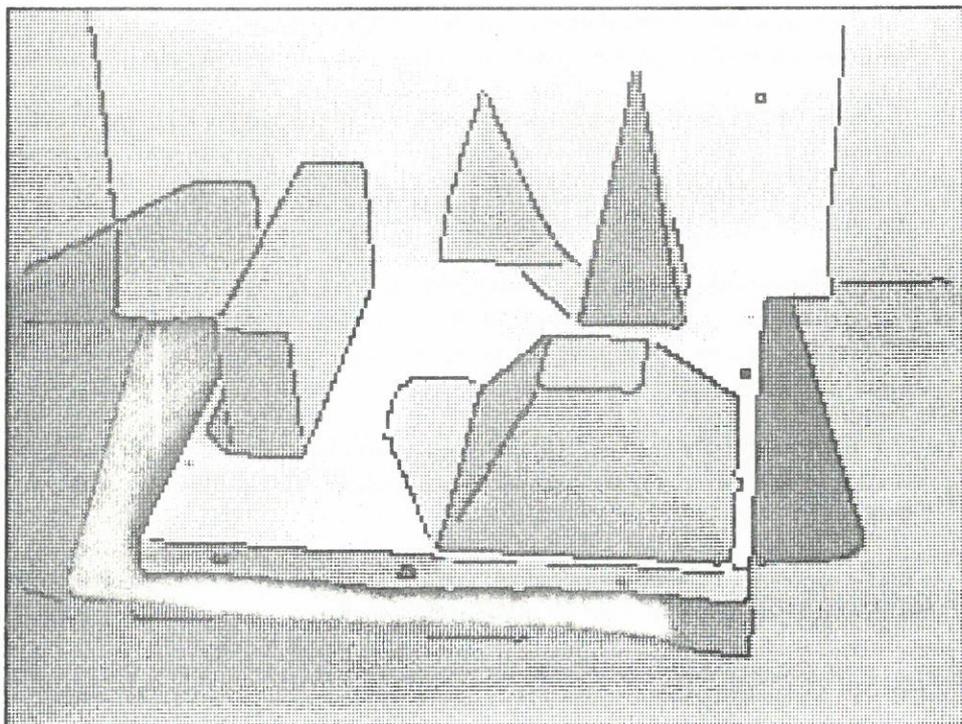


Bild 25: Ergebnis des Farbvektoransatzes für das Block-Bild (Stellung 10°)

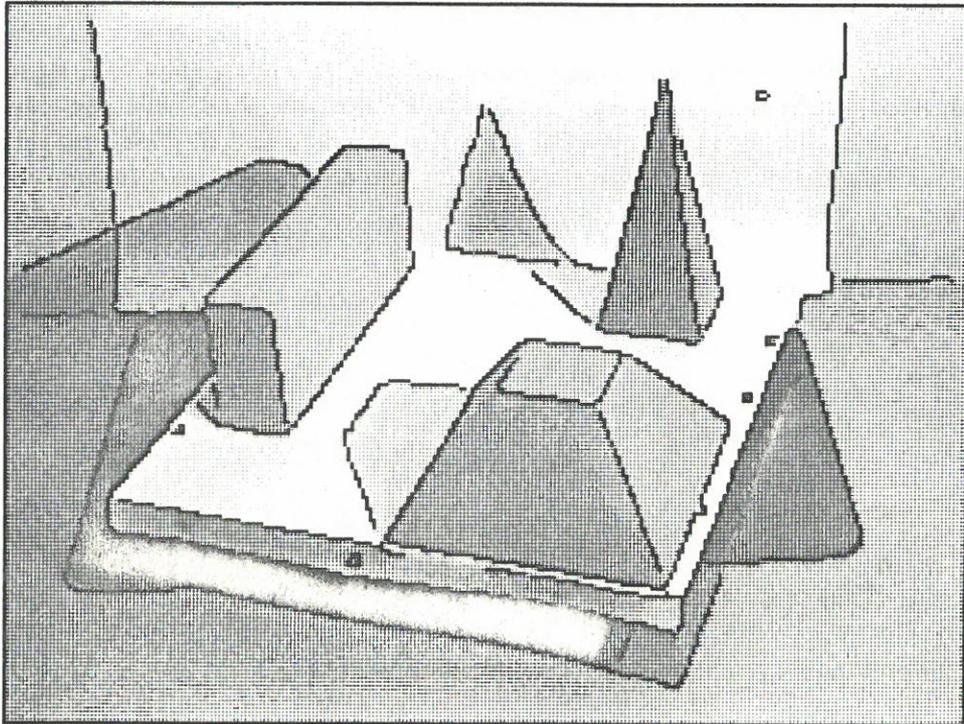


Bild 26: Ergebnis des Farbvektoransatzes für das Block-Bild (Stellung 20°)

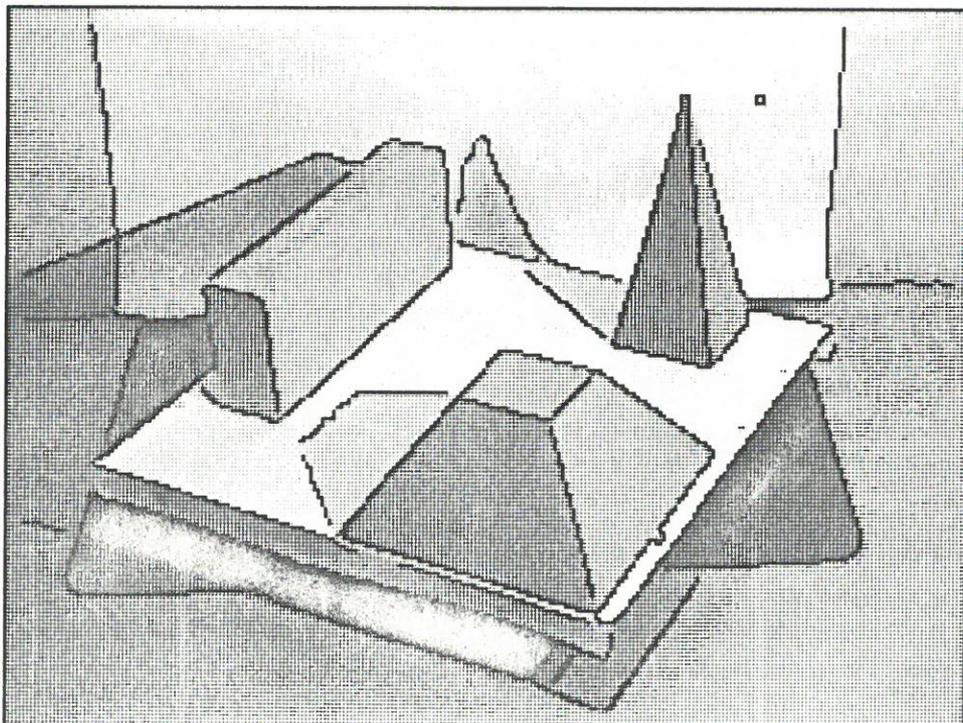


Bild 27: Ergebnis des Farbvektoransatzes für das Block-Bild (Stellung 30°)

- Schwächen zeigt der Farbvektoransatz als gradientenbasierter Ansatz erwartungsgemäß an Eck- und Kreuzungspunkten von Kanten.

Eine abschließende Beurteilung der Ergebnisse wird erst möglich sein, wenn die durch den Farbvektoransatz detektierten Kanten in der Korrespondenzanalyse eingesetzt werden. Dann wird es auch möglich sein, eine Antwort auf die Frage zu geben, ob der Einsatz von Farbe den größeren Rechenaufwand durch eine entscheidende Verbesserung der Gesamtergebnisse rechtfertigt.

Kapitel 9

Zusammenfassung und Ausblick

Der Schwerpunkt unserer Untersuchungen lag auf der Frage, ob es möglich ist, einen Ansatz zur Detektion von Kanten in Farbbildern zu entwickeln, der dem Stand der Kunst in der Grauwertbildverarbeitung gerecht wird. Wir haben dazu zunächst eine Einführung in die Thematik der Kantendetektion gegeben und sowohl die klassischen als auch aktuelle Grauwertansätze vorgestellt. Der anschließende Überblick über Ansätze zur Detektion von Kanten in Farbbildern zeigte, daß die bislang bekannten Farbansätze weder von ihrer Theorie noch von ihren Ergebnissen her zufriedenstellen können.

Die ersten Grundlagen für einen fundierten Farbansatz lieferten *Di Zenzo 86* und *Novak + Shafer 87*. Den von ihnen vorgeschlagenen Lösungsansatz haben wir ausführlich dargestellt und theoretisch diskutiert. Wir konnten zeigen, daß es auf dieser Basis möglich ist, auf formal fundierte Weise einen gradientenbasierten Farbansatz (Farbvektoransatz) zu entwickeln, der die Aspekte aktueller Grauwertverfahren berücksichtigt. In einer theoretischen Analyse wurden die Unterschiede zwischen dem Farbvektoransatz, dem Intensitätsansatz und verschiedenen alternativen Farbansätzen aufgezeigt. Die theoretisch zu erwartende Überlegenheit des Farbvektoransatzes gegenüber allen anderen untersuchten Detektionsansätzen konnte in verschiedenen Experimenten bestätigt werden. Für mehrere Bilder wurden mit dem Farbvektoransatz sehr gute Ergebnisse erzielt. Auch das von uns vorgestellte Selektionsverfahren konnte sich praktisch gut bewähren.

Der dem Farbvektoransatz zugrundeliegende Lösungsansatz liefert über die Anwendung zur Kantendetektion in Farbbildern hinaus die Grundlage für Verfahren zur Detektion von Diskontinuitäten in beliebigdimensionalen Bildfunktionen, deren Komponenten nicht auf gemessene Intensitäten beschränkt sind. Weitere Untersuchungen sollten die Anwendbarkeit dieses Ansatzes auf andere bildhafte Datenstrukturen an konkreten Beispielen überprüfen, sowie den Einfluß des zur Repräsentation von Farbbildern gewählten Farbsystems auf die Ergebnisse klären.

Auch die Implementation läßt einige Aufgaben für die Zukunft offen:

- Es muß untersucht werden, wie die Ergebnisse verschiedener Auflösungsebenen zu einem Gesamtergebnis integriert werden können.

- Das von uns verwendete Selektionsverfahren liefert zwar gute Ergebnisse; es fehlt jedoch bisher eine tiefgehende Analyse der mit der Segmentierung und Selektion von Kanten verbundenen Probleme.
- Insbesondere sollte geprüft werden, wie die in eine Nachverarbeitung eingehenden Schwellwerte automatisch in Abhängigkeit vom Bild bestimmt werden können.

Schließlich ist zu untersuchen, wie sich die gewonnene Kanteninformation in der Korrespondenzanalyse einsetzen läßt und ob es möglich ist, aufgrund lokaler Bildeigenschaften die detektierten Kanten nach ihrem physikalischen Ursprung zu klassifizieren.

Anhang A

Originalbilder

In diesem Anhang sind die Originalbilder der drei in den Experimenten verwendeten Aufnahmen abgebildet. Die Bilder wurden mit Hilfe einer RGB-Röhrenkamera aufgenommen und besitzen nach der Trennung der Halbbilder und einer geometrischen Entzerrung eine Auflösung von 192 Zeilen und 256 Spalten. Die Rot-, Grün- und Blauwerte werden jeweils durch 8 bit repräsentiert. Da farbige Darstellungen in diesem Bericht technisch nicht möglich sind, werden zu jedem Bild die drei Farbauszüge sowie das Intensitätsbild gezeigt.

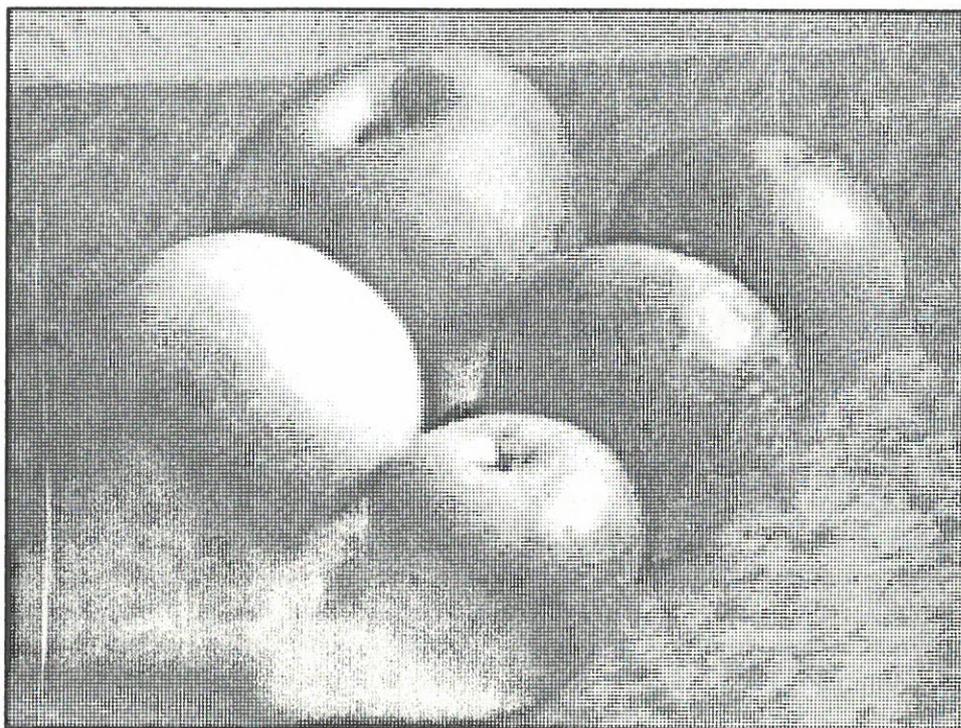


Bild 28: Apfelbild, Intensität

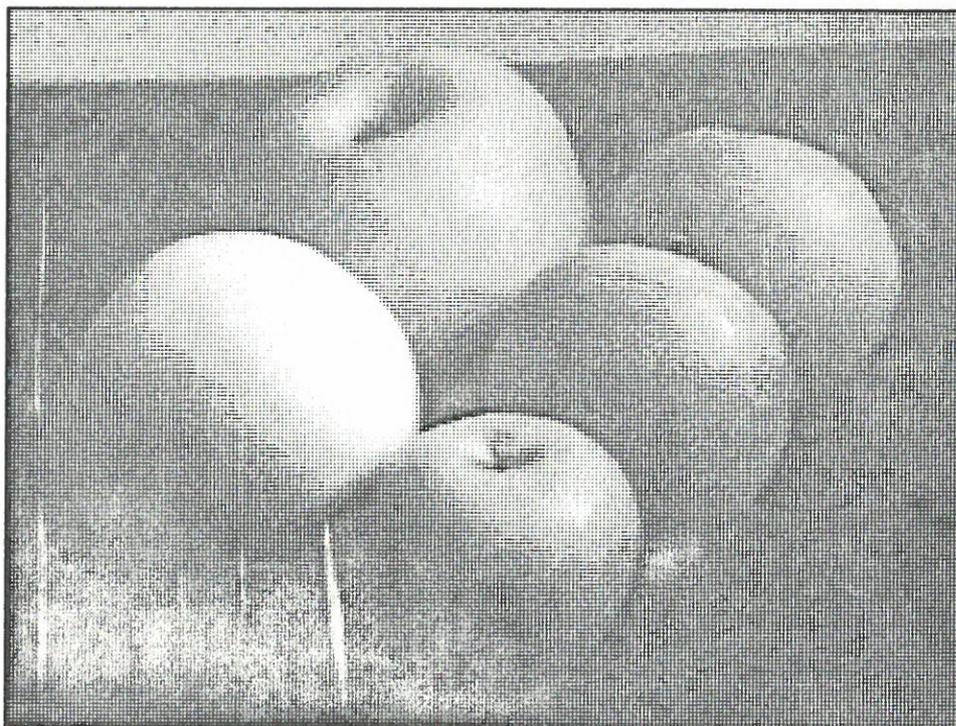


Bild 29: Apfelbild, Rotauszug



Bild 30: Apfelbild, Grünauszug

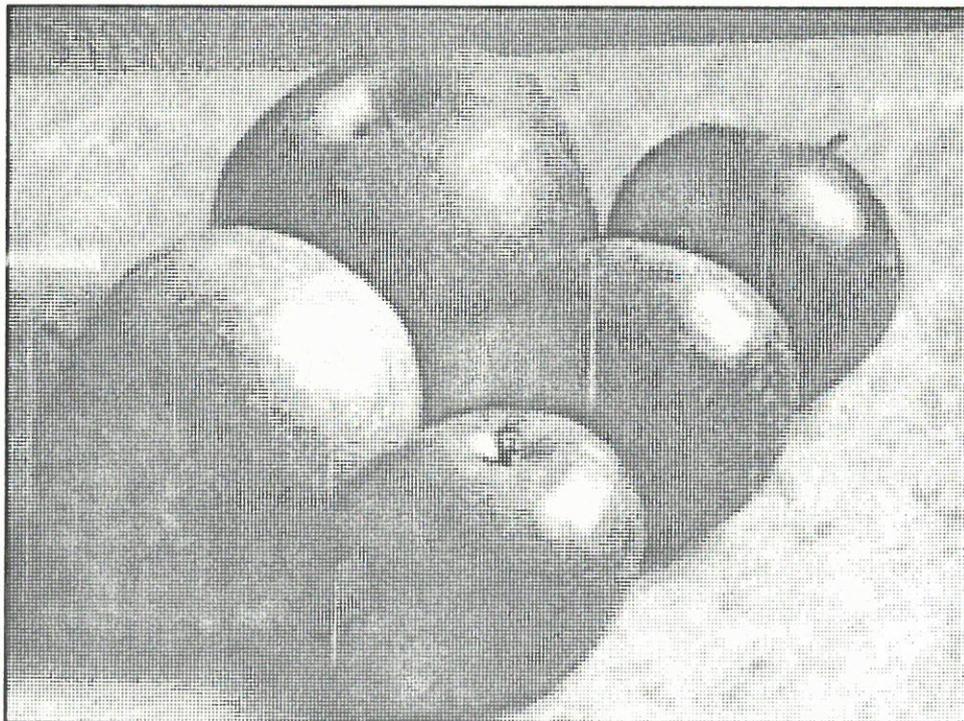


Bild 31: Apfelbild, Blauauszug

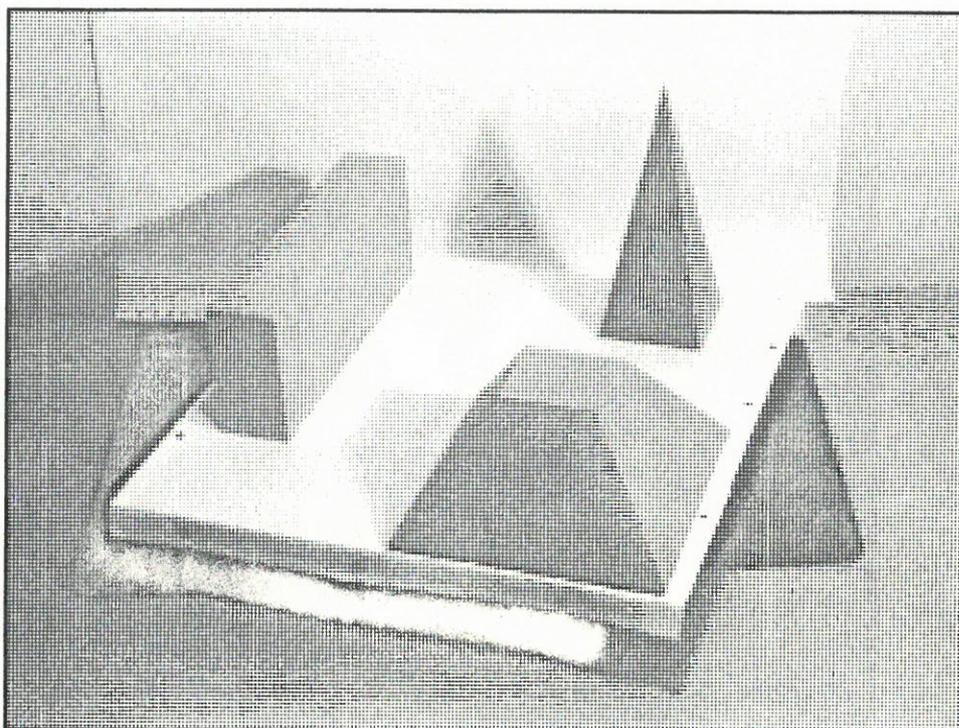


Bild 32: Blockbild, Intensität

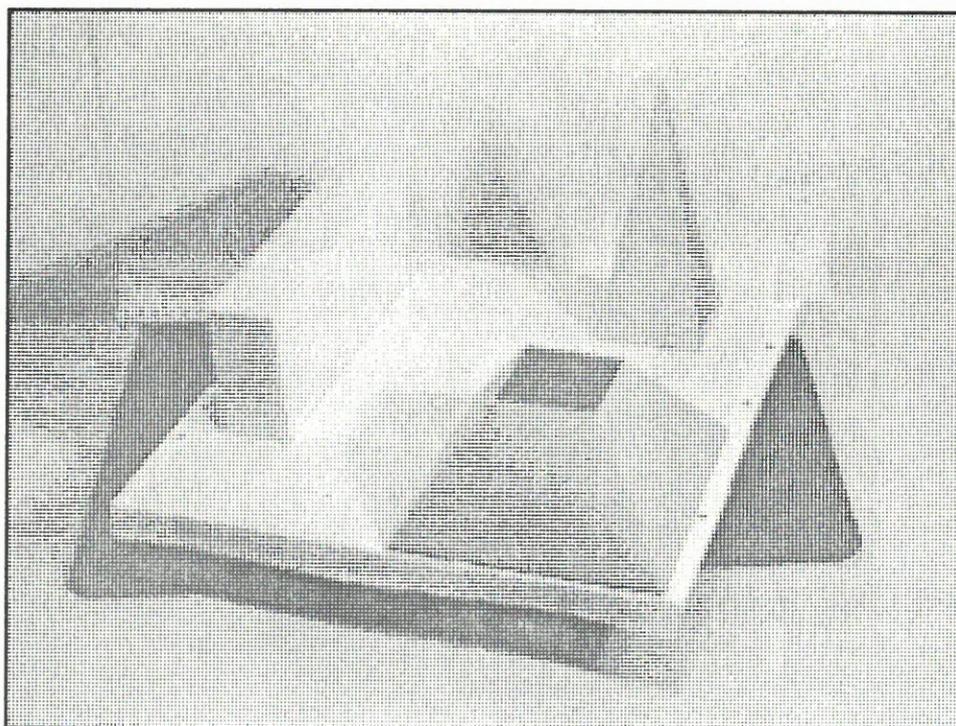


Bild 33: Blockbild, Rotauszug

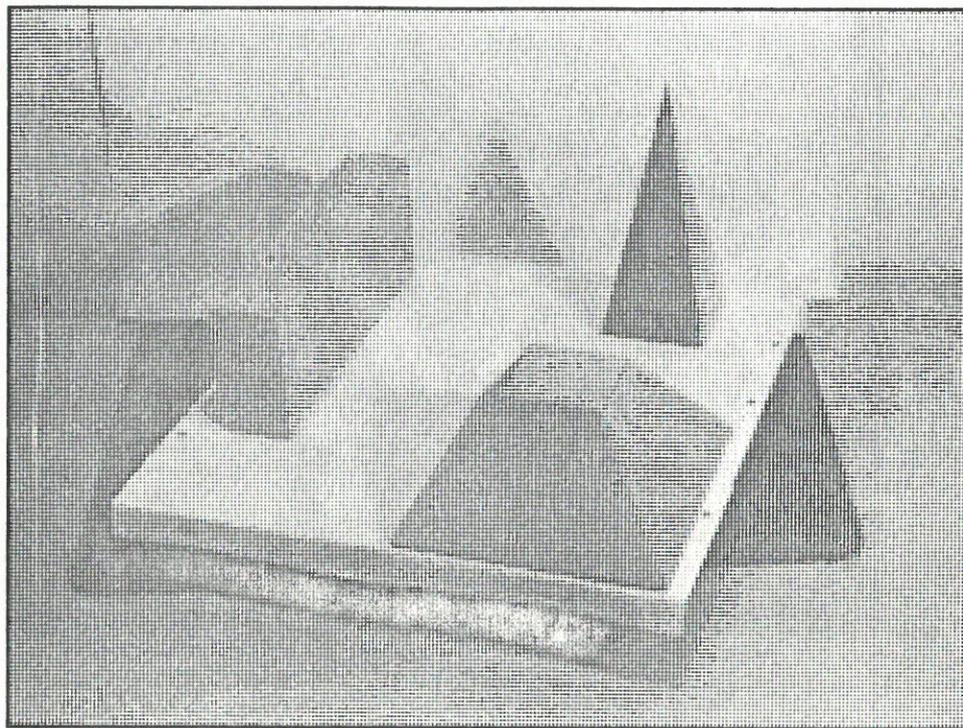


Bild 34: Blockbild, Grünauszug

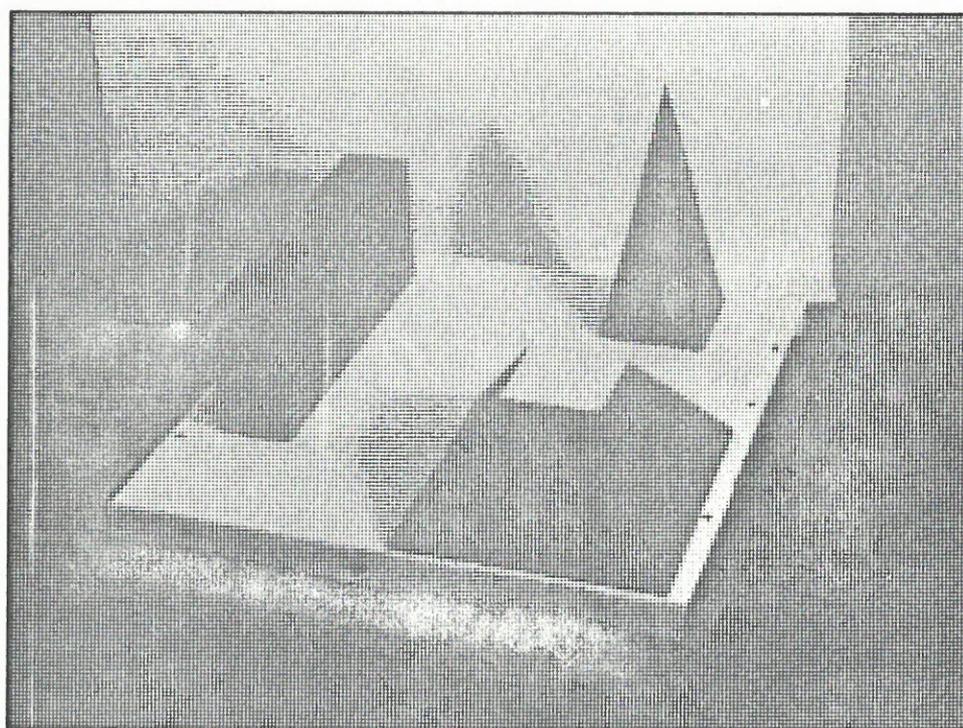


Bild 35: Blockbild, Blauauszug

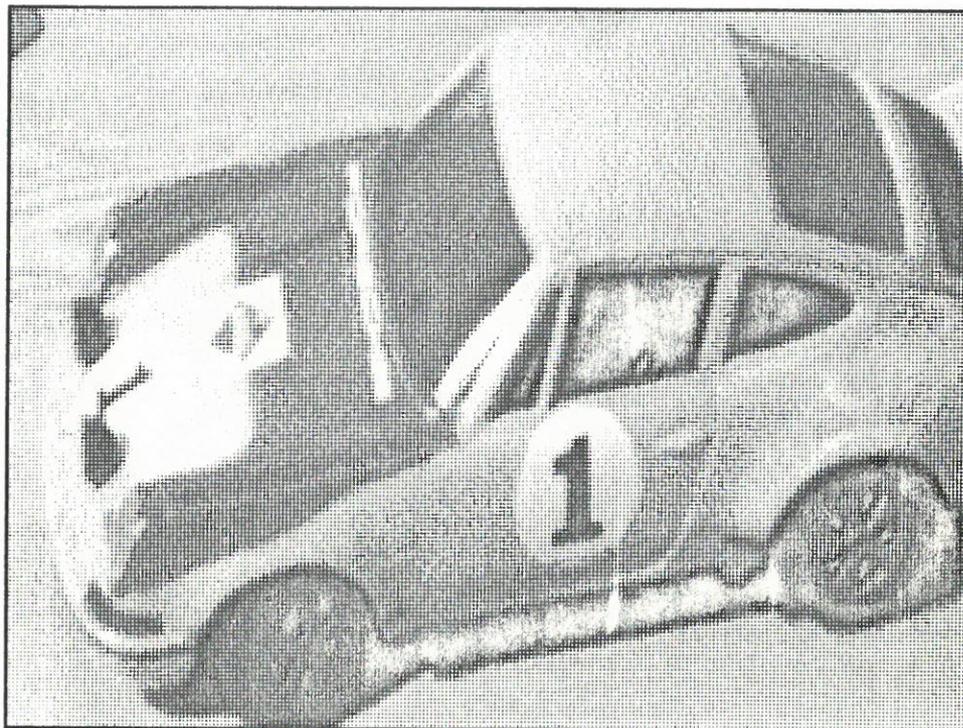


Bild 36: Autobild, Intensität

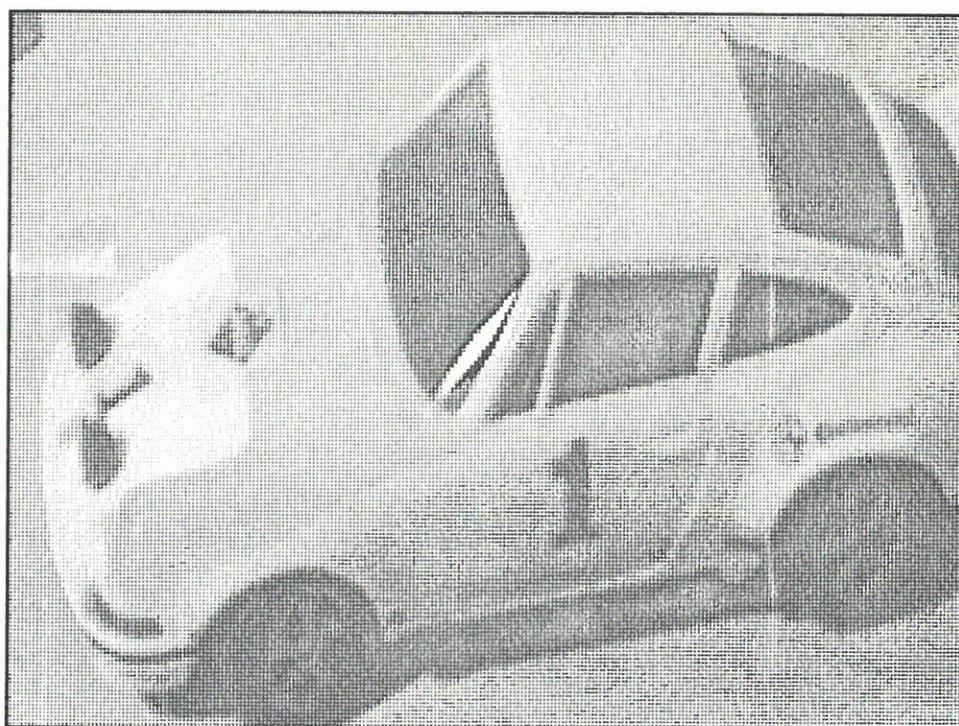


Bild 37: Autobild, Rotauszug

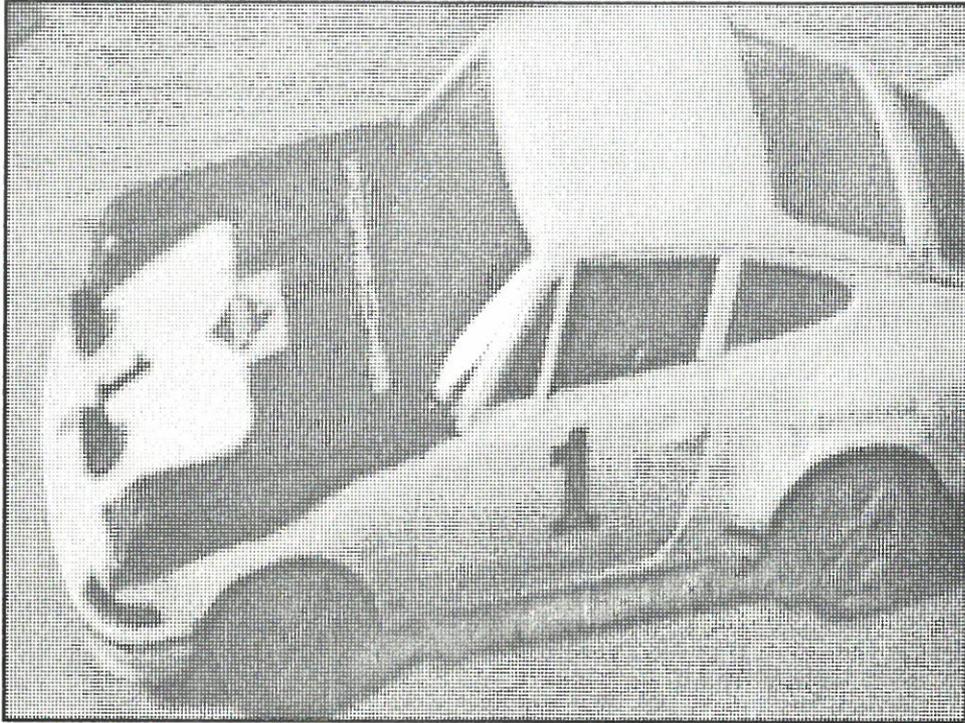


Bild 38: Autobild, Grünauszug

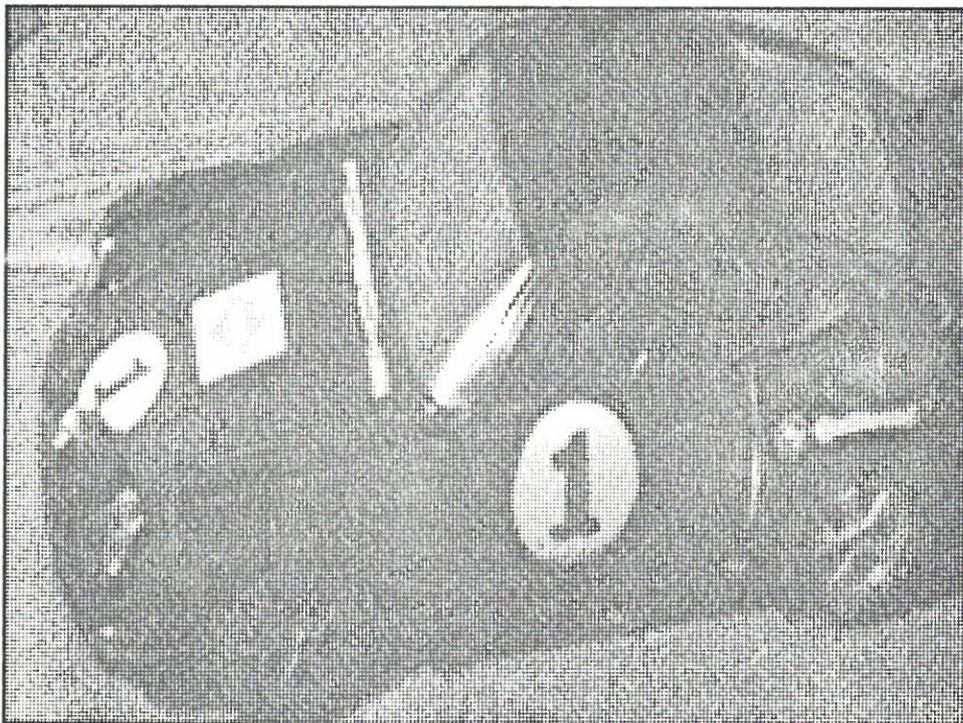


Bild 39: Autobild, Blauauszug

Literaturverzeichnis

Asada + Brady 86

The Curvature Primal Sketch.

Haruo Asada, Michael Brady.

IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-8** (1986) 2-14.

Ballard + Brown 82

Computer Vision.

Dana H. Ballard, Christopher M. Brown.

Prentice-Hall, Englewood Cliffs, New Jersey 1982.

Beaudet 78

Rotationally Invariant Image Operators.

Paul R. Beaudet.

Proc. 4th Int. Joint Conf. Pattern Recognition, Kyoto, Japan, November 1978, T. Sakai (Hrsg.), IEEE Computer Society Press, Washington, D.C. 1978, 579-583.

Binford 81

Inferring Surfaces from Images.

Thomas O. Binford.

Artificial Intelligence **17** (1981) 205-244.

Blicher 84

Edge Detection and Geometric Methods in Computer Vision.

Peter Blicher.

Ph.D. Dissertation, Dep. Math., Univ. California, Berkeley, CA, 1984.

Canny 83

Finding Edges and Lines in Images.

John F. Canny.

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

Canny 83b

A Variational Approach to Edge Detection.

John F. Canny.

Proc. National Conference on Artificial Intelligence AAAI-83, Washington, D.C., August 1983, William Kaufmann Inc., Los Altos, CA 1983, 54-58.

Canny 86

A Computational Approach to Edge Detection.

John F. Canny.

IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-8** (1986) 679-698.

Davis 75

A Survey of Edge Detection Techniques.

Larry S. Davis.

Computer Graphics and Image Processing **4** (1975) 248-270.

Di Zenzo 86

A Note on the Gradient of a Multi-Image.

Silvano Di Zenzo.

Computer Vision, Graphics, and Image Processing **33** (1986) 116-125.

Dreschler-Fischer 87

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

Leonie S. Dreschler-Fischer.

Proc. 11th German Workshop on Artificial Intelligence GWA1-87, Geseke, 28. September - 2. Oktober 1987, K. Morik (Hrsg.), *Informatik-Fachberichte* **152**, Springer Verlag, Berlin - Heidelberg - New York - Tokyo 1987, 1-15.

Gershon et al. 86

Ambient Illumination and the Determination of Material Changes.

Ron Gershon, Allan D. Jepson, John K. Tsotsos.

Journal of the Optical Society of America **A-3** (1986) 1700-1707.

Gershon et al. 87

Highlight Identification Using Chromatic Information.

Ron Gershon, Allan D. Jepson, John K. Tsotsos.

Proc. 1st International Conference on Computer Vision, London, Großbritannien, 8.-11. Juni 1987, J. Michael Brady und Azriel Rosenfeld (Hrsg.), IEEE Computer Society Press, Washington, D.C. 1987, 161-170.

Healey + Binford 87

The Role and Use of Color in a General Vision System.

Glenn Healey, Thomas O. Binford.

Proc. DARPA-Image Understanding Workshop, Los Angeles, CA, Februar 1987, Morgan Kaufmann Publishers Inc., Los Altos, CA 1987, 599-613.

Hille 82

Polynomiale Approximation von digitalen Bildfunktionen.

Gunter Hille.

Mitteilung IF1-III-M-98/82, Fachbereich Informatik, Universität Hamburg, 1982.

Horn 77

Understanding Image Intensities.

Berthold Klaus Paul Horn.

Artificial Intelligence **8** (1977) 201-231.

Hueckel 71

An Operator which Locates Edges in Digitized Pictures.

Manfred H. Hueckel.

Journal of the ACM **18** (1971) 113-125.

Huntsberger + Descalzi 85

Color Edge Detection.

T.L. Huntsberger, M.F. Descalzi.

Pattern Recognition Letters **3** (1985) 205-209.

Klinker et al. 87

Using a Color Reflection Model to Separate Highlights from Object Color.

Gudrun J. Klinker, Steven A. Shafer, Takeo Kanade.

Proc. 1st International Conference on Computer Vision, London, Großbritannien, 8.-11. Juni 1987, J. Michael Brady und Azriel Rosenfeld (Hrsg.), IEEE Computer Society Press, Washington, D.C. 1987, 145-150.

Klinker et al. 88

The Measurement of Highlights in Color Images.

Gudrun J. Klinker, Steven A. Shafer, Takeo Kanade.

International Journal of Computer Vision 2 (1988) 7-32.

Korn + Erdtel 85

Kombination verschiedener Filterkanäle zur Optimierung einer Merkmalsrepräsentation im Bildbereich.

Axel Korn, Carola Erdtel.

Proc. 7. DAGM-Symposium "Mustererkennung 1985", Erlangen, September 1985, H. Niemann (Hrsg.), *Informatik-Fachberichte* 107, Springer Verlag, Berlin - Heidelberg - New York - Tokyo 1985, 107-111.

Machuca + Phillips 83

Applications of Vector Fields to Image Processing.

Raul Machuca, Keith Phillips.

IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-5 (1983) 316-329.

Marr + Hildreth 80

A Theory of Edge Detection.

David Marr, Ellen Hildreth.

Proc. Royal Society of London B-207 (1980) 187-217.

Neumann 82

Knowledge Sources for Understanding and Describing Image Sequences.

Bernd Neumann.

Proc. 6th German Workshop on Artificial Intelligence GWA1-82, Bad Honnef, September 1982, W. Wahlster (Hrsg.), *Informatik-Fachberichte* 58, Springer Verlag, Berlin - Heidelberg - New York - Tokyo 1982, 1-21.

Nevatia 77

A Color Edge Detector and its Use in Scene Segmentation.

Ramakant Nevatia.

IEEE Transactions on Systems, Man, and Cybernetics SMC-7 (1977) 820-826.

Novak + Shafer 87

Color Edge Detection.

Carol L. Novak, Steven A. Shafer.

in: Proc. DARPA-Image Understanding Workshop, Los Angeles, CA, Februar 1987, Morgan Kaufmann Publishers Inc., Los Altos, CA 1987, Section 2 of the PI-Report *Image Understanding Research at CMU* by Takeo Kanade, 35-37.

Pietikainen + Harwood 86

Edge Information in Color Images based on Histograms of Differences.

Matti Pietikainen, David Harwood.

Proc. 8th Int. Conf. on Pattern Recognition, Paris, Frankreich, Oktober 1986, J.-P. Haton (Hrsg.), IEEE Computer Society Press, Washington, D.C. 1986, 594-596.

Poggio + Torre 84

Ill-Posed Problems and Regularization Analysis in Early Vision.

Tomaso Poggio, Vincent Torre.

Memo 773, Artificial Intelligence Laboratory, MIT, Cambridge, MA, 1984.

Poggio et al. 85

A Regularized Solution to Edge Detection.

Tomaso Poggio, H. Voorhees, Allan L. Yuille.

Memo 833, Artificial Intelligence Laboratory, MIT, Cambridge, MA, 1985.

Rice 45

Mathematical Analysis of Random Noise.

S.O. Rice.

Bell Systems Technical Journal **24** (1945) 46-156.

Robinson 76

Color Edge Detection.

Guner S. Robinson.

Proc. SPIE Symposium on Advances in Image Transmission Techniques, Vol. 87, 1976, 126-133.

Rosenfeld + Kak 82

Digital Picture Processing. Second Edition.

Azriel Rosenfeld, Avinash C. Kak.

Academic Press, Orlando 1982, 2 Bände.

Rosenfeld + Thurston 71

Edge and Curve Detection for Visual Scene Analysis.

Azriel Rosenfeld, Mark Thurston.

IEEE Transactions on Computers **C-20** (1971) 562-569.

Rubin + Richards 82

Color Vision and Image Intensities: When are Changes Material?.

John M. Rubin, W.A. Richards.

Biological Cybernetics **45** (1982) 215-226.

Sankar 78

Color Edge Detection: A Comparative Study.

P.V. Sankar.

Technical Report TR-666, Computer Science Center, Univ. of Maryland, 1978.

Schröder 87

Ein Ansatz zur Segmentation von Farbbildern durch hierarchische Ballungsanalyse.

Carsten Schröder.

Bericht FBI-III-B-132/87, Fachbereich Informatik, Universität Hamburg, 1987.

Sethi 85

A General Scheme for Discontinuity Detection.

I.K. Sethi.

Pattern Recognition Letters **3** (1985) 327-333.

Shafer 85

Using Color to Separate Reflection Components.

Steven A. Shafer.

COLOR research and application **10** (1985) 210-218.

Shiozaki 86

Edge Extraction using Entropy Operator.

Akira Shiozaki.

Computer Vision, Graphics, and Image Processing **36** (1986) 1-9.

Solinsky 85

The Use of Color in Machine Edge Detection.

James C. Solinsky.

Proc. VISION-85, Detroit, MI, 1985, 25-28.

Speck 83

Übersetzung von Linien- und Flächenstrukturen in kombinatorisch-relationale Datenstrukturen zur automatischen Mustererkennung in Digitalbildern.

Peter T. Speck.

Dissertation ETH Nr. 7508, Zürich, 1984.

Torre + Poggio 86

On Edge Detection.

Vincent Torre, Tomaso A. Poggio.

IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-8** (1986) 147-163.

Witkin 83

Scale-Space Filtering.

Andrew P. Witkin.

Proc. 7th Int. Joint Conf. on Artificial Intelligence, Karlsruhe, August 1983, Morgan Kaufmann Publishers Inc., Los Altos, CA 1983, 1019-1022.

Yakimovsky 76

Boundary and Object Detection in Real World Images.

Yoram Yakimovsky.

Journal of the ACM **23** (1976) 599-618.

Yuille + Poggio 86

Scaling Theorems for Zero-Crossings.

Alan L. Yuille, Tomaso A. Poggio.

IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-8** (1986) 15-25.

Danksagungen

An dieser Stelle möchte ich Prof. Dr. Leonie Dreschler-Fischer danken, die mich immer, wenn es nötig war, in meiner Arbeit unterstützt hat. Ihr Vorrat an Tee, Kaffee und Gummibärchen scheint schier unermesslich zu sein.

Bedanken möchte ich mich auch bei allen Mitarbeitern des Sissy-Projektes für die kameradschaftliche Zusammenarbeit, die über Jahre hinweg wesentlich zu der freundlichen Arbeitsatmosphäre beigetragen hat. Ganz besonders erwähnen möchte ich Carsten Schröder, der mir von unserem ersten Studententag an und so auch während meiner Mitarbeit im Sissy-Projekt mit viel Rat und Tat zur Seite stand. Auch Rainer Sprengel gebührt ein besonderer Dank. Er ließ mich immer wieder von seinem mathematischen Wissen profitieren.

Ingeborg Heer-Mück, Ursula Bauer und Hartmuth Krüger möchte ich für die so hervorragend funktionierende "Infrastruktur" am Arbeitsbereich KOGS danken. Erika Criegee gebührt Dank für ihren engagierten Einsatz bei der Beschaffung schwer zugänglicher Literatur.

Der größte Dank gebührt jedoch meinen Eltern. Sie haben mir mit viel Geduld eine langwierige Ausbildung ermöglicht und haben mir bei den Entscheidungen über meinen Werdegang stets freie Hand gelassen.

