

**Bereichsbasierte Verfahren zur
Straßenerkennung für die autonome
Führung von Fahrzeugen**

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

der Fakultät für Informatik
der Universität Karlsruhe (TH)

vorgelegte

D i s s e r t a t i o n

von

M.Sc. Jinyou Zhang

aus Hebei, China

Tag der mündlichen Prüfung:

Erster Gutachter:

Prof. Dr. H.-H. Nagel

Zweiter Gutachter:

Prof. Dr. A. Waibel

Februar 1994

Kurzfassung

Die autonome Führung von Straßenfahrzeugen hat in den letzten Jahren in den Forschungsgebieten Künstliche Intelligenz, Maschinensehen und Robotik zunehmend an Bedeutung gewonnen. Zur Straßenerkennung, Hindernisdetektion und Landmarkenerkennung wird ein Sichtsystem benötigt. Es ist ein entscheidender Bestandteil eines autonomen Systems, analog zum Sehvermögen des Autofahrers.

Die zuverlässige Erkennung von Straßen bildet die Voraussetzung für die Hindernisdetektion und die Landmarkenerkennung. Durch eine systematische Analyse der physikalischen und geometrischen Erscheinungsformen lassen sich Straßen in charakteristische Klassen einteilen. Nach dieser Untergliederung von Straßen und einem umfassenden Überblick über den Stand der Technik ist festzustellen, daß Straßen mit inhomogener Oberfläche und markierten Begrenzungen sowie Straßen mit inhomogener Oberfläche ohne markierte Begrenzungen mit existierenden Verfahren noch nicht zuverlässig erkannt werden können. Zur Lösung dieser Probleme werden in dieser Arbeit neue bereichsbasierte Verfahren vorgestellt.

Ausgehend von einer ausführlichen Diskussion der Theorie und der Verfahren der Texturanalyse stellen wir fest, daß gerichtete Textur ein aussagekräftiges Merkmal für die Segmentierung der Bilder von Straßen mit inhomogener Oberfläche, aber ohne markierte Begrenzungen ist. Zum Studium gerichteter Texturen führen wir eine Beschreibung durch das Orientierungsfeld der Texturen ein. Die lokale Texturrichtung und -anisotropie, welche das Orientierungsfeld charakterisieren, werden durch ein neu entwickeltes Verfahren aus den Kovarianzmatrizen der Grauwertänderungen berechnet.

In Straßenbildern erscheint die Texturdichte recht unterschiedlich. Ursache dafür ist sowohl der physikalische Unterschied von Texturen als auch die durch die perspektivische Projektion entstehende Änderung der Texturen. Mit Hilfe der durch eine Kamerakalibrierung ermittelten Abbildungsparameter der Kamera stellen wir ein neues Verfahren vor, einen optimalen Maßstab für jede Position in den perspektivisch abgebildeten Bildern systematisch zu bestimmen. Die mit optimalem Maßstab ermittelte Texturanisotropie wird dann als Merkmal für die Segmentierung von Straßenbildern in Bildfolgen eingesetzt.

Ein segmentiertes Straßenbild stellt lediglich die Aufteilung des Bildes dar. Für die Straßenerkennung müssen noch geometrische Modelle von Straßen dem Erkennungsprozeß bekannt sein. Wir modellieren in dieser Arbeit die geometrische Form von Straßen als eine Kombination von Geradensegmenten und Kreisbögen. Zur modellgestützten Erkennung von Straßen in Bildfolgen haben wir das System *SeTa* entwickelt, ein System zur Straßenerkennung mittels der Verifikation der Lagehypothese des geometrischen Straßen-

modells durch die im Bild extrahierten Merkmale. Die Erkennungsaufgaben im System *SeTa* werden in der Initialisierungs- und Verfolgungsphase unterschiedlich ausgeführt.

Das System *SeTa* stellt auch ein neues Konzept für die Segmentierung von Straßenbildern durch eine Kombination der Klassifikation mit unüberwachten Lernverfahren in der Initialisierungsphase und der Klassifikation mit überwachten Lernverfahren in der Verfolgungsphase vor.

Die Anwendungen der in dieser Arbeit vorgestellten Verfahren beschränken sich nicht nur auf Straßenbilder. Die Segmentierung von Bildern durch Texturauswertung ist auf eine Vielzahl unterschiedlicher Aufgaben übertragbar, wie z.B. bei der Sichtprüfung von Holzoberflächen oder von handgesponnenen Baumwollstoffen.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Entwicklung von autonomen Fahrzeugen	1
1.2	Aufgabenstellungen zur autonomen Fahrzeugführung	3
1.3	Sichtsysteme zur autonomen Fahrzeugführung	5
1.3.1	Sensoren in der Fahrzeugführung	6
1.3.2	Fahrbahnerkennung	6
1.3.3	Hindernisdetektion und Landmarkenerkennung	7
1.4	Motivation und Überblick der Arbeit	8
2	Analyse von Straßenszenen	11
2.1	Geometrische Aspekte von Straßen	12
2.1.1	Generische Geometriemodelle von Straßen	12
2.1.2	Verwendete Geometriemodelle von Straßen	17
2.1.3	Verfahren zur Ermittlung von Modellparametern	18
2.2	Physikalische Aspekte von Straßen	20
2.2.1	Untergliederung von Straßenszenen	20
2.2.2	Straßen der ersten Klasse	22
2.2.3	Straßen der zweiten Klasse	23
2.2.4	Straßen der dritten Klasse	25
2.2.5	Straßen der vierten Klasse	26
3	Bildauswerteverfahren zur Fahrzeugführung: Stand der Forschung	29
3.1	Versuchsträger zur autonomen Fahrzeugführung	30
3.2	Konturbasierte Verfahren zur Fahrbahnerkennung	34
3.3	Bereichsbasierte Verfahren zur Fahrbahnerkennung	38
3.4	Hindernisdetektion	42
3.5	Landmarkenerkennung	46
3.6	Auflistung der Forschungsaktivitäten zur autonomen Fahrzeugführung	49
4	Texturanalyse: Theorien und Verfahren	53
4.1	Texturwahrnehmung beim Menschen	54
4.1.1	Globale Grauwertstatistiken n -ter Ordnung	55
4.1.2	Statistiken lokaler Texturelemente	56

4.2	Statistische Verfahren zur Texturanalyse	57
4.2.1	Stochastische Modelle von Texturen	58
4.2.2	Punktbasierte Statistiken n -ter Ordnung	59
4.2.3	Globale Statistiken	63
4.2.4	Filterbasierte lokale Statistiken	64
4.3	Strukturelle Verfahren zur Texturanalyse	67
4.3.1	Deterministische Modelle von Texturen	67
4.3.2	Extraktion von Texteln	68
4.3.3	Extraktion der Texturen aus Texteln	70
4.4	Multikanalanalyse von Texturen	71
4.5	Schwerpunkte für weitere Forschungen zur Texturanalyse	73
5	Analyse gerichteter Texturen	75
5.1	Orientierungsfelder gerichteter Texturen	76
5.1.1	Detektion der Grauwertänderungen	76
5.1.2	Kovarianzmatrix der Grauwertänderungen	78
5.1.3	Orientierungsfelder von Texturen	83
5.2	Segmentierung von Orientierungsfeldern	89
5.2.1	Grundlage der Bildsegmentation	89
5.2.2	Klassifikation mit modifiziertem <i>Isodata</i> -Algorithmus als Lernverfahren	92
5.2.3	Experimentelle Ergebnisse	96
6	Segmentierung von Straßenbildern durch die Auswertung von Texturen	99
6.1	Bestimmung des optimalen Maßstabs von Filtern	100
6.1.1	Verzerrung eines Textels im perspektivisch abgebildeten Bild	100
6.1.2	Parameter der abgebildeten ellipsenförmigen Textel	101
6.1.3	Optimaler Maßstab von Filtern	104
6.2	Ermittlung von Texturanisotropien mit σ_{opt}	106
6.2.1	Erweiterung des Verfahrens zur Ermittlung von Texturanisotropien	107
6.2.2	Beispiele von ermittelten Texturanisotropien	108
6.3	Klassifizierung von Straßenbildern	115
6.3.1	<i>Isodata</i> -Algorithmus in der Initialisierungsphase	117
6.3.2	Statistische Klassenmodelle	117
6.3.3	Bayesscher Klassifikator in der Verfolgungsphase	117
6.3.4	Beispiele von Klassifizierungsergebnissen	118
6.3.5	Beispiele von Segmentierungsergebnissen durch grauwertbasierte Bereichsklassifikation sowie durch konturbasierte Bereichsanalyse	119
7	Modellgestützte Erkennung von Straßen aus segmentierten Bildern	125
7.1	Geometrisches Modell von Straßen	126
7.1.1	Modellierung von Straßen im Szenenbereich	126
7.1.2	Abbildung des GKG-Modells im Bildbereich	129
7.2	Bestimmung des Straßenbereiches	129

7.3	Schätzung der Parameter des GKG-Modells	131
7.3.1	Merkmalszuordnung	131
7.3.2	Bestimmung der Modelltransformation	132
7.3.3	Initialisierung der Modellposition	133
7.4	Beispiele von Straßenerkennungsergebnissen	135
8	Zusammenfassung und Ausblick	143
8.1	Beiträge dieser Arbeit	143
8.2	Einsetzbarkeit der Verfahren im Fahrzeug	145
	Anhang	147
A	Hauptachsentransformation	147
B	Die k-Mittelwert-Ballung	149
C	<i>Isodata</i>-Algorithmus mit dem Varianzkriterium	151
D	Der <i>Mahalanobis</i>-Abstand	153
E	Segmentierung von Straßenbildern durch konturbasierte Bereichsanalyse	155
E.1	Bestimmung von Konturpunkten	155
E.2	Verkettung von Konturpunkten	156
E.3	Bereichsmarkierung	159
E.4	Beispiele von Segmentierungsergebnissen	160
	Literaturverzeichnis	163
	Autorenverzeichnis	181

Kapitel 1

Einleitung

1.1 Entwicklung von autonomen Fahrzeugen

Autonome mobile Systeme sind seit einigen Jahren bedeutende Forschungsgegenstände der künstlichen Intelligenz, des Maschinensehens und der Robotik. Solche Systeme sollen zur Lösung von Aufgaben eingesetzt werden, die den Menschen wegen häufiger Wiederholung ermüden oder ihn Gefahren für Leben und Gesundheit aussetzen. Beispielhaft sind hier die Überwachung und Reparatur von Unterwasser-Erdölleitungen ([Blidberg 89], [Hallset 92]), die Ausführung von Aufgaben in einem Kern-Reaktor ([Fisher 85]), der Materialtransport in Fabrikations- und Lagerstätten ([Rembold 88]), die Exploration von weit entfernten Planeten ([Bares *et al.* 89]), die autonome Führung von Luftfahrzeugen ([Gilmore & Semeco 86]) und das Fahrerunterstützungssystem zur autonomen Führung von Straßenfahrzeugen ([Nagel 88], [Nagel 91b]) zu nennen.

Je nach Art der Durchführung der Aufgaben lassen sich die untersuchten Systeme in semi-autonome und autonome Systeme einteilen. Semi-autonome Systeme zeichnen sich entweder durch wenige Freiheitsgrade der Bewegung oder durch interaktive Eingriffe von Menschen während elementarer Operationen aus. Zur ersten Gruppe gehören Systeme, die sich entweder entlang eines Drahtes oder eines besonders markierten Streifens in einer gut strukturierten Umgebung bewegen, auch *AGV*¹ genannt. Typische Anwendungen solcher Systeme liegen im Bereich des Materialtransports in Fabrikationsstätten. Als nächste Generation solcher Systeme werden leitspurunabhängige Flurförderfahrzeuge für innerbetriebliche Transportaufgaben untersucht. Die zweite Gruppe von semi-autonomen Systemen sind von Menschen ferngesteuerte Systeme. Die Anwendungsumgebung solcher Systeme ist für heutige Sichtsysteme meist zu schwierig auszuwerten und für Menschen schwer zugänglich. Menschen übernehmen dabei die Ermittlung eines befahrbaren Weges und die Planung der Aktionsfolgen. Über eine Funkstrecke werden den Systemen die elementaren Operationen mitgeteilt, die von den Systemen autonom durchgeführt werden. Solange autonome Systeme die ihnen gestellten Aufgaben noch nicht befriedigend lösen können, werden semi-autonome Systeme mit langfristiger Perspektive die Anwendung

¹AGV: „automatic guided vehicle“.

dominieren.

Im Gegensatz zu semi-autonomen Systemen haben autonome Systeme die völlige Unabhängigkeit von interaktiven Eingriffen durch Menschen während einer Operation als Entwicklungsziel. Menschen brauchen nur vor dem Operationsbeginn dem autonomen System die durchzuführenden globalen Aufgaben einzugeben. Das autonome System soll die anderen verbleibenden Aufgaben selbst ausführen. Diese Thematik ist zum gegenwärtigen Zeitpunkt Ziel der Untersuchungen zahlreicher Forschungsgruppen in verschiedenen Ländern.

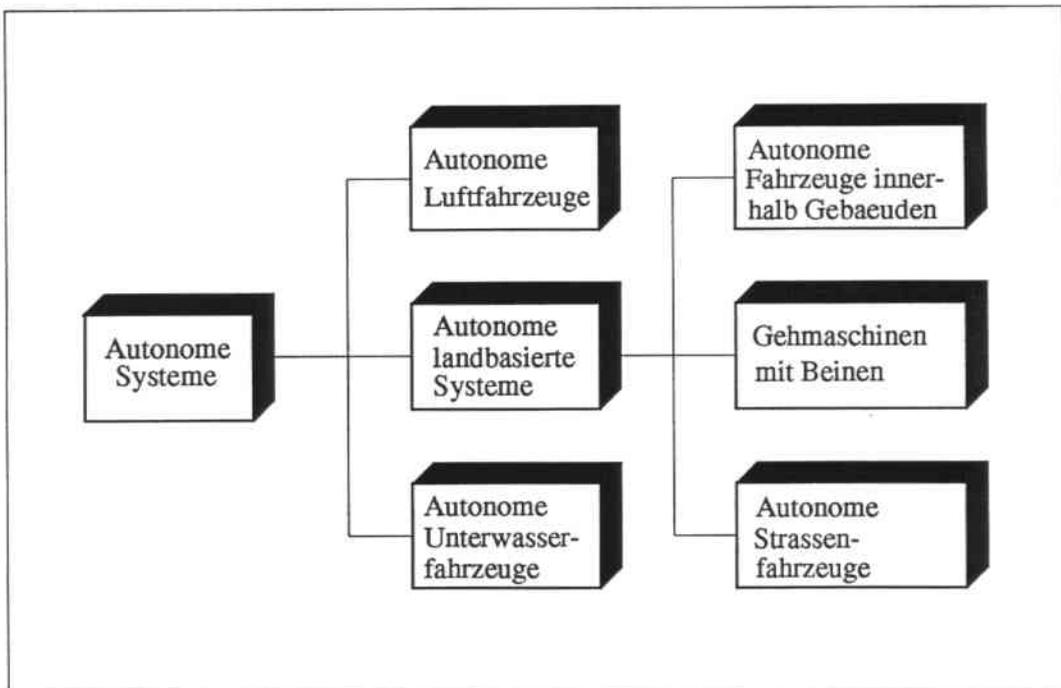


Abbildung 1.1: Eine Untergliederung von autonomen Systemen.

Die Untergliederung von autonomen Systemen nach Art der Durchführung von Operationen in semi-autonome und autonome Systeme ist sicher nicht die einzig mögliche. Entsprechend dem Unterschied der Arbeitsumgebung können autonome Systeme, wie in Abbildung 1.1 dargestellt, in die drei Hauptkategorien, autonome Luftfahrzeuge, autonome landbasierte Systeme sowie autonome Unterwasserfahrzeuge untergliedert werden. Im jeweiligen Kontext besitzen sie einen sehr unterschiedlichen Schwerpunkt bei der Entwicklung. Ein landbasiertes System (außer Gehmaschinen) hat wenige Freiheitsgrade der Bewegung und soll *ad hoc* in einer Laborumgebung oder einer natürlichen Umgebung autonom Aufgaben ausführen. Schwerpunkte dabei sind vorwiegend die weitere Verbesserung der sensorischen Fähigkeiten zur Umgebungserfassung. Bildgebende Sensoren und deren Verarbeitungseinheit spielen hier eine wichtige Rolle. Unter Wasser sind wegen schwacher Beleuchtung passive optische Sensoren kaum noch anwendbar. Aktive Sensoren, z.B. Ultraschall-Sensoren, deren Energien weniger von der dichten Umgebung

absorbiert werden, spielen eine entscheidende Rolle beim Erfassen der Umgebung. Da der Seeboden meist nicht so eben wie Festlandboden ist, müssen Unterwasser-Fahrzeuge alle 6 Freiheitsgrade der Bewegung besitzen. Bei einem autonomen Unterwasserfahrzeug besteht die zusätzliche Schwierigkeit darin, daß das gesamte System von der Umgebung isoliert sein muß. Unter autonomen Luftfahrzeugen sollen hier sowohl Flugzeuge, die mit einem Pilotenunterstützungssystem als Auto-Pilot ausgerüstet worden sind, als auch unbemannte Luftfahrzeuge für Spezialanwendungen verstanden werden. Zu den Aufgaben des Sichtsystems gehören in dieser Beispielkategorie die Terrainerkennung und auch die Objektverfolgung. Zum Erfassen der Informationen eignen sich hier sowohl Laserabtastsysteme als auch passive optische Sensoren. Als Schwerpunkte dabei sind die Zuordnung der durch Sensoren gewonnenen Informationen zur gespeicherten Terrainkarte, die Routenplanung und die Quer- bzw. Längsführung zu nennen. In einem autonomen Luftfahrzeug muß außerdem noch sichergestellt sein, daß ein eventuell auftretender Fehler im autonomen System das Luftfahrzeug nicht zum Absturz bringt.

Wegen der schwierigen Versuchsumgebungen wird die Forschung zur autonomen Führung von Luft- und Unterwasserfahrzeugen hauptsächlich auf der Simulationsebene durchgeführt. Die meisten Forschungsvorhaben beschäftigen sich zunächst mit der Führung von Landfahrzeugen. Die dabei gewonnenen Erfahrungen und die entwickelten Algorithmen können erweitert in den beiden anderen Forschungsgebieten verwendet werden.

Bevor auf diese These näher eingegangen wird, soll zunächst der Themenbereich dieser Arbeit auf die weitere Behandlung von autonomen Systemen und besonders von Sichtsystemen zur Führung von Land-Straßenfahrzeugen begrenzt werden. Einen sachgerechten Eindruck zu allgemeinen autonomen Systemen vermitteln beispielsweise die Konferenzbände [Kanade *et al.* 89] und [Groen *et al.* 93], das Sonderheft [Ras 91] der Fachzeitschrift „*Robotics and Autonomous Systems*“ sowie die Bücher [Cox & Wilfong 90], [Thorpe 90] und [Masaki 92].

1.2 Aufgabenstellungen zur autonomen Fahrzeugführung

Die Entwicklung von autonomen Systemen zur Führung von Landfahrzeugen hat eine große Bedeutung sowohl für die Modernisierung der Industrie und des Straßenverkehrs als auch für militärische Anwendungen. Ein solches System soll in Zukunft die Fähigkeit haben, Fahrzeuge autonom führen zu können, auch wenn es im aktuellen Einsatz als ein semi-autonomes oder Ko-Pilot-System vorgesehen ist. Zu den Aufgaben eines solchen Systems gehören sowohl die Interpretation der Signale über den Zustand der Fahrzeug-Umgebung, die Überwachung interner Fahrzeugzustände, die Fahrzeugführung als auch die Modellierung des Verhaltens eines menschlichen Fahrers. Als Diskursbereich für die folgenden Aufgabenstellungen zur autonomen Fahrzeugführung soll das Gesamtsystem *Fahrer - Fahrzeug - Fahrbahn*² verstanden werden. In Anlehnung an das Verhalten eines menschlichen Fahrzeugführers werden hier einige wichtige Aufgaben für die autonome

²Fahrbahn: in dieser Arbeit werden Fahrbahn und Straße als Synonyma verwendet.

Fahrzeugführung erläutert, wie z.B. Routenplanung, Navigation eines Fahrzeuges, Quer- und Längsführung eines Fahrzeuges, Fahrbahnerkennung, Hindernisdetektion und Landmarkenerkennung.

Routenplanung

Ein Mensch als Fahrzeugführer wird sich, bevor er losfährt, zunächst zumindest grob überlegen, auf welchen Straßen er fahren möchte. Diese Auswahl der zu befahrenden Straßen erfolgt aufgrund seiner Ortskenntnisse und evtl. einer Suche in einer Straßenkarte. Unter Berücksichtigung der Verkehrssituation und der Beschaffenheit von Straßen soll ein optimaler Weg selektiert werden. Dementsprechend soll ein maschinelles autonomes System einen Weg nach Vorgabe des Fahrtzieles bzw. der Zwischenziele automatisch erstellen. Diese Vorgehensweise wird als Routenplanung bezeichnet. Das Fahrtziel wird dem autonomen System über Ein-/Ausgabegeräte von Menschen vor dem Fahrtbeginn mitgeteilt. Dem autonomen System soll das zu berücksichtigende Straßennetz als Wissensbasis bekannt sein.

Quer- und Längsführung

Während der Führung eines Fahrzeugs schätzt der Fahrer mit Hilfe seines visuellen Wahrnehmungsvermögens den Abstand und die relative Geschwindigkeit des Fahrzeugs zu seiner Umgebung. Unter der Umgebung sollen hier sowohl andere Verkehrsteilnehmer als auch Fahrbahnen verstanden werden. Auf eine kritische Verkehrssituation soll er so schnell wie möglich reagieren und entsprechende Aktionen folgen lassen. Als Beobachtungsgegenstände eines Fahrers kommen Fahrbahnbegrenzungen, Fahrspurmarkierungen, andere Fahrzeuge, Verkehrszeichen usw. in Betracht. Zu den Aktionen gehören die Anpassung der Geschwindigkeit des Fahrzeugs an die aktuelle Verkehrssituation und die Stabilisierung des Fahrzeugs auf der Fahrbahn. Diese Aktionen werden Längsführung und Querführung eines Fahrzeugs genannt.

Fahrbahnerkennung

In einem maschinellen autonomen System wird die Querführung eines Fahrzeugs durch die Betätigung der Lenkung über ein Stellsystem realisiert. Der Lenkwinkel ist eine Funktion der Positionsabweichung zwischen der Istfahrbahn und der Sollfahrbahn sowie der Fahrzeugausrichtung zur Fahrbahn. Um Positionsfehler zu bestimmen, soll zunächst die Position der Fahrbahnbegrenzungen ermittelt werden. Diese Vorgehensweise wird Fahrbahnerkennung genannt.

Hindernisdetektion

Die Längsführung eines Fahrzeugs ist in erster Linie die Geschwindigkeitsanpassung an den Verlauf der Fahrbahn sowie an den Abstand und die Relativbewegung von Hindernissen. Sie wird durch die Betätigung von Gas und Bremse über ein Stellsystem realisiert.

Die Ermittlung der Position und evtl. auch der Geschwindigkeit der Gegenstände, die sich auf der Fahrbahn befinden, ist Aufgabe der sogenannten Hindernisdetektion.

Landmarkenerkennung

Bei der Längsführung eines Fahrzeugs genügt die Hindernisdetektion und die Bestimmung des Fahrbahnverlaufs im allgemeinen nicht, um das Fahrzeug zum geplanten Zielort zu bringen und den Zielort zu erkennen. Außerdem ist hier noch eine Landmarkenerkennung, z.B. die Erkennung von Straßeneinmündungen und -kreuzungen sowie von Verkehrszeichen, für die Navigation des Fahrzeugs nötig. Unter Navigation eines Fahrzeugs soll hier die Auswahl des nächsten Teilwegstückes verstanden werden.

Zur Lösung der oben gestellten Aufgaben im Hinblick auf eine autonome Fahrzeugführung benötigt ein autonomes System interne Sensoren zur Erfassung des aktuellen Fahrzeugzustandes, ein Sichtsystem zum Erfassen der Fahrzeugumgebung, ein Navigationssystem zur Planung seiner Wege, Regelungsmodule zur Quer- und Längsregelung sowie Stellsysteme für die Betätigung der Lenkung, der Längsbeschleunigung (z.B. Gas) und der Verzögerung (z.B. Bremse). Außerdem bietet ein solches autonomes Fahrzeug als kompletter Versuchsträger die Möglichkeit zum Test neuer Sensoren, zur Realisierung von Konzepten der Sensordatenfusion und der Systemarchitektur, zur Erprobung neuer Algorithmen für Bildauswertung, Navigation, Regelung usw..

Für eine erfolgreiche autonome Fahrzeugführung ist ein komplexes System nötig, das alle obengenannten Fahrzeugausrüstungskomponenten umfaßt. Darunter ist das Sichtsystem die größte technische Herausforderung bei der Entwicklung und auch entscheidender Bestandteil eines autonomen Systems. Der Grund liegt darin, daß die meisten Informationen über die Fahrzeug-Umgebung vom Sichtsystem erfaßt und interpretiert werden müssen.

1.3 Sichtsysteme zur autonomen Fahrzeugführung

Als ein Forschungsgebiet der Informatik beschäftigt sich das Maschinensehen mit der Aufnahme des Bildes einer Szene, der Extraktion der für eine Aufgabe signifikanten Strukturen aus einem Bild oder einer Folge von Bildern und der Interpretation der Szene ([Horn 86], [Orban & Nagel 92]). Das angestrebte Forschungsziel ist es, eine Maschine in die Lage zu versetzen, bildhafte Umgebungsinformation zur Lösung von Aufgaben zu verwenden.

Zur automatischen Führung eines Fahrzeuges auf der Straße muß ein autonomes System die Verkehrsumgebung erfassen und interpretieren. Ein Sichtsystem ist für ein autonomes System so wichtig wie das Sehvermögen für einen Autofahrer. Typische Aufgaben eines solchen Sichtsystems sind, wie schon im vorangegangenen Abschnitt angedeutet wurde,

- die Fahrbahnerkennung,

- die Hindernisdetektion und
- die Landmarkenerkennung.

Die Möglichkeiten zur Lösung der gestellten Aufgaben sollen im folgenden kurz erläutert werden.

1.3.1 Sensoren in der Fahrzeugführung

Zur Gewinnung der erforderlichen Information für eine detaillierte Beschreibung der Umgebungszustände um das Fahrzeug herum können sowohl aktive Sensoren wie Laser-Abtastsysteme, Radar- oder Ultraschall-Sensoren, als auch passive optische Sensoren wie monochrome Kameras, Farbkameras oder Infrarotkameras eingesetzt werden.

Mit einem aktiven Abstandssensor wird der Abstand zwischen dem abzubildenden Sensor und dem abgebildeten Objekt direkt erfaßt. Daraus kann die Lage eines 3D-Objektes relativ zum Sensor bestimmt werden. Bei der Hindernisdetektion und Landmarkenerkennung werden solche 3D-Daten benötigt. Aktive Sensoren weisen unbestreitbare Vorteile zur Umgebungserfassung gegenüber passiven optischen Sensoren besonders bei schwachem Umgebungslicht oder bei durch extremen Schattenwurf beeinflussten Bildstrukturen auf. Neben den angedeuteten Vorteilen ihres Einsatzes bei der direkten Gewinnung der Tiefeninformation haben aktive Sensoren aber Schwierigkeiten bei der Ermittlung von Fahrbahnbegrenzungen. Auf einer Straße sind die Fahrbahnbegrenzungen meistens nicht so aufgebaut, daß es eine signifikante 3D-Diskontinuität zwischen der Fahrbahn und der Fahrbahnbegrenzung gibt. Deshalb erscheint es nicht sehr vorteilhaft, bei der Erkennung von Fahrbahnen bzw. Fahrbahnbegrenzungen einen aktiven Sensor einzusetzen. Die Tiefendaten aus aktiven Sensoren werden daher wesentlich nur zur Hindernisdetektion verwendet.

Im Gegensatz zu aktiven Sensoren sind passive optische Sensoren geeigneter für die Erfassung von Informationen für die Fahrbahnerkennung. In einer gewöhnlichen Verkehrsumgebung unterscheidet sich die Fahrbahn von den Fahrbahnrändern grundsätzlich in Farbe und Grauwertstrukturen, welche durch eine Kamera aufgenommen werden können. Außerdem können die mit passiven optischen Sensoren gewonnenen Informationen bei der Hindernisdetektion sowie bei der Landmarkenerkennung verwendet werden. Ein weiterer erwähnenswerter Vorteil von passiven optischen Sensoren ist, daß sie im allgemeinen weniger Abtastzeit als aktive Sensoren brauchen. Deshalb werden die Daten aus passiven Sensoren sowohl zur Fahrbahnerkennung als auch zur Hindernisdetektion sowie zur Landmarkenerkennung verwendet.

1.3.2 Fahrbahnerkennung

Zur sichtsystemgestützten Ermittlung der Fahrbahnbegrenzung in einer Verkehrsumgebung werden konturbasierte und bereichsbasierte Verfahren angewendet. Beim ersten Verfahren wird vorausgesetzt, daß sich die Fahrbahnbegrenzung im Bild deutlich von der Fahrbahn abhebt, z.B. durch ein helles Markierungsband. Durch einen Kantenoperator

wird die Änderung der Grauwertverteilung als Funktion der Ortskoordinaten extrahiert. Dann werden die Konturpunkte, welche auf eine signifikante Diskontinuität im Grauwertverlauf hinweisen, durch eine Suche nach Extremwerten des Ableitungsoperators erster Ordnung oder nach Nulldurchgangsstellen des Ableitungsoperators zweiter Ordnung lokalisiert. Danach schließen sich die Verkettung der Konturpunkte, die Segmentation in geradlinige, kreisförmige, elliptische oder auch durch Splines beschriebene Segmente sowie die Gruppierung solcher Segmente zu größeren geometrischen Strukturen an. Die Fahrbahnbegrenzung läßt sich direkt aus den geometrischen Strukturen mit Hilfe von Straßenmodellen bestimmen.

Bei einer Vielzahl von Straßen (nicht bei Autobahnen und neueren Landstraßen) sind jedoch die Randmarkierungen nicht vorhanden oder unkenntlich geworden. In solchen Fällen funktioniert das oben genannte konturbasierte Verfahren nicht zuverlässig. Statt der signifikanten Diskontinuität im Grauwertverlauf sollen die Unterschiede in Farbe und Textur zwischen Straße und Umgebung zur Ermittlung der Fahrbahnbegrenzung verwendet werden. Bei einem bereichsbasierten Verfahren wird jeder einzelne Bildpunkt unter Berücksichtigung seiner lokalen Umgebung aufgrund spezieller Merkmale den Klassen „Fahrbahn“ oder „Nicht Fahrbahn“ zugeordnet. Durch eine Verfolgung der Randpunkte der Klasse „Fahrbahn“ kann die Fahrbahnbegrenzung bestimmt werden.

Bei der Ermittlung der Fahrbahnbegrenzungen sind geometrische Straßenmodelle sehr wichtig. Sie geben die Form der zu detektierenden Straßenbegrenzungen an. Außerdem sollen bei der Fahrbahnverfolgung die in den vorangegangenen Bildern und im aktuellen Bild ermittelten Informationen über die Position der Fahrbahn ausgenutzt werden, um den Fahrbahnverlauf zu schätzen.

1.3.3 Hindernisdetektion und Landmarkenerkennung

Ein auf der Fahrbahn befindliches Hindernis kann aus einem Abstandsbild ermittelt werden, das entweder direkt über einen aktiven Sensor oder durch ein Rekonstruktionsverfahren aus den Grauwertbildern gewonnen wurde. Als Beispiel zur Rekonstruktion der 3D-Information aus den 2D-Grauwertbildern sind hier die Verfahren „*Struktur aus Kontur*“ und „*Struktur aus Bewegung*“ zu nennen. Für die Hindernisdetektion sind die Lage, die Abmessungen und der Zustand (stehend oder bewegt) von Hindernissen signifikante Informationen. Der wesentliche Unterschied zwischen der Hindernisdetektion bei der Navigation und der konventionellen Identifikation eines Objektes besteht in der Fragestellung, ob alle Eigenschaften des Objektes genau erkannt oder lediglich seine Lage, Orientierung, Größe und Bewegung vermessen werden müssen.

Die Erkennung von Landmarken ist eine Voraussetzung für die Navigation eines autonomen Systems. Bei einem Navigationssystem für die „*Querfeldein*“-Fahrt muß ein in der Landkarte vorhergesagtes Objekt zusammen mit seiner Position vom Sichtsystem erkannt werden, so daß eine erfolgreiche Navigation durchgeführt werden kann. Durch einen Vergleich der abgespeicherten Position einer Landmarke mit der ermittelten Position wird die Information zur Formulierung der jeweils gültigen Fahraufgabe aktualisiert. Als Landmarke sind hier beispielhaft Straßeneinmündungen und -kreuzungen, Verkehrs-

zeichen neben und über der Fahrbahn, Gebäude neben der Fahrbahn usw. aufzuzählen. Für die Fahrzeugführung auf Straßen ist es jetzt möglich, kommerziell verfügbare Systeme (z.B. *Travelpilot* der Robert Bosch GmbH) zur Navigation eines Fahrzeuges zu verwenden. Mit der Information über die aktuelle Fahrzeugposition, die laufend vom Navigationssystem abgefragt wird, sowie den aus der Karte vorab bekannten Straßenverbindungen wird die Route zum ausgewählten Ziel gesucht.

1.4 Motivation und Überblick der Arbeit

Die allgemeine Aufgabenstellung für ein Sichtsystem zur autonomen Fahrzeugführung besteht, wie oben erwähnt, in der Straßenerkennung, der Hindernisdetektion und der Landmarkenerkennung. Die zuverlässige Erkennung von Straßen bildet die Voraussetzung für die Detektion der sich auf den Straßen befindenden Hindernisse und die Erkennung der Landmarken. Existierende Verfahren zur Straßenerkennung können nur auf bestimmten Straßen robuste Ergebnisse liefern. Daher ist es notwendig zu untersuchen, welcher Stand in der Straßenerkennung erreicht wurde und was noch zu lösen bleibt. Ausgehend von diesen Überlegungen als Motivation befaßt sich die vorliegende Arbeit mit der Entwicklung neuer Verfahren für die robuste Straßenerkennung.

Wir beginnen in Kapitel 2 mit einer systematischen Analyse der geometrischen und physikalischen Erscheinungsformen von Straßen. Aus einem sogenannten generischen geometrischen Straßenmodell werden alle möglichen Erscheinungsformen von Straßen in Videobildern hergeleitet. Anschließend werden die Verfahren zur Ermittlung der Modellparameter vorgeschlagen. Nach der Komplexität der Straßenoberfläche, der Straßenumgebung und der Straßenränder können Straßen in folgende vier charakteristische Klassen eingeteilt werden:

1. Straßen mit homogener Oberfläche und markierten Begrenzungen,
2. Straßen mit homogener Oberfläche ohne markierte Begrenzungen,
3. Straßen mit inhomogener Oberfläche und markierten Begrenzungen sowie
4. Straßen mit inhomogener Oberfläche ohne markierte Begrenzungen.

Für die Straßen jeder Klasse werden dann die brauchbaren Merkmale und ebenfalls die Verfahren zur Extraktion dieser Merkmale aus einem Bild besprochen. Das Ziel dieses Kapitels ist es, einen einheitlichen Beschreibungsrahmen für die weitere Behandlung dieser Problematik in der vorgelegten Arbeit zu schaffen.

In Kapitel 3 wird ein Überblick über den aktuellen Stand der Forschungstätigkeit zur autonomen Fahrzeugführung gegeben. Diese Übersicht beschreibt die wichtigen Ergebnisse zum Systemaufbau, der Straßenerkennung, der Hindernisdetektion und der Landmarkenerkennung für die autonome Fahrzeugführung. In Anlehnung an den in Kapitel 2 vereinbarten Beschreibungsrahmen werden besonders die Verfahren zur Straßenerkennung analysiert. Wir werden erkennen, daß die existierenden Verfahren zur Straßenerkennung

nur auf Straßen der ersten und zweiten Klasse zuverlässig funktionieren. Als zu lösende Aufgaben bei der Straßenerkennung verbleibt die Auswertung von Straßen der dritten und vierten Klasse.

Das wichtigste Merkmal in Bildern von Straßen der vierten Klasse ist der Unterschied in der Textur zwischen der Straßenoberfläche und ihrer Umgebung. Wir behandeln in Kapitel 4, 5 und 6 das Problem, die Bilder von Straßen der vierten Klasse durch Texturanalyse zu segmentieren.

Zum Beginn werden in Kapitel 4 sowohl Theorien zur Texturwahrnehmung beim Menschen als auch Verfahren zur maschinellen Texturanalyse beschrieben. Das Ziel dieses Kapitels besteht darin, durch die Analyse der klassischen Verfahren und die Zusammenfassung der aktuellen Forschungsergebnisse die Grundlage für die Analyse gerichteter Texturen in dieser Arbeit zu erarbeiten. In der Literatur gibt es bisher fast keine texturbasierten Verfahren für die Straßenerkennung. Im Abschluß dieses Kapitels werden zusätzlich vom Autor einige Schwerpunkte für weitere Forschungen zur Texturanalyse vorgeschlagen.

In Kapitel 5 beschreiben wir ein neues Verfahren zur Beschreibung gerichteter Texturen in Grauwertbildern mit dem Orientierungsfeld der Texturen. Das Orientierungsfeld läßt sich durch die lokale Richtung und Anisotropie der Textur, die aus den Kovarianzmatrizen der Grauwertänderungen berechnet werden, charakterisieren. Darauf aufbauend wird das ermittelte Orientierungsfeld mit Hilfe eines modifizierten *Isodata*-Algorithmus segmentiert. Die in diesem Kapitel beschriebene Ermittlung und Segmentierung gerichteter Texturen bilden eine Basis für die Auswertung von Straßenbildern.

Die Straßenbilder werden im allgemeinen durch eine perspektivische Projektion abgebildet. Die durch die perspektivische Verzerrung verursachte Änderung der Texturdichte spielt dabei eine wichtige Rolle. Mit Hilfe der durch eine Kamerakalibrierung gewonnenen Abbildungsparameter stellen wir in Kapitel 6 ein neues Verfahren zur Ermittlung des optimalen Maßstabs des Texturoperators vor. Mit dem optimalen Maßstab und einer neu eingeführten Definition für Texturanisotropie zeichnet sich die ermittelte Texturanisotropie als ein aussagekräftiges Merkmal für Straßenbilder der vierten Klasse aus. Anschließend werden in diesem Kapitel die Segmentierung der Bilder von Straßen der vierten Klasse durch die Klassifikation der Bildpunkte dargestellt.

Nach der Segmentierung von Straßenbildern kommen wir dann zur Straßenerkennung. Für die Straßenerkennung muß jede Straße rechnerintern durch ein Modell repräsentiert werden, dessen Merkmale bei dem Erkennungsprozeß mit den Bildmerkmalen verglichen werden. Wir modellieren in Kapitel 7 eine Straße mit zwei Geradensegmenten und einem dazwischen liegenden Kreisbogen in einer Ebene. Zur modellgestützten Straßenerkennung in Bildfolgen wird das System *SeTa* aufgebaut. Es besteht aus zwei Teilen. Der erste Teil behandelt die Straßenerkennung ausgehend von einer Initialisierungsphase auf der Grundlage eines modifizierten *Isodata*-Algorithmus. Der zweite Teil behandelt die Straßenerkennung mit einem Bayesschen Klassifikator in der Verfolgungsphase. Die Statistiken von Klassen werden dabei mit Hilfe von überwachten Lernverfahren ermittelt.

Der Schwerpunkt dieses Kapitels besteht darin, Straßen der vierten Klasse mit Hilfe ihres geometrischen Modells aus dem durch Texturauswertung segmentierten Bild zu erkennen. Nach der Schätzung der Position des Modells kann die aktuelle Lage des Fahrzeugs

ermittelt werden. Mit modellgestützten Vorgehensweisen lassen sich Positioniergenauigkeiten erreichen, die für die Fahrzeugführung ausreichen.

In Kapitel 8 diskutieren wir die hauptsächlichlichen Beiträge dieser Arbeit und die Erweiterungsmöglichkeit der vorgestellten Verfahren.

Kapitel 2

Analyse von Straßenszenen

Mit zunehmendem Interesse an der autonomen Führung von Fahrzeugen stieg auch die Anzahl verschiedener Verfahren zur Straßenerkennung. Bisher gibt es jedoch noch kein Verfahren, welches bei jeder Straßenszene robust funktioniert. Ursache dafür ist zum einen die Vielfalt der geometrischen Erscheinungsformen von Straßen und zum anderen die Vielfalt der physikalischen Erscheinungsformen von Straßen. Daher ist es notwendig, diese zwei Aspekte systematisch zu analysieren.

Die Vielfalt der physikalischen Erscheinungsformen von Straßen, die durch die Änderung der Beschaffenheit von Straßen, wie z.B. der Materialien ihrer Oberfläche, die Änderung der Beleuchtung usw. entstehen, erschwert die Extraktion von Straßenabbildungen aus Folgen von Videoaufnahmen. Ein bei jeder Straßenszene erfolgversprechendes Verfahren ist beim derzeitigen Stand des Maschinensehens nicht verfügbar. Es ist deshalb notwendig zu untersuchen, welche Verfahren für die Extraktion von Straßenbeschreibungen einer bestimmten Klasse geeignet sind. Dazu analysieren wir die physikalischen Erscheinungsformen von Straßen. Diese Analyse ermöglicht darüberhinaus eine Untergliederung von Straßenszenen und eine Zuordnung von Bildauswerteverfahren zu den charakterisierten Straßenklassen.

Die Wichtigkeit der Anwendung von Geometriemodellen von Objekten bei der Objekterkennung wird in letzter Zeit immer mehr herausgestellt ([Grimson 90]). Bei der Erkennung von Straßen sowohl im Bildbereich als auch im Szenenbereich spielen Geometriemodelle von Straßen ebenfalls eine wichtige Rolle. Jede Straße hat ihre eigene Form. Die Form einer Straße ändert sich auch von einem Ort zum anderen. Verfahren zur Straßenerkennung mit nur einem Geometriemodell versagen bei der Erkennung von Straßen, deren Form nicht dem Geometriemodell entspricht. Ein Ausweg aus dieser Schwierigkeit ist durch die Verwendung verschiedener Geometriemodelle für unterschiedliche Straßen möglich. Es ist deshalb sinnvoll, alle möglichen Formen von Straßen zu analysieren, die entsprechenden Geometriemodelle zu erstellen und die Verfahren zur Ermittlung von Modellparametern zu diskutieren.

Das Ziel dieses Kapitels ist, durch die Analyse von Geometriemodellen und der physikalischen Erscheinungsformen von Straßen ein einheitliches Kriterium für die weitere Behandlung dieser Problematik zu schaffen. Damit ergibt sich eine Möglichkeit zu beant-

worten, was erreicht wurde und was noch zu lösen ist.

2.1 Geometrische Aspekte von Straßen

2.1.1 Generische Geometriemodelle von Straßen

Eine Straße kann entweder durch ihre beiden Begrenzungen oder durch ihre Mittellinie zusammen mit Quersegmenten beschrieben werden. Solche analytischen Beschreibungen beruhen direkt auf Geometriemodellen von Straßen. Diese Modelle werden bei der Erkennung von Straßen sowohl im Bildbereich als auch im Szenenbereich benötigt.

In Abb. 2.1 wird ein generisches geometrisches Modell von Straßen wiedergegeben ([Ozawa & Rosenfeld 86] und [Sakurai *et al.* 87]). Eine Straße wird dabei durch ihre Mittellinie in Abb. 2.1 (a) und (b) sowie ihren Superelevationswinkel in Abb. 2.1 (c) dargestellt. Die Mittellinie einer Straße besteht wiederum aus einem Kreisbogen und zwei mit ihm verbundenen Liniensegmenten in horizontaler Richtung¹ sowie einer Parabel und zwei mit ihr verbundenen Liniensegmenten in vertikaler Richtung.

Die horizontale Form der Straße in Abb. 2.1 (a) läßt sich mit den folgenden Gleichungen beschreiben:

$$y = m_1x + n_1, \quad (2.1)$$

$$(x - p)^2 + (y - q)^2 = r^2, \quad (2.2)$$

$$y = m_2x + n_2. \quad (2.3)$$

Die beiden Liniensegmente treffen sich mit dem Kreisbogen an den Punkten $P_1(x_1, y_1, z_1)$ und $P_2(x_2, y_2, z_2)$ in den jeweils tangentialen Richtungen, damit die Übergänge zwischen den Liniensegmenten und dem Kreisbogen glatt verlaufen. Die horizontale Kurve bewirkt eine horizontale Änderung der Straßenrichtung. Der Radius r darf eine Schwelle nicht unterschreiten, so daß Fahrzeuge mit hoher zugelassener Geschwindigkeit in der Kurve nicht ins Schleudern kommen.

Die vertikale Form der Straße in Abb. 2.1 (b) besteht aus einer Parabel und zwei Liniensegmenten:

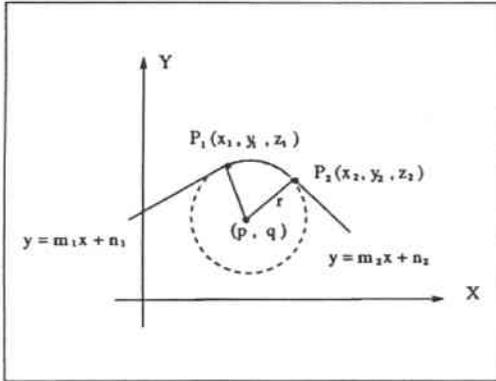
$$z = f_1x + g_1, \quad (2.4)$$

$$z = ax^2 + bx + c, \quad (2.5)$$

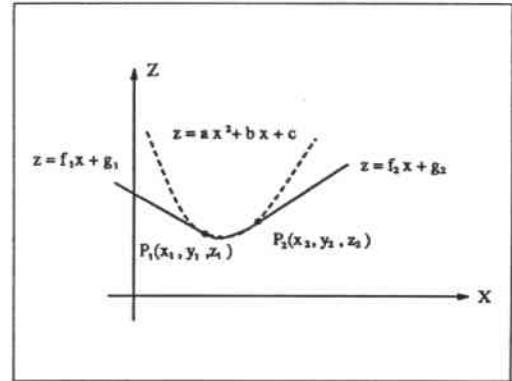
$$z = f_2x + g_2. \quad (2.6)$$

An den Punkten $P_1(x_1, y_1, z_1)$ und $P_2(x_2, y_2, z_2)$ lassen sich die Parabel und beide Liniensegmente verbinden. Die Neigung der Kurve wird beim Bau von Straßen berücksichtigt. Sie darf eine Schwelle nicht überschreiten.

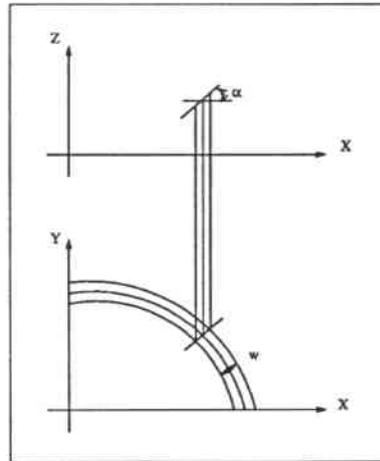
¹In [RAS-L-1 84] werden die Vorschriften für die Straßenkonstruktion in Deutschland beschrieben. Die Mittellinie von Straßen soll aus Geradensegmenten, Klotoiden und Kreisbögen bestehen. Die Klotoide vermittelt den Übergang zwischen der Geraden und dem Kreisbogen. Die Autobahn und Landstraße werden nach diesen Vorschriften aufgebaut. In der vorliegenden Arbeit werden hauptsächlich die Straßen der vierten Klasse (siehe Abschnitt 2.2.5) behandelt. Die Klotoide spielt in diesen Straßen keine wesentliche Rolle und wird daher nicht berücksichtigt.



(a) Ansicht einer Straßennittellinie von oben.



(b) Ansicht einer Straßennittellinie von der Seite.



(c) Der Superelevationswinkel einer Straße.

Abbildung 2.1: Darstellung eines generischen Straßenmodells mit einem Kreisbogen und zwei Liniensegmenten von oben in (a), einer Parabel und zwei Liniensegmenten von der Seite in (b) sowie einem Superelevationswinkel (Erläuterungen im Text) in (c).

Straßenbegrenzungen bestehen aus den Endpunkten der zur Mittellinie senkrechten Quersegmente. Die Quersegmente schneiden sich in ihren Mittelpunkten mit der Straßennittellinie und bilden einen vertikalen Superelevationswinkel α , wie in Abb. 2.1 (c) zu sehen ist. Durch einen positiven Superelevationswinkel wird die Fahrt von Fahrzeugen in Kurven begünstigt. Der Winkel α hängt vom Radius des Kreisbogens und der zugelassenen Geschwindigkeit von Fahrzeugen ab.

Durch Variation der Formparameter lassen sich verschiedene Spezialisierungen des Modells im Hinblick auf heute gebaute Straßen herstellen. In Tabelle 2.1 werden 12 mögliche

Straßenmodelle	Horizontale Form	Vertikale Form	Breite	α
	$m_1, n_1, p, q, r, m_2, n_2$	$f_1, g_1, a, b, c, f_2, g_2$		
Typ 1	Gerade	Gerade	k	0
	$m_1=m_2, n_1=n_2$	$f_1=f_2=0, g_1=g_2=k$		
Typ 2	Kreisbogen	Gerade	k	0
	p, q, r	$f_1=f_2=0, g_1=g_2=k$		
Typ 3	Kreisbogen, Geraden	Gerade	k	0
	$m_1, m_2, p, q, r, n_1, n_2$	$f_1=f_2=0, g_1=g_2=k$		
Typ 4	Gerade	Gerade	k	0
	$m_1=m_2, n_1=n_2$	$f_1=f_2 \neq 0, g_1=g_2=k$		
Typ 5	Kreisbogen	Gerade	k	0
	p, q, r	$f_1=f_2 \neq 0, g_1=g_2=k$		
Typ 6	Kreisbogen, Geraden	Gerade	k	0
	$m_1, m_2, p, q, r, n_1, n_2$	$f_1=f_2 \neq 0, g_1=g_2=k$		
Typ 7	Gerade	Parabel	k	0
	$m_1=m_2, n_1=n_2$	a, b, c		
Typ 8	Kreisbogen	Parabel	k	0
	p, q, r	a, b, c		
Typ 9	Kreisbogen, Geraden	Parabel	k	0
	$m_1, m_2, p, q, r, n_1, n_2$	a, b, c		
Typ 10	Gerade	Parabel, Geraden	k	0
	$m_1=m_2, n_1=n_2$	$f_1, g_1, a, b, c, f_2, g_2$		
Typ 11	Kreisbogen	Parabel, Geraden	k	0
	p, q, r	$f_1, g_1, a, b, c, f_2, g_2$		
Typ 12	Kreisbogen, Geraden	Parabel, Geraden	k	0
	$m_1, m_2, p, q, r, n_1, n_2$	$f_1, g_1, a, b, c, f_2, g_2$		

Tabelle 2.1: Diese Straßenmodelle im Szenenbereich entstehen durch die Beschränkung der freien Parameter in den Gleichungen (2.1)-(2.6). k bedeutet konstant (siehe auch Abb. 2.2). Die jeweils in einem Modell nicht vorkommenden Parameter werden nicht berücksichtigt.

geometrische Straßenmodelle im Szenenbereich durch die Beschränkung der freien Parameter in den Gleichungen (2.1)-(2.6) aufgelistet. Alleine durch die Parameter der Form sowohl in horizontaler als auch in vertikaler Richtung ist die Ausprägung des Modells aber noch nicht genügend beschrieben. Eine weitere wesentliche Größe ist der Superelevationswinkel, der aber bis heute bei der autonomen Fahrzeugführung noch nicht berücksichtigt wird. Die Breite der Straßen bleibt in Tabelle 2.1 unverändert.

Um einen illustrativen Eindruck von verschiedenen geometrischen Straßenmodellen zu gewinnen, werden in Abb. 2.2 die 2D-Abbildungen aller 12 Straßentypen aus Tabelle 2.1 dargestellt. Dabei werden die Begrenzungen von Straßen mit fettgedruckten Linien und die Mittellinien von Straßen mit unterbrochenen fettgedruckten Linien gezeichnet. Die Breite der Straßen wird nach wie vor als Konstante angenommen. Wie zu sehen ist, werden die Straßen in horizontaler Richtung als Linkskurve und Rechtskurve sowie

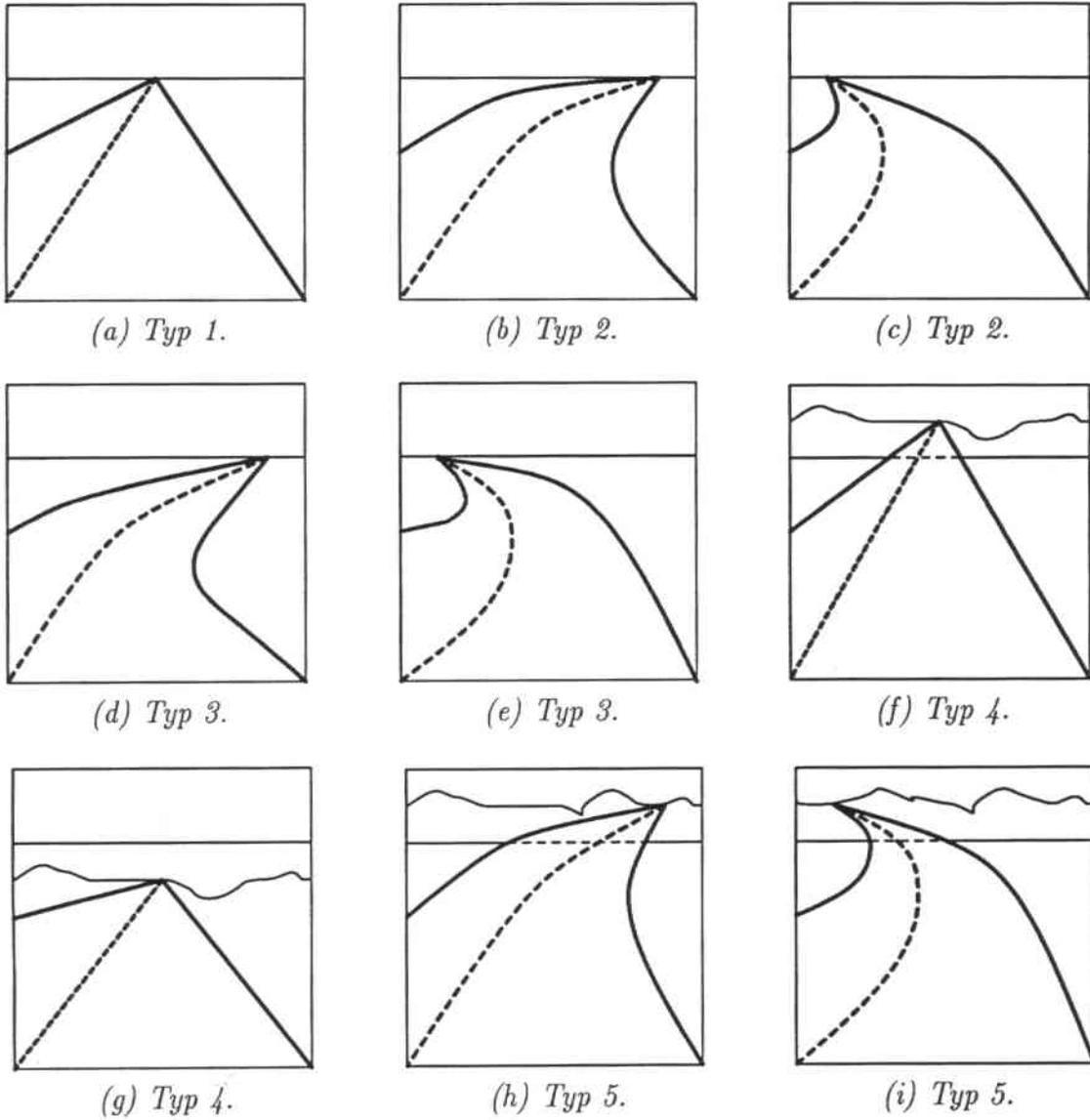
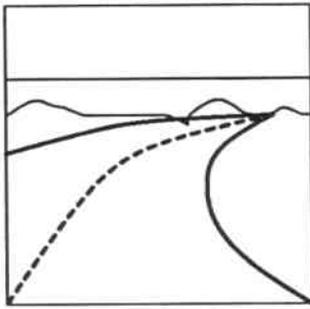
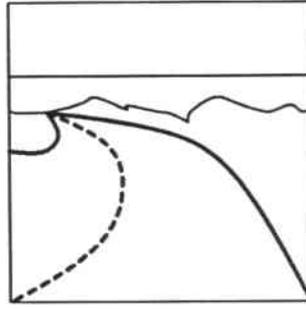


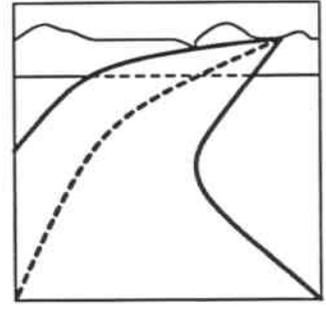
Abbildung 2.2: Darstellung der einzelnen Geometriemodelle von Straßen nach Tabelle 2.1: (a) Typ 1, geradlinige Straße auf einer Ebene; (b) und (c) Typ 2, Straßen mit kreisförmigen Rechts- und Linkskurven auf einer Ebene; (d) und (e) Typ 3, Straßen mit Rechts- und Linkskurven auf einer Ebene; (f) und (g) Typ 4, geradlinige Straßen auf bergauf- und bergabgerichteten Ebenen; (h)-(k) Typ 5, Straßen mit kreisförmigen Rechts- und Linkskurven auf bergauf- und bergabgerichteten Ebenen; (l)-(o) Typ 6, Straßen mit Rechts- und Linkskurve auf bergauf- und bergabgerichteten Ebenen; (p) und (q) Typ 7, geradlinige Straßen auf parabelförmigen Ebenen; (r)-(u) Typ 8-12, Straßen mit Kurven auf parabelförmigen Ebenen.



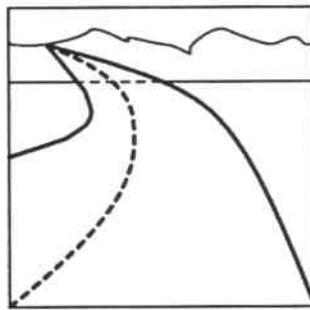
(j) Typ 5.



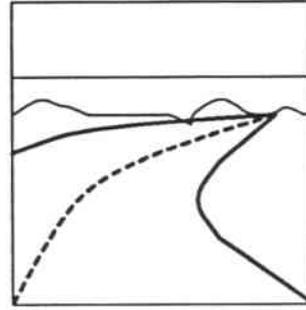
(k) Typ 5.



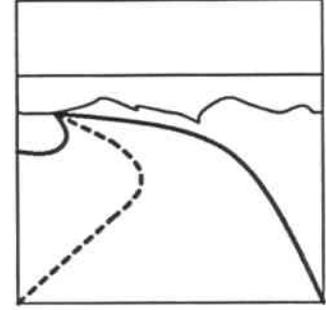
(l) Typ 6.



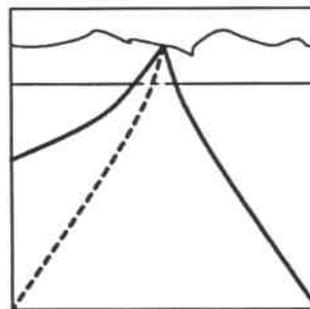
(m) Typ 6.



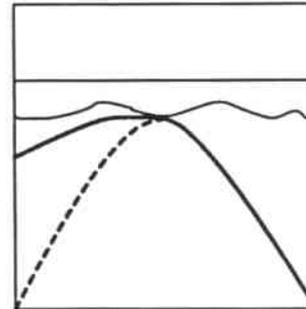
(n) Typ 6.



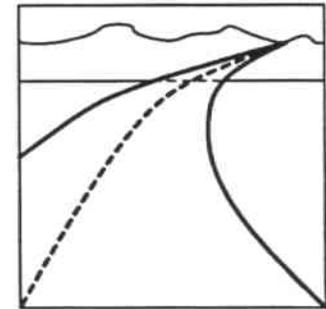
(o) Typ 6.



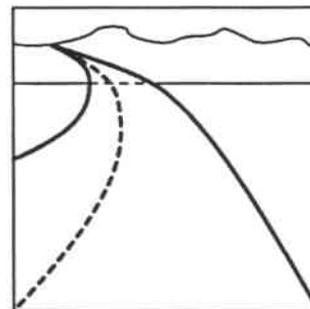
(p) Typ 7.



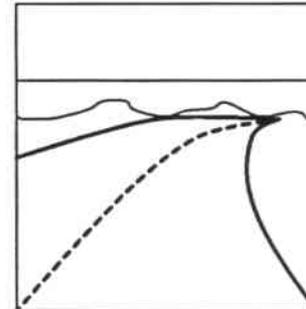
(q) Typ 7.



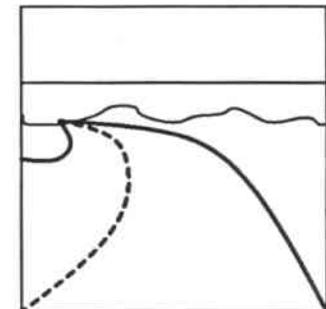
(r) Typ 8-12.



(s) Typ 8-12.



(t) Typ 8-12.



(u) Typ 8-12.

(Fortsetzung der vorherigen Seite)

in vertikaler Richtung als Bergaufkurve und Bergabkurve getrennt dargestellt. Für die Straßentypen 8-12 ist der Unterschied in den 2D-Abbildungen schwer erkennbar. Deshalb sind hier nur vier Beispiele gezeigt.

2.1.2 Verwendete Geometriemodelle von Straßen

Im folgenden sollen die am häufigsten verwendeten Geometriemodelle von Straßen sowohl im Szenenbereich als auch im Bildbereich anhand des in Abb. 2.1 dargestellten generischen Straßenmodells vorgestellt werden.

Ebenen-Modell

Das Ebenen-Modell einer Straße ist das einfachste Geometriemodell von Straßen im Szenenbereich (z.B. in [Turk *et al.* 88]). Dazu gehören die ersten drei Modelle in Tabelle 2.1. Sie werden in Abb. 2.2 (a) - (e) dargestellt. In einem Ebenen-Modell wird angenommen, daß die Straße in einer Ebene liegt, auf der sich das Fahrzeug bewegt. Es ist deutlich zu sehen, daß die Form von Straßen im Bildbereich mit diesem Modell nicht explizit definiert ist. Das Modell hat einige Vorteile gegenüber anderen Geometriemodellen von Straßen. Einer davon ist, daß das Modell für jeden Bildpunkt verwendet werden kann, der nicht unbedingt auf der Straßenbegrenzung liegen muß. Diese Eigenschaft bedeutet, daß mit diesem Modell auch Straßeneinmündungen bzw. -kreuzungen behandelt werden können. Die Anzahl der zu bestimmenden Parameter hängt davon ab, wie die Form der Straße in horizontaler Richtung im Szenenbereich angenommen wird.

Fluchtpunkt-Modell

Das Fluchtpunkt-Modell einer Straße besagt, daß sich die beiden Straßenbegrenzungen im 2D-Bild in einem Fluchtpunkt treffen (z.B. in [Liou & Jain 87]). Dieses Modell entspricht den Straßentypen 1 und 4 in Tabelle 2.1 und den Zeichnungen (a), (f) und (g) in Abb. 2.2. Dieses Modell ist offensichtlich im Bildbereich definiert. Solche Straßenbegrenzungen sind im 3D-Koordinatensystem gerade und zueinander parallel. Mit unterschiedlichen Neigungswinkeln der Kamera ändert sich auch die Position des Fluchtpunktes. Deshalb können auch bergabgerichtete und bergaufgerichtete Straßen wie Typ 4 in Tabelle 2.1 mit diesem Modell behandelt werden. Die Position des Fluchtpunktes stellt die Richtung des Straßenverlaufs dar. Ein Nachteil des Modells ist die Beschränkung seiner Anwendung auf Straßen mit parallelen geradlinigen Begrenzungen. Wenn die Breite der Straße bekannt ist, sind nur noch die Koordinaten des Fluchtpunktes im Bildbereich zu ermitteln.

Klotoiden-Modell

Eine Klotoide ist eine Kurve, deren Krümmung eine lineare Funktion der Bogenlänge ist. Das Klotoiden-Modell einer Straße besagt, daß eine Straßenbegrenzung im Szenenbereich eine Kombination von Geradensegmenten, Kreisbögen und glatten Kurven ist (z.B. in

[Zapp 88]). Eine glatte Kurve verbindet ein Geradensegment und einen Kreisbogen miteinander. Die Breite der Straßen ist auch in diesem Modell konstant. Das Straßenmodell kann als eine Modifikation des generischen Modells angesehen werden. Das Klotoiden-Modell entspricht eigentlich allen 12 Straßentypen in Tabelle 2.1 und Zeichnungen in Abb. 2.2. Das Modell eignet sich besonders für regelmäßig aufgebaute Straßen, wie z.B. Autobahnen.

Kreisbogen-Modell

Das Kreisbogen-Modell einer Straße beschreibt die Straßenbegrenzungen im Szenenbereich als zwei Kreisbögen mit demselben Mittelpunkt und unterschiedlichen Radien in einer Ebene ([Morgan *et al.* 90]). Mit anderen Worten ist das Kreisbogen-Modell eine vereinfachte Form des Ebenen-Modells mit der Beschränkung der Straßenform in horizontaler Richtung auf Kreisbögen (Typ 2 und 5 in Tabelle 2.1). Entsprechend den Gesetzen perspektivischer Abbildungen wird ein Kreis im Szenenbereich als Ellipse auf der Bildebene abgebildet. Ausprägungen dieses Modells werden in Abb. 2.2 (b)-(c) sowie (h)-(k) dargestellt. In diesem Modell können die Parameter der linken und rechten Straßenbegrenzungen getrennt bestimmt werden. Für jede Straßenbegrenzung sind 3 Parameter zu bestimmen. Das Modell erlaubt nur *eine* Krümmung einer Straßenbegrenzung in jeder Aufnahme.

Band-Modell mit variiertem Superelevationswinkel

[DeMenthon 87] modelliert die Straßenoberfläche im Szenenbereich als ein räumliches Band, das durch die Mittellinie und die dazu senkrecht stehenden Quersegmente beschrieben wird. Dies entspricht eigentlich dem generischen Modell von Straßen in [Ozawa & Rosenfeld 86] (Typ 1-12 in Tabelle 2.1 und Abb. 2.2 (a)-(u)).

2.1.3 Verfahren zur Ermittlung von Modellparametern

Entsprechend verschiedenen Geometriemodellen von Straßen sind ebenfalls verschiedene Parameter aus den segmentierten Daten sowohl im Bildbereich als auch im Szenenbereich zu bestimmen. Bei der Auswahl eines Verfahrens für die Bestimmung der Parameter eines Modells sollen vor allen die folgenden zwei Punkte beachtet werden. Zum einen spielt die Bearbeitungsgeschwindigkeit eines Verfahrens eine Rolle. Zum anderen ist die Robustheit des Verfahrens gegenüber dem in den Segmenten enthaltenen Rauschen zu beachten.

Im folgenden sollen vier wichtige Verfahren für die Bestimmung der Parameter der Geometriemodelle von Straßen beschrieben werden.

Ermittlung im Szenenbereich

Bei der Bestimmung der Parameter eines Geometriemodells können mit Hilfe der Parameter des Abbildungssystems zuerst die im Bild detektierten Kandidatenpunkte für Straßenbegrenzungen in den Szenenbereich rückprojiziert werden. Aus den Punkten im

Szenenbereich werden die Parameter des Modells ermittelt. Bei diesen Verfahren wird im allgemeinen vorausgesetzt, daß die Straße auf einer Ebene liegt. Die Position im Szenenbereich eines im 2D-Bild zu einer Straßenbegrenzung gehörigen Punktes kann durch die Lokalisierung des Schnittpunktes einer Ebene mit dem Vektor bestimmt werden, der vom Fokuspunkt der Kamera durch den Bildpunkt verläuft. Der Vorteil solcher Verfahren liegt in der Geschwindigkeit der Berechnung der Position eines abgebildeten Punktes im Szenenbereich. Diese Verfahren bringen aber zwei große Nachteile mit sich. Es ist sehr sensibel gegenüber Ungenauigkeiten des Neigungswinkels der Kamera. Außerdem ist das Ergebnis im Szenenbereich sehr stark vom Ergebnis im Bildbereich abhängig. Das Verfahren wird z.B. in [Turk *et al.* 88] zur Ermittlung der Parameter des Ebenen-Modells von Straßen verwendet.

Abstimmungsverfahren im Parameterraum

Abstimmungsverfahren zur Bestimmung der Parameter eines Geometriemodells bilden zuerst einen Parameterraum. Entsprechend dem jeweils eingesetzten Modell führt ein solches Verfahren einen Abstimmungsprozeß durch. Die Bestimmung der Parameter erfolgt dann durch eine Suche nach dem Maximum im Parameterraum. Ein Beispiel dafür ist die Bestimmung der Parameter des Fluchtpunkt-Modells in [Thorpe *et al.* 88]. Ein Vorteil solcher Verfahren besteht in ihrer Robustheit gegenüber Rauschen. Bei der Anwendung dieser Verfahren auf ein Modell mit vielen Parametern sieht man sofort einen ihrer Nachteile, nämlich daß die Suche nach dem Maximum in einem mehrdimensionalen Raum sehr zeitaufwendig ist.

Numerische Verfahren

Mit numerischen Verfahren wird die Bestimmung der Parameter eines Geometriemodells durch die Anpassung der detektierten Punkten an die Gleichungen der Straßenbegrenzungen realisiert. Die Schätzung erfolgt iterativ über eine Bildfolge. Beispiele dafür sind Kalman-Filter und Newtonsches Verfahren. Der Vorteil numerischer Verfahren liegt in ihrer Fähigkeit zur Bestimmung der Parameter komplizierter Geometriemodelle wie z.B. des Klotoiden-Modells.

„Zero-bank inverse perspective“-Verfahren

Basierend auf dem Band-Modell stellt [DeMenthon 87] einen iterativen Algorithmus zur Rekonstruktion der 3D-Straßengeometrie aus einem einzelnen Bild durch ein sogenanntes „zero-bank inverse perspective“-Verfahren vor. Das Verfahren wird in [DeMenthon & Davis 90] durch ein zusätzliches globales Optimierungsverfahren verbessert. Die Anwendung des Verfahrens auf reale Bilder ([Morgenthaler *et al.* 90]) zeigt vergleichbare Ergebnisse mit der aus Abstandsdaten ermittelten 3D-Straßengeometrie. Das Verfahren zur Rekonstruktion der 3D-Straßengeometrie eignet sich wegen seiner geringen Auswertegeschwindigkeit zur Zeit noch nicht für die Straßenverfolgung.

2.2 Physikalische Aspekte von Straßen

2.2.1 Untergliederung von Straßenszenen

Wir versuchen in diesem Abschnitt, Straßenszenen nach ihrer Komplexität in unterschiedliche Klassen zu untergliedern und geeignete Ansätze zur Extraktion von Straßenbeschreibungen in jeder Klasse zu diskutieren. Statt einer exakten Definition der Komplexität einer Straßenszene soll hier die Zusammensetzung der Szene herangezogen werden. Eine Straße besteht im allgemeinen aus drei Bereichen, nämlich der Straßenoberfläche, den Straßenbegrenzungen und der Straßenumgebung. Die Komplexität einer Straßenszene hängt von der Erscheinungsform aller drei Bereiche ab. Zur Analyse der Komplexität einer Straße sollen daher alle drei Bereiche in Betracht gezogen werden. Die Komplexität des Bildes einer Straßenszene hängt von zwei Faktoren ab. Der eine ist die Komplexität der Straße im Szenenbereich. Der andere ist der Typ der Sensoren für die Abbildung der Szenen. Da die Analyse von Sensoren zur Fahrzeugführung nicht den Schwerpunkt dieser Arbeit darstellt, gehen wir im folgenden davon aus, daß passive bildgebende Sensoren bei der Abbildung von Straßenszenen für die Extraktion von Straßenbeschreibungen verwendet werden. Somit steht nur die Komplexität von Straßen im Szenenbereich zur Diskussion.

Die Oberfläche einer Straße kann im Grauwert sowohl homogen als auch inhomogen sein. Mit einer homogenen Straßenoberfläche ist hier die Straße gemeint, die wenig Textur aufweist. Dabei werden Störungen ausgeschlossen, die durch andere Objekte oder Schattenwürfe verursacht werden. Unter einer inhomogenen Straßenoberfläche versteht man einerseits Bereiche, in denen Baumblätter, Schnee, Schattenwürfe, Kanaldeckel usw. auftreten. Andererseits gehören dazu auch Straßenoberflächen, die aus Materialien verschiedener Farbe bestehen.

Die Straßenbegrenzung ist eines der wichtigsten Merkmale von Straßen. Die Beschaffenheit der Straßenbegrenzungen spielt daher eine wichtige Rolle bei der Detektion von Straßenbeschreibungen aus einem Bild. Als Begrenzungen kommen sowohl farbige abgesetzte Markierungsbänder als auch die Bordsteinränder in Betracht. Allerdings gibt es auch Straßen, die keine begrenzenden Markierungen haben oder deren Markierungen nicht erkennbar sind.

Uns interessieren bei der Extraktion von Straßenbeschreibungen im wesentlichen nur die Straßenoberfläche und deren Begrenzungen. In vielen Fällen, insbesondere bei fehlenden Begrenzungen kommt aber auch die Analyse der Straßenumgebung ins Spiel. Zur Straßenumgebung gehören alle Gegenstände, die sich neben einer Straße befinden. Das sind z.B. Gebäude, Gras, Bäume usw. Sie besitzen im allgemeinen andere Grauwerte, Farben und Texturen als die Straßenoberfläche.

Ausgehend von diesen Überlegungen wird in Abb. 2.3 eine Untergliederung von Straßenszenen dargestellt. Dabei lassen sich Straßen nach der Komplexität der Straßenoberfläche, der Straßenumgebung und der Straßenränder in vier Klassen einteilen, und zwar

Klasse 1: Straßen mit homogener Oberfläche und markierten Begrenzungen,

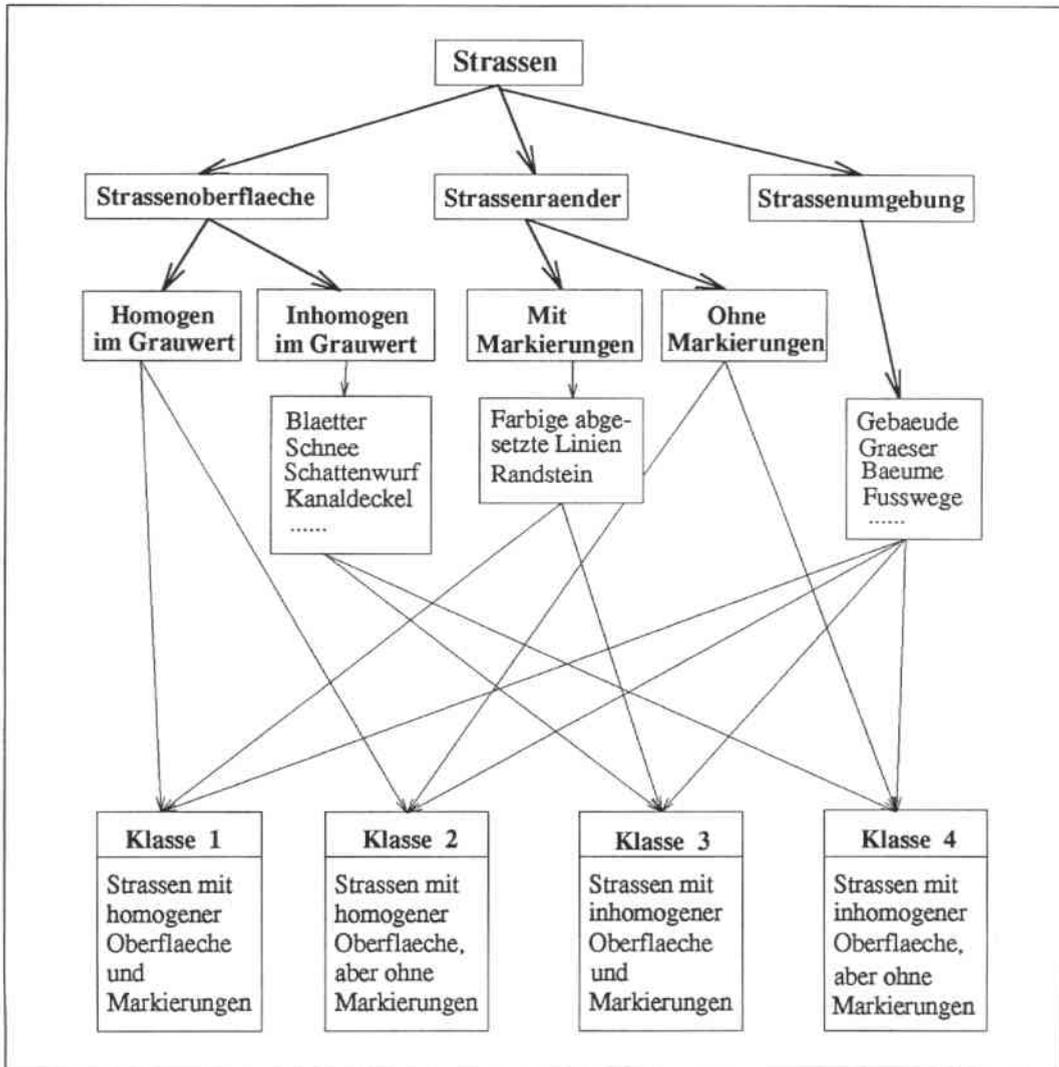


Abbildung 2.3: Untergliederung von Straßenszenen in vier Klassen je nach Komplexität der Straßenoberfläche, der Straßenumgebung und der Straßenränder.

Klasse 2: Straßen mit homogener Oberfläche ohne markierte Begrenzungen,

Klasse 3: Straßen mit inhomogener Oberfläche und markierten Begrenzungen sowie

Klasse 4: Straßen mit inhomogener Oberfläche ohne markierte Begrenzungen.

Im folgenden wollen wir Straßen jeder Klasse im einzelnen analysieren. Dies soll unter Berücksichtigung der folgenden Gesichtspunkte durchgeführt werden:

- Aufgabenbasierte Sichtsysteme spielen eine wichtige Rolle bei den derzeitigen Sichtsystemen. Es ist daher sinnvoll, die zu lösenden Aufgaben zuerst zu analysieren und abzugrenzen. Bei der Extraktion von Straßenbeschreibungen hat daher eine

Einordnung von Straßen für öffentlichen Verkehr, auf privatem Gelände sowie von Querfeldein-Wegen in die in dieser Arbeit eingeführten charakteristischen Klassen zu erfolgen.

- Die geometrische Form von Straßen unterscheidet zwischen verschiedenen Klassen. Es soll die Beziehung zwischen Form und Beschaffenheit jeder Straße untersucht werden.
- Es können eine Vielzahl von Grauwertstrukturen im Bild als mögliche Merkmale für die Beschreibung von Straßen in Betracht gezogen werden. Eine entscheidende Frage dabei ist, welche davon am wichtigsten für die jeweils zu lösende Aufgabe sind.
- Es gibt keine allgemein gültigen Bildauswerteverfahren, die bei jeder Straßenszene erfolgreich eingesetzt werden können. Für unterschiedliche Straßen sollen auch unterschiedliche Ansätze verwendet werden. Wo liegen die Leistungen und Grenzen verschiedener Verfahren?

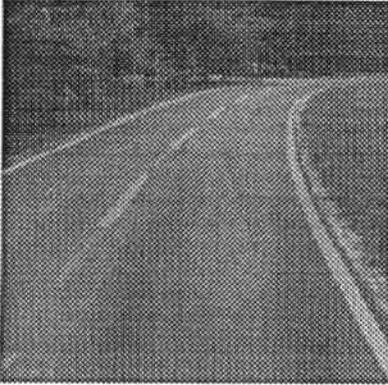
2.2.2 Straßen der ersten Klasse

Straßen der ersten Klasse sollen eine homogene Oberfläche und markierte Begrenzungen besitzen. Bevor wir solche Straßen analysieren, untersuchen wir zuerst die Frage, welche Straßen in Wirklichkeit in die erste Klasse eingeordnet werden können. In natürlichen Szenen gehören Autobahnen, gut markierte Landstraßen und Straßen des öffentlichen Verkehrs innerhalb von Ortschaften usw. zu dieser Klasse. Als Beispiele werden in Abb. 2.4 vier Straßenbilder der ersten Klasse dargestellt.

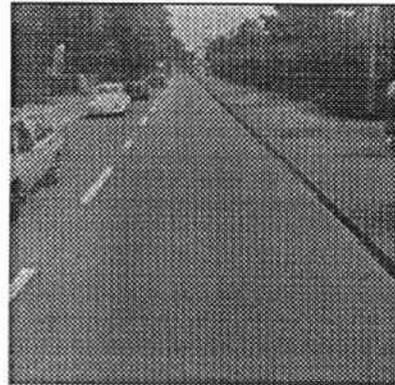
Straßen der ersten Klasse haben im allgemeinen klar erkennbare geometrische Formen. Sie wurden nach bestimmten Vorschriften, wie z.B. dem in Abb. 2.1 vorgestellten generischen Geometriemodell von Straßen, aufgebaut. Bei der Erkennung von Straßen der ersten Klasse können alle existierenden geometrischen Modelle in Betracht gezogen werden. Hier sind besonders explizite Modelle vorteilhaft.

Wie aus Abb. 2.4 erkennbar wird, erscheint die Oberfläche einer Straße der ersten Klasse im allgemeinen homogen im Grauwert. Der Unterschied im Grauwert zwischen der Straßenoberfläche und den Markierungsbändern ist im Bild leicht erkennbar. Aufgrund der Tatsache, daß die Diskontinuität der Grauwertfunktion zwischen der Straßenoberfläche und den Markierungen das wichtigste Merkmal ist, sind daher die Kantenelemente in Bildern solcher Szenen zu detektieren.

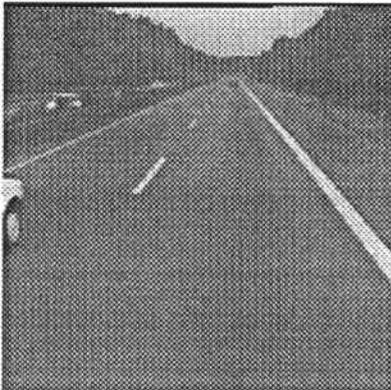
Zur Detektion der Straßenbegrenzungen aus den Bildern der ersten Klasse werden in erster Linie konturbasierte Verfahren eingesetzt. Mit Hilfe geometrischer Straßenmodelle können sogar nur ausgewählte kleine Bildausschnitte ausgewertet werden. Vorteil konturbasierter Verfahren ist vor allem ihre große Verarbeitungsgeschwindigkeit. Dies ist bei der Fahrzeugführung eine der wichtigsten Anforderungen. Diese Verfahren haben aber zweifellos den Nachteil, daß sie nicht robust gegenüber Szenenstörungen sind. Unter



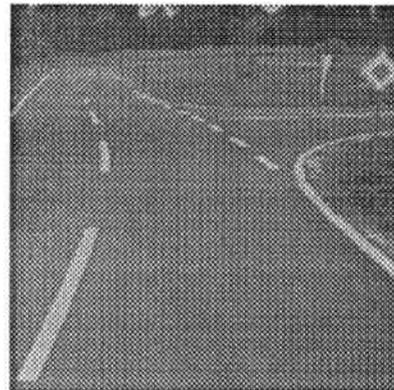
(a) Landstraße.



(b) Innerörtliche Straße.



(c) Autobahn.



(d) Innerörtliche Straße.

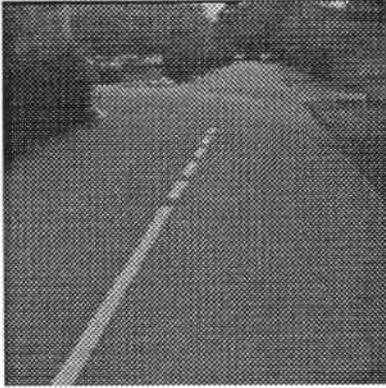
Abbildung 2.4: Beispiele von Straßen mit homogener Oberfläche und markierten Begrenzungen. Nach der Untergliederung von Straßenszenen in Abb. 2.3 gehören diese Straßen zur ersten Klasse.

Robustheit versteht man beim maschinellen Sehen die Eigenschaft, daß die Leistung eines Verfahrens durch Störungen nicht wesentlich gemindert wird.

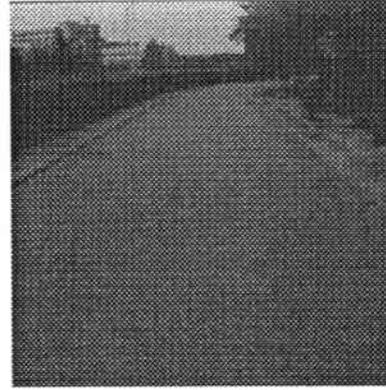
2.2.3 Straßen der zweiten Klasse

Als Straßen der zweiten Klasse sind Straßen mit homogener Oberfläche ohne markierte Begrenzungen gemeint. Zur zweiten Klasse von Straßen gehören Straßen in privatem Gelände und manche Landstraßen, deren farbig abgesetzte Markierungen nach einer gewissen Zeit nicht mehr zu erkennen sind. Wir zeigen in Abb. 2.5 vier Beispiele von Straßen der zweiten Klasse.

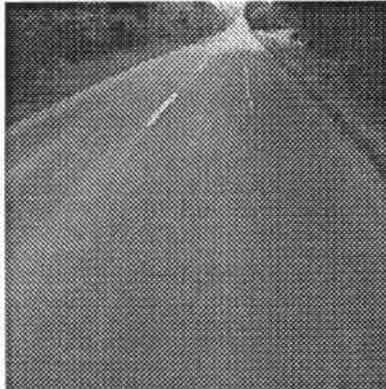
Straßen der zweiten Klasse, insbesondere Straßen auf privatem Gelände, haben keine strengen Vorschriften bezüglich ihrer Form. Dies bedeutet, daß die Verwendung impli-



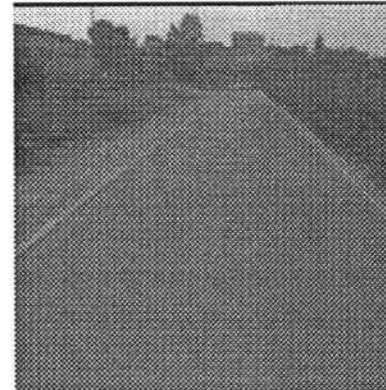
(a) Straße auf privatem Gelände.



(b) Straße auf privatem Gelände.



(c) Landstraße.



(d) Innerörtliche Straße.

Abbildung 2.5: Beispiele von Straßen mit homogener Oberfläche ohne markierte Begrenzungen. Nach der Untergliederung von Straßenszenen in Abb. 2.3 gehören diese Straßen zur zweiten Klasse.

ziter Geometriemodelle von Straßen, wie z.B. das Ebenen-Modell, bei deren Erkennung vorteilhaft ist.

Mit den oben eingeführten Annahmen über die Komplexität von Straßen der zweiten Klasse können in solchen Bildern sowohl die Diskontinuitäten der Grauwertfunktion als auch die Homogenität in Grauwert oder Farbe als Grundlage der Merkmalsextraktion herangezogen werden. Im allgemeinen ist die Homogenität der Straßenoberfläche in Grauwert oder Farbe das vorwiegende Merkmal.

Entsprechend dem ausgewählten Merkmal können sowohl konturbasierte als auch bereichsbasierte Verfahren eingesetzt werden. In vielen Verfahren versucht man, Probleme konturbasierter Verfahren, die auf die Bilder von Straßenszenen der ersten Klasse erfolgreich angewendet wurden, weitgehend zu umgehen. Bei konturbasierten Verfahren wird vorausgesetzt, daß die Diskontinuität der Grauwertfunktion zwischen Straßenoberfläche

und -begrenzungen signifikant ist. Dies bedeutet, daß die Ergebnisse solcher Verfahren sehr stark von der Komplexität der betreffenden Straßen abhängen. Der Vorteil der konturbasierten Verfahren ist nach wie vor ihre große Verarbeitungsgeschwindigkeit.

Die Verwendung bereichsbasierter Verfahren zur Extraktion von Straßenbegrenzungen bietet eine zweite Möglichkeit zur Lösung dieser Problematik. Statt der Grauwertkanten, die den Grauwertübergang in einem kleinen lokalen Bereich repräsentieren, versucht man ein Bild nach der Homogenität im Grauwert oder in der Farbe zu segmentieren. Mit bereichsbasierten Verfahren läßt sich die Extraktion von Straßen der zweiten Klasse grundsätzlich befriedigend lösen, da wir angenommen haben, daß Straßen dieser Klasse eine homogene Oberfläche haben.

2.2.4 Straßen der dritten Klasse

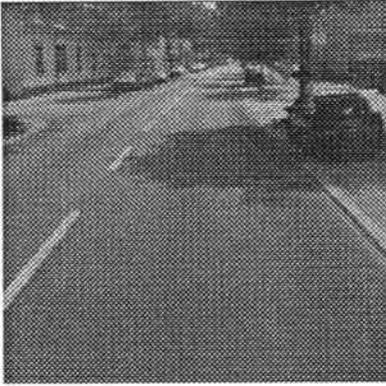
Straßen mit inhomogener Oberfläche und markierten Begrenzungen werden nach der Untergliederung in Abb. 2.3 als dritte Klasse bezeichnet. Straßen dieser Klasse sind meistens diejenigen Straßen der ersten Klasse, deren Oberfläche und evtl. auch Begrenzungen von anderen Objekten überlagert sind oder im Schatten anderer Objekte liegen. Dazu gehören Autobahnen, Landstraßen sowie innerörtliche Straßen. Als Beispiele werden vier Bilder solcher Straßen in Abb. 2.6 dargestellt.

Störungen auf Straßen der ersten Klasse ändern die physikalischen Erscheinungsformen von Straßen. Die Geometrie der Straßen wird dadurch aber nicht beeinflusst, d.h., daß bei der Erkennung von Straßen der dritten Klasse ebenfalls explizite Geometriemodelle, wie z.B. das Klotoiden-Modell oder Fluchtpunkt-Modell, verwendet werden können.

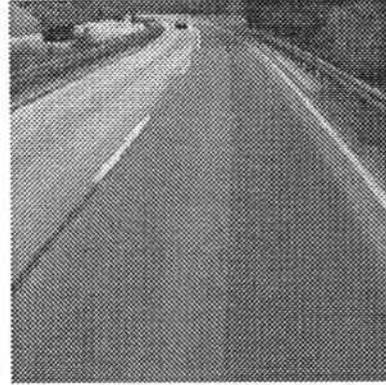
Analog zu den Überlegungen bei der Analyse von Straßen der ersten Klasse ist die Diskontinuität der Grauwertfunktion zwischen der Straßenoberfläche und den Markierungsbändern das wichtigste Merkmal. Wegen der Verdeckung der Straßenoberfläche durch andere Objekte oder Schattenwürfe ist dies in der Praxis nicht mehr der Fall. Auf der anderen Seite ist die Oberfläche solcher Straßen zwar inhomogen in Grauwert oder Farbe, Unterbereiche weisen aber Homogenität auf. Es läßt sich im allgemeinen nicht leicht zeigen, welches Merkmal im Bild für diese Aufgabe signifikanter ist. Aus dieser Überlegung ziehen wir den Schluß, daß sowohl die Diskontinuität der Grauwertfunktion in einer kleinen Umgebung als auch die Homogenität eines Bereiches in Grauwert oder Farbe für eine Merkmalsextraktion in Betracht gezogen werden sollen.

Zur Extraktion von Beschreibungen für Straßen der dritten Klasse können wie bei der Auswertung von Straßen der zweiten Klasse konturbasierte und bereichsbasierte Verfahren eingesetzt werden. Konturbasierte Verfahren orientieren sich an der Grauwertänderung zwischen Straßenoberfläche und deren Begrenzungen. Aus den obigen Überlegungen wissen wir, daß diese Methoden nur zum Erfolg führen, wenn die Störungen im Gegensatz zu den erwarteten Merkmalen zu vernachlässigen sind.

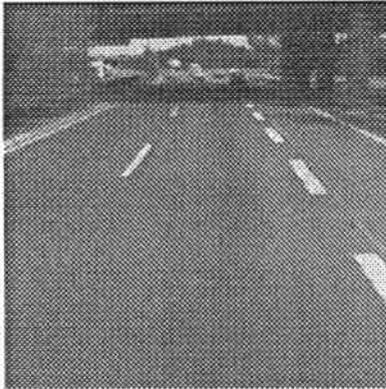
Bereichsbasierte Verfahren haben sich bei der Auswertung von Straßen der dritten Klasse als robuster erwiesen als konturbasierte Verfahren. Dabei wird die Straßenoberfläche möglicherweise zunächst in einige kleine Bereiche geteilt. Daraus wird der Straßebereich mit Hilfe geometrischer Straßenmodelle bestimmt.



(a) Innerörtliche Straße.



(b) Landstraße.



(c) Autobahn.



(d) Straße auf privatem Gelände.

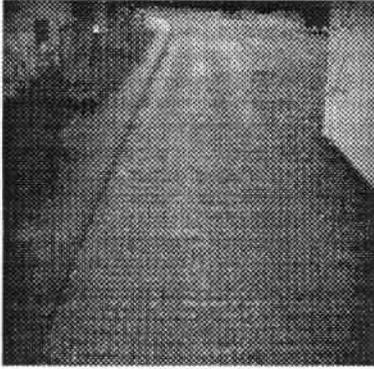
Abbildung 2.6: Beispiele von Straßen mit inhomogener Oberfläche und markierten Begrenzungen. Nach der Untergliederung von Straßenszenen in Abb. 2.3 gehören diese Straßen zur dritten Klasse.

Allein mit kontur- oder bereichsbasierten Verfahren läßt sich die Extraktion der Beschreibungen von Straßen der dritten Klasse nur teilweise lösen. Nicht zuletzt stören hierbei Schattenwürfe.

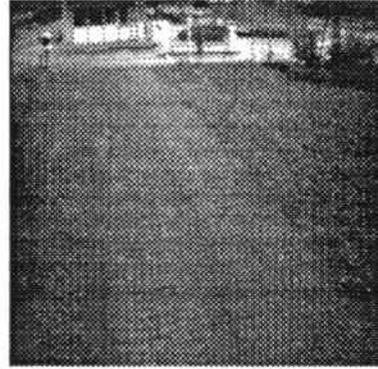
2.2.5 Straßen der vierten Klasse

Straßen, die nicht nur keine homogene Oberfläche, sondern auch keine markierten Begrenzungen haben, lassen sich nach Abb. 2.3 in die vierte Klasse einordnen. Im öffentlichen Verkehr kommen solche Straßen selten vor. Sie befinden sich zumeist auf privatem Gelände und Querfeldein-Wegen. In Abb. 2.7 werden vier Beispiele dafür gezeigt.

Straßen der vierten Klasse sind auch von Menschen oft nicht auf den ersten Blick als Straße zu erkennen. Der Grund dafür ist, daß sie nicht nach den Vorschriften für die



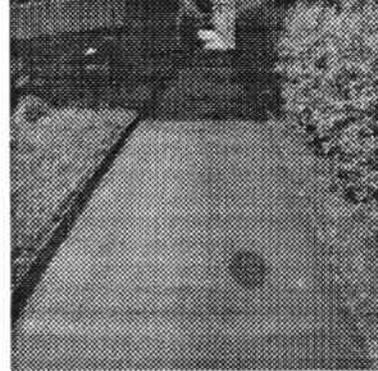
(a) Straße auf privatem Gelände.



(b) Straße auf privatem Gelände.



(c) Straße auf privatem Gelände.



(d) Straße auf privatem Gelände.

Abbildung 2.7: Beispiele von Straßen mit inhomogener Oberfläche ohne markierte Begrenzungen. Nach der Untergliederung von Straßenszenen in Abb. 2.3 gehören diese Straßen zur vierten Klasse.

Straßenkonstruktion gebaut wurden. Bei der Erkennung solcher Straßen sollen implizite Geometriemodelle, wie z.B. das Ebenen-Modell, eingesetzt werden.

In Bildern solcher Straßen ist die Grauwertänderung zwischen der Straßenoberfläche und der Umgebung nicht mehr überall signifikant. Außerdem besteht die Straßenoberfläche nicht mehr aus in Grauwert oder Farbe homogenen Bereichen. Textur ist hier das wichtigste Merkmal zur Unterscheidung zwischen der Straßenoberfläche und der Umgebung. Die Beschaffenheit der Oberfläche einer Straße der vierten Klasse läßt sich durch Oberflächen-textur beschreiben.

Um die Textur auf der Straßenoberfläche von derjenigen der Umgebung zu unterscheiden, können viele Texturmerkmale in Betracht gezogen werden. Konturbasierte Verfahren kommen hier nicht mehr in Frage. Das Problem der Extraktion von Straßenbeschreibungen läßt sich daher als Textursegmentierung formulieren.

Kapitel 3

Bildauswerteverfahren zur Fahrzeugführung: Stand der Forschung

Die Forschungstätigkeit zur autonomen Führung von Straßenfahrzeugen hat in den letzten Jahren durch militärische und industrielle Förderung weltweit stark zugenommen. Straßenfahrzeuge wurden mit Sensoren und Rechnern ausgerüstet und als Versuchsträger eingesetzt. Zur Quer- und Längsführung solcher Fahrzeuge wurden verschiedenartige Systeme entwickelt. Als ein wichtiger Bestandteil eines Regelkreises in solchen autonomen Systemen soll das Sichtsystem in der Lage sein, hilfreiche und präzise Informationen über den Umgebungszustand für Quer- und Längsführung anzubieten.

Im folgenden sollen die Sichtsysteme und Bildauswertungs-Algorithmen einiger wichtiger Forschungsprojekte zur autonomen Führung von Straßenfahrzeugen vorgestellt werden. Zur Abgrenzung der Aufgabenstellung bei der Straßenmodellierung, Fahrbahnerkennung, Hindernisdetektion sowie Landmarkenerkennung sollen zunächst die gewünschten Aufgaben verallgemeinert dargestellt werden. Anschließend werden die aufgestellten Verfahren zur Lösung bzw. partiellen Lösung solcher Probleme ausführlich beschrieben.

Die folgenden Aspekte sollen als gewünschte Eigenschaften eines autonomen Systems bei der Zusammenfassung der Literatur berücksichtigt werden:

Echtzeitverarbeitung : Während der Verarbeitung der Information bewegt sich ein Fahrzeug auf der Straße weiter. Es ist notwendig, die Interpretation einer Szene so schnell wie möglich durchzuführen, um rechtzeitig entsprechende Aktionen unternehmen zu können.

Allgemeingültigkeit : Ein Fahrzeug wird auf verschiedenartigen Straßen, wie z.B. Autobahnen, Landstraßen, innerörtlichen Straßen usw. fahren. Jeder Straßen-Typ hat seine eigene Anordnung von Einmündungen und Kreuzungen. Ein Sichtsystem zur autonomen Fahrzeugführung soll unabhängig vom Straßen-Typ immer seine Aufgabe ausführen können.

Robustheit : Die Oberflächenbeschaffenheit von Straßen ist sehr unterschiedlich. Außerdem kommen im Bild einer Straßenszene oft unerwünschte Intensitätsänderungen, wie z.B. Schattenflecke oder Nässe auf der Fahrbahn vor, die die Fahrbahnerkennung erschweren. Ein Sichtsystem soll bei solchen ungünstigen Szenenverhältnissen richtige Ergebnisse liefern.

Flexibilität : Ein Sichtsystem zur autonomen Fahrzeugführung wertet Bilder von natürlichen Szenen aus. Während der Führung eines Fahrzeugs kann sich die Beleuchtung stark ändern, wie z.B. bei Verdeckung der Sonne durch Wolken oder bei Einfahrt in eine bzw. Ausfahrt aus einer Unterführung. Das Sichtsystem muß seine Parameter aber auch an eine veränderte Szene entlang einer Straße anpassen, z.B. beim Abbiegen von einer Durchgangsstraße auf einen Feldweg oder in eine Hauseinfahrt.

3.1 Versuchsträger zur autonomen Fahrzeugführung

In diesem Abschnitt werden die dem Autor bekannt Versuchsträger mit den jeweiligen Zielsetzungen und dem erreichten Stand zur autonomen Führung von Straßenfahrzeugen vorgestellt. Die Darstellung der hier vorgestellten Fahrzeuge soll in erster Linie zur Gewinnung eines groben Eindrucks von autonomen Fahrzeugen dienen. Auf die Details der in diesen Fahrzeugen eingesetzten verschiedenen Verfahren wird in den nachfolgenden Abschnitten des Kapitels noch ausführlich eingegangen.

Die meisten Forschungsgruppen zur autonomen Führung von Straßenfahrzeugen in den USA haben zunächst bei der Entwicklung autonomer Roboter in Laborumgebungen Erfahrungen gesammelt. Sie legen den Schwerpunkt auf die Interpretation der Umgebung und verwenden Farbkameras und Laser-Abtastsysteme. Einige Projekte wurden als Teilprojekte des ALV¹, ein Projekt der DARPA², in der Zeit zwischen 1984 und 1988 gefördert. Darunter sind beispielsweise *Martin Marietta Denver Aerospace*, die *University of Maryland*, die *Carnegie-Mellon University (CMU)* und die *University of Massachusetts at Amherst* zu nennen.

Um die im Rahmen des Projektes ALV entwickelten Verfahren zu demonstrieren, wurde als Versuchsträger ein achträdiges Fahrzeug mit hydraulischem Antrieb bei *Martin Marietta Denver Aerospace* aufgebaut. Im Fahrzeug wurden zwei Bildauswertesysteme VICOM, eine parallele *Warp Maschine* der CMU, eine CCD-Farbkamera und ein Laser-Abtastsystem des Forschungsinstituts ERIM³ installiert. Das Versuchsfahrzeug steht allen am ALV beteiligten Forschungsgruppen zur Verfügung. Das Ziel des Projektes ALV ist die Entwicklung von Algorithmen zur autonomen Führung von Fahrzeugen auf den Straßen der zweiten und vierten Klasse. Das Fahrzeug konnte durch ein sichtsystemgestütztes Navigationssystem VITS⁴ ([Turk & Marra 86], [Turk et al. 88]) im Jahr 1986 mit einer Geschwindigkeit von 10 km/h automatisch fahren. Dabei wurden Hindernisse

¹ALV: „Autonomous Land Vehicle“.

²DARPA: „Defense Advanced Research Projects Agency“.

³ERIM: „Environmental Research Institute of Michigan“.

⁴VITS: „Vision Task Sequencer“.

detektiert. Auf einer hindernisfreien Straße erreicht das Fahrzeug Geschwindigkeiten bis zu 20 km/h. Die Teststrecken sind die Straßen der zweiten Klasse.

FMC Corporation hat ab 1982 parallel zum Projekt *ALV* ein eigenes Versuchsfahrzeug (einen Panzerwagen) entwickelt ([Kuan *et al.* 86+88]). Das autonome System zur Fahrzeugführung bei *FMC* ist ein missionsorientiertes System; es besteht aus einer Kombination von Fernbedienung und autonomer Führung. Zum Erfassen und zum Interpretieren des Umgebungszustandes werden eine Farbkamera, ein Ultraschall-Sensor, ein Infrarot-Sensor und ein Bildauswertesystem im Fahrzeug eingesetzt. Neben der autonomen Führung des Fahrzeugs besitzt das Projekt bei *FMC* noch die militärische Anwendung des Fahrzeugs als Entwicklungsziel. Bis zum Jahr 1986 konnte das Fahrzeug mit einer Geschwindigkeit von 8 km/h mit Hindernisdetektion und 19 km/h ohne Hindernisdetektion autonom fahren. Die Teststrecken in diesem Projekt sind ebenfalls die Straßen der zweiten Klasse.

An der *Carnegie-Mellon University (CMU)* werden seit 1984 Versuche in natürlicher Umgebung mit dem kleinen Fahrzeug *Terregator* ([Wallace *et al.* 86]) und dem *Navlab*⁵ ([Thorpe *et al.* 88]) durchgeführt. Der fahrende Navigationsträger *Navlab*, ein handelsüblicher Lastkraftwagen mit einem hydraulischen Antrieb und einer Elektrosteuerung, ist mit Farbkameras und einem Laser-Abtastsystem des *ERIM* ausgerüstet. Das Fahrzeug *Navlab* ist nur für die Vorführung der entwickelten Algorithmen zur autonomen Führung von Fahrzeugen gedacht. Die bis 1987 erzielten Ergebnisse wurden in [Goto & Stentz 87] zusammengefaßt. Das autonome System kann das *Navlab* mit einer Geschwindigkeit von 0.72 km/h (200 mm/s) auf den Straßen der zweiten und vierten Klasse automatisch führen. Dabei entdeckt es Hindernisse und erkennt Landmarken. Mit dem Verfahren *YARF*⁶ zur Straßenerkennung erreicht die Geschwindigkeit bis zu 24 km/h auf den Straßen der ersten Klasse ([Thorpe *et al.* 91]). Das System *ALVINN*⁷ führt sogar das *Navlab* mit einer Geschwindigkeit von 88 km/h eine 34 km lange Strecke auf einer Autobahn. Die Arbeiten an der *CMU* befassen sich mit beinahe allen Bereichen der Aufgaben für ein Sichtsystem zur autonomen Fahrzeugführung und spielen eine führende Rolle in der Entwicklung autonomer Systeme ([Thorpe 91] und [Thorpe *et al.* 92]).

Die in Europa entwickelten Sichtsysteme für die Fahrzeugführung zeichnen sich besonders dadurch aus, daß die speziellen Aspekte der Echtzeitverarbeitung berücksichtigt werden. Die meisten Sichtsysteme besitzen explizite geometrische Modelle der Fahrbahn, verwenden monochrome Kameras und werten Bildausschnitte aus. Die meisten Projekte werden durch das im Jahr 1986 von der Automobilindustrie initiierte *EUREKA*-Projekt *Prometheus*⁸ unterstützt. Ein Teilziel von *Prometheus* ist die Entwicklung eines Fahrer-Unterstützungssystems zur Führung von Straßenfahrzeugen ([Nagel 88] und [Franke 89]). Die Verantwortung für die Fahrzeugführung soll zunächst noch beim Mensch als Fahrer bleiben.

⁵ *Navlab*: „Navigation Laboratory“.

⁶ *YARF*: „Yet Another Road Follower“.

⁷ *ALVINN*: „Autonomous Land Vehicle In a Neural Network“.

⁸ *Prometheus*: „Programme for a european traffic with highest efficiency and unprecedented safety“.

An der Universität der Bundeswehr München (*UniBwM*) wurde das *VaMoRs*⁹, ein Kastenwagen vom Typ L508D der Mercedes-Benz AG mit Automatikgetriebe, als Versuchsträger ausgerüstet ([Dickmanns & Graefe 88b]). Zum Erfassen des Umgebungszustandes wird ein Doppelkameranystem mit einer Tele-Kamera für die Objekterkennung und einer Weitwinkel-Kamera für die Straßenverfolgung auf einer drehbaren Zweiachsplattform im Fahrzeug eingesetzt. Als Bildverarbeitungssystem dient zuerst das parallele Mehrrechnersystem *BVV 2 + 3* ([Kuhnert 86] und [Graefe 90]). Ab 1992 wird es durch Transputer-netze ersetzt ([Dickmanns *et al.* 93]). Entscheidende Vorteile gegenüber den Ergebnissen in den *USA* sind die Echtzeitverarbeitung und die Längsführung des Fahrzeuges. Das Versuchsfahrzeug *VaMoRs* kann mit einer Geschwindigkeit von bis zu 96 *km/h* auf freier Autobahn fahren. Auf den Straßen der zweiten und dritten Klasse sind die Fahrerergebnisse noch nicht in ähnlicher Weise dokumentiert wie für die Autobahnfahrt.

In September 1991 wurde bei der Vorführung von Ergebnissen aus verschiedenen *Prometheus*-Teilprojekten der Versuchsträger *VITA*¹⁰ der Daimler-Benz AG ([Ulmer 92]), ein Lastkraftwagen *D811* der Mercedes-Benz AG mit einem automatischen Getriebe, vorgestellt. Zum Erfassen des Umgebungszustandes wird das gleiche Doppelkameranystem wie in *VaMoRs* verwendet. An Versuchen mit dem *VITA*-Versuchsträger beteiligten sich neben dem Daimler-Benz Konzern auch weitere Institutionen, u.a. die *UniBwM*, das Fraunhofer-Institut für Informations- und Datenverarbeitung (*IITB*), die Universität Bochum, die Universität Paderborn. Bei der Vorführung wurden die folgenden Funktionen gezeigt: die Verfolgung der Straßenbegrenzungen, die Hindernisdetektion, die Verfolgung eines vorausfahrenden Fahrzeuges und die Erkennung von Straßenkreuzungen durch die Analyse der Form der Markierungsbänder. Im August 1991 hat das *VITA* eine Teststrecke von 130 *km* auf einer Autobahn mit einer Geschwindigkeit von bis zu 80 *km/h* gefahren ([Franke *et al.* 91]). Das ist die bisher längste Teststrecke für eine autonome Fahrzeugführung.

In der Volkswagen AG (*VW*) wurde das *ICAD*¹¹, ein Teilprojekt von *Prometheus*, entwickelt und in einem Versuchsträger eingesetzt ([Dreyer *et al.* 90]). Das Sichtsystem *OPTOPILOT II* dient dabei zur Erkennung von Straßenbegrenzungen durch die Farbbildauswertung ([Wilm 90]). Zur Hindernisdetektion wird zusätzlich ein Laser-Abtastsystem verwendet. Auf den Straßen der ersten Klasse erreicht das System eine Geschwindigkeit von bis zu 80 *km/h*.

Mit derselben Zielsetzung wie bei *Prometheus*, zur Erhöhung der Verkehrssicherheit und -effizienz, beginnt in Japan das gesamte Projekt *SSVS*¹² in 1990 ([Tsugawa *et al.* 92]). Einer der Schwerpunkte des Projektes ist die Entwicklung von Systemen zur autonomen Fahrzeugführung. In der ersten Phase des Projektes, die Analyse und Präzisierung der zu lösenden Aufgaben, wird festgehalten, daß solche autonomen Systeme erst in 20 - 30 Jahre am öffentlichen Verkehr teilnehmen können. In Japan wird fast in jeder Automobilfirma die autonome Führung von Fahrzeugen untersucht, z.B. *Toyota* ([Shigematu *et al.* 92]),

⁹ *VaMoRs*: Versuchsfahrzeug für autonome Mobilität und Rechnersehen.

¹⁰ *VITA*: „*Vision Technology Application*“.

¹¹ *ICAD*: „*Intelligent Computer Aided Driving*“.

¹² *SSVS*: „*Super Smart Vehicle System*“.

Daihatsu ([Ito *et al.* 92]), *Honda* ([Hashimoto *et al.* 92]), *Nissan* ([Aono 89]) and *Mazda* ([Okuno *et al.* 92]). Im folgenden sollen zwei Projekte, die jeweils einen eigenen Versuchsträger betreiben, vorgestellt werden. Die japanischen Systeme sind meistens sehr speziell für eigene Versuchsträger und Umgebung entwickelt. Sie sind echtzeitfähig, erscheinen aber nach den bisher verfügbaren Informationen nur begrenzt ausbaufähig.

In Zusammenarbeit mit der Fa. *Fujitsu* hat die Fa. *Nissan* das autonome System *PVS*¹³ erarbeitet. Im Jahr 1989 kann das System mit einer Geschwindigkeit von bis zu 30 km/h auf einer speziellen Teststrecke autonom fahren ([Aono 89]). In diesem gesamten Projekt werden Untersuchungen verschiedener Funktionen zur autonomen Fahrzeugführung unternommen. [Ozaki *et al.* 89] stellen das Sichtsystem *IDATEN* für die Straßenerkennung im Grauwertbild und Hindernisdetektion durch ein Stereoverfahren im Videotakt vor. Dabei werden 3 Kameras mit unterschiedlicher Ausrichtung für die Straßenerkennung und 2 Kameras für die Hindernisdetektion eingesetzt. In [Hattori *et al.* 92] wird die Erkennung gekrümmter Straßen der ersten Klasse unterschiedlicher Breite präsentiert. Dabei wird nur eine Kamera auf einer drehbaren Plattform verwendet. Zur Hinderniserkennung wird in [Sekine *et al.* 92] die Verwendung eines Laser-Abtastsystems im Fahrzeug berichtet. Mit diesem Verfahren wird dazu die Verfolgung eines vorausfahrenden Fahrzeugs erprobt. Die Teststrecken sind dabei nach wie vor die Straßen der ersten Klasse.

In der Fa. *Mazda* wurde das Fahrzeug *MOVER-2*¹⁴, ein handelsüblicher Lastkraftwagen, mit einer *CCD*-Kamera und Rechner ausgerüstet ([Okuno *et al.* 92]). Bei der autonomen Führung des Fahrzeugs erreicht die Geschwindigkeit bis zu 50 km/h auf einer geraden Straße der ersten Klasse, bis zu 30 km/h auf einer kreisförmigen Straße der ersten Klasse und bis zu 15 km/h auf einer S-förmigen Straße der ersten Klasse. Andere Funktionen, wie z.B. die Hindernisdetektion und Landmarkenerkennung, wurden noch nicht im Fahrzeug implementiert.

Um die vom Fraunhofer-Institut für Informations- und Datenverarbeitung (*IITB*) entwickelten Verfahren zur autonomen Führung eines Straßenfahrzeugs zu testen, wurde in Zusammenarbeit mit der Robert Bosch GmbH im Jahr 1988 ein Versuchsfahrzeug, ein Lastkraftwagen *D 609* der Mercedes-Benz AG mit einem automatischen Getriebe beschafft ([Geisler *et al.* 90], [Siegle *et al.* 92] und [Struck *et al.* 93]). In Abbildung 3.1 ist das Fahrzeug dargestellt. Es ist zur Zeit mit einem Leitreechner *EUROCOM 5* der Fa. *Eltec*, einem Bildauswertesystem *VISTA*¹⁵ des *IITB* ([Paul *et al.* 88]), einem Transputernetz mit zehn Transputern, drei Video-Kameras, einem System zur Unterstützung der Fahrzeugnavigation *Travelpilot* der Robert Bosch GmbH sowie Stellsystemen und anderen erforderlichen Einrichtungen ausgerüstet. Rechner, Sensoren und Stellsysteme werden von einer speziellen Einheit (einer Motor-Generator-Kombination) elektrisch versorgt, so daß ein autonomer Betrieb des Fahrzeuges mit seinen Einrichtungen möglich ist.

Die im *IITB* entwickelten Verfahren zeichnen sich im Vergleich mit oben beschriebenen Ansätzen zur autonomen Fahrzeugführung durch folgende Eigenschaften aus:

- Erkennung von einfach strukturierten Straßeneinmündungen und -kreuzungen;

¹³*PVS*: „Personal Vehicle System“.

¹⁴*MOVER*: „MOBILE Vision Experimental Robot“.

¹⁵*VISTA*: Visuelles Interpretations-System für Technische Anwendungen.



Abbildung 3.1: *Das Versuchsfahrzeug des Fraunhofer-Instituts für Informations- und Datenverarbeitung (IITB).*

- Kombination schneller modellgestützter Verfahren in Bildausschnitten zur Fahrspurverfolgung mit aufwendigeren datengetriebenen Verfahren im Gesamtbild zur Erkennung von Fahrbahnanordnungen;
- Kombination einer sichtsystemgestützten Umgebungserfassung mit einer kartenbasierten Navigation;
- Ausführung spezieller Manöver vor komplexen Fahrbahnanordnungen mit Unterstützung des Navigationssystems.

3.2 Konturbasierte Verfahren zur Fahrbahnerkennung

Bei der Erkennung von Fahrbahnen durch konturbasierte Verfahren wird vorausgesetzt, daß Diskontinuitäten in Grauwerten oder in Farbtönen die dominanten Merkmale zwischen Fahrbahn und Fahrbahnbegrenzung im Bild sind. Aus der Analyse von Straßen in Abschnitt 2.2 wissen wir, daß die Straßen der ersten Klasse diese Eigenschaft besitzen. Zur Fahrbahnerkennung in solchen Bildern sind konturbasierte Verfahren geeignet und robust. Wie in Abschnitt 2.2 angedeutet wurde, können die Begrenzungen in Bildern der Straßen der ersten und dritten Klasse teilweise mit konturbasierten Verfahren ermittelt werden, unter der Annahme, daß die Kantenelemente zwischen Fahrbahnen und Fahrbahnumgebungen in Bildern trotz der Störungen signifikant sind. Die konturbasierte

Fahrbahnerkennung wird im allgemeinen durch Teilbildauswertung, durch modellbasierte Gruppierung oder durch Auswertung von Abstandsbildern realisiert.

Fahrbahnerkennung durch Teilbildauswertung

Die an der *UniBwM* entwickelten Verfahren zur Fahrbahnerkennung sind typische Vertreter von Verfahren zur Fahrbahnerkennung durch Bildausschnittsauswertung. [Dickmanns & Graefe 88a+b] betrachten die autonome Führung eines Straßenfahrzeuges als eine Aufgabe mit Schwerpunkt auf der Regelungstechnik. Der zentrale Bestandteil der Fahrzeugführung durch Rechnersehen ist ein integriertes raum-/zeitliches 4D-Modell, die Modellierung der Szenenveränderungen im 3D-Raum und der Zeit. Die Architektur des im Fahrzeug eingesetzten parallelen Echtzeitbildverarbeitungssystems *BVV 2* bzw. *BVV 3* ist ebenfalls in Bezug auf dieses 4D-Modell aufgebaut ([Kuhnert 86] und [Graefe 90]).

Anhand der Modellierung von Fahrbahnbegrenzungen als Klotoide in der Ebene werden bei der Merkmalsextraktion Verfahren der gesteuerten Korrelation mit einer idealen stufenförmigen Kantenschablone nur in kleinen ausgewählten Teilbildern angewandt. Die Konturpunkte in einem Bildausschnitt werden zu einem Liniensegment zusammengefaßt. In [Dickmanns & Zapp 86] und [Zapp 88] wird ein krümmungsbasiertes Verfahren zur Fahrbahnverfolgung vorgestellt. Diskrete Kalman-Filter dienen hier zur Schätzung der Fahrbahnkrümmung und der Fahrzeugzustände. Die Modellierung von Fahrbahnbegrenzungen wird in [Behringer *et al.* 92] auf Klotoiden im Raum erweitert. Die Ermittlung von Fahrbahnbegrenzungen wird in [Behringer *et al.* 92] auf einen Transputernetz parallelisiert und in *VaMoRs* sowie *VITA* implementiert.

Eine Forschungsgruppe an der *University of Bristol* arbeitet seit dem Jahr 1986 auf dem Gebiet der autonomen Führung von Fahrzeugen. In [Morgan *et al.* 90] werden die Fahrbahnbegrenzungen als zwei Kreisbögen modelliert. Zur Extraktion linearer Strukturen aus jedem selektierten Ausschnitt wird ein gerichteter Kantendetektor verwendet. Das Verfahren hat Schwierigkeiten auf Straßen mit mehreren Fahrbahnbegrenzungen und unterbrochenen Markierungsbändern als Fahrbahnbegrenzungen. Dazu wird in [Schaaser & Thomas 92] ein Verfahren zur Ermittlung der Fahrbahnbegrenzung durch die Erkennung der weißen Markierungsbänder untersucht. Das Verfahren besteht aus der schwellwertbasierten Segmentation der weißen Markierungsbänder im Grauwertbild, der anschließenden Formanalyse und der Extraktion der Fahrbahnbegrenzungen aus den Markierungsbändern mit Hilfe des Kreisbogen-Modells von Straßen. Zur Erhöhung der Robustheit der Erkennung von Markierungsbändern haben [Campbell & Thomas 92] ein Modell mit 6 Parametern für ein Markierungsband definiert und ein modellbasiertes Erkennungsverfahren entwickelt. Das Verfahren zur Fahrbahnverfolgung funktioniert zuverlässig auf den Straßen der ersten Klasse.

[Franke 91] modelliert die Fahrbahn wie [Dickmanns & Zapp 86] als Klotoide. Zur Schätzung der Parameter des Modells wird statt eines Kalman-Filters ein iteratives Newtonsches Optimierungsverfahren verwendet. Bei der Ermittlung der Fahrbahnbegrenzung wird ein eindimensionaler stufenförmiger Kantendetektor innerhalb einiger selektierter Bildausschnitte verwendet. Zur Erhöhung der Verarbeitungsgeschwindigkeit wird

ein Transputernetz eingesetzt. Ähnlich wie die Verfahren der Forschungsgruppe an der *UniBwM* arbeitet das Verfahren auf den Straßen der ersten Klasse robust.

Die Volkswagen AG in Wolfsburg hat zur Führung von Fahrzeugen das Bildverarbeitungssystem *OPTOPILOT* entwickelt ([Zimdahl *et al.* 86]). Die Fahrbahnbegrenzung wird durch die Analyse weniger Zeilen innerhalb eines mit einer Farb-Fernsehkamera aufgenommenen Bildes ermittelt. Ein weiteres Beispiel für die Verfolgung des Straßenverlaufes in einer Farbbildfolge findet man in [Münkel & Welz 91].

Zur Erkennung von Straßen der ersten Klasse haben [Thorpe *et al.* 91] das Verfahren *YARF* entwickelt. Das Ziel dabei ist, mit Hilfe expliziter Geometriemodelle die Markierungsbänder in Bildausschnitten rasch zu detektieren und damit eine höhere Geschwindigkeit zu erreichen. Bei der Detektion der Markierungsbänder wird zuerst jedes Pixel in den ausgewählten Ausschnitten nach seinem Farbton binarisiert. Die geometrischen Eigenschaften der als Markierungsbänder detektierten Bereiche werden anschließend berechnet. *YARF* modelliert eine Straßenmittellinie als eine Parabel auf einer Szenenebene ([Kluge & Thorpe 92]). Die Breite der Straße ist konstant. Die Bestimmung der Parameter des Modells erfolgt durch eine Minimierung der quadratischen Abweichungen. Noch ein anderer erwähnenswerter Aspekt des Verfahrens besteht in der Modellierung der Fehler sowohl bei der Merkmalsdetektion als auch bei der Modellanpassung.

Die soeben vorgestellten Verfahren zur Teilbildauswertung verarbeiten Ausschnitte aus einem Bild und zeichnen sich durch ihre Echtzeitverarbeitungsfähigkeit aus. Die Auswahl der Position jedes Bildausschnittes spielt dabei eine entscheidende Rolle und hängt stark von der Modellierung der Straßengeometrie, der Schätzung des Fahrbahnverlaufs und der Schätzung der Fahrzeugbewegung ab. Die Hauptschwachstelle der Verfahren zur Fahrbahnerkennung durch Bildausschnittsauswertung liegt darin, daß sich in jedem Bildausschnitt zumindest ein dominantes Liniensegment befinden und dieses dominante Liniensegment ein Teil der Fahrbahnbegrenzung sein muß. Solche Verfahren verwenden explizite Modelle von Straßen und funktionieren robust auf den Straßen der ersten Klasse.

Fahrbahnerkennung durch modellbasierte Gruppierung

Zur Erkennung der Straßen der zweiten und dritten Klasse sowie zur Erkennung von Straßeneinmündungen und -kreuzungen genügen die Informationen innerhalb einiger kleiner Bildausschnitte bei konturbasierten Verfahren nicht mehr. Alle signifikanten Merkmale der Straße und von Objekten auf sowie neben der Straße sollten bei der Interpretation der Szene für die Fahrbahnerkennung ausgenützt werden. Hierzu werden Verfahren zur Fahrbahnerkennung durch die straßenmodellgesteuerte Gruppierung der lokalen Strukturen im Vollbild untersucht.

Zur Fahrbahnverfolgung werden in der *University of Maryland* konturbasierte Verfahren zur Auswertung von Bildern monochromer Kameras entwickelt ([Waxman *et al.* 85], [LeMoigne *et al.* 86] sowie [Davis *et al.* 86]). Nach der Faltung eines Grauwertbildes mit dem Sobel-Operator wird die Hough-Transformation zur Extraktion linearer Strukturen verwendet. Aus den detektierten Liniensegmenten werden dominante Geradensegmente, sogenannte „*pencils*“, gruppiert, die sich in einem Fluchtpunkt treffen und die den gera-

den Fahrbahnrandern im 3D-Koordinatensystem entsprechen. In einer Vorwärtskopplung werden vorherige Ergebnisse zusammen mit dem Fluchtpunkt-Modell der Straße und der Fahrzeugbewegung verwendet, um die mögliche Anfangsposition der Fahrbahnbegrenzungen im Folgebild vorauszusagen. Die Bewegung des Fahrzeugs muß sehr genau bekannt sein, um diese Position in der Vorwärtskopplungs-Phase richtig zu lokalisieren.

In [Kenue 89a] wird ein Verfahren zur Detektion der Straßenbegrenzungen vorgestellt, das ein Grauwertbild ebenfalls mit dem Sobel-Operator auswertet. Weitere Schritte des Verfahrens bestehen in der Extraktion der Liniensegmente durch die Hough-Transformation und in der Ermittlung des Fluchtpunktes der Straßenbegrenzungen.

[Liou & Jain 87] präsentieren ein Verfahren zur Extraktion der Fahrbahnbegrenzung durch die Verifikation der hypothetischen Fluchtpunkte im aktuellen Bild. Die hypothetischen Fluchtpunkte werden aus dem Ergebnis im vorangehenden Bild unter Berücksichtigung der Richtung und Geschwindigkeit der Fahrzeugbewegung generiert. Aufbauend auf diesem Ergebnis wird weiter ein krümmungsbasierter Algorithmus zur Fahrbahnerkennung in [Polk & Jain 92] vorgestellt. Statt eines Fluchtpunktes werden 6 Fluchtpunkte entsprechend 6 horizontalen Bildstreifen detektiert. Durch diese Vorgehensweise wird die Möglichkeit zur approximativen Beschreibung des Verlaufs von gekrümmten Straßen geschaffen infolge der Anwendung von Gruppierungsverfahren auf Kantenelemente im gesamten Bild.

Die Gruppierung lokaler Strukturen zu Fahrbahnbegrenzungen im Vollbild ist eine Kombination von modell- und datengesteuerten Verfahren. Solche Gruppierungsverfahren zur Fahrbahnerkennung haben bisher nur Fluchtpunkt-Modelle von Straßen verwendet. Die Verwendung anderer Straßenmodelle bei der Gruppierung lokaler Strukturen soll ebenfalls untersucht werden, um gekrümmte Straßen und evtl. auch Straßenkreuzungen zu erkennen. Eine effiziente Kombination des aus einem Geometriemodell ausgeprägten Wissens mit den aus dem Bild gruppierten Strukturen spielt eine wichtige Rolle in der modellbasierten Erkennung der Begrenzungen der Straßen der ersten und dritten Klasse. Ein Nachteil solcher Verfahren bei Anwendungen ist ihre geringe Verarbeitungsgeschwindigkeit infolge der Anwendung von Gruppierungsverfahren auf Kantenelemente im gesamten Bild.

Fahrbahnerkennung durch Auswertung von Abstandsbildern

Aktive Sensoren, wie z.B. Laser-Abtastsysteme oder Ultraschall-Sensoren, werden zum Erfassen von Umgebungszuständen bei der Fahrzeugführung eingesetzt. Die derart gewonnenen Daten enthalten Informationen über den Abstand zwischen dem abzubildenden Sensor und abgebildeten Objekten. Solche Daten werden nicht nur zur Hindernisdetektion, sondern auch zur Erkennung von Fahrbahnen, die nicht auf derselben Ebene wie Fahrbahnrande liegen, verwendet.

In [Sharma & Davis 86] wird die Fahrbahnerkennung durch die Auswertung der Abstandsdaten aus einem Laser-Abtastsystem untersucht. Die Abstandsdaten werden zuerst von einem sphärischen Koordinatensystem in ein kartesisches Koordinatensystem konvertiert. Darauf folgen die Anpassung jedes Punktes mit seiner lokalen Umgebung an eine

quadratische Fläche, die Berechnung der minimalen bzw. maximalen Hauptkrümmung für jeden Punkt, die Bestimmung von Konturpunkten durch einen Vergleich der Flächenkrümmung jedes Punktes mit einer Schwelle und die Detektion von Liniensegmenten aus Konturpunkten durch die Hough-Transformation. Wegen der Empfindlichkeit des Sensors gegen Richtungsänderungen der Objektoberfläche ist das Verfahren nicht sehr robust.

[Bergman & Cowan 86] untersuchen die Detektion der Fahrbahnbegrenzungen ebenfalls durch eine Auswertung von Abstandsdaten. Nach der Rauschunterdrückung in einem Abstandsbild wird bei der Fahrbahndetektion jede Zeile punktweise von links nach rechts durchlaufen, bis ein Punkt gefunden wird, dessen Wert innerhalb einstellbarer Schwellen liegt. Bei der Lokalisierung der Punkte von Fahrbahnbegrenzungen werden die in der vorangehenden Zeile ermittelten Positionen berücksichtigt.

Die Anwendung der Verfahren zur Fahrbahnerkennung durch die Auswertung von Abstandsdaten beschränken sich nur auf die Fahrbahnen, die nicht auf derselben Ebene wie Fahrbahnränder liegen. Ein besonderer erwähnenswerter Vorteil solcher Verfahren ist ihre Unempfindlichkeit gegenüber den durch Schattenwurf verursachten Störungen.

3.3 Bereichsbasierte Verfahren zur Fahrbahnerkennung

Bei den bereichsbasierten Verfahren zur Fahrbahnerkennung wird vorausgesetzt, daß Fahrbahnen bzw. auch Nicht-Fahrbahnen homogene Bereiche bezüglich der Eigenschaften

- Grauwert,
- Farbton oder
- Textur.

sind. Bereichsbasierte Verfahren werden bisher hauptsächlich bei Straßen der zweiten Klasse eingesetzt. Zur Bestimmung von Fahrbahnbegrenzungen wird ein Bild zunächst in Fahrbahnbereiche und Nicht-Fahrbahnbereiche segmentiert. Zu den existierenden Klassifikationsverfahren zur Fahrbahnerkennung gehören sowohl überwacht eingelernte Klassifikatoren¹⁶ als auch unüberwacht eingelernte Klassifikatoren¹⁷.

Überwacht eingelernte Klassifikatoren

Überwacht eingelernte Klassifikatoren verwenden Modelle der Merkmale (wie z.B. Farbmodell, Texturmodell) der Klassen „Fahrbahn“ und „Nicht Fahrbahn“. Die meisten existierenden überwacht eingelernten Klassifikatoren zur Fahrbahnerkennung verwenden den

¹⁶Überwacht eingelernte Klassifikatoren sind die Klassifikatoren, deren Parameter mit Hilfe der überwachten Lernverfahren ermittelt worden sind.

¹⁷Unüberwacht eingelernte Klassifikatoren sind die Klassifikatoren, deren Parameter mit Hilfe der unüberwachten Lernverfahren ermittelt worden sind.

Farbton als Merkmal jeder Klasse. Jedes Farbmodell einer Klasse enthält beispielweise den Mittelwert und die Standardabweichung eines Farbtons dieser Klasse.

Zur Fahrbahnverfolgung wurden an der *CMU* hauptsächlich eine Bereichsanalyse in Farbbildern zur Erkennung unmarkierter Fahrbahnen durchgeführt. Zur Verstärkung des Dynamikbereiches einer einzelnen Kamera wurden zwei Farbkameras mit unterschiedlicher Blendeneinstellung eingesetzt. [Crisman & Thorpe 88] präsentieren den Algorithmus *SCARF*¹⁸ zur Klassifikation der Bildpunkte in Farbbildern. Im Gegensatz zu [Turk *et al.* 88], die nur je ein Farbmodell für „Fahrbahn“ bzw. „Nicht Fahrbahn“ verwenden, greift *SCARF* auf mehrere Farbmodelle zurück (jeweils 4 Farbmodelle für die Klassen „Fahrbahn“ und „Nicht Fahrbahn“ in [Thorpe *et al.* 88]; jeweils 12 Farbmodelle für die Klassen „Fahrbahn“ und „Nicht Fahrbahn“ in [Crisman & Thorpe 88]). Bei der Klassifikation wird jeder einzelne Bildpunkt anhand seines Farbtons im 6D-Farbraum (zwei Farbbilder jeder Szene) der Klasse mit maximaler Mutmaßlichkeit („*Likelihood*“) zugeordnet. Nach der Klassifikation jedes Bildpunktes wird statt der Konturverfolgung in jedem Bereich die Fahrbahn unter Verwendung des Fluchtpunkt-Modells direkt durch ein bereichsbasiertes Abstimmungsverfahren detektiert. Die Klassifikations- und Abstimmungsverfahren werden in einem Bild mit reduzierter Auflösung 30×32 aus einem 480×512 Bild durchgeführt. Diese Reduzierung der Bilddimension bringt nicht nur den Vorteil höherer Verarbeitungsgeschwindigkeit mit sich, sondern unterdrückt auch Anomalieeffekte wie z.B. Risse auf der Straßenoberfläche.

Als eine Alternative zur Farbklassifikation wird in [Thorpe *et al.* 88] auch eine überwacht eingelernte Texturklassifikation durchgeführt. Als Texturmerkmal dient hier die Anzahl der Kantenelemente innerhalb eines Bildausschnittes. Nach wie vor wird hier die Texturklassifikation auch in einem 30×32 Pixel großen Bild durchgeführt. Das Ergebnis der Texturklassifikation wird mit der Farbklassifikation kombiniert und dem nächsten Modul für Bereichskonturverfolgung übergeben.

Die Ermittlung der Fahrbahnbegrenzungen wird in [Turk & Marra 86] und [Turk *et al.* 88] durch die Segmentierung eines Farbbildes und die anschließende Konturverfolgung des Fahrbahnbereiches erzielt. Es wird zuerst ein einkanaliges Bild durch eine linear gewichtete Kombination des R-, G- und B-Anteils jedes Farbbildpunktes aus dem RGB-Raum auf die Intensitätsachse projiziert. Basierend auf einer linearen Diskriminanten-Funktion wird jeder einzelne Bildpunkt des einkanaligen Bildes jeweils der Klasse „Fahrbahn“ oder „Nicht Fahrbahn“ zugeordnet. Die Schwelle für die Klasse „Fahrbahn“ wird aus einer Kombination des Mittelwertes und der Standardabweichung der Intensitätswerte innerhalb eines durch ein sogenanntes „*power windowing*“-Verfahren bestimmten Bildausschnittes berechnet. Nach der Segmentierung eines Bildes erfolgt die Ermittlung der Straßenbegrenzungen durch Randverfolgung in der Mitte des unteren Bildbereichs. Dabei wird angenommen, daß sich das Fahrzeug auf der Fahrbahn befindet.

In [Kuan *et al.* 86+88] und [Kuan & Sharma 87] werden Verfahren zur Fahrbahnverfolgung und Kreuzungserkennung präsentiert. Hier wird ähnlich wie in [Turk *et al.* 88] zuerst ein einkanaliges Bild aus dem RGB-Raum auf die Intensitätsachse projiziert. Jedes Pixel wird dann durch einen histogrammgesteuerten Klassifikator der Klasse „Fahrbahn“

¹⁸*SCARF*: „*Supervised Classification Applied to Road Following*“.

oder „Nicht Fahrbahn“ zugeordnet. Der Fahrbahnbereich im Bild wird nach bestimmten Kriterien selektiert und anschließend werden die Konturen des Fahrbahnbereiches durch ein Polygon beschrieben. Diese Beschreibung wird einem Geometrie-Modul übergeben, um Straßenränder und Straßenkreuzungen durch modellgesteuerte Verfahren zu identifizieren.

In [Wilm 90] wird das Bildverarbeitungssystem *OPTOPILOT II* vorgestellt. Zur Fahrbahnerkennung werden eine Farbkamera verwendet und Verfahren für die Farbsegmentierung entwickelt. Während des Einzugs eines Farbbildes wird jedes Pixel hinsichtlich seiner Farbe durch eine Tabelle (in 8 Ebenen) mit „0“ oder „1“ bewertet. Als Eingangssignal für die weitere Auswertung stehen bis zu 8 Binärbilder zur Verfügung, wobei in jedem dieser Binärbilder ein bestimmter Farbbereich ausgefiltert ist. Ein RGB-Binärbild wird durch eine logische Verknüpfung der Binärbilder aus 3 Kanälen generiert. Wie Fahrbahnbegrenzungen aus dem Binärbild bestimmt werden, läßt sich aus dieser Arbeit leider nicht genau ablesen.

Die Anwendungen der künstlichen neuronalen Netze bei Handschrift- und Spracherkennung haben eindrucksvolle Ergebnisse gezeigt ([Waibel *et al.* 89], [Hertz *et al.* 91]). Zur Fahrzeugführung präsentiert [Pomerleau 89] das Verfahren *ALVINN*, ein dreischichtiges künstliches neuronales Netz mit einem „Backpropagation“-Algorithmus. In der Trainingsphase werden dem Netz die Assoziationen zwischen den Farbverteilungen im Bild und den Steuerungsbefehlen gelehrt. Eine bemerkenswerte Trainingsmethode, ein sogenanntes „on-the-fly“-Verfahren, wird beim Training des Netzes verwendet. An die Eingänge des Netzes wird ein Mustervektor mit 960 Komponenten angelegt, wovon jeder dem kombinierten Farbwert eines Bildpunktes in einem Bild einer Straße mit einer Auflösung von 30×32 Pixeln entspricht. Die Ausgangsmuster entsprechen direkt einem der 45 vorgeesehenen Lenkwinkel (links abbiegen, ..., geradeaus fahren, ..., rechts abbiegen) für das Versuchsfahrzeug. Nachdem das Netz während einer manuell gesteuerten Fahrt des *Navlabs* mit ca. 3.5 km/h auf einer 200 Meter langen Teststrecke trainiert wurde, bewältigt es die autonome Führung mit dreimal höherer Geschwindigkeit als die anderen Systeme der *CMU*.

Einer der Nachteile künstlicher neuronaler Netze ist, daß ihre Leistung sehr stark von der Qualität der Trainingsmuster beeinflusst werden kann. Ein trainiertes Netz hat Schwierigkeiten bei der Klassifikation, wenn beim Einsatz signifikante Situationen in der Trainingsphase nicht vorkommen. Zur Überwindung des Problems verwendet [Pomerleau 92] strukturiertes Rauschen beim Training von *ALVINN*. Mit strukturierten Rauschen sollen die nicht-kooperativen Straßenszenen, wie z.B. fehlende Markierungen oder Erscheinen eines Fahrzeugs, modelliert werden. Zum quantitativen Verstehen der Leistung der verdeckten Knoten bezüglich der Änderung der Position und Orientation von Straßen wird in [Pomerleau & Touretzky 93] ein sogenanntes *HUSA*¹⁹-Verfahren verwendet. Nachdem das verbesserte *ALVINN* auf einer 3.2 km langen Strecke einer Autobahn trainiert worden war, kann das Netz mit einer Geschwindigkeit von 88 km/h das *Navlab* eine 34 km lange Strecke auf derselbe Autobahn autonom führen.

Zur Erhöhung der Robustheit bei der Führung von Fahrzeugen mit neuronalen Netzen

¹⁹*HUSA*: „Hidden Unit Sensitivity Analysis“

haben [Jochem *et al.* 93] ein Netz *MANIAC* mit zwei verdeckten Schichten entwickelt. Dabei besteht die erste verdeckte Schicht in *MANIAC* aus mehreren *ALVINN*-Netzen. Jedes davon ist getrennt für einen spezifischen Typ von Straßen, wie z.B. Straßen mit einer Spur oder zweien Spuren trainiert worden. Die Experimente zeigen, daß das Netz *MANIAC* auf den Straßen mit unterschiedlicher Spur bessere Ergebnisse als ein einzelnes Netz *ALVINN* liefert.

[Català *et al.* 92] benutzen ebenfalls ein dreischichtiges künstliches neuronales Netz mit einem „*Backpropagation*“-Trainingsalgorithmus für die Segmentierung von Straßenbildern. An die Eingänge des Netzes werden zwei Vektoren mit jeweils 784 Elementen angelegt. Der eine Vektor enthält den binarisierten Grauwertgradientenbetrag. Der andere Vektor enthält die absolute Grauwertdifferenz zwischen dem betrachteten Punkt und dem Mittelwert seiner kleinen Umgebung. Die verdeckte Schicht hat 224 Knoten. Diese Anzahl soll nach Angabe der Autoren optimal für diese Aufgabe sein. Die Ausgänge des Netzes sind 784 Elemente mit dem Wert eins oder null, wobei eins die Straße und null die Umgebung der Straße kodiert. Dies Verfahren wird mit den aufgezeichneten Bildern von Straßen der ersten Klasse erprobt.

[Fernandez-Maloigne *et al.* 92] verwenden die Statistiken aus den Grauwertlauflängenmatrizen als Merkmale für die Segmentierung von Straßenbildern. Erste Stufe ist die Grauwertnormierung durch eine Verschiebung des Grauwertistogramms und die Erhöhung des Grauwertkontrastes durch die Faltung mit einem Laplace-Operator. Die zweite Stufe besteht in der Berechnung der Grauwertlauflängenmatrizen und der Berechnung der fünf in [Galloway 75] eingeführten Textureigenschaften aus diesen Matrizen. Im letzten Schritt erfolgt dann die überwacht eingelernte Klassifikation der Pixel im Bild mit reduzierter Auflösung. Die Anwendung des Verfahrens zeigt, daß das Verfahren kleine isolierte Bereiche erzeugt und damit eine weitere Bereichverschmelzung bzw. -eliminierung erforderlich macht. Die Auseinandersetzung mit den Verfahren zur Texturanalyse im nächsten Kapitel wird zeigen, daß die Statistiken höher Ordnung keine signifikantere Informationen für die Textursegmentation als die Statistiken zweiter Ordnung enthalten. Wenn also überhaupt die globalen Statistiken der Grauwerte zur Texturanalyse herangezogen werden sollten, erscheinen die Statistiken aus Paarhäufigkeitsmatrizen am besten geeignet.

Die oben vorgestellten überwacht eingelernten Klassifikatoren verwenden Farbmodelle für die Klassen „*Fahrbahn*“ und „*Nicht Fahrbahn*“ bei der Klassifikation der Bildpunkte im Farbbild. Die Farbmodelle für die Klasse „*Fahrbahn*“ in [Turk *et al.* 88] und [Kuan *et al.* 88] werden in einem Ausschnitt des augenblicklichen Bildes ermittelt. Das setzt aber voraus, daß der ausgewählte Bildausschnitt ein Teil der Fahrbahn ist und der richtige Farbton der Fahrbahnoberfläche auftritt. Mit dem Wort „*richtig*“ ist gemeint, daß dieser Ausschnitt nicht gerade im Schatten oder in einem verschmutzten Bereich liegt. Diese beiden Verfahren verwenden jeweils nur eine Klasse für „*Fahrbahn*“ und „*Nicht Fahrbahn*“ und haben Schwierigkeiten bei Straßen mit Schatten und unterschiedlichen Farbtönen. Ein solches Verfahren kann nur bei denjenigen Straßen der zweiten Klasse erfolgreich verwendet werden, deren Oberfläche und Umgebung unterschiedliche Farbtöne aufweisen. Die Farbmodelle für die Klassen „*Fahrbahn*“ und „*Nicht Fahrbahn*“ in [Crisman & Thorpe 88] werden dagegen aus dem vorangegangenen Bild ermittelt, d.h. die Farbtöne der Straßen-

oberfläche und ihrer Umgebung sollen in aufeinanderfolgenden Bildern näherungsweise konstant bleiben. Dabei werden mehrere Farbmodelle für „*Fahrbahn*“ und „*Nicht Fahrbahn*“ verwendet, d.h. es werden Straßen wie auch Straßenumgebungen mit mehr als nur einem Farbton zugelassen. Um die Straßen der vierten Klasse zu detektieren, erscheint es notwendig, die Textur als Merkmale bei der Straßensegmentation zu verwenden.

Unüberwacht eingelernter Klassifikator

Crisman und Thorpe stellen in [Crisman & Thorpe 91] das Verfahren *UNSCARF*²⁰ zur Gruppierung der Bildpunkte im Farbbild mit einem modifizierten *Isodata*-Ballungsverfahren vor. Im Vergleich zu überwacht eingelernten Klassifikatoren benötigt *UNSCARF* keine Farbmodelle bei der Gruppierung der Bildpunkte und hat weniger Probleme bei einer temporären Farbtonänderung der Straße und ihrer Umgebung. Das Verfahren benötigt ca. 1 Minute für die Klassifikation von 30×32 Pixeln auf einer SUN-4 und ist geeignet zur Bestimmung der Straßenbegrenzungen in der Startphase. Nach Kenntnis des Autors ist das Verfahren *UNSCARF* das z.Z. einzige zur Fahrbahnerkennung verwendete unüberwacht eingelernte Klassifikationsverfahren.

3.4 Hindernisdetektion

Hindernisdetektion ist eine grundlegende Aufgabe eines Sichtsystems zur Fahrzeugführung. In seiner allgemeinen Bedeutung ist ein Hindernis ein Bereich, auf dem ein Fahrzeug nicht fahren kann oder soll. Die Unbefahrbarkeit eines Bereiches in einer Szene resultiert aus mehreren Gründen, wie z.B. der Materialeigenschaft des Bereiches, dem Verbot durch Verkehrszeichen oder der Form des Bereiches. Hier wird nur die Erkennung der aus dem zuletzt genannten Grund resultierenden Hindernisse behandelt. Solche Hindernisse können sowohl stationär als auch bewegt sein. Die Erkennung der formbedingten Hindernisse ist eigentlich ein Problem der Bestimmung der Tiefeninformation der Umgebung. Existierende Verfahren zur Hindernisdetektion gehören zu den folgenden vier Kategorien: Stereosehen durch Merkmalszuordnung, Tiefe aus Bewegung, Verwendung aktiver Sensoren und Detektion in einem einzelnen Bild.

Stereosehen durch Merkmalszuordnung und Tiefe aus Bewegung

Stereosehen ermittelt die 3D-Konfiguration mehrerer Punkte eines Objektes durch die Messung der Verschiebung von Abbildungen des gleichen 3D-Punktes in korrespondierenden Bildern. Korrespondierende Bilder sind die mit unterschiedlichen Kameras oder mit einer Kamera in unterschiedlichen Positionen aufgenommenen Bilder einer Szene. Solche Verfahren werden auch zur Hindernisdetektion bei der Fahrzeugführung eingesetzt. Als Vertreter solcher Verfahren soll hier *FIDO*²¹ genannt werden, das von [Moravec 83]

²⁰ *UNSCARF*: „*Unsupervised Classification Applied to Road Following*“.

²¹ *FIDO*: „*Find Instead of Destroy Objects*“.

ursprünglich entwickelt und von [Thorpe *et al.* 85] erweitert wurde. *FIDO* ist ein merkmalsbasiertes Verfahren und arbeitet mit den folgenden Schritten. Zunächst werden zwei Bilder einer Szene aufgenommen, ein linkes Bild und ein rechtes Bild. Beide Eingangsbilder werden dann durch einen Bildpyramidengenerator in der Größe reduziert. Eine Bildpyramidenkorrelation wird danach verwendet, um das erwünschte Merkmal „Grauwertecke“ eines Hindernisses im rechten Bild zu lokalisieren. Wenn ein solches Merkmal gefunden wird, sucht das Verfahren im linken Bild nach diesem Merkmal. Nachdem alle korrespondierenden Grauwertecken lokalisiert worden sind, berechnet das Verfahren die 3D-Ansicht des Hindernisses. Das Verfahren wurde auf der Warp Maschine ([Crisman & Webb 91]) implementiert und zur Führung eines kleinen Fahrzeuges eingesetzt. Das Verfahren braucht 3 Sekunden für jedes Aufnahmen-Paar.

[Ozaki *et al.* 89] verwenden auch ein Stereo-Verfahren zur Hindernisdetektion. Das Korrespondenzproblem wird gelöst durch die Zuordnung der horizontalen Kanten aus zwei Bildern, die mit zwei senkrecht übereinander angebrachten Kameras aufgenommen werden. Mit Unterstützung des Bildauswertesystems *IDATEN* sinkt die Zykluszeit des Verfahrens unter 100 ms.

Ein weiteres Beispiel für die Hindernisdetektion durch Merkmalszuordnung findet man in [Bruyelle & Postaire 92+93]. Das Verfahren benutzt zwei Zeilenkameras und führt eigentlich eine Auswertung von 1D-Daten durch. Der Vorteil der Nutzung von Zeilenkameras liegt in der geringeren Menge an zu verarbeitenden Daten.

Bei der Fahrzeugführung bewegt sich die abbildende Kamera relativ zur abgebildeten Umgebung. Aus dieser relativen Bewegung können die relativen Abstände zwischen den Objektpunkten und dem Projektionszentrum erschlossen werden. Diese Vorgehensweise wird Tiefe-aus-Bewegung genannt. Nur der translatorische Anteil der relativen Bewegung liefert einen Beitrag zur Tiefenschätzung. Der Rotationsanteil der Bewegung muß auch erkannt und vor der Auswertung der Tiefeninformation abgetrennt werden. Verfahren zur Tiefenschätzung aus Bewegung werden zur Hindernisdetektion bei der Fahrzeugführung eingesetzt.

[Enkelmann 90] hat ein Verfahren zur Detektion von Hindernissen auf der Straßenoberfläche durch einen Vergleich des aus einer Bildfolge berechneten optischen Flusses mit dem optischen Fluß untersucht, der sich aus dem Modell einer ebenen Fahrbahn vor dem Fahrzeug ergibt. Bei der Modellierung des optischen Flusses wird angenommen, daß die Bewegung der Kamera rein translatorisch erfolgt. Die Bewegungsparameter der Kamera werden dabei heuristisch bestimmt. Das Verfahren wird in [Enkelmann 91] darauf erweitert, daß die Bewegungsparameter der Kamera während der Fahrt zur Bildaufnahme aus dem Stellsystem des Versuchsfahrzeugs ermittelt werden. Anwendungen des Verfahrens zeigen die erfolgreiche Detektion sowohl stationärer als auch bewegter Hindernisse.

Ein Verfahren zur Hindernisdetektion durch Tiefen-Schätzung aus Verschiebungsvektoren wird in [Arkin *et al.* 87] vorgestellt. Verschiebungsvektoren werden nur an den Bildpunkten mit großer Grauwertkrümmung und Kontrast im Fahrbahnbereich ermittelt. Probleme des Verfahrens bei Anwendungen im Straßenverkehr zeigen sich bei der Suche nach korrespondierenden Punkten in aufeinander folgenden Bildern und bei der präzisen

Lokalisierung des Punktes *FOE*²².

Verwendung aktiver Sensoren

Mit Hilfe von aktiven Sensoren, wie Laser-Abtastsystemen, Ultraschall- oder Radar-Sensoren, lassen sich die Abstandsdaten zwischen abzubildendem Sensor und abgebildeten Objekten direkt ermitteln. Im Vergleich zu den indirekten Methoden des Maschinensehens sind die direkten Methoden langsamer bei der Aufnahme von Bildern, aber schneller und präziser bei der Gewinnung von 3D-Informationen. Zur Hindernisdetektion bei der Fahrzeugführung in natürlicher Umgebung, welche mit passiven optischen Sensoren wegen Schatten und schwacher Beleuchtung manchmal problematisch wird, sind aktive Sensoren von Vorteil. Die meisten bei der Fahrzeugführung eingesetzten Verfahren zur Hindernisdetektion verwenden Laser-Abtastsysteme.

Verfahren zur Hindernisdetektion in [Hebert & Kanade 86] und [Hebert 89] segmentieren Abstandsdaten, basierend auf lokalen geometrischen Eigenschaften der Objekt-oberfläche, der Krümmung und der Normalen der Oberfläche. [Veatch & Davis 90] berechnen die erste partielle Ableitung der Abstandsdaten direkt im sphärischen Koordinatensystem des Abstandssensors und anschließend detektieren sie Hindernisse in Abstandsbildern. Eine weitere direkte Methode ist die Segmentation der Abstandsbilder unter Verwendung der Elevation jedes Punktes (Vertikale Position jedes Punktes) als Merkmal ([Dunlay & Morgenthaler 86]). Die Position der Hindernisse wird vom sphärischen Koordinatensystem des Abstandssensors in ein mit dem Fahrzeug verbundenes kartesisches Koordinatensystem konvertiert.

In [Daily *et al.* 88] und [Olin & Tseng 91] wird dagegen zunächst ein kartesisches Elevationsbild aus einem Abstandsbild erzeugt. Ein 3D-Fahrzeugmodell mit einem virtuellen Sensor wird verwendet, um die Hindernisse im Elevationsbild zu lokalisieren. Statt einer symbolischen Beschreibung wird hier ein Hindernis definiert durch die möglichen Zustände des Fahrzeugmodells, wenn sich das Modell im Elevationsbild befände.

Um die Abtastzeit eines aktiven Sensors zu reduzieren, wird ein Laser-Abtastsystem in [Singh & Keller 91] horizontal zur Erdoberfläche angebracht. Unter der Annahme, daß sich das Fahrzeug in einer 2D-Welt bewegt, wertet das Verfahren nur ein Abstandsprofil aus. Das vereinfachte Verfahren arbeitet schnell, aber hat Schwierigkeiten, wenn das Fahrzeug in senkrechter Richtung eine Bewegung ausführt. Eine Alternative besteht darin, zunächst den Fahrbahnbereich zu detektieren und die Hindernisdetektion nur darauf zu beschränken ([Dunlay 88]).

Zur Überwindung der Geschwindigkeitsprobleme bei der Bildaufnahme mit aktiven Sensoren werden Hindernisse in [Turk & Marra 86] durch die Segmentation eines Farbbildes detektiert. Die Grundannahmen über Hindernisse und Fahrbahn beruhen darauf, daß die Hindernisse von der Fahrbahn im Farbbild segmentiert werden können und die Fahrbahnoberfläche so wenig Texturen wie möglich aufweist. Bei der Gewinnung der 3D-Beschreibung eines als Hindernis detektierten Fahrbahnbereiches wird ein Laser-Abtastsystem verwendet. Mit ähnlicher Weise arbeitet das Verfahren zur Hindernisdetek-

²² *FOE*: „focus of expansion“.

tion in [Xie *et al.* 93]. Statt einer Farbkamera benutzen [Xie *et al.* 93] eine monochrome CCD-Kamera. Unter der Annahme, daß die Hindernisse auf der Straßenoberfläche dominante geradlinige Strukturen haben, erfolgt die Lokalisierung von Hinderniskandidaten zuerst im Grauwertbild. Die mit dem Laser-Abtastsystem gewonnenen Daten dienen hier zur Verifikation von Hindernissen und geben dabei die 3D-Information der detektierten Hindernisse.

Detektion in einem einzelnen Grauwertbild

Unter bestimmten Voraussetzungen können Hindernisse auch in 2D-Bildern ohne Verwendung der 3D-Information detektiert werden. In [Dickmanns & Christians 89] und [Dickmanns & Mysliwetz 92] wird ein solches Verfahren zur Hindernisdetektion unter Verwendung eines 4D-Modells vorgestellt. Das Verfahren besteht aus den folgenden drei Schritten: Bestimmung der Kandidaten für Hindernisse durch die Detektion der Inhomogenität der Grauwerte, Lokalisierung der Hindernisse durch die Detektion der horizontalen und vertikalen Kanten und Schätzung der relativen 3D-Zustände mit Hilfe des 4D-Modells. Das Verfahren wird innerhalb eines rechteckigen Bildausschnittes im Fahrbahnbereich jedes Bildes durchgeführt. Zu den Zuständen gehören Abmessung und Position der Hindernisse sowie ihre räumliche Beziehung zum Fahrzeug.

Mit den ermittelten Fahrbahnbegrenzungen versuchen [Solder & Graefe 90] die Hindernisse auf der Fahrspur durch ein heuristisches Verfahren zu detektieren. Diese Idee geht auf [Graefe *et al.* 88] und [Dickmanns & Christians 89] zurück. In [Efenberger *et al.* 92] wird das Verfahren auf die Detektion der sich auf den Nebenfahrbahnen dem Versuchsfahrzeug von hinten nähernden Fahrzeuge erweitert. Dabei wird das Bild einer Straßenszene mit einer nach hinten ausgerichteten, im Fahrzeug montierten Kamera aufgenommen. Statt der direkten Ermittlung der Begrenzungen von Nebenfahrbahnen werden diese durch eine Extrapolation der Begrenzungen der Eigenfahrbahn bestimmt. Ein Fahrzeug wird in diesem Verfahren einfach als ein rechteckiger Bereich modelliert. Die Detektion von Objekten erfolgt bei diesem Verfahren nach Aussage der Autoren mit Hilfe eines heuristischen Ansatzes. Vorteil des Verfahrens ist seine hohe Verarbeitungsgeschwindigkeit (40 ms).

[Schwarzinger *et al.* 92a] beschreiben ein Verfahren zur Detektion, Verfolgung und Identifizierung eines vorausfahrenden Fahrzeugs in einem Grauwertbild. Bei der Detektion und Verfolgung eines Fahrzeugs wird die symmetrische Eigenschaft der Hinterseite des Fahrzeugs ausgenutzt. Es wird dabei angenommen, daß die Hinterseite eines Fahrzeugs im Bild um eine vertikale Linie symmetrisch ist. Parallel zur Verfolgung eines Fahrzeugs führt das Verfahren noch zusätzlich die 2D-modellbasierte Klassifikation des Fahrzeugs durch ([Schwarzinger *et al.* 92b]). Die verwendeten Merkmale sind Linienelemente und Grauwertecken im Bild. Die Zuordnung der aus dem Bild extrahierten Merkmale zu den Modellmerkmalen wird durch ein sogenanntes *elastisches* künstliches neuronales Netz ([Durbin & Willshaw 87]) iterativ realisiert. Der erste Teil des Verfahrens, die Detektion und Verfolgung eines Fahrzeugs, wurde im Versuchsträger der Volkswagen AG (VW) implementiert, während der zweite Teil des Verfahrens wegen des Zeitbedarfs für die

iterative Arbeitsweise bisher nur mit den aufgezeichneten Bildern erprobt worden ist.

Weitere Beispiele für die Detektion von Hindernissen durch heuristisches Suchen im Bildbereich der Straßenoberfläche findet man bei [Kenue 89b] und [Yu *et al.* 92].

3.5 Landmarkenerkennung

Bei der kartenbasierten autonomen Fahrzeugführung kann eine lange Fahrt zu signifikanten Positionsfehlern des Fahrzeuges bei einem odometriebasierten Navigationssystem führen. Um diese Abweichungen zu korrigieren, wird die Landmarkenerkennung eingesetzt. Sie setzt die Erkennung von Objekten in der aktuellen Szene und die Zuordnung eines dieser Objekte zu einer bestimmten Landmarke in einer geographischen Karte voraus. Bei Erkennungsaufgaben sind ausgedehntes Kartenwissen und Modelle von Landmarken fundamental. Verfahren zur Landmarkenerkennung bestehen normalerweise aus den folgenden drei Schritten: Hypothesengenerierung, Objektdetektion und Landmarkenerkennung. Hypothesen enthalten die Erwartungen über die geographische Position und geometrische Beschreibung von Landmarken. Mit solchen Erwartungen lassen sich die Suchräume bei der Objektdetektion im Bild drastisch reduzieren. Die Landmarkenerkennung benötigt detaillierte merkmalsbasierte Objektmodelle, die Informationen über lokale Merkmale wie z.B. Farbton, Textur, Grauwert, Abmessung usw. enthalten.

Als etablierter Ansatz, Objekte durch Modellzuordnung zu erkennen, gilt das Sichtsystem *VISIONS* ([Hanson & Riseman 78]). Das System *VISIONS* selbst wird hauptsächlich zur Interpretation der Bilder von Gebäuden eingesetzt. Die Philosophie in der Systemstruktur von *VISIONS* wird zur Landmarkenerkennung bei der Fahrzeugführung benutzt. Ein Beispiel dafür bildet das in [Lawton *et al.* 86] und [Levitt & Lawton 89+90] vorgestellte Sichtsystem zur Terrainerkennung durch Zuordnung der auf der Basis von Kartenwissen vorausgesagten Landmarken zu den aus Bildern extrahierten Objekthypothesen. Diese Vorgehensweise wird auch als Hypothesenüberwachung bezeichnet. Das System besteht aus generischen Objektmodellen, Algorithmen zur Bildstrukturextraktion, einem *STM*²³, einem *LTM*²⁴ und verschiedenen Schlußfolgerungsprozessen. Das *STM* ist ein dynamischer temporärer Speicher für Hypothesen über Objekte, die extrahierte Bildstruktur und die gewonnene 3D-Information. Im *LTM* befinden sich *A-priori*-Terrainbeschreibungen und -hypothesen sowie Wissen über spezifische Landmarken. Sobald genügend Hinweise auf eine Hypothese im *STM* akkumuliert worden sind, wird diese in das *LTM* übertragen. *STM* und *LTM* sind hierarchisch aufgebaut. Die Kontrollstruktur des Systems generiert sowohl datengesteuerte als auch modellgesteuerte Hypothesen. An der *University of Massachusetts* wird ein System *AuRA*²⁵ ([Arkin 87]), eine Erweiterung des Systems *VISIONS*, zur Roboternavigation aufgebaut.

In [Dickinson & Davis 90] und [Davis *et al.* 92] wird ein modellbasiertes Sichtsystem zur Straßenverfolgung durch eine Verifikation von Hypothesen über Objekte, die aus ei-

²³*STM*: „short term memory“.

²⁴*LTM*: „long term memory“.

²⁵*AuRA*: „Autonomous Robot Architecture“.

ner Karte generiert werden, mit den über Sensoren ermittelten Daten vorgestellt. Das Szenenmodell wird durch ein semantisches Netzwerk repräsentiert. Die Kontroll-Struktur ist ein durch Produktionssysteme unterstütztes Tafelsystem (*Blackboard*). Das System wurde nur mit aufgezeichneten Bildern erprobt und nicht im Fahrzeug implementiert. [Nasr & Bhanu 88] verwenden auch ein modellgestütztes Verfahren bei der Erkennung von Landmarken für autonome mobile Roboter. Ein semantisches Netzwerk wird bei der Repräsentation der Kartenprimitive verwendet. Die aus dem Bild extrahierten Informationen werden in einem Tafelsystem abgespeichert. Im Vergleich zu den obenerwähnten Arbeiten wird hier explizites Wissen über die zu erkennende Landmarke bei der Generierung eines Szenenmodells benutzt. Es wird weiter gezeigt, daß das Verfahren auch eine Rückkopplung zu den Algorithmen der Bildverarbeitung bietet, um Parameter bei der Merkmalsextraktion und -gruppierung zu adaptieren.

[Kristen & Munkelt 92] verwenden die Theorie der Markoff-Felder für eine kontext-sensitive Interpretation von Straßenbildern. Mit Hilfe der lokalen Charakteristiken eines allgemeinen Markoff-Feldes (MF) werden die räumlichen Beziehungen disjunkter Regionen eines Straßenbildes durch eine symmetrische Relation relativer topologischer Tiefen modelliert. Bildregionen und relationale Beziehungen bilden einen Graphen, der das Nachbarschaftssystem des MF festlegt, und dessen Knotenbewertungen als Zustände des MF den konkreten Interpretationen der Regionen entsprechen. Die Markoff-Feld-Datenbasen werden dabei mit einem „Lernen aus Beispielen“-Verfahren aus Beispielbildern erzeugt. Die Interpretation von Straßenbildern zeigt vielversprechende Ergebnisse.

Als eine der wichtigsten Aufgaben bei der Landmarkenerkennung wurde in letzter Zeit die Erkennung von Verkehrszeichen untersucht. Verkehrszeichen lassen sich grundsätzlich in Gefahr-, Vorschrift- und Richtzeichen untergliedern. Entscheidende Informationen erhalten wir durch ihre Farbe, Form, Symbole und den Text. In einer Straßenszene finden sich Verkehrszeichen sowohl neben oder über der Straße als auch auf der Straßenoberfläche. Die Anzahl der Verkehrszeichen ist sehr groß. Nach der deutschen Straßenverkehrsordnung gibt es beispielweise in Deutschland mehr als 400 verschiedene Verkehrszeichen. Die Untersuchungen werden bisher nur mit einem kleinen Anteil davon durchgeführt.

[Mori 92] verwendet Verkehrszeichen auf der Fahrbahn als Landmarken für die Navigation des mobilen Fahrzeugs „*Harunobu-4*“. Sie wird durch die Verifikation der gespeicherten Verkehrszeichen mit den aus einem Farbbild extrahierten Objekten realisiert. Zur Verkürzung der Verarbeitungszeit und Vergrößerung der Reichweite der Kamera wird in [Mori 90] ein sogenanntes aktives Meßverfahren („*active sensing*“) verwendet. Aufgrund der im vorangehenden Bild gewonnenen Ergebnisse und des Wissens aus dem Straßenmodell wird das Verfahren realisiert durch die dynamische Ausrichtung der Kamera nach den wichtigen Teilen der Umgebung und die dynamische Positionierung des auszuwertenden Bildausschnittes.

[Seitz *et al.* 91] präsentieren ein Verfahren zur Erkennung von Verkehrszeichen in einer Pyramidenstruktur von Grauwertbildern. Als Merkmale werden Kantenelemente im Bild detektiert. Die Bestimmung von Verkehrszeichen erfolgt durch die Analyse der geometrischen Beziehungen zwischen den detektierten Merkmalen. In [Seitz *et al.* 91] ist allerdings nur die Lokalisierung von Verkehrszeichen realisiert worden. Ein weiterer Schritt, die In-

terpretation der Symbole und Texte innerhalb der detektierten Verkehrszeichen, fehlt dabei noch.

Ausgehend von der Tatsache, daß die Farbe von Verkehrszeichen sich deutlich von derjenigen ihrer Umgebung unterscheidet, klassifiziert [Ritter 92] zuerst mit einem überwacht lernenden künstlichen neuronalen Netz die Pixel des Farbbildes einer Straßenszene in vier Klassen, nämlich Rot, Gelb, Blau und Hintergrund. Eine Bereichsanalyse wird anschließend durchgeführt. Aus den Ergebnissen der Bereichsanalyse werden mit Hilfe von A-Priori-Wissen über Verkehrszeichen die Kandidatenbereiche für Verkehrszeichen selektiert. Nach der Form eines Kandidatenbereiches sowie seiner Lage bezüglich der Straße wird eine Hypothese darüber generiert, welches Verkehrszeichen in diesem Bereich sein könnte. Das System zur Erkennung von Verkehrszeichen wird zur Zeit nur bis zur Generierung von Hypothesen entwickelt. Die Verifikation von Hypothesen soll noch untersucht werden. Die Hypothesengenerierung über Verkehrszeichen wurde nur mit 15 Verkehrszeichen erprobt. Die Erweiterung auf andere Verkehrszeichen soll ebenfalls erfolgen.

Als eine Alternative zu [Ritter 92] wird in [Hartmann & Mertsching 92] ein zweistufiges hierarchisches Verfahren zur Erkennung von Verkehrszeichen vorgestellt. Das Verfahren führt zuerst die Faltung eines Grauwertbildes mit einem Laplace-Operator durch. Daraus lassen sich hierarchische Struktur-Kodes aufbauen. Basierend auf solcher symbolischen Beschreibungen von Bildstrukturen wird weitergehend eine Formanalyse durchgeführt. Die detektierten Formelemente werden anschließend einem Graphsuchen-Algorithmus übergeben, der die Klassifizierung von Verkehrszeichen als Ergebnisse ausgibt. Wie in [Büker *et al.* 93] berichtet wird, wird das Verfahren bereits an einem Satz von über 300 Verkehrsszenen erfolgreich getestet.

Straßeneinmündungen und -kreuzungen sind die einfachsten Landmarken. Deren Erkennung ist eine grundsätzliche Voraussetzung für die autonome Führung von Fahrzeugen auf den Straßen mit Einmündungen und Kreuzungen. Wir stellen im folgenden zwei Verfahren für die Lösung dieser Aufgaben dar.

Das Verfahren zur Kreuzungsdetektion in *YARF* ([Kluge & Thorpe 93]) wertet die ausgewählte Bildausschnitte aus. Falls keine Markierungsbänder in fünf aufeinanderfolgenden Bildausschnitten gefunden werden und der Abstand dazwischen zumindest 2 m beträgt, wird angenommen, daß eine Kreuzung vorliegt. Mit Hilfe der geometrischen Kreuzungsmodelle wird eine Hypothese über die Position der Kreuzungsmerkmale generiert. Die Verifikation der Hypothese erfolgt durch die Detektion der Merkmale in den entsprechenden Ausschnitten. Zur Gewinnung der erforderlichen Informationen für die Kreuzungserkennung und evtl. eine anschließende Abbiegung wird in [Kluge & Thorpe 93] eine zusätzliche Kamera eingesetzt.

[Struck *et al.* 93] verwenden zwei zusätzliche Kameras mit unterschiedlichen Blickrichtungen (jeweils nach rechts und links) für die Detektion und Verfolgung von Kreuzungen. Dabei werden die als geradlinig angenommene Straße sowie in diese Straße einmündende Fahrspuren detektiert. Daraufhin werden von einem parallel arbeitenden zweiten Verfahren Bildausschnitte auf relevante Merkmale von Straße und Einmündung gesetzt und diese modellgestützt verfolgt. Dabei wird aus einer Anordnung von drei Kameras, die einen Beobachtungswinkel von ca. $\pm 70^\circ$ vor dem Fahrzeug abdecken, das jeweils momen-

tan geeignete Bild ausgewählt. Das Navigationssystem *Travelpilot* gibt dabei Hinweise über die relative Anordnung der Fahrbahnmittelachsen, welche sich in einer Kreuzung schneiden, und dient zur Plausibilitätskontrolle der Bildauswerteergebnisse.

3.6 Auflistung der Forschungsaktivitäten zur autonomen Fahrzeugführung

Die oben aufgeführten Sichtsysteme und Bildauswerteverfahren haben zur Entwicklung autonomer Systeme beigetragen, funktionieren aber nur auf Straßen mit bestimmten Randbedingungen (markierte Fahrbahn oder unterschiedliche Farbtöne von Fahrbahn und ihrer Umgebung) zuverlässig. Bevor ein wirklich autonomes System am öffentlichen Straßenverkehr teilnehmen kann, müssen noch viele grundlegende Probleme gelöst werden. Die Sichtsysteme zur autonomen Fahrzeugführung wie auch für andere Einsatzgebiete des Maschinensehens sollten durch weitere Experimente verbessert werden ([Jain & Binford 91], [Petkovic & Wilder 91], [Negahdaripour & Jain 91+92], [Pavlidis 92]).

Um einen illustrativen Eindruck zu den in diesem Kapitel besprochenen Systemen und Verfahren zu gewinnen, werden in Tabelle 3.1 diejenige Forschungsprojekte zur Fahrzeugführung, die einen Versuchsträger besitzen, nach ihrem Aktivitätsbereich aufgelistet. Im folgenden sollen anhand der Tabelle 3.1 die erzielten Ergebnisse und die noch zu lösenden Probleme zur Erkennung von Straßen verschiedener Klassen in Anlehnung an die Untergliederung von Straßen in Kapitel 2 angegeben werden.

Befriedigend gelöst: Wie in Tabelle 3.1 zu sehen ist, ist die Erkennung der Straßen der ersten Klasse mit konturbasierten Verfahren gut gelöst. Die Straßen der ersten Klasse haben deutliche Markierungen und homogene Oberfläche. Solche Straßen haben auch eine bestimmte Form. Für die Erkennung solcher Straßen wurden explizite geometrische Modelle benutzt, deren Informationen als quantitative Daten mit gewissen Toleranzen beschrieben wurden. Die am häufigsten verwendeten Merkmale sind die Begrenzungsmarkierungen der Fahrbahn. Sie lassen sich durch konturbasierte Verfahren aus einem Bild extrahieren. Durch die Verwendung expliziter Modelle wurde die Fahrbahnerkennung erleichtert, da nur die als wichtig angesehenen Ausschnitte eines Bildes ausgewertet werden müssen.

Ebenfalls befriedigend ist die Erkennung von denjenigen Straßen der zweiten Klasse gelöst, deren Oberfläche unterschiedliche Farbtöne im Vergleich zu denjenigen ihrer Umgebungen aufweist. Implizite Modelle sind geeignet für solche Straßen. Implizite Modelle haben typische Formulierungen wie: Fahrbahn ist lokal gerade oder eben. Beispielfhaft sind hier das Ebenen-Modell im Szenenbereich und das Fluchtpunkt-Modell im Bildbereich zu nennen. Solche Modelle enthalten keine konkreten Daten über die Fahrbahn. Die Straßen der zweiten Klasse unterscheiden sich im Farbton von ihren Umgebungen. Sie wurden als Merkmale in bereichsbasierten Segmentationsverfahren bei der Fahrbahnerkennung verwendet.

Es ist zur Zeit Stand der Technik, ein autonomes Fahrzeug auf Straßen der ersten

Sensoren und Forschungsaktivitäten	Forschungsgruppe									
	VITS	FMC	Navlab	PVS	MV	VaMoRs	UB	VITA	ICAD	IITB
<u>Sensoren</u>										
Farbkamera	•	•	•	•	○	○	○	○	•	○
Monochrome Kamera	○	○	○	•	•	•	•	•	○	•
Laser-Abtastsystem	•	•	•	•	○	○	○	○	•	•
<u>Funktionen</u>										
Querführung	•	•	•	•	•	•	•	•	•	•
Längsführung	•	•	•	•	•	•	○	•	•	•
<u>Fahrbahnerkennung</u>										
- durch Teilbildauswertung	○	○	•	•	•	•	•	•	•	•
- durch Gruppierung	○	○	○	○	○	○	○	○	○	•
- im Abstandsbild	○	○	○	○	○	○	○	○	○	○
- durch Farbsegmentation	•	•	•	•	○	○	○	○	•	○
- durch Texturanalyse	○	○	•	○	○	○	○	○	○	○
<u>Geometrische Straßenmodelle</u>										
- Ebenen-Modell	•	•	○	•	•	○	○	○	○	•
- Fluchtpunkt-Modell	○	○	•	○	○	○	○	○	○	○
- Klothoiden-Modell	○	○	○	○	○	•	○	•	○	•
- Kreisbogen-Modell	○	○	○	○	○	○	•	○	○	○
- Kurven-Modell	○	○	•	○	○	○	○	○	○	○
- Band-Modell	○	○	○	○	○	○	○	○	○	○
<u>Komplexität von Straßen</u>										
- Straßen erster Klasse	▷	▷	★	★	★	★	★	★	★	★
- Straßen zweiter Klasse	★	★	★	▷	▷	▷	▷	▷	▷	▷
- Straßen dritter Klasse	▷	▷	▷	▷	▷	▷	▷	▷	▷	▷
- Straßen vierter Klasse	○	○	▷	○	○	○	○	○	○	○
<u>Hindernisdetektion</u>										
- mit Laser-Abtastsystem	•	•	•	•	○	○	○	○	•	•
- in einem Grauwertbild	○	○	○	○	○	•	○	•	○	*
- durch Tiefe-aus-Bewegung	○	○	○	○	○	○	○	○	○	*
- durch Stereo-Verfahren	○	○	•	•	○	○	○	○	○	○
<u>Landmarkenerkennung</u>										
- Kreuzung und Einmündung	○	•	•	○	•	○	○	•	○	•
- Verkehrszeichen	○	○	○	○	○	○	○	*	○	○

Tabelle 3.1: Eine Auflistung der Forschungsaktivitäten einiger wichtiger Projekte zur Fahrzeugführung (VITS : Vision Task Sequencer; FMC : FMC Corporation; Navlab : Navigation Laboratory; PVS : Personal Vehicle System; MV (MOVER) : Mobile Vision Experimental Robot; VaMoRs : Versuchsfahrzeug für autonome Mobilität und Rechnersehen; UB : University of Bristol; VITA : Vision Technolog Application; ICAD : Intelligent Computer Aided Driving; IITB : Fraunhofer-Institut für Informations- und Datenverarbeitung) : • = im Fahrzeug eingesetzt; * = untersucht mit aufgezeichneten Bildern; ★ = robust funktionierend; ▷ = teilweise funktionierend; ○ = keine Aktivität.

Klasse mit einer Geschwindigkeit bis ca. 100 km/h zu führen. Dieser Erfolg beim Einsatz auf Straßen der ersten Klasse beruht weniger auf der Leistungsfähigkeit von Sichtsystemen, sondern stärker auf der Einbindung von Ansätzen der Systemdynamik.

Teilweise gelöst: Die bisherigen Verfahren zeigen, daß die Erkennung von Straßen der dritten Klasse nur teilweise unter bestimmter Bedingungen gelöst sind. Bei der Erkennung solcher Straßen wurden meistens explizite Modelle eingesetzt. Prinzipiell können alle Verfahren zur Extraktion von Straßen der ersten Klasse für die Straßenextraktion verwendet werden. Sie sind allerdings für solche Einsatzfälle bis heute noch nicht hinreichend zuverlässig.

Nicht gelöst: Bisherige Verfahren haben Schwierigkeit bei der Erkennung der Straßen der vierten Klasse. Bei der oben vorgestellten bereichsbasierten Klassifikation werden die unterschiedlichen Farbtöne auf der Fahrbahn und ihrer Umgebung als Merkmale verwendet. Die Unterschiede in Farbtönen in einer kleinen Region sind für die Segmentation der Straßen der vierten Klasse nicht ausreichend. Die Unterschiede der örtlichen Einordnung der Grauwerte in einer kleinen Region, nämlich die Unterschiede der Texturen, sollten daher ebenfalls berücksichtigt werden.

Jedes Verfahren zur Fahrzeugführung benötigt Modelle von Fahrbahnen. Solche Modelle enthalten die Erwartung über die Form und den räumlichen Verlauf von Fahrbahnen. Sie unterstützen die Fahrbahnerkennung durch Vorhersagen, wie und wo eine Fahrbahn zu finden ist und welcher Algorithmus zur Detektion und Verfolgung verwendet werden soll. Die meisten oben vorgestellten Verfahren verwenden nur *ein* Modell von Fahrbahnen. Damit wird die Teststrecke nur auf Straßen beschränkt, die diesem Modell entsprechen. Zu lösende Probleme bestehen in der weiteren Erstellung von Modellen und Erkennungsalgorithmen für verschiedene Fahrbahnen einschließlich von Fahrbahnkreuzungen und -einmündungen.

Die Erkennung von Straßen der vierten Klasse versuchen wir in dieser Arbeit durch die Textursegmentation zu lösen.

Kapitel 4

Texturanalyse: Theorien und Verfahren

Der Begriff „Textur“ hat verschiedene Bedeutungen in unterschiedlichen Bereichen der Naturwissenschaften. Beim Maschinensehen versteht man darunter visuelle Textur. Visuelle Textur ist eine Oberflächeneigenschaft. Sie kann durch eine physikalische Oberflächenänderung, eine Markierung auf einer Oberfläche, eine Sammlung von Objekten usw. entstehen. [Julesz & Bergen 83] definieren visuelle Textur als eine Gesamtmenge kleiner Elemente, wobei die Elemente entweder Punkte mit unterschiedlicher Farbe oder einzelne Flecke sind. Nach der Definition von [Haralick & Shapiro 91] ist visuelle Textur durch die räumliche Verteilung von Bildgrauwerten oder diskreten Merkmalen zusammen mit deren Beziehungen untereinander charakterisiert. [Tomita & Tsuji 90] formulieren visuelle Textur als die auf Objektflächen homogen verteilten Muster. Wegen der Vielfalt von Texturen gibt es bis heute noch keine exakte Definition und allgemeingültige Beschreibung für den Begriff „visuelle Textur“. Statt des Versuches einer strengen Definition für visuelle Texturen, was nicht der Schwerpunkt dieser Arbeit ist, übernehmen wir den umgangssprachlichen Gebrauch und betrachten visuelle Texturen als die visuell wahrnehmbaren Eigenschaften der Oberflächen von Objekten. Statt visueller Textur wird im weiteren Verlauf dieser Arbeit kürzer als von Textur gesprochen.

Seit über dreißig Jahren wird Textur im Maschinensehen untersucht. Automatische Texturanalyse stellt ein wichtiges Forschungsgebiet der Bildauswertung dar. Je nach den Texturarten und den Aufgabenstellungen werden verschiedene Verfahren zur Texturanalyse herangezogen, die wir in die folgenden drei Klassen unterteilen:

- Beschreibung und Erkennung von Texturen,
- Segmentation von Bildern durch die Auswertung von Texturen und
- Gewinnung von Informationen über Abstand und Orientierung eines Oberflächenelementes durch die Auswertung von Texturen.

In dieser Arbeit werden die Verfahren zur Lösung der ersten zwei Aufgaben behandelt. Die dritte Aufgabe, die sogenannte Struktur-aus-Textur-Ermittlung, wird hier nicht berück-

sichtigt. Eine Übersicht über Struktur-aus-Textur-Verfahren geben die Arbeiten von [Witkin 81], [Aloimonos & Swain 85] und [Blake & Marinos 90].

Die Verteilung von Texturen kann deterministisch, stochastisch oder kombiniert deterministisch-stochastisch sein. Verfahren zur Texturanalyse können daher auch in Anlehnung an die Texturarten klassifiziert werden. Die in der Fachliteratur beschriebenen Texturanalyseverfahren lassen sich einteilen in

- statistische und
- strukturelle Methoden.

Beim ersten Ansatz wird Textur durch statistische Werte beschrieben, die aus örtlichen Eigenschaften der Grauwerte oder lokalen Merkmale im Bild extrahiert werden, überwiegend Statistiken erster und zweiter Ordnung in den Grauwerten sowie Verteilungen lokaler Texturelemente. Statistische Texturanalyseverfahren messen die lokalen Statistiken von Texturen und sind geeignet für feine Mikrot Texturen. Beim strukturellen Ansatz wird Textur durch Texel¹ und deren geometrische Beziehungen untereinander beschrieben. Strukturelle Verfahren zur Texturanalyse ermitteln eine detaillierte strukturelle Beschreibung von Texturelementen und Aufbauregeln. Sie sind geeignet für Makrot Texturen.

Zur Veranschaulichung werden 4 Beispiele von Texturen aus [Brodatz 66] in Abbildung 4.1 dargestellt. Die Texturen in Abb. 4.1 (a) und (b) sind geeignet für eine Auswertung durch statistische Ansätze. Dagegen sind die Texturen in Abb. 4.1 (c) und (d) eher für eine Auswertung durch strukturelle Verfahren geeignet. In diesem Kapitel werden die in der Fachliteratur beschriebenen Verfahren nach Texturarten gegliedert und dargestellt.

Eine Übersicht über die Arbeiten zum Thema Texturanalyse geben die Artikel von [Haralick 79], [Wechsler 80], [Gool *et al.* 85] und [Reed & Buf 93] sowie die Bücher von [Rao 90] und [Tomita & Tsuji 90]. Da die Verfahren zur Texturanalyse so vielseitig sind, kann in diesem Kapitel keine vollständige Darstellung dieses Themenkreises gegeben werden. Stattdessen besteht das Ziel dieses Kapitels besonders darin, die Grundlagen der Texturanalyse zu erarbeiten, die klassischen Verfahren zu analysieren und die aktuellen Forschungsergebnisse zusammenzufassen.

4.1 Texturwahrnehmung beim Menschen

Auf dem Forschungsgebiet der Psychophysik ist die Texturwahrnehmung beim Menschen seit vielen Jahren weltweit ein intensiv untersuchtes Thema. Der Schwerpunkt der Untersuchungen befaßt sich mit der Frage, welche Eigenschaften von Texturen - globale Grauwertstatistiken bis zu n -ter Ordnung oder Statistiken erster Ordnung von lokalen Texturelementen - für die Texturunterscheidung beim Menschen entscheidend sind. Mit der Ordnung der Statistik ist hier die herangezogene Anzahl der Variablen oder Freiheitsgrade bei der Beschreibung einer Struktur gemeint. Zahlreiche Untersuchungen mit unterschiedlichen Stimuli wurden zu diesem Thema durchgeführt. In diesem Abschnitt

¹Texel: Abkürzung von Texturelement.

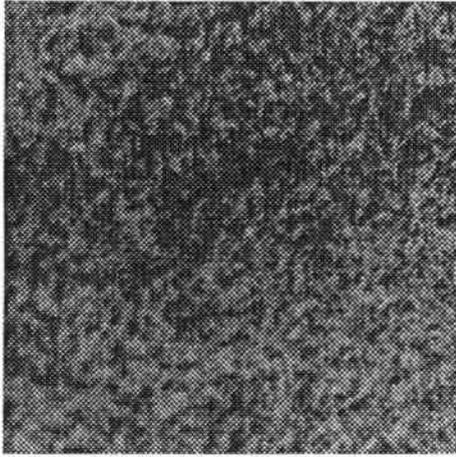
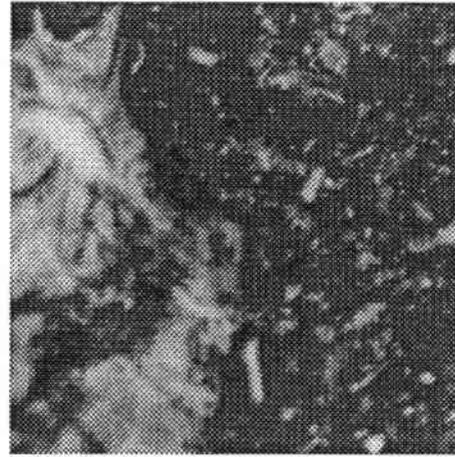
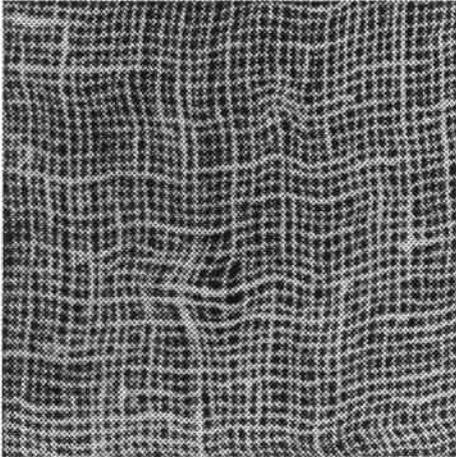
(a) *Gras.*(b) *Marmor.*(c) *Leinengewebe.*(d) *Baumstumpf.*

Abbildung 4.1: Beispiele von Texturen aus [Brodatz 66]; Gras in (a) (D9 in [Brodatz 66]), Marmor in (b) (D58), Leinengewebe in (c) (D104) und Baumstumpf in (d) (D97).

sollen wichtige Arbeiten zusammen mit den gewonnenen Ergebnissen vorgestellt werden. Weil die Fähigkeiten des menschlichen Gehirns und des visuellen Systems sicher noch nicht komplett analysiert wurden, können wir die erzielten Schlußfolgerungen nur als Hypothesen bezeichnen.

4.1.1 Globale Grauwertstatistiken n -ter Ordnung

[Julesz 75] teilt die Wahrnehmung beim Menschen in zwei getrennte Prozesse: reine Wahrnehmung und Wahrnehmung mit Kognition. Die Aufgaben reiner Wahrnehmung können mühelos, parallel, ohne Unterstützung durch kognitive Verarbeitungsstufen des Gehirns ausgeführt werden. Die kognitiven Aufgaben des visuellen Systems müssen mit besonderer

Aufmerksamkeit und genauem Hinsehen ausgeführt werden. Die Texturunterscheidung beim Menschen soll als eine Art der reinen Wahrnehmung angesehen werden.

Zur Texturunterscheidung werden Texturen mit unterschiedlichen Statistiken erzeugt. Die Untersuchungen zeigen, daß Texturen mit unterschiedlichen Statistiken erster und zweiter Ordnung vom Menschen noch unterschieden werden können. Dies gilt jedoch nicht mehr für Texturen, die in ihren Statistiken dritter und höherer Ordnung voneinander abweichen. Julesz behauptet, daß die Leistungsfähigkeit des visuellen Systems vermutlich bei der statistischen Beschreibung dritter und höherer Ordnung überfordert ist.

Obwohl man einige eher schwache Gegenbeispiele gefunden hat, besteht die Vermutung, daß im wesentlichen Statistiken zweiter Ordnung zur Texturunterscheidung genügen. Die im nächsten Abschnitt vorzustellenden grauwertbasierten statistischen Verfahren zur Texturanalyse benutzen meist Statistiken zweiter Ordnung als Texturmerkmale.

Ausgehend von der oben genannten Hypothese hat Julesz mit Unterstützung von Forschungsergebnissen weiter vermutet, daß die Texturunterscheidung nur von den globalen statistischen Eigenschaften der Texturen abhängig ist.

4.1.2 Statistiken lokaler Texturelemente

Die Kernhypothese über Texturwahrnehmung beim Menschen nach Julesz, wonach die globalen Grauwertstatistiken für die Texturunterscheidung entscheidend seien, wurde später von ihm selber und anderen Wissenschaftlern widerlegt. Es wurde festgestellt, daß die Texturwahrnehmung des visuellen Systems ein lokaler Prozeß ist. Daraus entstand eine neue Hypothese über die Texturwahrnehmung, wonach die Statistiken erster Ordnung lokaler Texturelemente für die Texturunterscheidung am wichtigsten seien.

In seiner *groben Skizze*²-Theorie beschreibt [Marr 82] Texturen mit symbolischen Geometrieformen, die aus den Grauwertänderungen im Bild ermittelt werden können. Als Geometrieformen schlägt er vor, Kanten und kompakte Flecken sowie lineare Flecken zu verwenden. Die Texturunterscheidung soll daher durch die Auswertung der Statistiken erster Ordnung der Attribute von Geometrieformen, wie z.B. Grauwert, Farbe, Größe, Ausrichtung, Endposition usw. erfolgen.

[Beck 83] präsentiert experimentelle Ergebnisse mit spezifischen Stimuli zur Unterstützung seiner Behauptung, daß die Texturunterscheidung vorwiegend aufgrund der Eigenschaften von lokalen Texturelementen erfolgt. Er hat festgestellt, daß die Statistiken lokaler Texturelemente wichtiger als die globalen Grauwertstatistiken zweiter Ordnung für die Textursegmentierung sind. Obwohl er diese Vermutung schon im Jahr 1972 geäußert hat, hat er nicht wie [Marr 82] eine explizite Definition von Texturelementen angegeben.

Nach vielen Überlegungen und Untersuchungen stoßen [Julesz & Bergen 83] die ursprüngliche Hypothese von [Julesz 75] um, wonach die Texturunterscheidung allein nach unterschiedlichen globalen Grauwertstatistiken erfolgt, und stellen eine neue Hypothese vor. Mit Versuchsergebnissen zeigen [Julesz & Bergen 83], daß die Statistiken erster Ordnung von *Textons*³ geeigneter als die globalen Grauwertstatistiken zweiter Ordnung für

² *grobe Skizze*: „*primal sketch*“.

³ *Texton*: *Texton* wird definiert als Fleck (wie z.B. rechteckige Fläche, Ellipse oder Liniensegment)

die Texturunterscheidung bei reiner Texturwahrnehmung sind. Sie haben mit Beispielen gezeigt, daß zwei unterschiedliche Texturen, die die gleichen globalen Grauwertstatistiken zweiter Ordnung haben, aber sich in ihren lokalen Statistiken erster Ordnung von *Textons* unterscheiden, vom visuellen System unterschieden werden können. Julesz und seine Mitarbeiter stellten danach die neue Vermutung auf, daß die *Textons* zusammen mit ihren Attributen die fundamentalen Elemente für die Texturunterscheidung sind ([Julesz 86]). [Williams & Wechsler 92] geben eine Zusammenfassung über die neuen Forschungsergebnisse in der Texturunterscheidung an.

Die Hypothesen aus den oben genannten vier Arbeiten stimmen fast überein. Aus diesen Vermutungen darf zumindest derzeit geschlossen werden, daß die auf lokalen Merkmalen basierenden Statistiken aussagekräftiger zur Texturunterscheidung sind als punktbaasierte globale Statistiken. Ein Grund dafür besteht darin, daß bei der Extraktion lokaler Merkmale mehrere Bildpunkte herangezogen worden sind. Die Untersuchungen in den oben genannten Arbeiten werden nur mit künstlich erzeugten Bildmustern durchgeführt. Die Extraktion von Texturelementen in realen Bildern, welche eine wichtige und schwierige Aufgabe ist, wird dort nicht berücksichtigt. Solche Verfahren werden in Abschnitt 4.3 dieses Kapitels vorgestellt.

4.2 Statistische Verfahren zur Texturanalyse

Statistische Verfahren werden zur Analyse derjenigen Texturen verwendet, die weder wiederholt noch orientiert organisiert sind. Zu den statistischen Verfahren gehören sowohl punktbaasierte als auch auf lokalen sowie globalen Merkmalen basierende Verfahren. Bei den punktbaasierten Ansätzen werden zunächst lokale Beziehungen benachbarter Punkte im Bild berechnet und anschließend die Statistiken aus solchen Beziehungen extrahiert. Nach der an der lokalen Beziehungsberechnung beteiligten Punktanzahl werden Statistiken in Statistiken erster Ordnung, zweiter und höherer Ordnung unterteilt. Andere Statistiken eines Grauwertbildes wie z.B. sein Leistungsspektrum werden auch zur Texturanalyse genutzt. Bei den auf lokalen Merkmalen basierenden Ansätzen werden oft Filter zur Extraktion lokaler Grauwertänderungen eingesetzt. Statt der Detektion der Texturelemente werden Statistiken lokaler Verteilungen von Kantenelementen oder Flecken berechnet. Als Texturmerkmale stehen den Segmentierungsverfahren solche Statistiken zur Verfügung.

Statistische Verfahren zur Texturanalyse werden in [Haralick 79], [Wechsler 80] und [Gool *et al.* 85] ausführlich aufgezählt. Deshalb wollen wir hier nicht auf alle in der Literatur beschriebenen Verfahren eingehen. Statt dessen sollen hier die vielversprechendsten klassischen Verfahren beschrieben und die neuen Verfahren zusammengefaßt werden.

zusammen mit seinen Attributen, wie z.B. Grauwert, Ausrichtung, Breite, Länge usw..

4.2.1 Stochastische Modelle von Texturen

Ein Texturmodell ist ein mathematisches Verfahren, mit dem ein texturiertes Bild erzeugt und beschrieben werden kann. Nach der stochastischen Auffassung über Texturentstehung ist Textur eine Zufallsstichprobe der Wahrscheinlichkeitsverteilung im Bildraum. Texturmodelle sind wichtig für die Texturanalyse. Mit einer expliziten Modellierung von Texturen kann festgestellt werden, was im Bild zu ermitteln ist und welche Parameter dabei bestimmt werden sollen. Obwohl Texturmodelle für die Texturbeschreibung so wichtig sind, sind die meisten Verfahren zur Texturanalyse texturmerkmalsbasiert und verwenden keine Modelle. In diesem Abschnitt werden drei stochastische Modelle, Markoff-Zufallsfeld, Autoregressives Modell und Fraktale Dimension, dargestellt. Deterministische Texturmodelle werden in Abschnitt 4.3.1 behandelt.

Markoff-Zufallsfeld

Das Markoff-Zufallsfeld-Modell kann durch Markoff-Prozesse erzeugt werden (z.B. [Cross & Jain 83]). In einem Markoff-Prozeß wird die Entscheidung, ob eine Zelle in einer Reihe schwarz, weiß oder grau sein soll, nach einer gegebenen mathematischen Formel durch die Farbe einer bestimmten Anzahl vorausgehender Zellen bestimmt. Eine für Experimente wichtige Eigenschaft dieser Texturmodelle besteht darin, daß mit einer bestimmten Klasse von Markoff-Prozessen die Texturen erzeugt werden können, die identische Statistiken erster, zweiter und n -ter Ordnung besitzen, sich aber in ihren Statistiken $(n+1)$ -ter Ordnung unterscheiden. Die Texturen bei der Untersuchung der Leistungsfähigkeiten des visuellen Systems des Menschen sind meistens mit Markoff-Prozessen erzeugt worden. Wegen der Zufälligkeit seiner Verteilung ist das Markoff-Zufallsfeld-Modell geeignet für die Beschreibung von Mikrot Texturen, aber nicht für die Beschreibung regulärer Makrotexturen.

Autoregressives Modell

Das autoregressive Modell eines Bildes wird ebenfalls zur Modellierung von Texturen verwendet. Sei $f(x, y)$ der Grauwert am Punkt (x, y) . Nach dem autoregressiven Modell kann $f(x, y)$ durch eine lineare Kombination der Grauwerte seiner Nachbarn mit additivem weißen Rauschen $w(x, y)$ erzeugt werden:

$$f(x, y) = \sum_{p=-M}^M \sum_{q=-N}^N a_{pq} f(x-p, y-q) + w(x, y), \quad (4.1)$$

wobei gilt $(p, q) \neq (0, 0)$. Die Koeffizienten $\{a_{pq} | p = -M, \dots, M, q = -N, \dots, N, (p, q) \neq (0, 0)\}$ charakterisieren die Texturen. Sie geben den funktionalen Zusammenhang zwischen einem Punkt und seinen lokalen Nachbarn an. Mit wachsender Fenstergröße nimmt der Rechenaufwand stark zu. Deswegen eignet sich das Verfahren nur für die Beschreibung von Mikrot Texturen. Als ein Beispiel wird das Modell in [Khotanzad & Chen 89] für die Charakterisierung von Texturen verwendet. Dabei werden als Texturmerkmale 6 Koeffizienten an jedem Punkt aus einem Bild geschätzt und für die Segmentation des Bildes eingesetzt.

Fraktale Dimension

Die Fraktale Dimension ist der wichtigste Parameter eines fraktalen Brown'schen Modelles. Mit fraktalen Beschreibungen wurden erhebliche Fortschritte in der Erzeugung natürlich aussehender Objekte in der Computergraphik erzielt. [Pentland 84] zeigt, daß die fraktale Dimension zur Beschreibung von Grobheiten der texturierten Oberfläche verwendet werden kann. Die fraktale Beschreibung natürlicher Texturen ist eines der grundlegenden mathematischen Hilfsmittel zum Verständnis von Objektoberflächen.

Bei der fraktalen Beschreibung von Texturen ist die fraktale Dimension ein Wert zwischen 2 (ebene Oberfläche) und 3 (Oberfläche mit höheren Spitzen). Je irregulärer eine Textur ist, desto größer ist ihre fraktale Dimension. Eine wichtige Eigenschaft der fraktalen Beschreibung von Texturen ist die sogenannte Selbst-Ähnlichkeit. Damit ist gemeint, daß die fraktale Beschreibung einer Textur invariant gegenüber den linearen Transformationen der Daten sowie skaleninvariant ist. In [Pentland 84] wird gezeigt, daß eine fraktale Oberfläche auch fraktale Grauwerte hat. Die fraktale Dimension einer Oberfläche wird durch die Auswertung der Grauwerte im Bild ermittelt. Die fraktale Dimension eignet sich besonders für die Beschreibung irregulärer Texturen. Für weitere Anwendungen fraktaler Dimensionen in der Bildverarbeitung siehe zum Beispiel [Rueff 89], [Zahlten *et al.* 93] und [Paulus *et al.* 93].

4.2.2 Punktbasierte Statistiken n -ter Ordnung

Statistiken erster Ordnung

Statistiken erster Ordnung sind Statistiken von einzelnen Bildpunkten. Sie beschreiben die Wahrscheinlichkeit, mit der ein einzelner Punkt im Texturbild eine bestimmte Helligkeit oder, um es genau zu sagen, eine bestimmte Leuchtdichte hat. Texturmerkmale werden meistens aus dem Histogramm der Grauwerte im Bild berechnet. Sei $h(i)$ die Häufigkeit eines Grauwertes i und μ der mittlere Grauwert, so kann man die folgenden Statistiken als Texturmerkmale verwenden:

$$\text{Mittelwert:} \quad \mu = \frac{1}{N} \sum_{i=0}^{N-1} ih(i) \quad (4.2)$$

$$\text{Standardabweichung:} \quad \sum_{i=0}^{N-1} (i - \mu)^2 h(i) \quad (4.3)$$

$$\text{Entropie:} \quad - \sum_{i=0}^{N-1} h(i) \log h(i) \quad (4.4)$$

$$\text{Drittes Moment :} \quad \sum_{i=0}^{N-1} (i - \mu)^3 h(i) \quad (4.5)$$

Dabei ist leicht einzusehen, daß solche Statistiken nur mit dem Grauwert zu tun haben. Die lokale örtliche Beziehung benachbarter Punkte im Bild wird damit nicht berücksichtigt. Verfahren zur Texturanalyse mit Grauwertstatistiken erster Ordnung kommen

heutzutage sehr selten vor. Der Grund dafür besteht darin, daß zwei Texturen, die in ihren Statistiken zweiter, dritter oder n -ter Ordnung übereinstimmen, auch in allen Statistiken niedrigerer Ordnung gleich sind. Damit ist gemeint, daß aus einer Statistik höherer Ordnung die Statistiken jeder niedrigeren Ordnung eindeutig bestimmt werden können. Deswegen werden bei der Texturanalyse meist Statistiken zweiter Ordnung verwendet.

Obwohl Grauwertstatistiken erster Ordnung wenig bedeutend für die Texturbeschreibung sind, wird festgestellt, daß Statistiken erster Ordnung von lokalen Merkmalen und Texturelementen wichtige Texturmerkmale sind. Der Grund liegt daran, daß Texturelemente eine nicht lokale Beschreibung darstellen. Im Abschnitt 4.2.4 und 4.3 werden solche Verfahren behandelt.

Statistiken zweiter Ordnung

Statistiken zweiter Ordnung sind die Statistiken von Punktpaaren. Texturanalyse mit Statistiken zweiter Ordnung dominieren die statistischen Verfahren. Sie können als eine Obermenge der Statistiken erster Ordnung betrachtet werden. Wenn die Statistiken zweiter Ordnung von zwei Texturen identisch sind, dann müssen die beiden Texturen auch identische Statistiken erster Ordnung haben. Die meisten statistischen Verfahren zur Texturanalyse verwenden die Texturmerkmale aus folgenden Methoden:

Grauwertpaarhäufigkeitsmatrix: Die Grauwertpaarhäufigkeit $f(i, j|d, \theta)$ beschreibt die Wahrscheinlichkeit, daß ein Bildpunkt mit dem Grauwert i in der Richtung θ mit einem Abstand d einen Nachbar mit dem Grauwert j hat. Wir nehmen an, daß ein Bild N Grauwertstufen besitzt. Wir können dann jedes $f(i, j|d, \theta)$ ($i, j = 0, \dots, N - 1$) als ein Element einer $N \times N$ Matrix betrachten. Die Matrix $(f(i, j|d, \theta))_{N \times N}$ wird Grauwertpaarhäufigkeitsmatrix⁴ in Richtung θ mit dem Abstand d genannt. Sie wird im weiteren mit GPM abgekürzt. Eine GPM wird auch als ein 2D-Histogramm der Grauwerte bezeichnet. [Haralick *et al.* 73] leiten aus der GPM 14 Texturmerkmale ab. Darauf werden wir in diesem Abschnitt noch einmal zurückkommen.

Grauwertdifferenzfeld: Die Grauwertdifferenz $P_\delta(k)$ beschreibt die Wahrscheinlichkeit, daß die Grauwertdifferenz zwischen zwei Bildpunkten mit einem Abstand δ den Wert k ($k = 0, 1, \dots, N - 1$) aufweist. Das Grauwertdifferenzfeld $(P_\delta(k))_N$ ist eine Untermenge der Grauwertpaarhäufigkeitsmatrix. Die Statistiken der Grauwertdifferenz werden z.B. in [Weszka *et al.* 76] zur Texturklassifikation verwendet.

Im folgenden werden die 5 am häufigsten verwendeten Statistiken aufgezählt:

$$\text{Energie:} \quad \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} [f(i, j|d, \theta)]^2 \quad (4.6)$$

$$\text{Entropie:} \quad - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j|d, \theta) \log f(i, j|d, \theta) \quad (4.7)$$

⁴Grauwertpaarhäufigkeitsmatrix: engl. *gray level co-occurrence matrix*.

$$\text{Korrelation:} \quad \frac{1}{\sigma_x \sigma_y} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - \mu_x)(j - \mu_y) f(i, j|d, \theta) \quad (4.8)$$

$$\text{Lokale Homogenität:} \quad \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1 + (i - j)^2} f(i, j|d, \theta) \quad (4.9)$$

$$\text{Trägheit:} \quad \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - j)^2 f(i, j|d, \theta)^2. \quad (4.10)$$

dabei gilt:

$$\mu_x = \sum_{i=0}^{N-1} i \sum_{j=0}^{N-1} f(i, j|d, \theta)$$

$$\mu_y = \sum_{j=0}^{N-1} j \sum_{i=0}^{N-1} f(i, j|d, \theta)$$

$$\sigma_x^2 = \sum_{i=0}^{N-1} (i - \mu_x)^2 \sum_{j=0}^{N-1} f(i, j|d, \theta)$$

$$\sigma_y^2 = \sum_{j=0}^{N-1} (j - \mu_y)^2 \sum_{i=0}^{N-1} f(i, j|d, \theta)$$

Da die Texturmerkmale aus der originalen GPM die Grauwertstatistiken von Punktpaaren sind, lassen sich mit ihnen nur Mikrot Texturen gut beschreiben. Bei der Beschreibung von Makrot Texturen mit der GPM ergeben sich Schwierigkeiten. Zur Überwindung des Problems verallgemeinern [Davis *et al.* 79] die GPM auch auf andere lokale Merkmale, wie z.B. Grauwertkanten. Damit können auch Statistiken von Beziehungen zwischen lokalen Merkmalen berechnet werden. Die Untersuchungen in [Davis *et al.* 79] zeigen, daß Statistiken von Paarhäufigkeitsmatrizen zwischen Kanten für die Texturklassifizierung aussagekräftiger als Statistiken aus der originalen GPM sind. Weitere Beispiele dafür findet man in [Pong *et al.* 92] und [Haddon *et al.* 93].

Die von [Haralick *et al.* 73] vorgeschlagenen 14 Texturmerkmale können in zwei Gruppen, die rein statistischen Merkmale und die Merkmale nach der Informationstheorie, untergliedert werden. Es ist nötig festzustellen, welche Merkmale für die Texturklassifizierung wichtiger sind. Ein Verfahren in dieser Richtung ist in [Gotlieb & Kreyszig 90] dargestellt. Statt einzelner Merkmale werden dort kombinierte Statistiken verwendet. Aus 14 Merkmalen werden 6 von den Autoren als repräsentativ angesehene Statistiken selektiert und alle 63 möglichen Kombinationen gebildet. Anwendungsergebnisse zeigen, daß die kombinierten Statistiken leistungsfähiger als die einzelnen Statistiken sind. Die Autoren weisen weiter darauf hin, daß Kombinationen höherer als vierter Ordnung keine signifikant besseren Leistungen bringen werden.

Die Texturanalyse mit den Texturmerkmalen der GPM gehört prinzipiell zu den sogenannten merkmalsbasierten Verfahren. Die Verwendung von Texturmodellen in einem solchen Verfahren wird in [Chetverikov 91] erprobt. Er präsentiert ein Verfahren, in dem

die Kontrastkurve von Texturen mit dem ersten Moment der GPM modelliert wird. Mit diesem Modell können, nach der Meinung des Autors, sowohl stochastische als auch deterministische Texturen beschrieben werden. Dabei wird die Kontrastkurve definiert als eine Funktion des Abstandes d in $f(i, j | \theta, d)$. Wie der Autor gezeigt hat, ist das Modell nicht geeignet für die Beschreibung von Texturen mit hierarchischen Strukturen und Texturen mit großer Grauwertänderung zwischen den Texeln.

Anders als die obengenannten GPM-basierten Verfahren, die die aus der GPM abgeleiteten Statistiken als Merkmale verwenden, greift [Lohmann 93] direkt auf die GPM zurück. Mit den GPM als Merkmalsvektoren wird ein Verfahren zur Segmentierung von Texturen vorgestellt. Dabei wird angenommen, daß diese Merkmalsvektoren Polynomialverteilung aufweisen.

Der Aufbau der GPM und die Berechnung der Statistiken daraus sind sehr rechenzeit-aufwendig und brauchen viel Speicherplatz. Diese Eigenschaften der GPM beschränken machmal ihre Anwendung. Deshalb erfolgen Untersuchungen zur Reduzierung der Rechenzeit und des Speicherbedarfs. Mit dieser Überlegung versuchen [Unser 86] und [Peckinpough 91], die Berechnung von Statistiken aus der GPM durch die Auswertung von Histogrammen zu ersetzen. Diese Vereinfachung ist allerdings nur für einige Statistiken möglich. Anwendungsbeispiele zeigen, daß Rechenzeit und Speicherbedarf dadurch wesentlich reduziert werden können.

Statistik höherer Ordnung

Die meisten Statistiken höherer Ordnung werden aus Grauwertlauflängenmatrizen berechnet. Die Grauwertlauflänge ⁵ $P_{\theta}(g, k)$ ($g = 1, \dots, m, k = 1, \dots, n$) beschreibt die Anzahl der Punkte, die in Richtung θ k Nachbarn mit dem Grauwert g haben. Statistiken höherer Ordnung werden heute kaum noch bei der Texturanalyse untersucht. Ein frühes Beispiel für die Auswertung von Statistiken höherer Ordnung ist in [Galloway 75] zu finden. Sowohl die theoretische Untersuchung in [Connors & Harlow 80a] als auch das Experiment in [Weszka *et al.* 76] weisen darauf hin, daß Statistiken höherer Ordnung keine signifikant besseren Ergebnisse für die Texturklassifizierung als Statistiken zweiter Ordnung bringen. Diese Aussage stimmt mit den Ergebnissen der Untersuchung zur Texturunterscheidung beim Menschen überein. Daher erscheint die genaue Beschreibung von Texturen mit Statistiken höherer Ordnung zur Zeit nicht nötig.

Welche Statistik ist wichtiger ?

In den sechziger und siebziger Jahren sind viele statistische Verfahren zur Texturanalyse entwickelt worden. [Weszka *et al.* 76] haben einen experimentellen Vergleich von Statistiken zwischen den bekanntesten Methoden - GPM, Grauwertlauflängenmatrix, Grauwertdifferenzfeld und Leistungsspektrum - durchgeführt. Das Vergleichskriterium ist das Verhältnis der richtigen Klassifikation von realen Texturen mit unterschiedlichen Statistiken. Zum Schluß wird festgestellt, daß die Statistiken der GPM die besten sind. Vier Jahre

⁵Grauwertlauflänge: engl. *gray level run length*.

später führten [Connors & Harlow 80a] einen theoretischen Vergleich mit den Statistiken aus denselben vier Methoden durch. In Anlehnung an die enthaltenen Texturinformationen werden auch die Statistiken der GPM als die besten bewertet. Statistiken zweiter Ordnung aus anderen Methoden können prinzipiell aus den Statistiken der GPM mathematisch abgeleitet werden. Die Texturanalyse durch Auswertung von Statistiken der GPM ist das einzige heute noch intensiv untersuchte, punkt-basierte statistische Verfahren.

4.2.3 Globale Statistiken

Leistungsspektrum

Die Analyse eines Bildes durch die Auswertung seiner Fouriertransformierten im Ortsfrequenzraum wird auch bei der Texturanalyse verwendet. Bei der Darstellung eines Bildes im Ortsfrequenzraum können einige Eigenschaften des Bildes verdeutlicht werden, die im Ortsraum nicht leicht zu erkennen sind. Mit der Fouriertransformation ist die Möglichkeit geschaffen worden, das Spektrum eines Bildes zu messen und zu analysieren. Aus den Spektren werden Statistiken zur Texturklassifizierung berechnet. Das Leistungsspektrum eines Bildes wird wie folgt definiert:

Leistungsspektrum: Seien $f(x, y)$ ein texturiertes Bild der Größe $I \times J$. Seine diskrete 2D-Fouriertransformation wird definiert als

$$F(u, v) = \sum_{x=0}^{I-1} \sum_{y=0}^{J-1} f(x, y) \exp^{-2\pi\sqrt{-1}(ux/I+vy/J)}. \quad (4.11)$$

Das Leistungsspektrum, $|F(u, v)|^2$, beschreibt die Stärke jeder Ortsfrequenz der Texturen. [Lim & Malik 81] beschreiben ein Anwendungsbeispiel für die Verwendung des Leistungsspektrums in der Texturklassifikation.

Entsprechend den in [Haralick 79] beschriebenen 14 Merkmalen werden in [Liu & Jernigan 90] 28 Texturmerkmale aus dem Leistungs- und Phasenspektrum im Ortsfrequenzraum definiert. [Liu & Jernigan 90] bemühen sich darum, bessere Texturmerkmale im Ortsfrequenzraum zu finden. Eine Kombination von vier aus 28 Merkmalen wird als das beste Merkmal für die Bildklassifizierung bestimmt. Solche Texturmerkmale aus den Spektren sind, wie die Autoren mit den Anwendungsbeispielen gezeigt haben, im allgemeinen wenig sensibel gegenüber additivem Rauschen und geeignet für die Klassifizierung natürlicher Texturen. Die Merkmale in [Liu & Jernigan 90] müssen in einem großen Ausschnitt berechnet werden, um stabile Werte zu gewährleisten. Wie bei allen spektralbasierten Verfahren ist auch hier Bildsegmentierung problematisch.

Grauwertbeziehungshistogramm

Ausgehend von Beziehungen benachbarter Punkte definieren [He & Wang 91] eine sogenannte Texturspektraldomäne für die Texturanalyse. Es wird zunächst eine Beziehung zwischen jedem Bildpunkt und seinen 8 Nachbarn berechnet. Ein Nachbar wird je nach

dem Verhältnis zwischen seinem Grauwert und dem Grauwert des Zentralpunktes als 2 (größer), 1 (gleich) oder 0 (kleiner) bewertet. Es gibt für 8 Nachbarn mit 3 unterschiedlichen Werten insgesamt 3^8 (6561) mögliche Beziehungen. Solche Beziehungen werden mit sogenannten Textureinheiten beschrieben. Die Häufigkeit aller Beziehungen in einem Bild wird von den Autoren als Texturspektrum bezeichnet. Das Texturspektrum ist eigentlich nichts anders als das Histogramm der Grauwertbeziehungen aller Punkten zu ihren 8 Nachbarn. Die von den Autoren genannten Texturfilter sind auch die Operationen bei Histogrammen.

Grauwertdifferenzhistogramm

Entsprechend der Vorgehensweise bei der Texturwahrnehmung beim Menschen stellen [Amadasun & King 89] 5 Texturmerkmale für maschinelle Texturunterscheidung vor: Texturgrobheit, Texturkontrast, Texturänderungsfrequenz, Texturkomplexität und Texturstärke. Mit diesen Merkmalen sollen sowohl Mikro- als auch Makrotexturen richtig beschrieben werden. Ein Grauwertbild wird dabei zuerst in ein sogenanntes Grauwertdifferenzhistogramm überführt. Der i -te Wert des Histogrammes ist die Summation der Grauwertdifferenzen zwischen den Punkten mit dem Grauwert i und ihren lokalen Nachbarn. Die Texturmerkmale werden aus diesem Histogramm berechnet. Zur Bewertung der Leistungsfähigkeit der aufgestellten Merkmale haben die Autoren einige interessante Experimente durchgeführt. Sie lassen 88 Menschen 10 Texturen aus [Brodatz 66] nach den 5 vorgestellten Textureigenschaften eine Reihenfolge zuordnen. Aus denselben Bildern werden diese 5 Texturmerkmale berechnet und es wird dann nach der Größe des Ergebnswerts ebenfalls eine Reihenfolge zugeordnet. Die berechneten Reihenfolgen stimmen mit den von Menschen subjektiv zugeordneten Reihenfolgen fast vollständig überein.

4.2.4 Filterbasierte lokale Statistiken

Die in den beiden vorangehenden Abschnitten besprochenen Statistiken sind grauwertbasierte globale Statistiken. Im allgemeinen lassen sich Grauwertänderungen besser mit lokalen Merkmalen darstellen, wie z.B. Kanten, lokalen Maxima, lokalen Krümmungen usw.. Solche lokalen Merkmale werden meistens durch die Faltung eines Bildes mit bestimmten Filtern ermittelt. Die am häufigsten verwendeten Texturmerkmale aus einem gefilterten Bild sind die Dichte lokaler Merkmale, der Mittelwert lokaler Merkmale, die Entropie des Histogramms lokaler Merkmale und die Entropie des Richtungshistogramms lokaler Merkmale.

Lokale Merkmale

[Pietikäinen & Rosenfeld 82] verwenden Statistiken erster Ordnung von Kantenpaaren als Texturmerkmale. Zur Extraktion von Kanten wird ein Bild mit 8 Filtern entsprechend 8 Richtungen gefaltet. Der Filter mit der größten Antwort an jedem Punkt wird selektiert und seine Richtung als Gradientenrichtung des Grauwertes an diesem Punkt definiert. Entlang der Gradientenrichtung jeder Kante wird bis zu einem bestimmten Abstand

eine andere Kante mit entgegengesetzter Gradientenrichtung gesucht. Hierbei wird angenommen, daß die beiden Kanten zu demselben Texturelement gehören. Statistiken von Kantenpaaren werden berechnet und als Texturmerkmale verwendet. Solche Statistiken entsprechen eigentlich den Statistiken aus verallgemeinerten Paarhäufigkeitsmatrizen in [Davis *et al.* 79].

Gradientenrichtungen von Grauwerten in einem texturierten Bild beschreiben die Richtungen der Grauwertänderungen zwischen den Texturen. Diese Information wird in [Yu *et al.* 91] bei der Texturanalyse ausgenutzt. Sie verwenden die Richtungsenergie als Texturmerkmal für die Textursegmentierung. Es werden zuerst die absoluten Differenzen der Grauwerte eines Punktes und seiner Nachbarn in 8 Richtungen berechnet. Die Summation der Differenzen in jeder Richtung innerhalb eines Ausschnittes wird von den Autoren als Richtungsenergie bezeichnet. Für jeden Bildpunkt wird ein Vektor mit den Texturenergien in 8 Richtungen bestimmt. Die Segmentation eines Bildes wird realisiert durch die Kantendetektion im Richtungsenergiebild.

Ausgehend von der Definition über Szenentexturen modelliert [Vistnes 89] Szenentexturen als eine Sammlung von Unterstrukturen. Eine Unterstruktur wird weiter charakterisiert durch ihre Zusammensetzung, ihre Form sowie ihre Position und Orientierung. Eine Bildtextur ist die Projektion einer Szenentextur. Für die Extraktion von Texturmerkmalen wird ein Satz von Schablonen mit unterschiedlicher Größe und Orientierung verwendet. Statt der Bestimmung von isolierten Unterstrukturen werden hier die Statistiken der Texturen aus der größten Schablonenantwort direkt ermittelt. Der statistische t -Test dient hier dazu, die Grenzlinie zweier texturierter Regionen zu bestimmen.

[Tuceryan 92] berechnet lokale Texturmerkmale aus Momenten von Bildgrauwerten. Dabei werden die Bildmomente bis zu dritter Ordnung ($m_{00}, m_{10}, m_{01}, m_{20}, m_{11}, m_{02}$) verwendet. Die Anwendung solcher Texturmerkmale für die Bildsegmentierung zeigt gute Ergebnisse.

Texturenergie

Eine der Schwierigkeiten in der Texturmessung liegt in der Anpassung von Form, Ausrichtung und Größe von Filtern an die betrachteten Texturen. Zur Überwindung der ersten zwei Probleme hat [Laws 80] einige Faltungsmasken mit verschiedener Form eingeführt. Vier der wichtigsten 5×5 -Masken sind in Abb. 4.2 dargestellt. Aus den Faltungsergebnissen wird die sogenannte Texturenergie berechnet. Die Texturenergie in jedem Punkt ist die Summe der Quadrate jedes Pixelwertes der gefalteten Bilder. Der Autor behauptet, daß seine Texturenergien für die Textursegmentation besser als die Statistiken der GPM sind. Diese Aussage wird später auch von [Pietikäinen *et al.* 82] mit Anwendungsergebnissen unterstützt. Sie weisen weiter darauf hin, daß bessere Ergebnisse eigentlich von der Form der Filter abhängen. Die in Abb. 4.2 dargestellten 4 Filter werden auch in [Hsiao & Sawchuk 89] zur Messung der Texturenergie für die überwacht eingelernte Texturklassifikation mit dem klassischen Bayes'schen Klassifikator benutzt.

Es wird spekuliert, daß Texturen hierarchisch organisiert sind und auch hierarchisch analysiert werden sollten. Die von [Laws 80] eingeführten Filter haben verschiedene Form

$$\begin{array}{ccccc} -1 & -2 & 0 & 2 & 1 \\ -4 & -8 & 0 & 8 & 4 \\ -6 & -12 & 0 & 12 & 6 \\ -4 & -8 & 0 & 8 & 4 \\ -1 & -2 & 0 & 2 & 1 \end{array}$$

(a) *L5E5*.

$$\begin{array}{ccccc} -1 & 0 & 2 & 0 & -1 \\ -4 & 0 & 8 & 0 & -4 \\ -6 & 0 & 12 & 0 & -6 \\ -4 & 0 & 8 & 0 & -4 \\ -1 & 0 & 2 & 0 & -1 \end{array}$$

(c) *L5S5*.

$$\begin{array}{ccccc} -1 & 0 & 2 & 0 & -1 \\ -2 & 0 & 4 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & -4 & 0 & 2 \\ 1 & 0 & -2 & 0 & 1 \end{array}$$

(b) *E5S5*.

$$\begin{array}{ccccc} 1 & -4 & 6 & -4 & 1 \\ -4 & 16 & -24 & 16 & -4 \\ 6 & -24 & 36 & -24 & 6 \\ -4 & 16 & -24 & 16 & -4 \\ 1 & -4 & 6 & -4 & 1 \end{array}$$

(d) *R5R5*.

Abbildung 4.2: Vier der Filter in [Laws 80] für die Berechnung der Texturenergie im Bild: eine vertikale Grauwertkante *L5E5* in (a), ein V-förmiges Grauwerteck *E5S5* in (b), eine vertikale Grauwertlinie *L5S5* in (c) und ein Grauwertpunkt *R5R5* mit hoher Orts-Frequenz in (d).

und Ausrichtung. Eine offene Frage ist, wie die Größe der Filter auch an die Texturen angepaßt werden kann. Untersuchungen zu diesem Thema werden nun im folgenden vorgestellt. Solche Verfahren werden meist durch die Auswertung der Texturenergie in unterschiedlichen Ebenen einer Pyramide realisiert.

Zuerst haben [Larkin & Burt 83] diese Idee bei Gaußpyramiden von Grauwertbildern implementiert. Dazu wird zuerst eine Gaußpyramide eines Grauwertbildes aufgebaut ([Burt 83]). Die Texturenergie wird auf jeder Ebene der Pyramide berechnet. Darauf folgend wird eine Energiepyramide basierend auf der Texturenergie aus jeder Ebene der Gaußpyramide aufgebaut. Dadurch soll die Möglichkeit geschaffen werden, daß Texturenergien von Texturen mit unterschiedlicher Größe gemessen werden können. Die Autoren geben dabei leider nicht an, auf welcher Ebene der Pyramide die Klassifizierung von Texturen stattgefunden hat.

Mit ähnlicher Überlegung verwenden [Sher & Rosenfeld 89] Texturenergie zur Extraktion kompakter texturierter Regionen im Bild und implementieren die Texturklassifikation in einer Pyramide. Zu Beginn wird ein Grauwertbild mit 8 in [Laws 80] gegebenen Filtern gefaltet. Für jedes gefaltete Bild wird die Texturenergie berechnet und darauf eine Pyramide aufgebaut. Dadurch entstehen insgesamt 8 Energiepyramiden für jedes Grauwertbild. In jeder Pyramide wird die Detektion einer kompakten Region durch die Lokalisierung eines isolierten Punktes auf einer höheren Ebene und anschließend die Verbindung des Punktes mit den Punkten auf den tiefsten Ebenen realisiert.

Bei den oben vorgestellten pyramidenbasierten Arbeiten wird Texturenergie dazu benutzt, ein Bild durch iterative Aufspaltung in Unterbereiche zu segmentieren. [Unser & Eden 89+90] präsentieren ein Verfahren zur Texturklassifizierung mit der Texturenergie als Texturmerkmal in einer Pyramidenstruktur. Auf der tiefsten Ebene wird die Texturenergie jedes Bildpunktes mit einer Filterbank berechnet. Die Texturmerkmale auf den höheren Ebenen werden berechnet durch iterative Gaußglättung der Texturenergie auf der tiefsten Ebene. In [Unser & Eden 89+90] werden sowohl überwacht eingelernte als auch unüberwacht eingelernte Klassifikatoren auf die Texturmerkmale einer höheren Pyramidenebene angewendet. Eine Schwäche dieses Verfahrens ist seine Ungenauigkeit bei der Bildsegmentierung. Ein grob-zu-fein Prozeß, der ausgehend von der Segmentierung eines Bildes auf höheren Pyramidenebenen die Segmentierung des Bildes auf tieferen Ebenen erzielt, erfolgt hier nicht.

4.3 Strukturelle Verfahren zur Texturanalyse

In diesem Abschnitt sollen strukturelle Verfahren zur Analyse deterministischer Texturen vorgestellt werden. Deterministische Texturen bestehen allgemein aus einzelnen kompakten Strukturen. Sie werden beschrieben mit

- Texeln und
- Aufbauregeln.

Ein Texel ist ein primitives Texturelement mit einer Reihe von Attributen. Dabei ist das triviale Texel ein einzelner Punkt. Die Aufbauregel einer Textur beschreibt die geometrische Anordnung von Texeln. Eine Textur ist nur vollständig beschrieben, wenn die Texel und ihre Beziehungen zueinander erfaßt werden. Eine der Schwierigkeiten bei der Analyse deterministischer Texturen besteht darin, daß die Texel und die Aufbauregel voneinander abhängig sind. Sie bauen prinzipiell einen Zyklus auf: zur Ermittlung eines Texels soll die Aufbauregel bekannt sein, die Bestimmung der Aufbauregel ist aber durch die Extraktion des Texels bedingt. Außerdem sind strukturelle Verfahren im allgemeinen auf die betrachteten Aufgaben abgestimmt und weisen wenig Flexibilität auf. Die meisten in der Fachliteratur beschriebenen Verfahren zur Analyse deterministischer Texturen werden in zwei Schritte untergliedert, die Ermittlung von Texeln und die Bestimmung der Aufbauregel.

4.3.1 Deterministische Modelle von Texturen

Es gibt in der Fachliteratur nur sehr wenige Arbeiten, die die Modellierung von deterministischen Texturen behandeln. Mit einem deterministischen Texturmodell ist gemeint, daß sowohl die Texel als auch die Aufbauregel der Texel modelliert werden.

Kachelmodell

Ausgehend von der Vorstellung über ideale Texturen betrachtet [Zucker 76a] reale Texturen als Deformierung von idealen Texturen. Er verwendet ein Kachelmodell von Texturen. Ein Kachelmodell von Texturen besteht aus Texeln, idealen Texturen, Transformationen und beobachtbaren Texturen. Ideale Texturen entstehen aus einem kompakten Aufbau von Texeln. Sie werden durch Graphen beschrieben. In jedem Knoten des Graphen liegt ein Texel. Die Kanten des Graphen beschreiben die Aufbauregel. Natürliche Texturen entstehen durch eine Transformation der idealen Texturen. Z.B. kann die perspektivische Deformierung von Texturen als Ergebnis der Transformation von idealen Texturen betrachtet werden. In [Zucker 76a] werden nur synthetische Texturen untersucht. Anwendungen auf reale Bilder erfolgen nicht.

Mosaikmodelle

[Ahuja & Rosenfeld 81] stellen eine Klasse mosaikartiger Texturmodelle vor. Dabei unterscheiden sie zwei Arten von Modellen: Zellstrukturmodelle und Belagmodelle. In einem Zellstrukturmodell wird angenommen, daß Texturen durch eine Mosaikteilung einer Oberfläche in miteinander verbundene konvexe Polygone entstehen. Unterschiedliche Texturmodelle unterscheiden sich in der Art der Mosaikteilung. Ein mosaikartiges Belagmodell betrachtet Textur als eine zufällige Anordnung einer Gruppe geometrischer Figuren. Mit einem Belagmodell kann ein texturiertes Bild durch eine zufällige Verteilung geometrischer Figuren erzeugt werden. Zu den geometrischen Figuren gehören Liniensegmente, Kreise, Ellipsen, Rechtecke usw..

Mit den mosaikartigen Texturmodellen kann man lokale Eigenschaften, wie z.B. die Dichte der Zellen oder die Anzahl der Figuren aus einem modellierten Texturbild berechnen. Aus einem natürlichen Texturbild sind ebenfalls solche Eigenschaften zu extrahieren. Durch einen Vergleich von Eigenschaften aus beiden Bildern werden die Texturen im Bild identifiziert.

4.3.2 Extraktion von Texeln

Im allgemeinen werden mit den strukturellen Verfahren zur Texturanalyse zunächst die Texel im Bild lokalisiert und dann die Aufbauregel bestimmt. Zur Extraktion von Texeln wurden sowohl konturbasierte als auch bereichsbasierte Verfahren entwickelt. Die mit den konturbasierten Verfahren detektierten Texel werden durch ihre Konturen beschrieben. Die bereichsbasierten Texel, wie z.B. Flecke (engl. *blobs*), haben Bereichseigenschaften als Attribute. Die Attribute bereichsbasierter Texel enthalten daher mehr Informationen als die Attribute konturbasierter Texel. Die meisten als primitive Texturelemente im Bild extrahierten Texel sind bereichsbasiert.

Konturbasierte Verfahren

Bei einem konturbasierten Verfahren zur Extraktion von Texeln werden zu Beginn alle Kantenelemente detektiert. Nur eine Teilmenge davon sind eigentlich Konturpunkte von Texeln. Zur Beschreibung eines Texel müssen daher zunächst Konturpunkte selektiert, miteinander verbunden und zu geschlossenen Texelkonturen zusammengefaßt werden.

Ein Beispiel dafür wird in der Arbeit von [Vilnrotter *et al.* 86] beschrieben. Sie verwenden die Informationen über Kanten bei der Texelextraktion im Bild. Das Verfahren läuft folgendermaßen ab: Es werden zunächst 6 sogenannte Kantenwiederholungsfelder entsprechend 6 Richtungen in einem Kantenbild gebildet. Ein Feld ist eigentlich nichts anders als das Histogramm der Auftretenshäufigkeit von Kantenpaaren mit ähnlicher Richtung, aber unterschiedlichen Abständen (Vergleiche mit der Paarhäufigkeitsmatrix der Kanten in [Davis *et al.* 79]). Für wiederholte Texturen haben solche Felder lokale Maxima mit einem konstanten Abstand. Durch die Analyse solcher Maxima kann festgestellt werden, mit welchem Abstand sich die Kanten im Bild am häufigsten wiederholen. Die Texel werden lokalisiert durch das Auffinden ihrer geschlossenen Konturpunktketten. Der Abstand zwischen den Texeln läßt sich daher aus dem Wiederholungsfeld bestimmen. In diesem Verfahren werden die Texturen nur in 6 Richtungen analysiert. Dies ist im allgemeinen zu grob für die Analyse natürlicher Texturen.

Bereichsbasierte Verfahren

Bereiche mit homogenem Grauwert sind ein geeignetes primitives Element zur Texelbeschreibung. Zur Extraktion bereichsbasierter Texel wird oft die Segmentierung eines texturierten Bildes in Unterbereiche eingesetzt.

Mit der Definition von Texeln als homogener Bereich im Grauwert detektieren [Tomita *et al.* 82] die Texel durch ein Bereichswachstumsverfahren im Grauwertbild. Nach der Texeldetektion folgt die Berechnung der Statistiken jedes Texels und die Klassifizierung von Texeln mit einem histogrammbasierten Klassifikator. Bei der Ermittlung von Aufbauregeln werden zunächst die genauen geometrischen Eigenschaften eines als typisch angesehenen Texels berechnet. Mit den Anwendungsbeispielen wird gezeigt, daß das Verfahren Probleme bei der Klassifizierung von Texeln unterschiedlicher Größe hat. Die Texel in einem realen Bild sind aber meistens unterschiedlich in ihrer Größe.

In [Leu & Wee 85] wird versucht, Texel aus einem Binärbild zu extrahieren. Als erster Schritt des Verfahrens wird ein Binärbild mit einem Schwellenverfahren aus einem Grauwertbild generiert. Die Texel werden dann durch die Bestimmung der zusammenhängenden Regionen im Binärbild extrahiert. Die weiteren Verarbeitungsschritte des Verfahrens, nämlich die Formanalyse jedes Texels, die Klassifikation von Texeln nach ihren Statistiken und die Ermittlung von Aufbauregeln verlaufen ähnlich wie in [Tomita *et al.* 82]. Ein Problem bei diesem Verfahren ist die Auswahl einer optimalen Schwelle zur Erzeugung von Binärbildern. Verfahren zur Auswahl von Schwellen werden in [Sahoo *et al.* 88] ausführlich diskutiert.

Nach der Vorstellung der *Textons* in [Julesz & Bergen 83] als Flecken präsentieren [Voorhees & Poggio 87] ein Verfahren zur Extraktion von Flecken (eine Art von *Textons*)

im Bild. Anstatt der Bestimmung von Nulldurchgangsstellen in einem mit dem LoG⁶-Operator gefalteten Bild, was für die Kantendetektion wichtig ist, werden Flecke durch adaptive Schwellen extrahiert. Adaptive Schwellen lassen sich aus der Schätzung des Rauschens bestimmen. Statistiken erster Ordnung von Flecken werden bei der Bestimmung der Begrenzung zwischen unterschiedlichen Flecken verwendet. Als ein Vorteil gegenüber den obengenannten Verfahren können in diesem Verfahren Texel mit unterschiedlicher Größe detektiert werden.

[Blostein & Ahuja 89b] beschreiben die Texel als kreisförmige Scheiben mit konstantem Grauwert. Die kreisförmigen Scheiben im Bild werden detektiert durch die Auswertung von $\nabla^2 G * I$ und $\frac{\partial}{\partial \sigma} \nabla^2 G * I$. Um Texel mit unterschiedlicher Größe zu extrahieren, wird ein Grauwertbild mit einem Satz von 6 LoG-Filtern mit unterschiedlicher Größe (4×4 , 8×8 , ..., 24×24) gefaltet. In jedem gefalteten Bild werden die kreisförmigen Scheiben an jedem Maximum bzw. Minimum der Faltungsantworten ermittelt. Statt eine optimale Größe für jede Scheibe zu suchen, wird die Bestimmung der richtigen Größe durch die Schätzung der Form in der texturierten Objektoberfläche ersetzt. Ein solches Verfahren mit Filtern verschiedener Größe stellt eine Alternative für das Verfahren mit konstanter Filtergröße, aber mehreren Schwellen in [Voorhees & Poggio 87] dar. In der Arbeit von [Blostein & Ahuja 89a] wird die durch perspektivische Projektion entstandene Texturformänderung berücksichtigt. Eine detaillierte Beschreibung der Extraktion von Texeln findet man in [Blostein & Ahuja 89b]. Eine Schwäche des Verfahrens ist die Beschränkung der Texelform auf kreisförmige Scheiben. Wichtige Informationen über Texel, wie ihre Orientierung und Länge, gehen bei diesem Verfahren verloren.

Zur Extraktion von Texeln mit unterschiedlicher Größe verwendet das Verfahren in [Wen & Fryer 91] ebenfalls einen Satz von LoG-Filtern. Anstatt kreisförmiger Scheiben wie in [Blostein & Ahuja 89b] werden lineare Balken (engl. *bar*) als Texel detektiert. Ein Grauwertbild wird zunächst mit 11 LoG-Filtern (3×3 , 5×5 , ..., 23×23) gefaltet und anschließend werden die Maxima bzw. Minima der Faltungsantworten in jedem Bild bestimmt. Wenn sich ein Maximum bzw. Minimum an derselben Stelle in zwei gefilterten Bildern befindet und die Differenz der Filtergröße von zwei Bildern mehr als 8 Pixel beträgt, wird das kleinere σ als das für einen Balken an diesem Punkt geeignetere ausgewählt. Wenn die Punkte mit einem Abstand von bis zu $2\sqrt{3}\sigma$ um ein Maximum bzw. ein Minimum eine Faltungsantwort größer bzw. kleiner als Null haben, werden sie als die Punkte von hellen bzw. dunklen Balken betrachtet. Die Statistiken der detektierten Balken sollen zur Segmentation von realen Bildern eingesetzt werden.

4.3.3 Extraktion der Texturen aus Texeln

Deterministische Texturen können auch durch die Bestimmung von Aufbauregeln und die anschließende Extraktion von Texeln analysiert werden. Im folgenden sollen Verfahren mit dieser Vorgehensweise zur Extraktion von Strukturen aus Texeln vorgestellt werden.

⁶LoG: „Laplacian of Gaussian“.

Syntaktisches Verfahren

[Lu & Fu 78] verwenden ein syntaktisches Verfahren für die Beschreibung struktureller Texturen. Als Texel wird dort ein einzelnes Pixel definiert. Die Aufbauregeln lassen sich mit Grammatiken beschreiben. Ein Bild wird zunächst in einige Ausschnitte geteilt. In jedem Ausschnitt läßt sich ein Baum von Pixeln aufbauen. Dieser Baum wird dann mit einem durch eine vorher definierte Grammatik erzeugten Baum verglichen. Das Verfahren wurde nur mit Binärbildern erprobt. Die Beschreibung realer Texturen mit einer solchen Grammatik ist sehr kompliziert und unpraktisch, wie auch die Autoren zugeben.

GPM-Basierte Verfahren

[Conners & Harlow 80b] präsentieren ein Verfahren zur Analyse deterministischer Texturen durch die Auswertung der GPM. Anstatt der Berechnung von Statistiken werden Periode und Größe von Texturen aus der GPM extrahiert. Eine kleine Periode weist darauf hin, daß die Texturen schnell wiederholte feine Strukturen sind. Die Autoren erweitern das Kachelmodell von [Zucker 76a] mit der Einführung von mathematischen Definitionen. Das Verfahren ist sehr aufwendig. Es ist geeigneter für periodische Texturen.

Die oben vorgestellten Verfahren werden realisiert durch die globale Analyse eines Bildes. Sie wurden erfolgreich verwendet bei der Extraktion wiederholter Texturen in kleinen Bildausschnitten mit denselben Texeln. Sie haben aber Probleme mit Texturen realer Szenen. Die Gründe dafür sind: die Texturen realer Szenen enthalten mehr als eine Art von Texeln, und die Frequenz von Texturen ändert sich innerhalb einer Szene.

4.4 Multikanalalanalyse von Texturen

Zu Beginn der siebziger Jahre vermuteten Psychophysiker, daß die Unterscheidung der Stimuli im visuellen Wahrnehmungssystem lokal in einigen voneinander quasi-unabhängigen Kanälen stattfinden könnte. Ausgehend von dieser Spekulation wurde versucht, texturierte Bilder in mehrere Ortsfrequenzbänder zu zerlegen. Die multikanalbasierte Texturanalyse besteht dabei im allgemeinen aus folgenden Schritten:

- Charakterisierung der Kanäle und Bestimmung der entsprechenden Filter,
- Extraktion von Texturmerkmalen aus jedem gefilterten Bild und
- Integration der Texturmerkmale unterschiedlicher Kanäle für die Bildklassifikation.

[Faugeras 78] stellte ein Modell des menschlichen Wahrnehmungssystems für die Texturanalyse vor. Entsprechend diesem Modell verwenden [Coggins & Jain 85] lokale Spektrumenergie im Ortsfrequenzraum als Texturmerkmale für die Texturklassifikation. Zur Extraktion der Texturinformation unterschiedlicher Kanäle werden 11 Filter, bestehend aus 4 Orientierungskanälen und 7 Ortsfrequenzkanälen, eingesetzt. Die Filteroperationen

werden realisiert durch die Multiplikation der Filterkoeffizienten mit der Fouriertransformierten eines Grauwertbildes. Die lokale Spektrumenergie für jeden Kanal wird berechnet aus dem Grauerthistogramm eines Ausschnittes des gefilterten Bildes im Ortsfrequenzraum. Es wird dann ein Merkmalsraum mit 11 Texturmerkmalen für die Klassifikation aufgebaut. Der Unterschied der Spektrumenergie zu der Texturenergie in [Laws 80] ist, daß er die Filter nur aufgrund einer heuristischen Überlegung aufstellt, während die Multikanal-Filter nach einem menschlichen Texturwahrnehmungsmodell entworfen werden.

Die am häufigsten verwendeten Multikanalfilter sind Gaborfilter. Eine ausführliche Darstellung über die Anwendung der Gaborfilter in der Bildverarbeitung ist bei [Porat & Zeevi 88] zu finden. Im folgenden sollen nur die Anwendungen der Gaborfilter zur Texturanalyse vorgestellt werden.

[Perry & Lowe 89] verwenden eine modifizierte Gaborfunktion als Multikanalfilter. Damit haben die Filter unterschiedliche Auflösung im Orts- und Ortsfrequenzraum. Dabei werden insgesamt 48 Filter (3 Frequenz-, 8 Orientierungs- und 2 Phasenfilter) verwendet. Aus den Filterantworten werden die Filtergröße mit der maximalen Summation der Antworten in 8 Orientierungen sowie die Orientierung mit der maximalen Summation der Antworten in unterschiedlichen Filtergrößen selektiert. Als Texturmerkmale werden die Filterantworten mit der optimalen Filtergröße und Orientierung selektiert. Ausgehend von diesen Merkmalen wird die Bildsegmentation durch ein Bereichswachstumsverfahren erzielt.

[Bovik *et al.* 90] präsentieren ebenfalls ein Verfahren zur Texturanalyse in Multikanälen. Unter der Annahme, daß unterschiedliche Texturen signifikante Differenzen in ihrer dominanten Frequenz haben, wird ein Bild in mehrere enge Ortsfrequenz- und Orientierungskanäle kodiert. Als Multikanal-Filter werden 2D-Gaborfilter, die allgemein bekanntesten Quadraturfilterpaare, verwendet. Durch einen Vergleich der Amplitudenantworten unterschiedlicher Kanäle wird die Begrenzung zwischen unterschiedlichen Texturen detektiert. Die Diskontinuität der Texturphasen wird durch die Lokalisation der größten Änderung der Kanalphasenantworten bestimmt. Dabei besteht die Schwierigkeit, die richtige Anzahl der Filterpaare in einer solchen Quadraturfiltermethode auszuwählen.

[Jain & Farrokhnia 91] verwenden eine Filterbank von gerade-symmetrischen Gaborfiltern zur Gewinnung mehrerer Kanäle eines Bildes. Basierend auf der Rekonstruktion des Originalbildes aus den gefilterten Bildern wird eine systematische Vorgehensweise für die Auswahl der Filter benutzt. Als Texturmerkmal wird die Texturenergie in lokaler Umgebung eines Punktes in jedem gefilterten Bild berechnet. Die Texturenergien aus allen Kanälen bilden einen Merkmalsraum für einen Klassifikator. Die Besonderheit dieses Verfahrens ist die Aufstellung einer Methode für die Auswahl der Filter.

4.5 Schwerpunkte für weitere Forschungen zur Texturanalyse

Texturanalyse ist eines der grundlegenden Themen der Bildauswertung. Texturen können sehr viel Information über die Eigenschaften von Objektoberflächen enthalten. Wir haben in diesem Kapitel sowohl Theorien zur Texturwahrnehmung beim Menschen als auch Verfahren zur maschinellen Texturanalyse beschrieben. Als Schlußfolgerung sind im folgenden einige vom Autor als wichtig angesehene Schwerpunkte für weitere Forschungen zur Texturanalyse aufgezählt.

Klassische Statistische Verfahren

In Hinsicht auf klassische Statistiken wird festgestellt, daß die Statistiken aus den Paarhäufigkeitsmatrizen (PM) von Grauwerten oder Kanten die zur Zeit besten Texturmerkmale für die Texturklassifizierung darstellen. Die Statistiken der GPM sind bereits in zahlreichen Arbeiten theoretisch detailliert analysiert worden. Die Anwendungen der aus der GPM gewonnenen Statistiken auf reale Probleme sollten näher untersucht werden. Ein inhärenter Nachteil der GPM ist, daß der Aufbau für alle Richtungen und Abstände sehr zeitaufwendig ist und viel Speicherplatz braucht. Die Vereinfachung der Berechnungen beim Aufbau der GPM und die Reduzierung des Speicherbedarfs sowie ihre Implementierung als Spezielschaltung sollten weiter untersucht werden.

Fraktale Beschreibung

Die Fraktale Geometrie beinhaltet eine Familie von mathematischen Funktionen für die Beschreibung natürlicher Phänomene und Formen. Als Parameter wird oft die fraktale Dimension zur Modellierung komplexer natürlicher Objekte verwendet. Im Gegensatz dazu wird in der Bildauswertung noch wenig untersucht, wie die fraktale Dimension aus einem realen Bild zu extrahieren ist. Als eine vielversprechende Methode sollten die fraktale Beschreibung und fraktalbasierte Verfahren zur Texturanalyse weiter untersucht werden.

Multikanal- und Multiauflösungsdetektion von Texeln

Die meisten in der Fachliteratur beschriebenen strukturellen Verfahren bestehen aus der Extraktion von Texeln und der Bestimmung von Aufbauregeln. Im Bild einer natürlichen Szene sind Texel komplizierter. Sie können von unterschiedlicher Größe oder teilweise verdeckt sein. Dementsprechend sollten Texel mit Multikanalfiltern bzw. Filtern verschiedener Größe extrahiert werden. Andererseits können die Texel durch die Aufnahme mit einer Kamera perspektivisch deformiert werden. Solche Probleme erschweren die Bestimmung von Aufbauregeln. Im Gegensatz dazu sind Statistiken von Texeln hier ein besseres Mittel. Aus diesen Überlegungen ziehen wir die Konsequenz, daß die Multikanalanalyse von Texeln noch weiter untersucht werden muß.

Analyse gerichteter Texturen

Gerichtete Texturen kommen oft in Bildern natürlicher Szenen vor und stellen eine wichtige Klasse von Texturen dar. Sie zeichnen sich durch die lokale Selektivität ihrer Orientierung aus. In der Literatur finden sich noch sehr wenige Arbeiten in dieser Richtung. Im nächsten Kapitel stellen wir ein neues Verfahren zur Analyse gerichteter Texturen dar.

Anwendung der bereits entwickelten Verfahren

Es treten zwei unterschiedliche Situationen bei Anwendungen der Texturanalyse auf. In der ersten Situation existiert nur eine Art von Texturen in einem Bild. Die Aufgabe ist, die Textur in eine der bekannten Kategorien zu klassifizieren. In der zweiten Situation bestehen die Texturen in einem Bild aus mehreren Texturarten. Die Aufgabe hier besteht darin, Regionen mit derselben Textur im Bild voneinander zu segmentieren oder die Begrenzung zwischen unterschiedlichen Texturen zu finden. Zur Texturanalyse wird meistens die Klassifikation der Texturen durchgeführt. Die Bildsegmentation durch die Auswertung von Texturen sollte berücksichtigt werden.

Kapitel 5

Analyse gerichteter Texturen

Gerichtete Texturen kommen oft in Bildern natürlicher Szenen vor. Sie bilden eine wichtige Klasse von Texturen. Gerichtete Texturen werden durch eine lokale Ausrichtung von Textur-elementen gekennzeichnet. Die Untersuchungen in der Psychophysik zeigen, daß das Wahrnehmungssystem beim Menschen besondere Orientierungsselektionskanäle besitzt. Im Forschungsgebiet der Bildauswertung wird erst in letzter Zeit erprobt, wie diese Aspekte bei der maschinellen Auswertung gerichteter Texturen auszunutzen sind.

Die lokale Orientierung gerichteter Texturen kann sich zufällig in einem Bild ändern. Sie enthält wichtige Informationen über Texturen. Wenn eine gerichtete Textur vorliegt, ist deren Orientierung ein wesentliches Merkmal. Durch diese Überlegungen wird deutlich, daß die Ermittlung der lokalen Orientierung von Texturen ein erster Schritt zur Beschreibung gerichteter Texturen ist. Darüber hinaus ist die lokale Orientierung der Bewegung zwischen Objekten und Kameras ein entscheidender Begriff in der Bewegungsanalyse bei Bildfolgen.

Die Grauwertänderung gerichteter Texturen ist stark anisotrop. Die lokale Richtung der Texturen gibt lediglich an, in welcher Richtung sich der Grauwert der Textur relativ schwach ändert. Sie enthält keinerlei Aussage darüber, wie schwach die Grauwertänderung in dieser Richtung im Verhältnis zur Grauwertänderung in der dazu senkrechten Richtung eigentlich ist. Deshalb ist die Stärke der Anisotropie von Texturen ebenfalls ein entscheidender Faktor für die Texturbeschreibung.

Zur Beschreibung gerichteter Texturen verwenden wir in der vorliegenden Arbeit das Orientierungsfeld eines texturierten Bildes. Das Orientierungsfeld von Texturen wird in Anlehnung an das Verschiebungsfeld der Bewegung in Bildfolgen ([Nagel 83]) benannt. Jeder Punkt eines Verschiebungsfeldes besteht aus zwei Elementen: Dem Betrag und der Richtung der Bewegung. Jeder Punkt im Orientierungsfeld eines Bildes ist ein Vektor mit zwei Komponenten: Die erste Komponente ist die lokale Orientierung der Textur an diesem Punkt, die zweite Komponente beschreibt die Stärke der Anisotropie der Textur. Die Stärke der Anisotropie von Texturen wird in dieser Arbeit mit Texturanisotropie abgekürzt.

Die Auswertung der Orientierung von gerichteten Texturen wird erst in letzter Zeit berücksichtigt. Die erste Arbeit in dieser Richtung stammt von [Kass & Witkin 87].

Die Autoren beschreiben gerichtete Texturen mit sogenannten Flußfeldern. Dort wird ein Grauwertbild zunächst mit einem gerichteten Bandpaßfilter gefaltet. Es wird dann angenommen, daß der Mittelwert der Antworten von Bandpaßfiltern null ist. Unter dieser Voraussetzung läßt sich die lokal dominante Richtung aus den Mittelwerten der Varianzen der Filterantworten in einem Fenster ermitteln. Die Kohärenz der Texturen wird als der Mittelwert des Gradientenbetrages definiert. Ein Fluß-Koordinatensystem wird in diesem Verfahren zur Beschreibung der Texturanisotropie gebildet.

Zur Analyse von gerichteten Texturen berechnen [Rao & Schunck 89] Orientierungsfelder aus den Grauwertgradienten eines texturierten Bildes. Das Orientierungsfeld besteht aus zwei Bildern: Einem Orientierungsbild und einem Kohärenzbild. Diese Idee geht eigentlich auf [Kass & Witkin 87] zurück. Das komplette Verfahren besteht aus den folgenden fünf Schritten: Die Faltung eines Bildes mit den x - und y -Ableitungen der Gaußfunktion, die Berechnung der Gradientenrichtung, die Berechnung des Gradientenbetrages, die Schätzung der lokalen Richtung durch die Mittelwertbildung der lokalen Gradientenrichtungen und die Schätzung der Kohärenz der Texturen aus den Gradientenbeträgen.

Wir stellen in diesem Kapitel ein neues Verfahren zur Ermittlung von Orientierungsfeldern vor. Das neuentwickelte Verfahren berechnet zunächst die beiden ersten Ableitungen der Bildfunktion in x - und y -Richtung. Darauf folgend werden ihre Kovarianzen in einer lokalen Umgebung berechnet. Mit den Kovarianzen bzw. den Varianzen läßt sich die Kovarianzmatrix bilden und deren Hauptachsen ermitteln. Die Richtung der Hauptachse mit dem kleineren Eigenwert stellt die Richtung der Textur dar. Die beiden Eigenwerte beschreiben die Grauwertänderungen in den Richtungen der Hauptachsen. Daraus leiten wir ein Maß für die Stärke der Texturanisotropie ab.

Die weitere Auswertung von Orientierungsfeldern richtet sich nach dem Ziel der Bildauswertung. Dabei unterscheidet man zwei Aufgabenstellungen. Bei der ersten Aufgabenstellung wird die Grenzlinie zwischen zwei gerichteten Texturen gesucht. Im zweiten Fall geht es darum, eine symbolische Beschreibung der Texturen zu extrahieren und diese Beschreibung für die weiteren Anwendungen zu übernehmen. In diesem Kapitel wird zunächst die Ermittlung der Orientierungsfelder von Texturen behandelt.

5.1 Orientierungsfelder gerichteter Texturen

5.1.1 Detektion der Grauwertänderungen

Zur Extraktion eines Orientierungsfeldes betrachten wir zunächst die Detektion von Grauwertänderungen. Dazu bieten sich zwei Möglichkeiten. Zum einen kann die erste Ableitung der Bildfunktion berechnet werden. Damit wird der Anteil der hohen Ortsfrequenzen in der Bildfunktion extrahiert. Eine zweite Methode besteht darin, die zweite Ableitung der Bildfunktion zu berechnen. Das entspricht der Filterung eines bestimmten Ortsfrequenzbereiches aus einem Bild. Es ist eine allgemeine Konvention, die Ableitungen der Bildfunktion durch eine Faltung der Bildfunktion mit den Ableitungen der Gaußfunktion zu approximieren. Der mathematische Grund liegt darin, daß die Faltungsoperation und

die Ableitung einer Funktion miteinander vertauschbar sind. Mathematisch ausgedrückt:

$$(g(x, y) * I(x, y))^{(n)} = g(x, y)^{(n)} * I(x, y),$$

dabei ist $I(x, y)$ die Bildfunktion und $g(x, y)$ die Gaußfunktion, der hochgestellte geklammerte Index bezeichnet die Ordnung der Ableitung.

Gleichung (5.1) gibt eine allgemeine Gaußfunktion mit vektorwertigen Koordinaten wieder:

$$g(\vec{x}, \vec{x}_0, \Sigma) = \frac{1}{2\pi |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\vec{x}-\vec{x}_0)^T \Sigma^{-1} (\vec{x}-\vec{x}_0)}, \quad (5.1)$$

dabei sind $\vec{x}_0 = (x_0, y_0)^T$ ein Vektor, der den Zentralwert der Gaußfunktion beschreibt und Σ die Kovarianzmatrix mit:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}. \quad (5.2)$$

Die Größe und Orientierung der Operatoren werden durch die Kovarianzmatrix bestimmt. Vergrößern wir die Varianzen und Kovarianzen, so vergrößert sich die Größe der Operatoren entsprechend.

Wir betrachten zuerst die isotrope Gaußfunktion: wir setzen für den Vektor $\vec{x}_0 = (x_0, y_0)^T = (0, 0)^T$ und für die Kovarianzmatrix $\sigma_{11} = \sigma_{22} = \sigma^2$ sowie $\sigma_{12} = \sigma_{21} = 0$ ein.

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2} \frac{x^2+y^2}{\sigma^2}} \quad (5.3)$$

Im folgenden wollen wir nur die isotrope Gaußfunktion (5.3) benutzen.

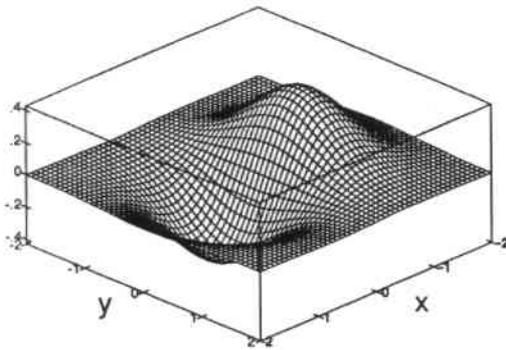
Zur Detektion der lokalen Grauwertänderungen benutzen wir in diesem Kapitel die beiden ersten Ableitungen der Gaußfunktion in x- und y-Richtung:

$$\frac{\partial g(x, y)}{\partial x} = \frac{-x}{\sigma^2} \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2} \frac{x^2+y^2}{\sigma^2}}, \quad (5.4)$$

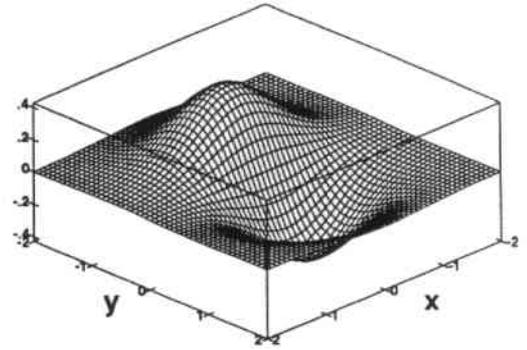
$$\frac{\partial g(x, y)}{\partial y} = \frac{-y}{\sigma^2} \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2} \frac{x^2+y^2}{\sigma^2}}. \quad (5.5)$$

Zur Veranschaulichung werden die beiden Operatoren in Abb. 5.1 für $\sigma = 1$ dargestellt. Diese beiden Operatoren werden oft bei der Detektion von Grauwertkantenelementen eingesetzt. Die Detektion von Kantenelementen im Bild ist eine der grundlegenden Methoden zur Extraktion lokaler Strukturen in der Bildauswertung. Zur Lösung des Problems werden oft die Grauwertgradienten eines Bildes berechnet. Sei $f(x, y) = g(x, y) * I(x, y)$. Die Steigung der Kante ergibt sich aus dem Betrag des Gradienten

$$\|\nabla\| = \sqrt{\left(\frac{\partial f(x, y)}{\partial x}\right)^2 + \left(\frac{\partial f(x, y)}{\partial y}\right)^2} \quad (5.6)$$



(a) Die erste Ableitung der Gaußfunktion in x-Richtung.



(b) Die erste Ableitung der Gaußfunktion in y-Richtung.

Abbildung 5.1: Die ersten Ableitungen der Gaußfunktion in x- und y-Richtung.

und die Richtung der Kantensteigung zu

$$\alpha = \arctan \left(\frac{\partial f(x, y) / \partial y}{\partial f(x, y) / \partial x} \right). \quad (5.7)$$

Das lokale Maximum bzw. Minimum von Gradienten in der Gradientenrichtung stellt die Position des Kantenelementes dar. Für eine ausführliche Behandlung der Kantendetektion durch die Auswertung des Grauwertgradienten sei auf [Canny 86] und [Korn 88] verwiesen.

Der Grauwertgradient enthält im wesentlichen Informationen über die Grauwertänderung an einem Punkt. In unserer Betrachtung besteht eine Textur aber aus mehreren Punkten. Mit der Gradientenrichtung läßt sich daher die Orientierung gerichteter Texturen nicht zuverlässig darstellen. Außerdem enthält der Grauwertgradient keine Informationen über die Anisotropie der Grauwertänderungen von Texturen. Zur Beschreibung von Texturen benötigt man daher Grauwertänderungen innerhalb eines größeren lokalen Bereiches. Nicht nur Texturorientierung, sondern auch eine Angabe zur Stärke der Texturanisotropie ist für die Beschreibung von Texturen notwendig.

Die soeben gemachte Aussage über die Unterschiede zwischen der Kantendetektion und der Bestimmung des Texturorientierungsfeldes kann auch so formuliert werden, daß bei der Kantendetektion die Position exakt lokalisiert werden muß, während bei Orientierungsfelddetektion Stärke der Anisotropie und Orientierung der Texturen bestimmt werden müssen.

5.1.2 Kovarianzmatrix der Grauwertänderungen

Die Kovarianz zweier Variablen enthält eine der wichtigsten statistischen Aussagen über die Verknüpfung dieser Variablen. Sie läßt sich durch mehrere Messungen unter gleichen Bedingungen abschätzen. Die Kovarianz zwischen zwei verschiedenen Variablen ist null,

wenn es keine Beziehung zwischen beiden Variablen gibt, d.h. wenn sie völlig unabhängig voneinander sind. Die Kovarianz für zwei gleiche Variable ergibt dann die Varianz für diese Variable.

Zum Studium der Grauwertänderungen von Texturen untersuchen wir die Kovarianz der beiden Ableitungen der Bildfunktion: $f_x = \frac{\partial f}{\partial x}$ und $f_y = \frac{\partial f}{\partial y}$. Eine Abschätzung des Mittelwertes erhalten wir, wenn wir eine Anzahl von Bildpunkten innerhalb eines lokalen Ausschnittes in Betracht ziehen, d.h. f_x und f_y an benachbarten Punkten berechnen und deren Mittelwert bilden:

$$\mu_x = \sum_{i=-n}^n \sum_{j=-n}^n w(i, j) f_x(x + i, y + j), \quad (5.8)$$

$$\mu_y = \sum_{i=-n}^n \sum_{j=-n}^n w(i, j) f_y(x + i, y + j). \quad (5.9)$$

Dabei sind $w(i, j)$ die Gewichtungsfunktion und $2n+1$ die Breite des Integrationsfensters. In diesem Kapitel nehmen wir $w(i, j)$ wie folgt an:

$$w(i, j) = \frac{1}{(2n+1)^2}.$$

Damit sind die Varianzen und Kovarianzen der Funktionen f_x und f_y definiert als:

$$\sigma_{xx} = \sum_{i=-n}^n \sum_{j=-n}^n w(i, j) f_x^2(x + i, y + j) - \mu_x^2, \quad (5.10)$$

$$\sigma_{xy} = \sigma_{yx} = \sum_{i=-n}^n \sum_{j=-n}^n w(i, j) f_x(x + i, y + j) f_y(x + i, y + j) - \mu_x \mu_y \quad (5.11)$$

und

$$\sigma_{yy} = \sum_{i=-n}^n \sum_{j=-n}^n w(i, j) f_y^2(x + i, y + j) - \mu_y^2. \quad (5.12)$$

Mit den berechneten Varianzen und Kovarianzen läßt sich die Kovarianzmatrix bilden als:

$$\Sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix}. \quad (5.13)$$

Die Kovarianzmatrix Σ in (5.13) ist stets symmetrisch und positiv semi-definit.

Durch die in Anhang A beschriebene Hauptachsentransformation für die Matrix (5.13) erhalten wir

$$\Sigma' = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \quad (5.14)$$

Dabei sind λ_1 und λ_2 die Eigenwerte der Matrix. Wir sehen sofort, daß die Funktionen f_x und f_y im Hauptachsensystem unkorreliert sind. Die Hauptachsentransformation einer Matrix kann auch als eine Drehung des Koordinatensystems betrachtet werden. Das alte

Koordinatensystem wird so gedreht, daß seine neuen Achsen genauso ausgerichtet sind wie die Hauptachsen der Matrix. Die Berechnung der Richtung der Hauptachse einer Matrix ist ebenfalls in Anhang A beschrieben. Wir erhalten hier als Ergebnis:

$$\lambda_{1,2} = \frac{1}{2} \left(\sigma_{xx} + \sigma_{yy} \pm \sqrt{(\sigma_{xx} - \sigma_{yy})^2 + 4(\sigma_{xy})^2} \right) \quad (5.15)$$

und

$$\theta = \frac{1}{2} \arctan \frac{2\sigma_{xy}}{\sigma_{xx} - \sigma_{yy}}. \quad (5.16)$$

Die Richtung θ der Hauptachse stellt die Richtung der stärksten Grauwertänderung von Texturen dar.

Um den Zusammenhang zwischen der Grauwertänderung und der Größe der beiden Eigenwerte zu verdeutlichen, betrachten wir einige einfache, aber illustrative Beispiele von Grauwertübergängen in Abb. 5.2.

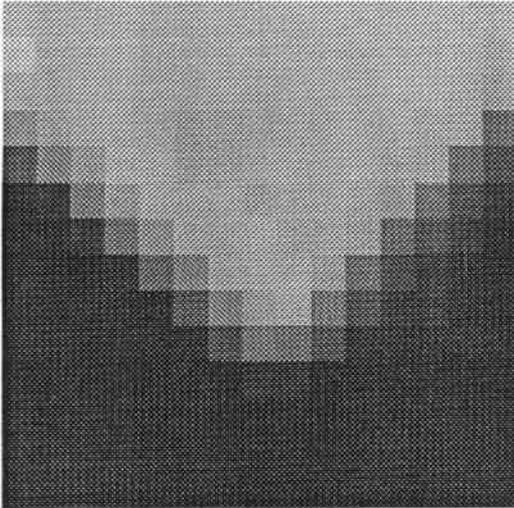
In Abb. 5.2 werden zwei Bildausschnitte mit einer Größe von 15×15 Pixeln dargestellt. In Abb. 5.2 (a) und (c) befindet sich eine Grauwertecke bzw. eine Grauwertkante. Bei der Berechnung der ersten Ableitungen der Bildfunktion haben wir in (5.4) und (5.5) $\sigma = 0.8$ gewählt. Mit diesem Wert beträgt die Größe der Filter 5×5 Pixel. Die Berechnung der Varianzen bzw. Kovarianzen wird innerhalb eines 5×5 Fensters durchgeführt.

In Abb. 5.2 (b) und (d) werden die ermittelten Kovarianzmatrizen als Ellipsen dargestellt. Dabei ist die Länge der beiden Achsen einer Ellipse den Eigenwerten direkt proportional. Die Richtung des Eigenvektors mit dem größten Eigenwert stellt die Richtung der ersten Hauptachse dar. Bei der Darstellung der Kovarianzmatrizen ist zu beachten, daß die Eigenwerte auf den größten Eigenwert des betrachteten Ausschnitts normiert werden. Auf diese Weise sind die Ellipsen in Bereichen mit annähernd homogenem Grauwert nicht erkennbar, da ihre Fläche sehr klein wird. An der Grauwertkante verhalten sich die Ellipsen wie Linien.

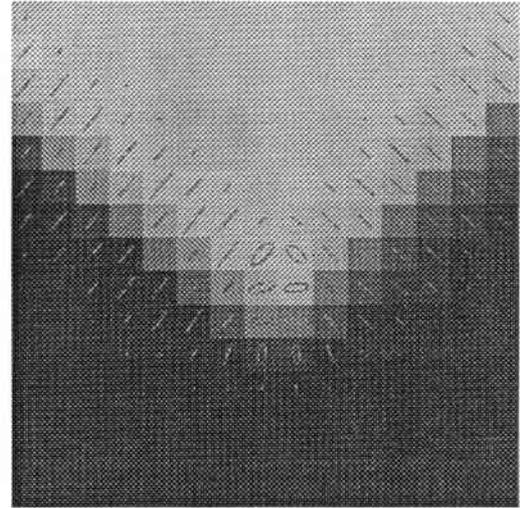
Um die Kovarianzmatrix der Grauwertänderungen quantitativ zu studieren, stellen wir die mit den Gleichungen (5.15) und (5.16) ermittelten Eigenwerte und den Winkel der ersten Hauptachse in Tabelle 5.1 dar. Aus Tabelle 5.1 wird deutlich, daß die Eigenwerte in beiden Richtungen der Hauptachsen sehr unterschiedlich sind. Dieser Sachverhalt weist darauf hin, daß die Texturen in diesem Bildausschnitt sehr anisotrop sind.

Der Zusammenhang zwischen der Grauwertänderung und der Größe der beiden Eigenwerte kann mit folgenden Worten beschrieben werden:

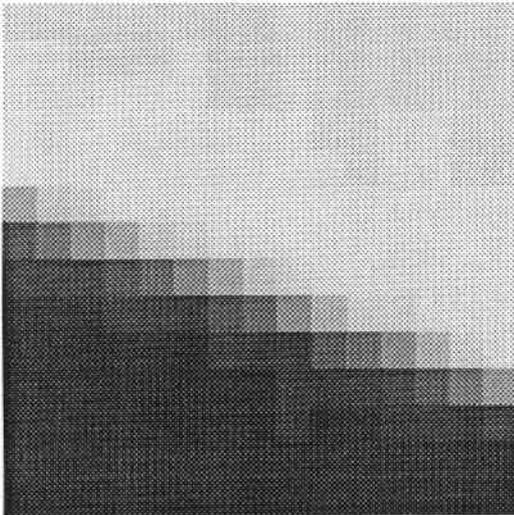
- Beide Eigenwerte sind sehr unterschiedlich: Die Grauwertstruktur weist eine starke Anisotropie auf. Die Orientierung der Grauwertänderung stimmt mit der Richtung des Eigenvektors mit dem größten Eigenwert überein.
- Beide Eigenwerte sind gleich und klein: Der Punkt befindet sich in einem Bereich mit annähernd konstantem Grauwert.
- Beide Eigenwerte sind gleich und groß: Dies weist darauf hin, daß sich hier eine Grauwertecke befindet.



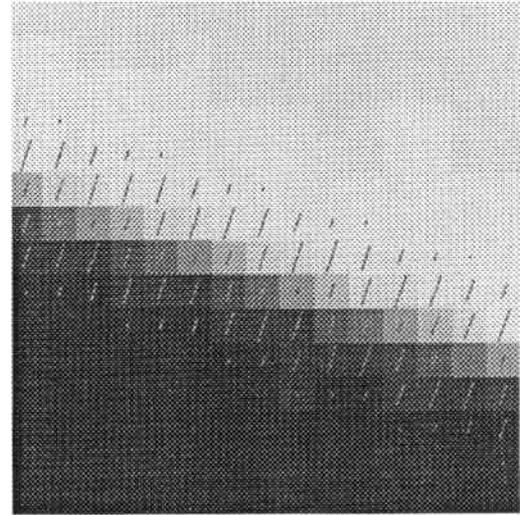
(a) Eine Grauwertecke.



(b) Die Kovarianzmatrizen zu (a).



(c) Eine Grauwertkante.



(d) Die Kovarianzmatrizen zu (c).

Abbildung 5.2: Darstellung der Kovarianzmatrix der Grauwertänderung von zwei elementaren Strukturen als Ellipsen: (a) Eine Grauwertecke, (b) die als Ellipsen dargestellten Kovarianzmatrizen, (c) eine Grauwertkante und (d) die als Ellipsen dargestellten Kovarianzmatrizen an der Grauwertkante. Die Größe der Filter sowie des Fensters bei der Berechnung von Kovarianzen beträgt 5×5 Pixel.

444	53	1	1	2	3	6	8	4	3	2	1	18	201	933
1593	584	85	3	1	4	5	9	5	3	2	17	194	948	1684
2034	1799	779	155	9	5	3	5	4	3	18	179	901	1675	1307
1079	1924	1941	994	231	17	6	5	2	16	185	862	1598	1303	820
1010	929	1701	1999	1189	329	30	4	14	178	876	1547	1238	807	1203
1724	1181	855	1557	2114	1498	516	58	175	884	1606	1226	782	1167	1402
1202	1747	1323	827	1398	2187	1817	722	878	1623	1282	798	1195	1388	712
339	1058	1740	1470	841	1188	2097	1946	1621	1315	829	1232	1441	720	161
49	268	924	1686	1669	974	891	1846	1623	715	1226	1464	732	153	14
6	41	223	780	1610	1879	1240	1474	1432	1185	1504	771	154	12	1
6	8	31	147	555	1277	1628	1522	1391	1247	740	176	13	1	0
3	1	2	10	60	255	507	579	510	331	121	16	1	1	1
1	1	1	1	2	13	32	38	30	18	5	1	1	1	3
2	2	1	2	2	3	4	5	1	1	1	0	1	1	2
0	1	1	2	2	1	1	2	0	1	1	1	0	0	1

(a) Der größere Eigenwert.

3	3	1	1	1	3	1	6	2	1	0	0	0	1	0
1	1	1	1	1	3	2	4	2	1	0	0	0	0	0
1	0	1	1	1	0	1	0	1	1	0	0	0	0	0
1	0	0	1	1	2	2	1	1	0	0	0	0	1	0
0	1	1	1	2	3	2	1	2	0	0	1	1	1	0
1	2	2	1	1	1	1	15	2	1	0	1	1	1	1
1	1	0	0	0	0	16	177	98	2	0	0	0	0	0
0	0	0	0	0	5	120	802	810	99	1	0	0	0	0
0	0	0	1	0	10	272	762	760	305	15	0	1	0	0
1	1	4	5	2	9	279	566	585	289	16	0	1	1	0
0	1	4	4	1	3	137	628	587	123	5	0	0	0	0
0	0	0	1	0	1	24	99	90	18	0	0	0	0	0
0	0	0	1	1	1	4	10	7	1	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(b) Der kleinere Eigenwert.

129	129	148	123	25	117	97	158	77	4	4	58	43	43	44
130	129	127	128	57	140	89	146	131	11	28	50	45	44	44
129	129	128	128	125	155	156	164	154	36	48	47	44	44	44
129	128	128	129	130	128	111	102	84	40	45	46	44	44	45
129	129	129	130	130	131	134	105	46	42	43	45	44	44	45
128	129	129	129	129	129	130	137	45	44	44	45	44	45	45
128	129	129	129	128	128	129	133	41	44	44	45	45	45	46
129	129	129	129	128	127	128	132	42	48	44	44	45	45	47
126	127	129	128	127	125	130	163	172	54	44	44	44	44	45
104	118	125	126	126	123	119	174	176	53	45	43	44	44	85
73	89	117	124	124	121	110	96	83	66	50	45	47	28	44
80	74	175	141	127	119	106	95	90	75	58	49	19	19	46
76	87	74	175	126	125	110	98	101	88	64	45	13	49	53
84	95	85	76	72	65	50	38	54	104	88	72	52	53	45
77	86	79	76	70	64	42	29	87	100	95	94	88	59	66

(c) Die Richtung (in Grad) des zum größeren Eigenwert gehörenden Eigenvektors.

Tabelle 5.1: Der größere Eigenwert in (a), der kleinere Eigenwert in (b) und die Richtung (in Grad) des zum größeren Eigenwert gehörenden Eigenvektors in (c) der Kovarianzmatrizen der Grauwertänderungen an der Grauwertecke in Abb. 5.2(a).

In den zwei letztgenannten Fällen können wir den Schluß ziehen, daß die vorliegende Grauwertänderung schwach anisotrop ist. Eine sorgfältige Betrachtung dieser Probleme ist auch Grundvoraussetzung für eine genaue Bestimmung der richtigen Größe und Orientierung von optimalen Filtern. Im Kapitel 6 dieser Arbeit wird dieses Problem behandelt. Dort leiten wir aus diesen Überlegungen ab, wie optimale Filter für die Texturdetektion zu erzeugen sind.

5.1.3 Orientierungsfelder von Texturen

Bisher haben wir nur über die Berechnung der Richtung und der Stärke der Grauwertänderungen von Texturen besprochen. Jetzt untersuchen wir die Frage der Bestimmung der Orientierungsfelder von Texturen. Aus der zu Beginn des Kapitels eingeführten Definition für das Orientierungsfeld von Texturen sind die lokale Orientierung und die Stärke der Anisotropie zu ermitteln. Diese beiden Größen sind eng mit der Kovarianzmatrix verknüpft. Die gesuchte Richtung von Texturen kann einfach aus (5.16) abgeleitet werden. Sie ist die Richtung des Eigenvektors mit dem kleineren Eigenwert. Mathematisch ausgedrückt:

$$\alpha = \theta + \frac{\pi}{2}, \quad (5.17)$$

dabei ist θ in (5.16) definiert.

Die Texturrichtung in (5.17) stellt nur die Richtung der Texturänderungen dar. Sie enthält keinerlei Aussage darüber, wie stark die Anisotropie der Texturen ist. Wir müssen noch die Größe der Texturanisotropie bestimmen, damit wir Texturen mit ähnlicher Richtung, aber unterschiedlichem Kontrast auch unterscheiden können. Dazu kommen wir auf die Eigenwerte λ_1 und λ_2 in Gleichung (5.15) zurück. Diese beiden Eigenwerte geben uns eigentlich direkt an, wie groß die Grauwertänderungen unabhängig voneinander in den Richtungen der Hauptachsen sind. Viel mehr interessiert uns aber, wie das Verhältnis der beiden Eigenwerte aussieht. Aus diesen Überlegungen läßt sich die Stärke der Anisotropie von Texturen definieren als:

$$s = 1 - \frac{\lambda_2}{\lambda_1}, \quad (5.18)$$

dabei ist λ_1 der größere Eigenwert und λ_2 der kleinere. Die Texturanisotropie s hat einen Wert zwischen 0 und 1. Je größer s ist, desto stärker ist die Anisotropie der Texturen.

Bei Gleichung (5.18) müssen wir noch zwei Effekte berücksichtigen. Zum einen ist das Verhältnis der beiden Eigenwerte an einer Grauwertecke gleich dem Verhältnis im Bereich mit homogenem Grauwert. Zum anderen ist s bei starker Grauwertänderung gleich groß wie bei schwacher Grauwertänderung. Als Alternative zur Definition der Stärke der Anisotropie aus (5.18) können die Gradientenbeträge einbezogen werden. Wir erhalten so:

$$s' = \frac{\sqrt{f_x^2 + f_y^2}}{C} \left(1 - \frac{\lambda_2}{\lambda_1} \right), \quad (5.19)$$

wobei C einen Normierungsfaktor bezeichnet.

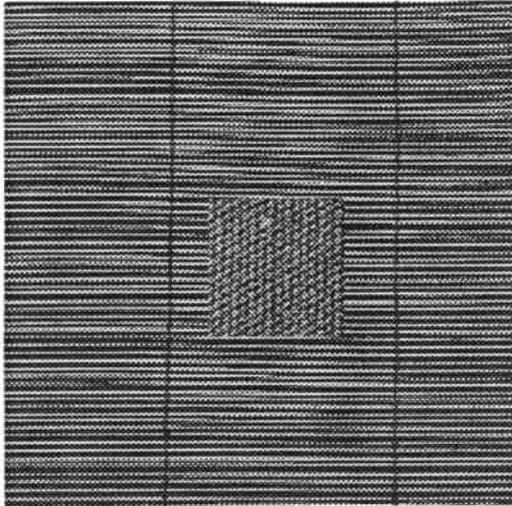
Im folgenden sollen einige Beispiele zur Veranschaulichung der ermittelten Orientierungsfelder dargestellt werden. Das Bild in Abb. 5.3 (a) besteht aus zwei unterschiedlichen Mustern: Strohhalme im äußeren Bereich und eine Baumwollleinwand im inneren Bereich (D49 und D77 aus [Brodatz 66]). Leicht ist zu erkennen, daß die Strohhalme im äußeren Bereich gerichtet angeordnet sind, während die Baumwollleinwand im inneren Bereich unterschiedliche dominante Richtungen aufweist. Die Strohhalme werden noch zusätzlich mit zwei senkrecht stehenden Liniensegmenten geteilt. Das Bild ist ein gutes Beispiel, um die Leistungsfähigkeit unseres Verfahrens bei der Unterscheidung zwischen gerichteten und anderen Texturen zu zeigen. Die Ergebnisse werden in Abb. 5.3 (b), (c) und (d) dargestellt. Dabei beträgt die Größe der Gaußfilter sowie des Fensters bei der Berechnung von Kovarianzen 7×7 Pixel. Die Texturorientierung, die an jedem Punkt berechnet wird, wird in (b) nur an den abgetasteten Punkten durch ein weißes Liniensegment dargestellt. Deutlich ist der Unterschied der Orientierung zwischen beiden Texturen zu erkennen. Wegen des Abtastens sind die Orientierungen der beiden vertikalen Liniensegmente in (b) leider nicht zu sehen. Die Texturanisotropie nach Gleichung (5.18) ist in (c) mit dem Vergrößerungsfaktor 255 hell als Intensitätswert kodiert. Mit diesem Bild wird deutlich, daß die Texturanisotropie im äußeren Bereich stärker als die im inneren Bereich ist.

Die nach (5.19) berechneten Größen s' werden in (d) gezeigt. Es ist einfach zu erkennen, daß s' nach (5.19) einen ganz anderen Wert als s nach (5.18) besitzt. Je nach dem Ziel der Bildauswertung haben beide Definitionen ihre Vor- und Nachteile. Daher werden wir beide benutzen. In (c) und (d) sind die beiden vertikalen Liniensegmente deutlich zu erkennen. Die Bilder in (c) and (d) stehen als Basis für eine Segmentation des Bildes zur Verfügung. Dieses Problem wird im nächsten Abschnitt behandelt werden.

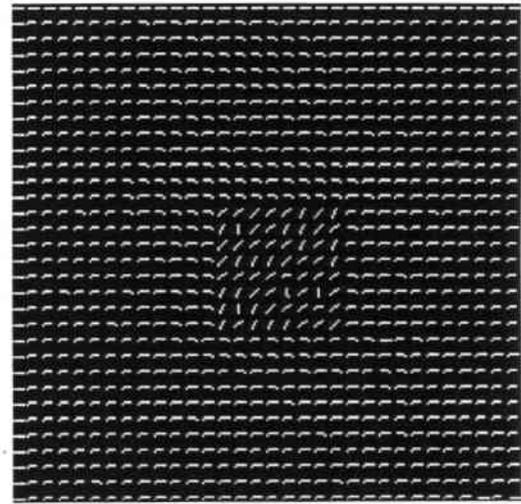
Wir zeigen mit dem Beispiel in Abb. 5.4 die Wichtigkeit der Texturorientierung für die Texturbeschreibung. Abb. 5.4 (a) besteht aus vier Ausschnitten. Sie sind die Rotationen eines Bildausschnittes einer Holzoberfläche (D68 aus [Brodatz 66]). Sie lassen sich nur in der Richtung voneinander unterscheiden. In (b) wird die Texturorientierung dargestellt. Zum Vergleich wird die Texturanisotropie mit dem Gradientenbetrag als Gewicht als Hintergrund auch in (b) dargestellt. Deutlich ist in diesem Bild zu erkennen, daß die Texturanisotropie in allen vier Bereichen fast gleich ist. Dieser Effekt ist auch im Bild der Texturanisotropie ohne Gewicht in (c) zu erkennen. Überlappend wird auch die Texturorientierung auf dem Bild in (b) dargestellt.

In Abb. 5.5 wird das Orientierungsfeld von handgesponnenem Baumwollstoffen (D11 aus [Brodatz 66]) dargestellt. Das Bild besteht aus mehreren vertikalen Streifen. Fünf davon bestehen wiederum aus gerichteten Untermustern. Zur Beschreibung von Texturen soll das Bild zunächst segmentiert werden. In (b) wird das ermittelte Orientierungsfeld gezeigt. Die weißen Liniensegmente in den Bereichen mit Untermustern stellen die Orientierungen der gerichteten Untermuster dar. In (c) haben die Texturanisotropien in den Bereichen mit Untermustern einen größeren Wert als die in anderen Streifen.

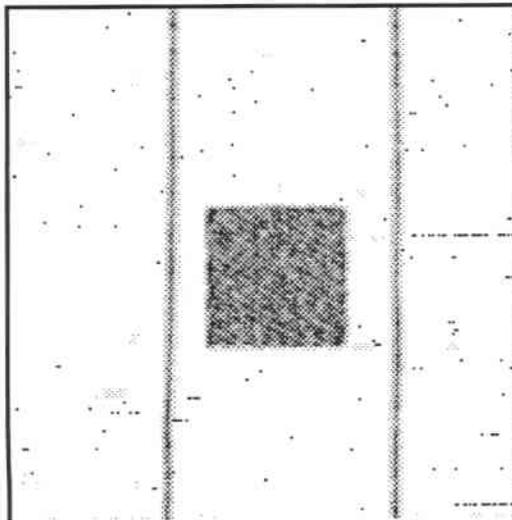
In Abb. 5.6 (a) wird ein Straßenbild dargestellt. Die Verfolgung der Fahrbahnbegrenzungen ist Voraussetzung für eine autonome Fahrzeugführung. Mit diesem Beispiel wollen wir zeigen, daß die Texturanisotropie von gerichteten Grauwertübergängen am Straßenrand eine wichtige Information für die Detektion von Straßenbegrenzungen ist.



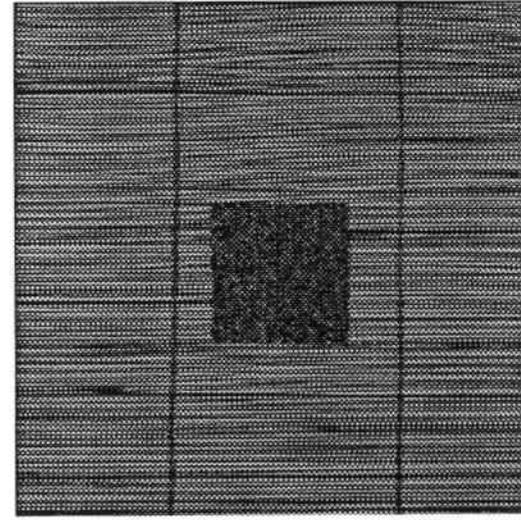
(a) Das Grauwertbild.



(b) Die Texturorientierung.

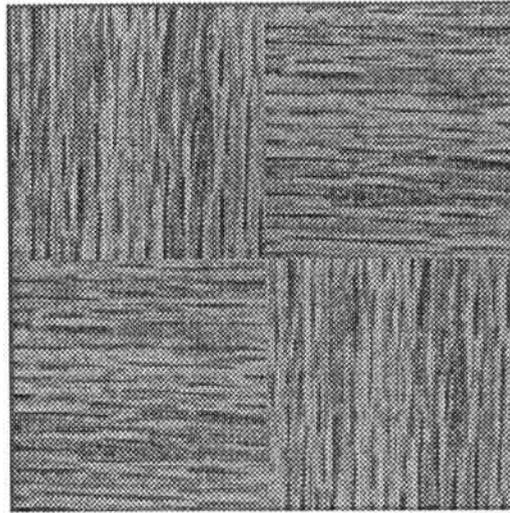


(c) Die Texturanisotropie ohne Gewichtung.

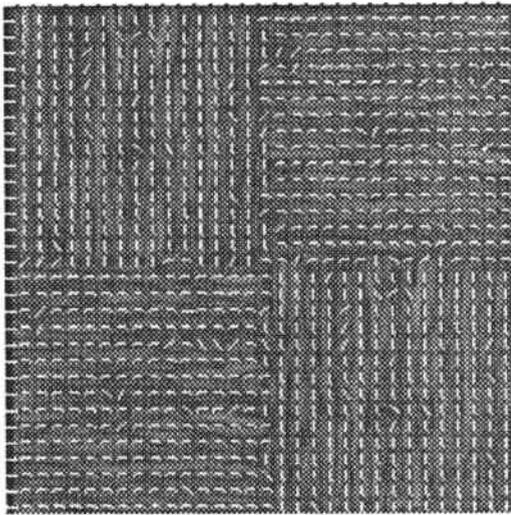


(d) Die Texturanisotropie mit dem Gradientenbetrag als Gewicht.

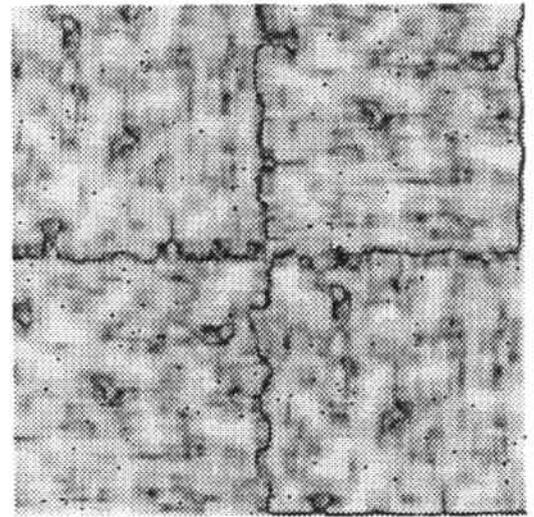
Abbildung 5.3: Darstellung des Orientierungsfeldes von Strohhalmen im äußeren Bereich und Baumwollleinwand im inneren Bereich (D49 und D77 aus [Brodatz 66]). Dargestellt sind im einzelnen: (a) Das Grauwertbild, (b) die lokalen Orientierung, (c) die Texturanisotropie ohne Gewichtung und (d) die Texturanisotropie mit dem Gradientenbetrag als Gewicht. Helle Bereiche entsprechen Bildpositionen mit starker Texturanisotropie. Die Größe der Filter beträgt 7×7 Pixel und die Größe des Fensters bei der Berechnung der Kovarianzen beträgt ebenfalls 7×7 Pixel.



(a) Das Grauwertbild.

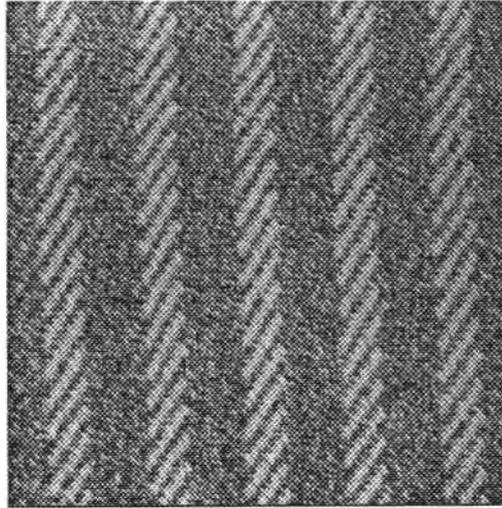


(b) Die Texturorientierung und -anisotropie mit Gewichtung.

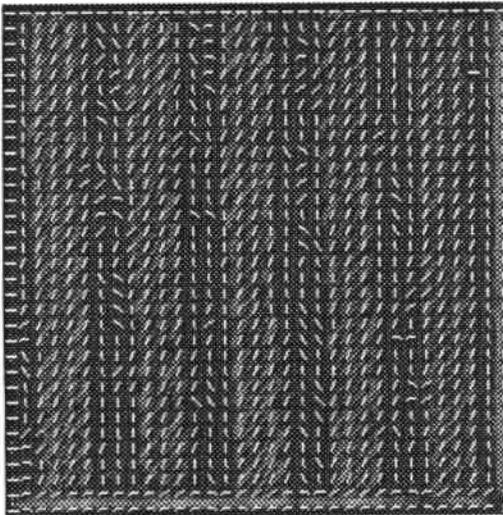


(c) Die Texturanisotropie ohne Gewichtung.

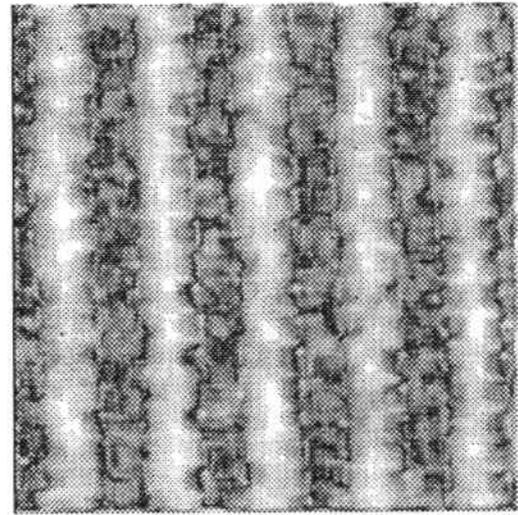
Abbildung 5.4: Darstellung des Orientierungsfeldes einer Holzoberfläche (D68 aus [Brodatz 66]): (a) Das Grauwertbild, (b) die Texturorientierung und -anisotropie mit dem Gradientenbetrag als Gewicht und (c) die Texturanisotropie ohne Gewichtung. Die Größe der Filter beträgt 9×9 Pixel und die Größe des Fensters bei der Berechnung der Kovarianzen beträgt 11×11 Pixel.



(a) Das Grauwertbild.

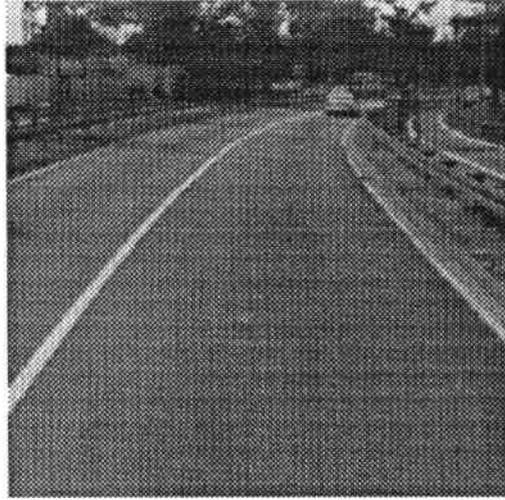


(b) Die Texturorientierung und -anisotropie mit Gewichtung.

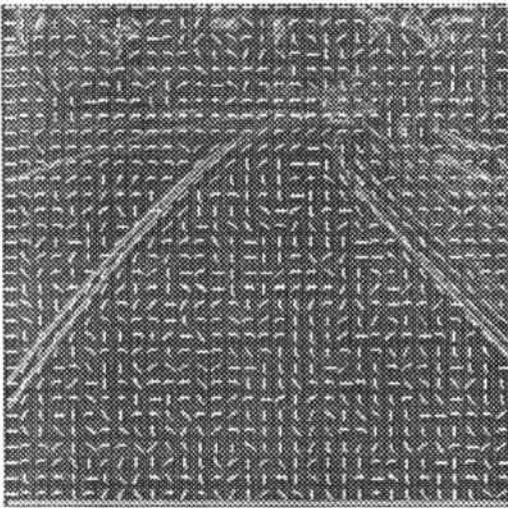


(c) Die Texturanisotropie ohne Gewichtung.

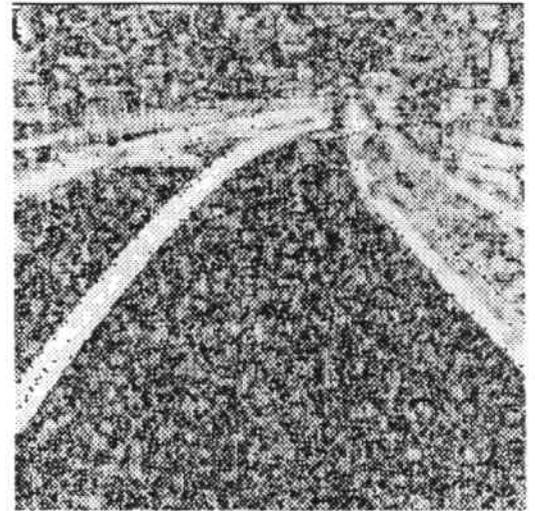
Abbildung 5.5: Darstellung des Orientierungsfeldes handgesponnener Baumwollstoffe (D11 aus [Brodatz 66]): (a) Das Grauwertbild, (b) die Texturorientierung und -anisotropie mit dem Gradientenbetrag als Gewicht und (c) die Texturanisotropie ohne Gewichtung. Die Größe der Filter beträgt 9×9 Pixel und die Größe des Fensters bei der Berechnung der Kovarianzen beträgt 11×11 Pixel.



(a) Das Grauwertbild.



(b) Die Texturorientierung und -anisotropie mit Gewichtung.



(c) Die Texturanisotropie ohne Gewichtung.

Abbildung 5.6: Darstellung des Orientierungsfeldes eines Straßenbildes: (a) Das Grauwertbild, (b) die Texturorientierung und -anisotropie mit dem Gradientenbetrag als Gewicht und (c) die Texturanisotropie ohne Gewichtung. Die Größe der Filter beträgt 9×9 Pixel und die Größe des Fensters bei der Berechnung der Kovarianzen beträgt 11×11 Pixel.

Die Fahrbahnbegrenzungen in (a) bestehen aus weißen Markierungen. Die Fahrbahnbegrenzungen haben eine eindeutig dominantere Orientierung als die Fahrbahnoberfläche. In (b) wird die lokale Orientierung und die Texturanisotropie mit Gewichtung dargestellt. Die Texturanisotropie ohne Gewichtung ist in (c) dargestellt.

5.2 Segmentierung von Orientierungsfeldern

Im vorherigen Abschnitt haben wir die Ermittlung von Texturorientierungsfeldern vorgestellt. Diese waren in der Lage, gerichtete Texturen zu charakterisieren. Die Orientierungsfelder texturierter Bilder enthalten explizitere Informationen über die Texturen als Grauwertbilder.

5.2.1 Grundlage der Bildsegmentation

Die Segmentierung eines Bildes in Bereiche mit homogenen Attributwerten ist eines der Kernprobleme in der Bildauswertung. Um die erforderlichen Kenntnisse über die Bildsegmentation bereitzustellen, sollen hier zunächst die Grundlagen und die bekanntesten Verfahren der Bildsegmentation vorgestellt werden. An dieser Stelle möchten wir darauf hinweisen, daß die Diskussionen über die Bildsegmentierung in diesem Abschnitt als Grundlage nicht nur für die Segmentierung von Texturorientierungsfeldern, sondern auch für die Bildsegmentierungsverfahren in den nachfolgenden zwei Kapiteln gelten. Eine formale Definition der Bildsegmentation nach [Pavlidis 77] wird im folgenden angegeben:

Definition 5.1 (Bildsegmentation) *Unter der Segmentierung eines diskreten Bildes $f(i, j)$ mit $\{0 \leq i \leq M - 1 \cap 0 \leq j \leq N - 1\}$ versteht man die Unterteilung von f in disjunkte, nichtleere Teilmengen f_1, f_2, \dots, f_P so, daß mit einem zu definierenden Einheitlichkeitskriterium E gilt:*

- $\bigcup_{i=1}^P f_i = f$.
- f_i ist zusammenhängend $\forall i$ mit $i = 1, \dots, P$.
- Für alle f_i ist das Einheitlichkeitskriterium $E(f_i)$ erfüllt.
- Für jede Vereinigungsmenge zweier benachbarter f_i, f_j ist $E(f_i \cup f_j)$ nicht erfüllt.

□

Bei der Bildsegmentierung unterscheidet man prinzipiell zwischen kantenorientierten, bereichsorientierten und klassifikationsorientierten Verfahren ([Fu & Mui 81], [Haralick & Shapiro 85]). Die Auswahl eines geeigneten Verfahrens bei Anwendungen richtet sich nach der zu lösenden Aufgabe.

Kantenorientierte Verfahren

Die kantenorientierten Verfahren detektieren zunächst die Kantenelemente zwischen verschiedenen Bereichen. Ausgehend von einem Kantenelementbild werden anschließend

die Grenzpunkte jedes Bereiches bestimmt. Ein solches kantenorientiertes Verfahren besteht im allgemeinen aus den folgenden Schritten: Zunächst wird ein Grauwertbild mit Filter-Masken gefaltet. Je nach Art der Filter sind die Maxima bzw. Minima oder Null-durchgangsstellen im gefalteten Bild zu suchen. Im Prinzip sollten die Begrenzungen zwischen verschiedenen Bereichen mit den bestimmten Konturen vollständig dargestellt werden können. In der Praxis sieht es aber ganz anders aus. In vielen Fällen, insbesondere bei Bildern natürlicher Szenen, können Lücken in den Trennlinien zweier Bereiche oder Kantenelemente innerhalb eines zusammenhängenden Bereiches auftreten. Eine erfolgreiche Segmentierung von Bildern kann nur durch zusätzliche Verarbeitungsschritte, wie z.B. Lückenschließung oder Kantenelement-Eliminierung, erreicht werden. Auf dieses Thema kommen wir noch in Kapitel 7 zurück. Wir präsentieren dort ein Verfahren zur Segmentierung von Straßenbildern durch Verfolgung von Bereichskonturen.

Bereichsorientierte Verfahren

Eine alternative Vorgehensweise zu kantenorientierten Verfahren sind bereichsorientierte Verfahren, bei denen die Nachbarschaftsbeziehungen von Punkten bei der Segmentierung berücksichtigt werden. In einem bereichsorientierten Verfahren läßt sich zunächst ein Kriterium definieren, gemäß dessen die Teilbereiche einheitliche Eigenschaften haben sollen. Als Beispiele sind Grauwert-, Farb- und Texturhomogenität zu nennen. Die Wahl geeigneter Einheitlichkeitseigenschaften hängt von der Natur der zu segmentierenden Bilder sowie von der zu lösenden Aufgabe ab.

Ein typisches bereichsorientiertes Verfahren ist das sogenannte Bereichswachstumsverfahren. Bei diesem Verfahren wird ein Punkt entsprechend der Beziehungen zu seinen Nachbarn einem Unterbereich zugeordnet. Eine entscheidende Voraussetzung für die erfolgreiche Segmentierung verlangt, daß die Bereiche im Sinne der Einheitlichkeitseigenschaft zusammenhängend sind. Ein Bereichswachstumsverfahren arbeitet im allgemeinen iterativ in mehreren Schritten. Ausgehend von geeigneten Anfangspunkten wird ein iterativer Prozeß gestartet, bei dem die benachbarten Punkte mit ähnlichen Eigenschaften demselben Unterbereich zugeordnet werden. Idealerweise sollte für jeden Unterbereich genau ein Anfangspunkt bestimmt werden. Bei diesen Verfahren können einzelne isolierte Punkte oder kleine Unterbereiche entstehen. Sie können dann mit einer nachfolgenden Bereichverschmelzung wieder den Bereichen mit ähnlichen Eigenschaften zugeordnet werden. In der vorliegenden Arbeit werden Bereichswachstumsverfahren nicht mehr behandelt. Für eine ausführliche Diskussion darüber sei auf [Zucker 76b], [Monga 87], [Pavlidis & Liow 90] und [Zhang 91] verwiesen.

Anders als bei Bereichswachstumsverfahren kann auch die Segmentierung von Bildern durch die Auswertung von Histogrammen der Grauwert- oder Merkmalsverteilung realisiert werden. Solche Verfahren werden in der Literatur auch als punktorientierte Verfahren bezeichnet. Bei den histogrammbasierten Verfahren wird vorausgesetzt, daß die Objekte vom Hintergrund mit Hilfe einfacher Schwellwertoperationen getrennt werden können. Darüber hinaus besteht das Kernproblem eines solchen Verfahren in der richtigen Wahl einer oder mehrerer Schwellen ([Otsu 78], [Weszka & Rosenfeld 79], [Sahoo *et*

al. 88], [Kurita *et al.* 92]). Histogrammbasierte Verfahren sind für die Segmentierung von Bildern mit signifikant unterschiedlichen Intensitäten (z.B. Bilder polyedrischer Szenen) geeignet, da sich signifikante maximale Werte im Histogramm solcher Bilder bilden. Für Bilder natürlicher Szenen bietet ein solches Verfahren große Schwierigkeiten bei der Wahl der Schwellen.

Klassifikationsorientierte Verfahren

Allgemein können Schwellenverfahren in N -dimensionalen Histogrammen als Trennung von Objekten durch $(N-1)$ -dimensionale Hyperebenen angesehen werden. Daher können sie wie die Klassifikationsverfahren in der Mustererkennung behandelt werden. Man spricht hier von Bildsegmentation durch Klassifikation oder Bereichsklassifikation. Bei der Segmentierung eines Bildes durch Klassifizierung im Merkmalsraum können verschiedene Informationen über denselben Bereich bei der Merkmalsbildung herangezogen werden. Diese Merkmale spannen einen mehrdimensionalen Raum auf, der als Merkmalsraum bezeichnet wird. Mit dieser Vorgehensweise lassen sich die Bereiche im Bild durch Ballungen im Merkmalsraum darstellen. Die zu beantwortende Fragestellung lautet also: Wie können die Merkmalsvektoren im Merkmalsraum voneinander getrennt werden?

Bei der Bereichsklassifizierung gibt es zwei grundsätzlich verschiedene Situationen. Bei einer ist eine vorklassifizierte Stichprobe von Objekten vorhanden. In der Regel wird auch die Klassifikatorstruktur vorgegeben. Mit einer vorklassifizierten Stichprobe werden anschließend die Parameter von Klassifikatoren ermittelt. Man spricht hier von einem überwacht eingelernten Klassifikator. Bei der anderen Situation sind sowohl die Statistiken der Merkmale jeder Klasse als auch die Anzahl der Klassen unbekannt. Die Klassifikationsverfahren bilden aus den Eingabemustern eine Unterteilung eines Bildes in Bereiche. Die auf diese Weise arbeitenden Verfahren werden als unüberwacht eingelernte Klassifikatoren bezeichnet.

Bildsegmentierung ist eine der schwierigsten und auch wichtigsten Aufgaben in vielen Anwendungsgebieten des Maschinensehens. Der Unterschied in den Grauwerten oder Farbtönen allein ist oft nicht ausreichend für die Segmentierung der Abbildungen komplizierter Szenen. Der Unterschied in den Beziehungen zwischen benachbarten Punkten, nämlich der Unterschied in Texturen, muß im allgemeinen auch mit herangezogen werden. Mit den im vorherigen Abschnitt ermittelten Texturorientierungsfeldern als Texturmerkmale präsentieren wir hier die Textursegmentierung. Bei der Texturanalyse ist die Stichprobe von Texturen wegen der Vielfalt der Texturen schwer zu ermitteln. Unüberwacht eingelernte Klassifikatoren sind daher vorteilhaft für die Textursegmentation. Dieses Problem wird im nachfolgenden Abschnitt anhand der Segmentierung der Texturorientierungsfelder durch einen modifizierten *Isodata*-Algorithmus ausführlich diskutiert.

5.2.2 Klassifikation mit modifiziertem *Isodata*-Algorithmus als Lernverfahren

Ballungsanalyse spielt eine wichtige Rolle bei der unüberwacht eingelernten Mustererkennung und Bildsegmentierung ([Zhang & Wu 87], [Phillips & Rosenfeld 88], [Jolion *et al.* 91], [Chan & Cheung 92], [Pappas 92]). Eine solche Ballungsanalyse ist besonders dann ein hilfreiches Werkzeug für die Musteranalyse, wenn wenig *A-Priori-Informationen* über die zu klassifizierenden Daten zur Verfügung stehen. Die k -Mittelwert-Ballung und der *Isodata*-Algorithmus ([Ball & Hall 65]) sind zwei der bekanntesten Verfahren zur automatischen Gruppierung gegebener Daten in Klassen. Bei der k -Mittelwert-Ballung wird die Anzahl der zu bildenden Klassen vorgegeben ([Coleman & Andrew 79], [Huntsberger *et al.* 85]). Die Zuordnung eines Merkmalsvektors zu einer Klasse erfolgt über die Bestimmung eines minimalen Abstandes. Der *Isodata*-Algorithmus klassifiziert die Bildpunkte in Unterbereiche durch iterative Verbesserungen einer Anfangsklassifikation im Sinne von vorgegebenen Gütekriterien. Der *Isodata*-Algorithmus wird oft als verallgemeinerte k -Mittelwert-Ballung bezeichnet. Er basiert auf der k -Mittelwert-Ballung und zusätzlichen Forderungen an die Qualität der Klassifikationsergebnisse. In diesem Sinne wird die k -Mittelwert-Ballung auch als der *Basic-Isodata*-Algorithmus angesehen. Die Unterschiede bestehen darin, daß beim *Isodata*-Algorithmus weitere heuristische Verfahren nach der Ballung angewendet werden, um die klassifizierten Daten wieder aufzuspalten oder weiter zu verschmelzen.

In Anhang *B* und Anhang *C* werden die k -Mittelwert-Ballung und der originale *Isodata*-Algorithmus beschrieben. Im folgenden präsentieren wir die Modifikation des *Isodata*-Algorithmus für unsere Aufgabenstellung. Dabei werden die möglichen Varianten des konventionellen *Isodata*-Algorithmus ausführlich diskutiert.

Der Merkmalsraum

Die Klassifizierung eines Bildes in unterschiedliche Bereiche kann formuliert werden als die Ballung der Bildpunkte in signifikante Klassen im Merkmalsraum. Die Eingabevektoren eines Klassifikators sind die im Bild extrahierten Bildmerkmale. Die Qualität der Klassifizierung hängt stark davon ab, welche Merkmale ausgewählt worden sind. Mit unterschiedlichen Merkmalen werden auch entsprechende Bildpunkte in unterschiedliche Bereiche klassifiziert. Die Wahl von Merkmalen ist hauptsächlich von der betreffenden Aufgabenstellung abhängig.

Das einfachste Merkmal eines Bildpunktes ist sein Grauwert. Mit ihm wird ein eindimensionaler Merkmalsraum gebildet. Die Klassifizierung der Punkte im Merkmalsraum resultiert in einer Teilung des Histogramms in einige Intervalle. Für ein RGB-Farbbild wird ein dreidimensionaler Merkmalsraum mit

$$\vec{x}_i = (\text{rot}, \text{grün}, \text{blau})^T$$

gebildet. Eine Klasse im Merkmalsraum entspricht Bildpunkten mit ähnlicher Farbe. Sowohl der Grauwert als auch die Farbtöne enthalten keine Informationen bezüglich des Ortes.

Für ein texturiertes Bild ist die Textureigenschaft unbedingt zu berücksichtigen. Für die Aufgabenstellung in diesem Abschnitt, nämlich die Segmentierung von Texturorientierungsfeldern, sollen die im vorangegangenen Abschnitt ermittelte Texturrichtung und -anisotropie als Merkmale ausgenutzt werden. Aus dieser Betrachtung erhalten wir zwei neue Merkmale für ein Grauwertbild:

$$\vec{x}_i = (\text{Grauwert}, \text{Texturrichtung}, \text{Texturanisotropie})^T.$$

Die Bildpunkte, die eine Ballung im Merkmalsraum bilden, müssen im 2D-Bild nicht zusammenhängen. Daher werden oft weitere Verarbeitungsschritte benötigt, um die endgültige Segmentierung zu erzielen. Als Alternative kann man die räumlichen Informationen (beide Koordinaten) direkt bei der Ballung im Merkmalsraum einbeziehen. Mit dieser Vorgehensweise ist es möglich, daß die im Merkmalsraum zu derselbe Klasse gebildeten Punkte in einem Bildbereich ebenfalls zusammenhängend sind. Jeder Punkt im Eingabebild hat 5 Merkmale:

$$\vec{x}_i = (\text{Grauwert}, \text{Texturrichtung}, \text{Texturanisotropie}, \text{Zeile}, \text{Spalte})^T \quad (5.20)$$

oder symbolisch abgekürzt

$$\vec{x}_i = (g_i, \alpha_i, s_i, x_i, y_i)^T. \quad (5.21)$$

Die Einbeziehung der Position in den Ballungsprozeß hat den Vorteil der Eliminierung kleiner Bereiche bzw. einzelner isolierter Punkte. Dabei wird vorausgesetzt, daß Unterbereiche im Bild zusammenhängend sind. Wir sollten hier klar sagen, daß diese Merkmale in (5.20) nur für die Klassifizierung von gerichteten Texturen geeignet sind. Für die Klassifizierung anderer Bilder müssen entsprechend andere Merkmale verwendet werden.

Das Abstandsmaß

Das Grundproblem der automatischen Klassifikation läßt sich so formulieren, daß ähnliche Objekte möglichst zu gleichen, unähnliche Objekte zu verschiedenen Klassen gehören sollen. Es stellt sich dabei die Frage, wie die „Ähnlichkeit“ bzw. „Unähnlichkeit“ (Abstand) zweier Objektbeschreibungen quantitativ definiert werden soll. Man braucht ein Abstandsmaß um festzustellen, welcher unter den bekannten Fällen die größte Übereinstimmung mit dem aktuellen Muster, d.h. den geringsten Abstand dazu hat. Als eines der wichtigsten Probleme diskutieren wir hier, wie der Abstand zweier Objektbeschreibungen bei unserem Klassifikationsverfahren gewählt wird.

Bei der Auswahl eines Abstandsmaßes muß zunächst die reale Bedeutung der Eingabemuster berücksichtigt werden. Wegen der Vielfalt von Mustern soll das Ähnlichkeitsmaß am besten invariant gegenüber Skalenänderungen sein, d.h. das Ergebnis eines Klassifikators soll nicht von der physikalischen Einheit der Eingabemuster abhängig sein. Eigentlich sind Invarianzeigenschaften von Merkmalen wie z.B. Translations-, Rotations- und Skaleninvarianz bei der Mustererkennung immer wünschenswert ([Hu 62], [Li 92]).

Eines der am häufigsten verwendeten Abstandsmaße zweier Objekte ist der Euklid'sche Abstand. Der Euklid'sche Abstand ist nicht skaleninvariant. Deshalb müssen die Eingabemuster, die nicht miteinander vergleichbar groß sind, vor der Ballung normiert werden. Beim Euklid'schen Abstand spielen nur die Klassenzentren eine Rolle. Mit anderen Worten, ein Eingabemuster wird einer Klasse zugeordnet, zu deren Zentrum es den geringsten Abstand hat. Das hat zur Folge, daß die Muster im Merkmalsraum kugelförmige Ballungen bilden.

Eine Alternative zum Euklid'schen Abstand ist das Abstandsmaß beim sogenannten Kosinus-Klassifikator. Im betrachteten Fall wird dabei der Kosinus des Winkels zwischen dem Eingabevektor und einem der Klassenzentrumsvektoren als Ähnlichkeitsmaß gewählt. Mathematisch ausgedrückt ergibt dies:

$$\cos \phi_{ij} = \frac{\vec{x}_i \cdot \vec{z}_j}{\|\vec{x}_i\| \cdot \|\vec{z}_j\|}. \quad (5.22)$$

Der Kosinus-Abstand hat auf der einen Seite den Vorteil im Vergleich zum Euklid'schen Abstand, daß er skaleninvariant ist. Auf der anderen Seite ist an diesem Ähnlichkeitsmaß wie am Euklid'schen Abstand unbefriedigend, daß er nicht translationsinvariant ist.

Wie oben erwähnt, haben wir bei der Merkmalsauswahl beide Koordinaten berücksichtigt. Das bedeutet, daß die resultierenden Unterbereiche im Bild sich in etwa kreisförmig bilden. Das widerspräche den meisten zu bildenden Klassen. Zur Überwindung dieses Problems sollen andere Abstandsmaße verwendet werden, damit ellipsenförmige und sogar angenähert lineare Unterbereiche im Bild entstehen können.

Der sogenannte *Mahalanobis-Abstand* ist ein Abstandsmaß, das unsere Anforderungen erfüllen kann. Er wird definiert als

$$r_{ij}^2 = (\vec{x}_i - \vec{z}_j)^T \Sigma^{-1} (\vec{x}_i - \vec{z}_j). \quad (5.23)$$

Dabei ist Σ eine $s \times s$ Kovarianzmatrix, \vec{x}_i ein Eingabevektor und \vec{z}_j ein Klassenzentrum. Aus (5.23) sehen wir sofort, daß die Entscheidung nicht nur von den Klassenzentren, sondern auch von den Kovarianzen zwischen verschiedenen Werten aller Klassen abhängig ist. Der *Mahalanobis-Abstand* ist invariant bzgl. aller Parallelverschiebungen und allen linearen, nichtsingulären Transformationen der Eingabemuster. Insbesondere ist hier die Skaleninvarianz gesichert. In Anhang D werden die Eigenschaften des *Mahalanobis-Abstandes* und eine Vereinfachung seiner Berechnung beschrieben. Die vereinfachte Form ist

$$r_{ij}^2 = \vec{z}_j^T \Sigma^{-1} \vec{z}_j - \vec{x}_i^T \vec{w}_j, \quad (5.24)$$

wobei der Vektor \vec{w}_j die Ersetzung von $2\Sigma^{-1}\vec{z}_j$ bezeichnet.

In der Praxis kann der Zeitaufwand noch weiter reduziert werden. Mit der vereinfachten Form des *Mahalanobis-Abstandes* suchen wir eigentlich nur den geringsten Abstand zwischen einem Eingabepunkt und den Klassenzentren. Die Berechnung des *Mahalanobis-Abstandes* eines Eingabevektors zu einem Klassenzentrum kann sofort abgebrochen werden, wenn die im Lauf der Berechnung akkumulierte Summe der Teilabstände im Merkmalsraum das Minimum der Abstände des Eingabevektors zu allen bisherigen Klassen überschreitet.

Die Wahl der Anfangsklassifikation

Der erste Schritt bei einem Klassifikationsverfahren besteht in der Wahl einer Anfangsklassifikation. Davon hängen sowohl das Ergebnis als auch die Effizienz der Klassifikation entscheidend ab. Es gibt viele Möglichkeiten für die Wahl dieser Anfangsklassifikation. Im einfachsten Fall wählt man die ersten k Eingabevektoren als Klassenzentren. Eine Alternative ist die zufällige Auswahl der Klassenzentren aus den Eingabevektoren.

Die Anfangsklassenzentren können auch überwacht bestimmt werden. Eine Möglichkeit besteht darin, Klassifikationsergebnisse aus dem vorangehenden Bild als Klassenzentren anzunehmen. Dies ist in vielen Fällen vorteilhaft für die Konvergenz der Klassifikation, besonders bei der Segmentierung von Bildern in einer Bildfolge. Zur Beschleunigung der Konvergenz verwenden wir die Mittelwerte der Eingabevektoren in k Intervallen als Anfangsklassenzentren. In nächsten Abschnitt werden wir anhand der Klassifikationsergebnisse dieses Problem weiter diskutieren.

Die optimale Anzahl von Klassen

Der *Isodata*-Algorithmus basiert auf der k -Mittelwert-Ballung und zusätzlichen Bedingungen, die den Klassifikationsprozeß steuern. Für eine solche Klassifikation ist die Bestimmung der richtigen Anzahl der Klassen sehr wichtig. Das Problem ist im allgemeinen nur durch Probieren lösbar. Eine konventionelle Vorgehensweise besteht darin, zusätzliche quantitative Gütekriterien für die Qualität der Klassifikation zu definieren. Wir stellen hier eine weitere Möglichkeit dar ([Coleman & Andrew 79]). Sie beruht auf der Streumatrix der Interklassenabstände und Intraklassenabstände. Die Streumatrix der Intraklassenabstände beschreibt die Streuung der Punkte vom Klassenzentrum innerhalb einer Klasse. Sie wird definiert als

$$S_w = \frac{1}{k} \sum_{j=1}^k \left(\frac{1}{m_j} \sum_{\vec{x}_i \in C_j} (\vec{x}_i - \vec{z}_j)(\vec{x}_i - \vec{z}_j)^T \right). \quad (5.25)$$

Dabei ist \vec{z}_j das Zentrum der j -ten Klasse, m_j die Punktanzahl in der j -ten Klasse und k die Anzahl der Klassen. In ähnlicher Weise wird die Streumatrix der Interklassenabstände angegeben:

$$S_b = \frac{1}{k} \sum_{j=1}^k (\vec{z}_j - \vec{z}_0)(\vec{z}_j - \vec{z}_0)^T. \quad (5.26)$$

Dabei ist \vec{z}_0 das Zentrum aller Punkte. Eines der am häufigsten verwendeten Gütemaße ist das Produkt der Spuren beider Matrizen. Mathematisch ausgedrückt:

$$\beta = \text{Spur}(S_w) \cdot (\text{Spur} S_b). \quad (5.27)$$

Wenn $k = 1$ ist, haben wir $S_b = 0$. Setzen wir $k = N$ (N ist die Anzahl der Datenpunkte), erhalten wir $S_w = 0$. Aus dieser Überlegung ziehen wir den Schluß, daß mindesten ein $k \in [2, N - 1]$ existiert, welches das Gütemaß β in (5.27) maximiert. Das Gütekriterium in (5.27) wird im 6. Schritt des *Isodata*-Algorithmus eingesetzt.

Beim *Isodata*-Algorithmus läßt sich die Wahl der optimalen Klassenanzahl indirekt durch iterative Verbesserung der Klassifikation lösen. Dann hängt die resultierende Klassenanzahl von Steuerungsparametern η und δ ab. Im folgenden erläutern wir die Bestimmung beider Parameter.

Beim originalen *Isodata*-Algorithmus ist der Aufspaltungsparameter η die maximal zugelassene Varianz für eine Klasse. Es ist allgemein schwer festzustellen, in welcher Größenordnung sie angenommen werden soll. Zur Überwindung dieser Schwierigkeit benutzen wir in der vorliegenden Arbeit η als einen dynamischen Skalierungsparameter zwischen der zugelassenen maximalen Varianz und der mittleren Varianz. Es wird zunächst die Varianz für jede Dimension einer Klasse berechnet. Daraus läßt sich der Mittelwert für jede Dimension aus allen Klassen bilden. Wenn die maximale Varianz einer Klasse, dividiert durch die Mittelvarianz in der entsprechenden Dimension, größer als η ist, wird diese Klasse in der betrachteten Dimension in zwei Klassen aufgespalten. Während der Iteration läßt sich der Parameter η auch schrittweise vergrößern. Je größer der Aufspaltungsparameter η ist, desto weniger Klassen entstehen.

Analog zum Aufspaltungsparameter verwenden wir auch das Verhältnis vom zugelassenen maximalen Abstand zweier Klassen zum mittleren Abstand als den dynamischen Verschmelzungsparameter δ . Wenn das Verhältnis des Abstandes zweier Klassen zum mittleren Abstand kleiner als der Parameter δ ist, lassen sich beide Klassen verschmelzen. Während der Iteration läßt sich der Verschmelzungsparameter δ schrittweise vergrößern. Mit einem größeren δ lassen sich nicht nur mehr Klassen verschmelzen, sondern auch die Klassifikation beschleunigen.

5.2.3 Experimentelle Ergebnisse

In diesem Abschnitt werden die experimentellen Ergebnisse bei der Segmentierung von Texturorientierungsfeldern mit der im vorangegangenen Abschnitt beschriebenen modifizierten *k*-Mittelwert-Ballung und dem *Isodata*-Algorithmus dargestellt.

In den Bildern sowohl der Texturorientierung als auch der Texturanisotropie ist oft Rauschen vorhanden. Um das Orientierungsfeld, evtl. auch das Grauwertbild zu glätten, ohne gleichzeitig die abrupten Wertübergänge zwischen verschiedenen Bildbereichen zu verschmieren, wird die Medianfilteroperation mit einem quadratischen Einzugsbereich von 5×5 Bildpunkten als Vorverarbeitung der Merkmale angewendet. Die Experimente zur Bildsegmentierung wurden auf einem SUN 4 Rechner implementiert. In Abb. 5.7 werden die Segmentierungsergebnisse von den Bildern in Abb. 5.3, 5.4, 5.5 und 5.6 dargestellt. Wie in Abb. 5.7 zu sehen ist, haben wir die Experimente mit der Absicht durchgeführt, Bilder verschiedener Art auszuwerten. Die Aufgaben bestehen dabei darin, daß die Grenzlinien zwischen gerichteten Texturen in (b), zwischen gerichteten Texturen und Texturen ohne dominante Richtung in (a) und (c) sowie zwischen gerichteten Texturen und homogeneren Bereichen in (d) ermittelt werden. Entsprechend der verschiedenen Art von Bildern werden ebenfalls verschiedene Merkmale für die Segmentierung der Bilder ausgewählt. Ergebnisse erscheinen trotz starker Unterschiede in den Ausgangsbildern von vergleichbarer Qualität. Im folgenden wird dies im einzelnen beschrieben.

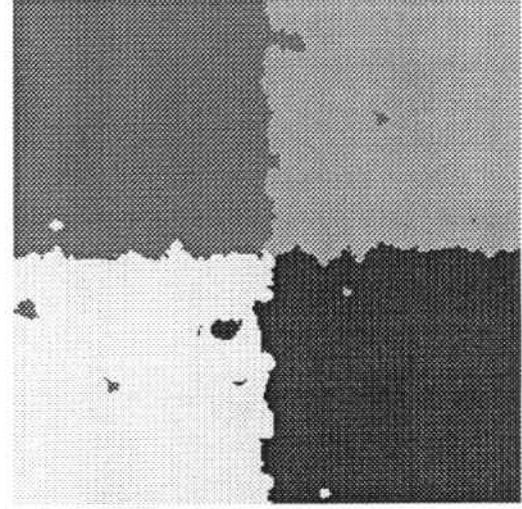
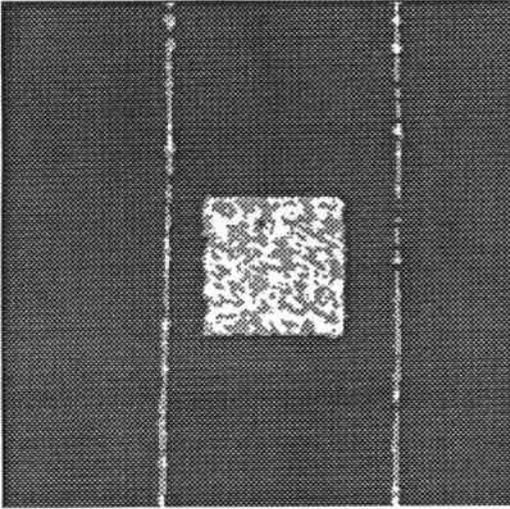
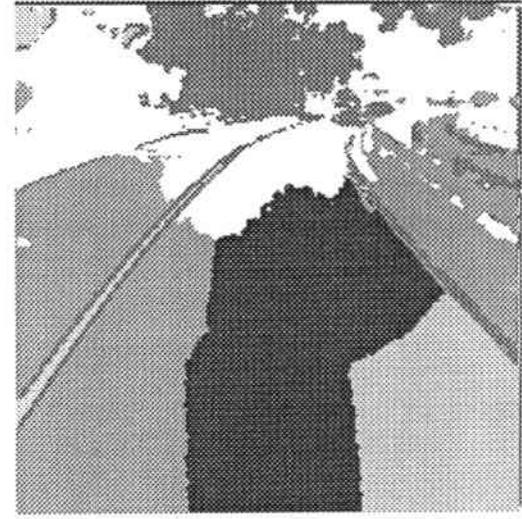
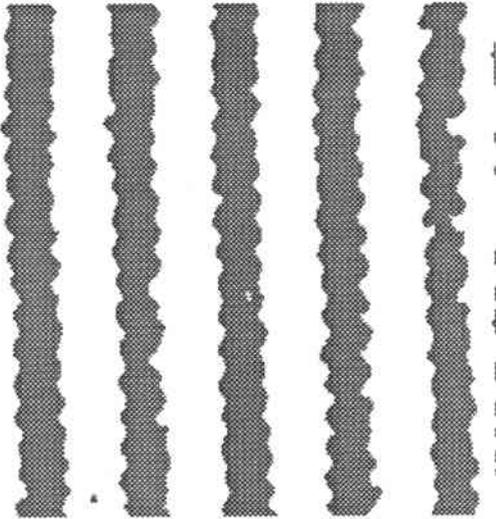
(a) *Segmentation des Bildes aus Abb. 5.3 (a).*(b) *Segmentation des Bildes in Abb. 5.4 (a).*(c) *Segmentation des Bildes in Abb. 5.5 (a).*(d) *Segmentation des Bildes in Abb. 5.6 (a).*

Abbildung 5.7: Segmentierungsergebnisse von Texturorientierungsfeldern: (a) Die Segmentierung des Bildes in Abb. 5.3(a) mit $\vec{x}_i = (\alpha_i, s'_i, x_i, y_i)^T$ als Merkmale für $k = 3$, (b) die Segmentierung des Bildes in Abb. 5.4(a) mit $\vec{x}_i = (\alpha_i, x_i, y_i)^T$ als Merkmale für $k = 4$, (c) die Segmentierung des Bildes in Abb. 5.5(a) mit $\vec{x}_i = (\alpha_i, s'_i, x_i, y_i)^T$ als Merkmale für $k = 2$ und (d) die Segmentierung des Bildes in Abb. 5.6(a) mit $\vec{x}_i = (g_i, s'_i, x_i, y_i)^T$ als Merkmale für $k = 20$.

In Abb. 5.7 (a) wird die Segmentierung des Bildes in Abb. 5.3 (a) dargestellt. Wir haben als Merkmale die Texturorientierung, -anisotropie und beide Koordinaten verwendet. Das Ergebnis in Abb. 5.7 (a) ergibt sich aus der k -Mittelwert-Ballung mit $k = 3$. Mit den gleichen Merkmalen und dem gleichen Ballungsverfahren ergibt sich das Bild in Abb. 5.7 (c) die Segmentierung für das Bildes in Abb. 5.5. Zu diesem Bild müssen wir sagen, daß die Übereinstimmung sehr ermutigend ist. Eine weitere Segmentierung der Unterstrukturen in Abb. 5.7 (c) ist in dieser Arbeit nicht in Betracht gezogen worden.

Ein weiteres Beispiel in Abb. 5.7 (b) macht deutlich, daß die Texturorientierung ein aussagekräftiges Merkmal für die Klassifizierung ist. Es ist leicht zu erkennen, daß es keine großen Unterschiede in Grauwerten und Texturanisotropien zwischen den vier Bildausschnitten gibt. Daher wurde hier lediglich die Texturorientierung zusammen mit beiden Koordinaten als Merkmalsvektor gewählt. Das Bild in Abb. 5.7 (b) ergibt sich mit der k -Mittelwert-Ballung für $k = 4$.

In den oben erwähnten drei Beispielen wurden die Texturorientierungsfelder zusammen mit beiden Koordinaten als Merkmale gewählt. Für das Bild in Abb. 5.6 sind diese Merkmale nicht mehr geeignet. Die Unterbereiche haben unterschiedliche Grauwerte. Einige Unterbereiche sind angenähert homogen. Für das Bild benutzen wir Grauwert, Texturanisotropie und beide Koordinaten als Merkmale. Da in Bildern natürlicher Szenen die Klassenanzahl schwer vorauszusagen ist, erscheint hier der *Isodata*-Algorithmus vorteilhaft. Wir verwenden hier deshalb den *Isodata*-Algorithmus für die Klassifizierung. Als Parameter wurden $k = 5$ für die Anfangsklassenanzahl sowie $\eta = 1.1$ und $\delta = 0.6$ für die Steuerung der iterativen Aufspaltungen und Verschmelzungen von Klassen gewählt. Das Bild in Abb. 5.7 (d) ergibt sich nach zehn Iterationen mit $k = 20$. In diesem Bild ist es zu sehen, daß das Straßenbild nicht zufriedenstellend automatisch segmentiert wird. Die Segmentierung von Straßenbildern wird im nächsten Kapitel behandelt.

Kapitel 6

Segmentierung von Straßenbildern durch die Auswertung von Texturen

Da der Unterschied in der Textur zwischen der Straenoberflche und ihrer Umgebung die entscheidende Information fr die Bildsegmentierung trgt, versuchen wir in diesem Kapitel die Bilder von Straen der vierten Klasse durch die Auswertung von Texturen zu segmentieren.

Die Extraktion von Texturen in Bildern natrlicher Szenen ist eine der schwierigsten Aufgaben in der Texturanalyse. Eine der Schwierigkeiten besteht darin, da die Dichte der Texturen recht unterschiedlich sein kann. Die Extraktion von Texturen aus einem solchen Bild mit dem gleichen Filter an allen Stellen des Bildes htte zur Folge, da die vorliegende Grauwertstruktur nicht bercksichtigt wird. Kleine Filter knnen zwar feinere Texturen extrahieren, sie sind aber sehr sensibel gegenber Rauschen. Ausgedehnte Filter hingegen unterdrcken Rauschen effizienter. Diese Form der Rauschunterdrckung geht allerdings immer auf Kosten einer Detailauflsung. Daher ist es wnschenswert, Filter mit optimalem Mastab je nach der lokalen Grauwertstruktur zu verwenden.

Die Bestimmung eines optimalen Filters wird blicherweise mit der Verwendung eines ganzen Satzes von Filtern durchgefhrt. Bei dieser Vorgehensweise werden Filter unterschiedlicher Gre und evtl. auch Richtung zur Faltung eines Grauwertbildes eingesetzt. Davon wird dann ein geeignetes fr jede Bildposition ausgewhlt. Die Idee dieser Anstze sttzt sich auf die Annahme, da die nderung Texturdichte lediglich ein lokales Ereignis ist.

Wir wissen, da ein Objekt durch eine perspektivische Projektion verzerrt abgebildet wird. Die nderung kann dabei analytisch ermittelt werden. Diese Eigenschaft ermglicht es uns, die durch perspektivische Projektion verursachte nderung der Dichte sowie der Gre von Texturen systematisch zu bestimmen. Es setzt allerdings voraus, da die Abbildungsparameter der Kamera bereits durch eine Kamerakalibrierung ermittelt worden sind.

Die Straenbilder werden im allgemeinen mit perspektivisch abbildenden Kameras aufgenommen. Die nderung der Dichte und Gre von Texturen auf der Straenoberflche wird hauptschlich durch die perspektivische Verzerrung verursacht. Die Straenoberflche

ist meistens homogener als ihre Umgebung. Mit dieser Überlegung als Motivation stellen wir ein Verfahren zur Bestimmung des optimalen Maßstabs von Filtern in Straßenbildern vor. Anschließend zeigen wir die Anwendungsergebnisse des so bestimmten optimalen Maßstabs in der Berechnung von Texturanisotropien.

Mit der ermittelten Texturanisotropie und beiden Bildkoordinaten als Merkmale untersuchen wir nun die Segmentierung von Straßenbildern. Dabei lassen sich die Bildsegmentierungen in der Initialisierungs- und Verfolgungsphase voneinander unterscheiden. Dafür werden sowohl unüberwacht eingelernte als auch adaptive überwacht eingelernte Klassifikatoren in zwei unterschiedlichen Phasen verwendet.

In der Initialisierungsphase benutzen wir den im fünften Kapitel beschriebenen modifizierten *Isodata*-Algorithmus für die Klassifizierung der Straßenbilder. Aus Gründen der Effizienz der Verarbeitung ist uns daran gelegen, die Straßenbilder in der Verfolgungsphase mit einem überwacht eingelernten Klassifikator zu segmentieren. Bei überwacht eingelernten Klassifikatoren werden die Statistiken von Klassen bzw. die Parameter der Klassifikatoren durch eine Stichprobe mit bekannten Objekten geschätzt. Als überwacht eingelernter Klassifikator wird ein Bayesscher Klassifikator in der Verfolgungsphase eingesetzt. Aufgrund der örtlichen bzw. zeitlichen Änderungen der Straßenbilder in einer Bildfolge versagt manchmal ein überwacht eingelernter Klassifikator, der mit fest definierten Prototypen arbeitet. Daher ist ein adaptiver Klassifikator erwünscht. Wir realisieren diese Idee durch die Erneuerung der Klassenstatistiken nach jedem Laufzyklus in der Verfolgungsphase.

6.1 Bestimmung des optimalen Maßstabs von Filtern

Mit der perspektivischen Abbildung einer texturierten Oberfläche lassen sich die Form, Fläche und Dichte von Texeln systematisch ändern. Diese Änderungen wurden als wichtige Informationen beim Struktur-aus-Textur-Verfahren ausgenutzt. Wir behandeln in diesem Abschnitt das Problem, wie diese Verzerrungseffekte bei der Ermittlung von Texturen im Bild berücksichtigt werden können.

6.1.1 Verzerrung eines Texels im perspektivisch abgebildeten Bild

Ein Bild ist die Projektion eines 3D-Welt-Ausschnittes auf eine 2D-Bildebene. Dieser Abbildungsprozeß bringt verschiedene Verzerrungen der Objektform mit sich. Die Verzerrungen können dabei im allgemeinen in folgende zwei Arten unterteilt werden:

Abstandseffekt: Je näher ein Objekt in der Szene vor der Bildebene liegt, desto größer erscheint seine Abbildung im Bild.

Verkürzungseffekt: Die Verzerrung eines Objektes wird auch vom Winkel zwischen der Blicklinie der Kamera und der Oberflächennormale im betrachteten Objektpunkt

beeinflußt.

Die Abbildungssysteme der meisten Kameras können entweder als perspektivisch oder orthographisch modelliert werden. Dabei zeigt die perspektivische Projektion alle zwei der obengenannten Effekte. Wir gehen im weiteren davon aus, daß die Bilder bei uns durch die perspektivische Projektion abgebildet werden.

Bei der perspektivischen Projektion ist bekannt, daß ein Kreis in der Szene durch eine Ellipse im Bild abgebildet werden kann. Zum Studium der Änderungen der perspektivisch abgebildeten Texel nehmen wir an, daß die Texel in der Szene kreisförmig sind. Zur Unterscheidung von den Bildtexeln bezeichnen wir die Texel in der Szene als physikalische Texel. Die physikalischen Texel sind untereinander geometrisch identisch. Wir nehmen an, daß die kreisförmigen Texel sich in der Szene regelmäßig auf einer Ebene verteilen. Die Form und Größe der entsprechenden ellipsenförmigen abgebildeten Texel ändern sich im Bild von Ort zu Ort (Abb. 6.1). Zur Detektion solcher Texel aus einem Bild werden

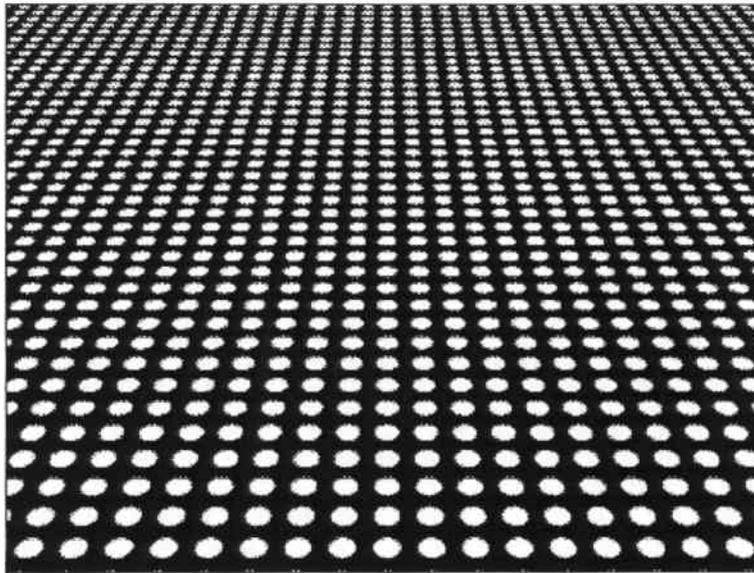


Abbildung 6.1: Die Abbildung von auf einer Szenenebene regelmäßig verteilten kreisförmigen Texeln.

Filter mit variierendem Maßstab und auch variierender Orientierung je nach Bildposition als notwendig und wünschenswert angesehen.

6.1.2 Parameter der abgebildeten ellipsenförmigen Texel

Das Grundprinzip unseres Verfahrens besteht darin, zunächst die Parameter der im Bild abgebildeten Texel in jedem Bildpunkt aus den Parametern der physikalischen Texel und den Parametern des Abbildungssystems abzuleiten und anschließend den Maßstab und

evtl. auch die Orientierung der optimalen Filter zur Ermittlung von Texeln im Bildbereich zu bestimmen. Dazu definieren wir zunächst im Weltkoordinatensystem ein Texel als

$$(x_w - x_{w0})^2 + (y_w - y_{w0})^2 = r^2.$$

Da wir angenommen haben, daß die physikalischen Texel auf einer bekannten Szenenebene liegen, brauchen die z_w -Koordinaten nicht berücksichtigt werden. Durch die Anwendung von homogenen Koordinaten läßt sich diese Kreisfunktion auch in Matrixschreibweise darstellen:

$$\begin{pmatrix} x_w \\ y_w \\ 1 \end{pmatrix}^T \mathbf{P} \begin{pmatrix} x_w \\ y_w \\ 1 \end{pmatrix} = 0 \quad (6.1)$$

mit

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & -x_{w0} \\ 0 & 1 & -y_{w0} \\ -x_{w0} & -y_{w0} & x_{w0}^2 + y_{w0}^2 - r^2 \end{pmatrix}. \quad (6.2)$$

Im folgenden werden wir nur noch homogene Koordinaten betrachten.

Die Abbildung eines Szenenpunktes in einen Bildpunkt kann durch einen Satz von Transformationen beschrieben werden. Wir wollen hier nicht auf die einzelnen Transformationen eingehen. Die Teiltransformationen lassen sich jedoch in homogenen Koordinaten zu folgender Form zusammenfassen:

$$k \begin{pmatrix} x_w \\ y_w \\ 1 \end{pmatrix} = \mathbf{M} \begin{pmatrix} x_b \\ y_b \\ 1 \end{pmatrix}, \quad (6.3)$$

dabei sind \mathbf{M} die Transformationsmatrix, (x_b, y_b) die Koordinaten im Bild und k ein von x_w und y_w abhängiger Parameter ($k \neq 0$). Im folgenden gehen wir davon aus, daß sich die Koeffizienten der Transformationsmatrix \mathbf{M} durch eine Kamerakalibrierung bestimmen lassen. Wir setzen nun die Abbildungstransformation (6.3) in die Kreisfunktion (6.1) ein und erhalten im Bild die Funktion:

$$k^{-2} \begin{pmatrix} x_b \\ y_b \\ 1 \end{pmatrix}^T \mathbf{M}^T \mathbf{P} \mathbf{M} \begin{pmatrix} x_b \\ y_b \\ 1 \end{pmatrix} = 0. \quad (6.4)$$

Bei Division dieser Gleichung durch den Faktor k^{-2} entfällt dieser und braucht daher nicht weiter berücksichtigt werden. Wir wollen nun aus der Funktion (6.4) die Parameter der Bildtexel berechnen. Dazu fassen wir $\mathbf{M}^T \mathbf{P} \mathbf{M}$ zusammen als:

$$\mathbf{C} = \mathbf{M}^T \mathbf{P} \mathbf{M} = \begin{pmatrix} c_A & c_B & c_D \\ c_B & c_C & c_E \\ c_D & c_E & c_F \end{pmatrix}. \quad (6.5)$$

Die 2D-Rotation eines Bildtexels verändert seine geometrische Form nicht. Daher können wir das Koordinatensystem um einen Winkel

$$\theta = \frac{1}{2} \arctan \frac{2c_B}{c_A - c_C} \quad (6.6)$$

zur alten x_b -Achse so drehen, daß die Funktion (6.4) im dadurch entstehenden Koordinatensystem eine Ellipsenfunktion darstellt (Anhang A). Wir bezeichnen diese Rotation als $\mathbf{R}(\theta)$. Durch die Rotation des Koordinatensystems ergibt sich aus der Matrix \mathbf{C} die folgende Koeffizientenmatrix für eine Ellipse:

$$\begin{pmatrix} \lambda_1 & 0 & c'_D \\ 0 & \lambda_2 & c'_E \\ c'_D & c'_E & c_F \end{pmatrix}$$

mit

$$\begin{pmatrix} c'_D \\ c'_E \end{pmatrix} = \mathbf{R}^T(\theta) \begin{pmatrix} c_D \\ c_E \end{pmatrix}$$

und

$$\lambda_{1,2} = \frac{1}{2} \left((c_A + c_C) \pm \sqrt{(c_A + c_C)^2 - 4(c_A c_C - c_B^2)} \right). \quad (6.7)$$

Die Ermittlung der beiden Eigenwerte λ_1 und λ_2 ist ebenfalls in Anhang A beschrieben.

Im neuen Koordinatensystem erhalten wir die folgende Ellipsenfunktion:

$$\lambda_1 \left(x'_b + \frac{c'_D}{\lambda_1} \right)^2 + \lambda_2 \left(y'_b + \frac{c'_E}{\lambda_2} \right)^2 = \frac{c'^2_D}{\lambda_1} + \frac{c'^2_E}{\lambda_2} - c_F. \quad (6.8)$$

Dabei gilt

$$\begin{pmatrix} x'_b \\ y'_b \end{pmatrix} = \mathbf{R}^T(\theta) \begin{pmatrix} x_b \\ y_b \end{pmatrix}.$$

Mit der Abkürzung

$$\begin{aligned} \mu &= \frac{c'^2_D}{\lambda_1} + \frac{c'^2_E}{\lambda_2} - c_F \\ &= \frac{c_C c_D^2 - 2c_B c_D c_E + c_A c_E^2}{c_A c_C - c_B^2} - c_F = \frac{-\det(\mathbf{C})}{c_A c_C - c_B^2} \\ &= \frac{-\det(\mathbf{P})(\det(\mathbf{M}))^2}{c_A c_C - c_B^2} = \frac{r^2 (\det(\mathbf{M}))^2}{c_A c_C - c_B^2} \end{aligned} \quad (6.9)$$

und

$$a = \sqrt{\frac{\mu}{\lambda_1}}, \quad b = \sqrt{\frac{\mu}{\lambda_2}}$$

sowie

$$\begin{cases} x'_{b0} = -\frac{c'_D}{\lambda_1} \\ y'_{b0} = -\frac{c'_E}{\lambda_2} \end{cases}$$

läßt sich die Ellipsenfunktion in einfacher Form darstellen:

$$\frac{(x'_b - x'_{b0})^2}{a^2} + \frac{(y'_b - y'_{b0})^2}{b^2} = 1. \tag{6.10}$$

6.1.3 Optimaler Maßstab von Filtern

Anhand der Parameter der abgebildeten Texel versuchen wir nun, einen optimalen Maßstab in jedem Bildpunkt zu bestimmen. Als ein Hilfsmittel berechnen wir zuerst die Fläche der Texel. Sie ist

$$S = \pi ab = \pi \frac{\mu}{\sqrt{c_{ACC} - c_B^2}} = \frac{\pi r^2 (\det(\mathbf{M}))^2}{\left(\sqrt{c_{ACC} - c_B^2}\right)^3}. \tag{6.11}$$

Aus dieser Gleichung wollen wir nun die Beziehung der Fläche eines Bildtexels zur Fläche der Abbildung des sich im Ursprung des Weltkoordinatensystems befindenden physikalischen Texels herleiten. Dazu kommen wir auf die Matrix \mathbf{P} zurück. Die Anzahl der Additionen in Matrix \mathbf{P} läßt sich weiter reduzieren, indem man die Matrix in zwei Teilmatrizen \mathbf{P}' und \mathbf{P}'' zerlegen:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -r^2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -x_{w0} \\ 0 & 0 & -y_{w0} \\ -x_{w0} & -y_{w0} & x_{w0}^2 + y_{w0}^2 \end{pmatrix} = \mathbf{P}' + \mathbf{P}''.$$

Durch diese Zerlegung ergeben sich eine diagonale Matrix \mathbf{P}' und eine singuläre Matrix \mathbf{P}'' . Es ist dabei anzumerken, daß \mathbf{P}' die Matrixschreibweise eines sich im Ursprung befindenden Kreises ist und \mathbf{P}'' die Lage des betrachteten Texels darstellt. Entsprechend dieser Zerlegung erhalten wir ebenfalls zwei Teilmatrizen von \mathbf{C} :

$$\mathbf{C} = \mathbf{C}' + \mathbf{C}''.$$

Damit ergibt sich

$$\begin{aligned} c_{ACC} - c_B^2 &= (c'_A + c''_A)(c'_C + c''_C) - (c'_B + c''_B)^2 \\ &= \underbrace{c'_A c'_C - (c'_B)^2}_{d_1} + \underbrace{c'_A c''_C + c''_A c'_C + c''_A c''_C - 2c'_B c''_B - (c''_B)^2}_{d_2} \\ &= d_1 + d_2. \end{aligned}$$

Die Elemente der Matrizen C' und C'' in d_1 und d_2 lassen sich wie folgt aus Gleichungen (6.2) und (6.5) berechnen:

$$\begin{cases} c'_A = m_{11}^2 + m_{21}^2 - r^2 m_{31}^2 \\ c'_B = m_{11}m_{12} + m_{21}m_{22} - r^2 m_{31}m_{32} \\ c'_C = m_{12}^2 + m_{22}^2 - r^2 m_{32}^2 \\ c''_A = -m_{31}(2m_{11}x_{w0} + 2m_{21}y_{w0} - m_{31}x_{w0}^2 - m_{31}y_{w0}^2) \\ c''_B = -m_{31}(m_{12}x_{w0} + m_{22}y_{w0}) - m_{32}(m_{11}x_{w0} + m_{21}y_{w0}) \\ \quad + m_{31}m_{32}x_{w0}^2 + m_{31}m_{32}y_{w0}^2 \\ c''_C = -m_{32}(2m_{12}x_{w0} + 2m_{22}y_{w0} - m_{32}x_{w0}^2 - m_{32}y_{w0}^2). \end{cases}$$

Die obigen Gleichungen zeigen, daß d_1 einen bemerkenswerten Unterschied gegenüber d_2 aufweist: Während d_2 abhängig von der Lage (x_{w0}, y_{w0}) ist, bleibt d_1 eine Konstante.

Wir setzen d_1 und d_2 in Gleichung (6.11) ein und erhalten

$$S = \frac{\pi r^2 (\det(\mathbf{M}))^2}{(\sqrt{d_1})^3} \frac{1}{\left(\sqrt{1 + \frac{d_2}{d_1}}\right)^3} = S_0 \frac{1}{\left(\sqrt{1 + \frac{d_2}{d_1}}\right)^3}. \quad (6.12)$$

Dabei ist S_0 die Fläche der Abbildung des sich im Ursprung befindenden Texels. Mit der Gleichung (6.12) sind wir in der Lage, die Fläche der Abbildung eines physikalischen Texels in jedem Bildpunkt zu ermitteln.

Die Flächen der nebeneinander abgebildeten Texel unterscheiden sich nicht sehr voneinander, wenn die x-Achse des Bildkoordinatensystems etwa parallel zu der betrachteten Ebene liegt. Daher ist es sinnvoll, die Gleichung (6.12) nur zeilenweise oder an einigen ausgewählten Bildpunkten zu berechnen. In Abb. 6.2 wird das Verhältnis S/S_0 entlang einer Linie mit konstanter Bildspalte (x_b) beispielhaft dargestellt. Auf der horizontalen Achse sind die Bildzeilen aufgetragen.

Um den Maßstab von Filtern zu ermitteln, approximieren wir zuerst ein ellipsenförmiges Bildtexel durch ein kreisförmiges. Der Radius des kreisförmigen Texels läßt sich einfach aus der Fläche der Ellipse berechnen: $r = \sqrt{S}/\pi$. Wir nehmen an, daß die Filter aus den Ableitungen der Gaußfunktion

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}\left(\frac{x^2+y^2}{\sigma^2}\right)}$$

bestehen. Der Radius r ist dabei proportional zum σ der Gaußfunktion. Daher ergibt sich die folgende Beziehung

$$\sigma_{opt} = \sigma_0 \sqrt{\frac{S}{S_0}}, \quad (6.13)$$

wobei σ_0 die größte Standardabweichung dieser Gaußfunktion im Bild ist.

Die Gleichung (6.13) läßt sich in verschiedene Filter einsetzen. Es ist dabei anzumerken, daß der Wert von σ_0 im allgemeinen nicht bekannt ist. Daher gilt der optimale Maßstab der Filter nur im Sinne der perspektivischen Projektion. σ_0 ist dabei die einzige freie Variable.

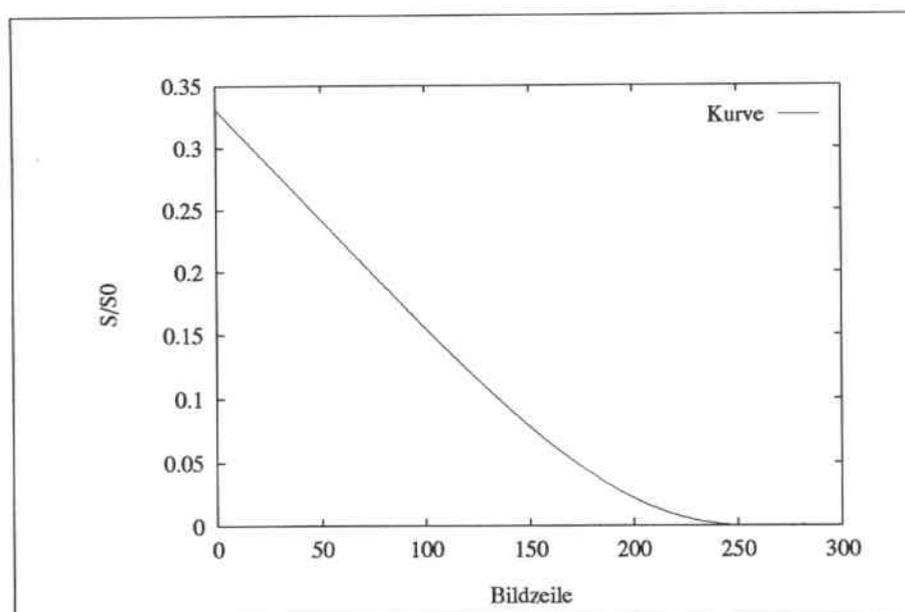


Abbildung 6.2: Das Verhältnis der Fläche der sich entlang einer Bildspalte befindenden Texel zur Fläche der Abbildung des sich im Ursprung befindenden physikalischen Texels.

6.2 Ermittlung von Texturanisotropien mit σ_{opt}

Die Bestimmung des optimalen Maßstabs in perspektivisch abgebildeten Bildern wird im vorangegangenen Abschnitt anhand der Texel beschrieben. Wir wissen, daß es nur dann sinnvoll ist, die Texel als Merkmale für die Bildsegmentation einzusetzen, wenn sie im Bild regelmäßig verteilt sind. Die Texturen in Straßenbildern sind jedoch im allgemeinen geometrisch vielseitiger. Obwohl sie im Bild identifizierbar sind, kommen sie selten mit einer bestimmten, sich wiederholenden geometrischen Form vor. Daher müssen wir auf die Extraktion von einzelnen Texeln verzichten. Die Aufgabe besteht nun darin, eine andere aussagekräftige Beschreibung von Texturen zu finden.

Nachdem wir die Straßenbilder näher untersucht haben, stellen wir fest, daß die Texturen die folgenden 2 Eigenschaften aufweisen:

- Die Umgebung von Straßen aus Rasen, Steinen und Bäumen enthält mehr deutlich gerichtete Texturen als die Straßenoberfläche.
- Auch ohne Fahrbahnmarkierung lassen sich am Übergang zwischen Straßenoberfläche und -umgebung gerichtete Texturen feststellen.

Das bedeutet, daß gerichtete Texturen als vielversprechende Bildmerkmale erscheinen.

Die Beschreibung und die Ermittlung gerichteter Texturen wurden bereits im vorangegangenen Kapitel ausführlich diskutiert. Im allgemeinen sind die Texturen in Straßenbildern weniger regulär als diejenigen aus [Brodatz 66]. Daher läßt sich das Verfahren ohne zusätzliche Änderung schwer auf Straßenbilder übertragen.

6.2.1 Erweiterung des Verfahrens zur Ermittlung von Texturanisotropien

Wir diskutieren nun die Erweiterung des in Kapitel 5 beschriebenen Verfahrens für die Ermittlung von Texturanisotropien aus Straßenbildern. Diese Erweiterung besteht im wesentlichen aus den folgenden 3 Aspekten:

- Verwendung eines optimalen Maßstabs in der Berechnung der Texturanisotropie,
- Einführung einer neuen Definition für Texturanisotropie und
- Verzicht auf die Schwelle für Texturanisotropie.

Im folgenden werden sie detailliert beschrieben.

Optimaler Maßstab zur Berechnung der Grauwertgradienten f_x und f_y sowie der Wichtungsfunktion $w(i, j)$

Die Ermittlung von Texturanisotropie wird in zwei Stufen durchgeführt. Die erste Stufe ist die Berechnung des Grauwertgradienten. Die zweite Stufe besteht aus der Berechnung der Kovarianzmatrix der Grauwertgradienten und anschließend der Ermittlung der Anisotropie. In beiden Schritten muß das Problem des optimalen Maßstabs betrachtet werden.

Die Berechnung der Ableitungen f_x und f_y erfolgt durch die Faltung des Grauwertbildes mit g_x und g_y in Gleichungen (5.4) und (5.5). Die Standardabweichung σ in g_x und g_y bestimmt die Größe der entsprechenden Filter. Wir setzen das in Gleichung (6.13) ermittelte σ_{opt} in diesen beiden Gleichungen ein. Damit ist die Aufgabe in der ersten Stufe bereits gelöst.

Die Gewichtungsmaske $w(i, j)$ in Gleichungen (5.8) - (5.12) hat eine konstante Größe und identische Koeffizienten. Statt einer Rechteckmaske, die im allgemeinen kein gutes Glättungsfilter darstellt, nehmen wir die Gaußsche Funktion mit σ_{opt} als Gewichtung des Integrationsfensters.

So haben wir nach Gleichung (6.13) nur eine freie Variable σ_0 zur Berechnung des Gradienten- und des Wichtungsfilters.

Bei der derzeitigen Implementierung wird nur ein optimaler Maßstab für jede Bildzeile ermittelt. Nachdem σ_0 festgelegt ist, verkleinert sich σ_{opt} im Bild von unten nach oben. Dabei wird die kleinste Maske mit 3×3 Pixeln angenommen. Die Koeffizienten aller Filtermasken in der vorliegenden Arbeit sind auf 1 normiert.

Eine neue Definition für Texturanisotropie

In Kapitel 5 wurde die Texturanisotropie s in Gleichung (5.18) in der Umgebung eines Bildpunktes (x_b, y_b) mit dem Betrag des Gradienten in diesem Punkt gewichtet. Das so erweiterte Texturmaß s' in Gleichung (5.19) berücksichtigte damit auch die Kontraststärke der Textur.

Auf Bildern realer Straßenoberflächen mit einer i.d.R. eher feinkörnigen Textur weist allerdings der Gradientenbetrag ein nicht zu vernachlässigendes Rauschen auf. Um aber auch hier die Kontraststärke in der Texturanisotropie mit zu bewerten führen wir folgendes Maß ein:

$$\begin{aligned} s'' &= \frac{1}{C} \text{spur}(\Sigma) \left(1 - \frac{\lambda_2}{\lambda_1}\right) \\ &= \frac{1}{C} (\lambda_1 + \lambda_2) \left(1 - \frac{\lambda_2}{\lambda_1}\right) \\ &= \frac{1}{C} \lambda_1 e^2. \end{aligned} \tag{6.14}$$

Dabei sind

$$e = \frac{\sqrt{\lambda_1^2 - \lambda_2^2}}{\lambda_1}$$

die Exzentrizität der mit der Kovarianzmatrix Σ beschriebenen Ellipse und C ein Normierungsfaktor. Die Spur der Kovarianzmatrix Σ der Grauwertgradienten innerhalb der Integrationsmaske ist ein zuverlässigeres, weil weniger verrauschtes Maß für den Kontrastumfang in der Umgebung des betrachteten Punktes, als der Gradientenbetrag an dieser Stelle selbst.

Null-Schwelle

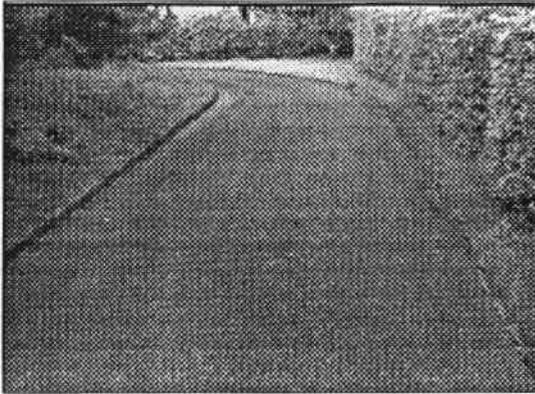
Zur Vereinfachung des Verfahrens bei der Ermittlung von Texturanisotropien aus verschiedenen Straßenbildern versuchen wir auf die Schwelle für die Texturanisotropie zu verzichten.

Die Texturen, die sich weiter entfernt von der Kamera befinden, werden mit kleineren Filtern gefaltet. Ihre Anisotropie wird nicht so stark von der Filtermaske unterdrückt. An den sich unmittelbar vor der Kamera befindenden Stellen haben die Untertexturen signifikant größere Grauwertänderungen. Durch die Faltung mit einer größeren Filtermaske lassen sie sich jedoch effektiver glätten. Daher erscheint eine Schwelle nicht mehr nötig, wenn der jeweils optimale Maßstab bei der Berechnung von Texturanisotropien eingesetzt wird.

6.2.2 Beispiele von ermittelten Texturanisotropien

Wir haben das neue Verfahren auf verschiedene Straßenbilder bzw. -bildfolgen angewendet. Einige der Untersuchungsergebnisse werden in Abb. 6.3 - 6.6 dargestellt. Alle gezeigten Bilder sind aus dem Versuchsfahrzeug des IITB heraus aufgenommen. Jedes Bild ist 380×282 Pixel groß.

Um den durch die Anwendung eines optimalen Maßstabs geschaffenen Vorteil zu verdeutlichen, zeigen wir in Abb. 6.3 zuerst die mit unterschiedlichen σ und Größe des



(a) Ein Straßenbild.

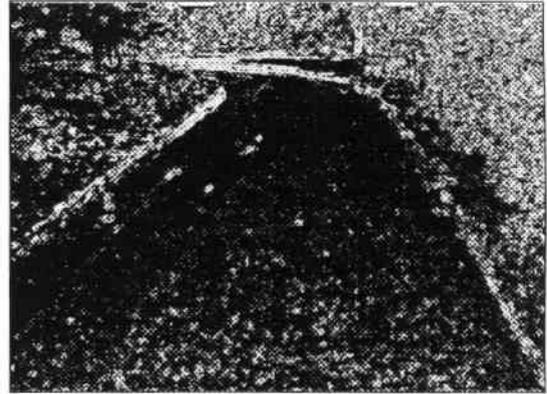
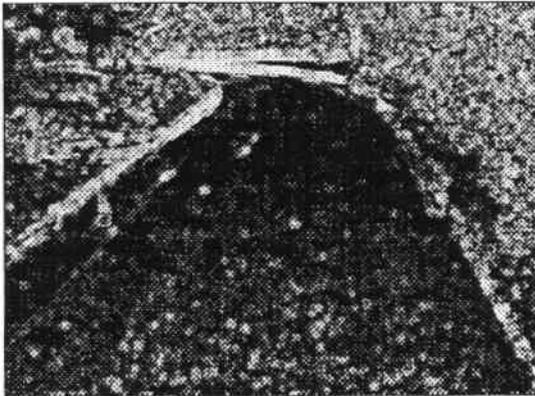
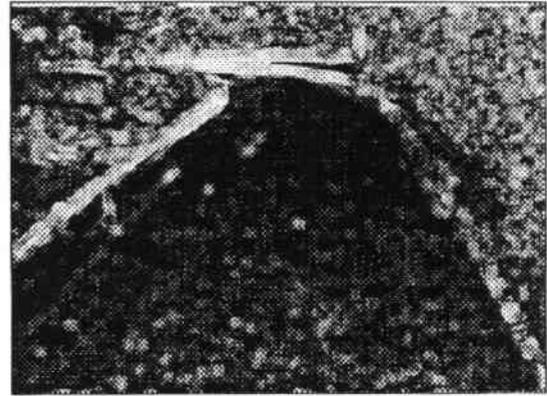
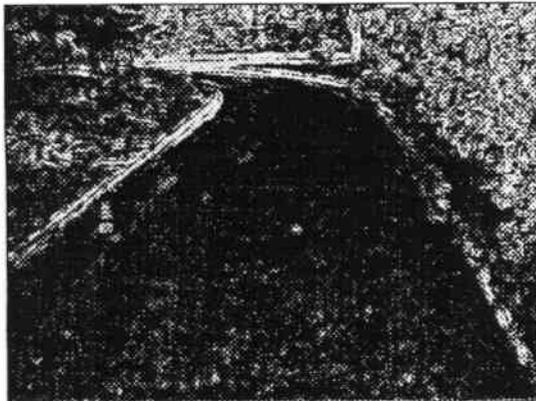
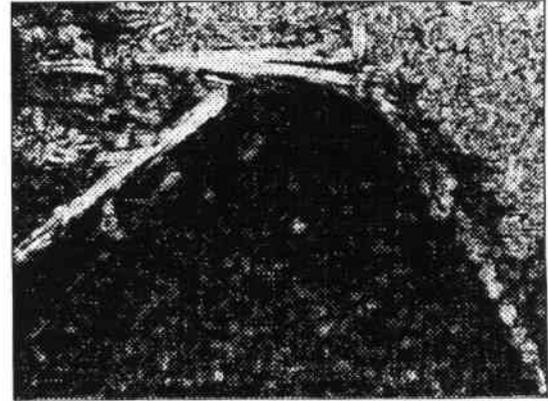
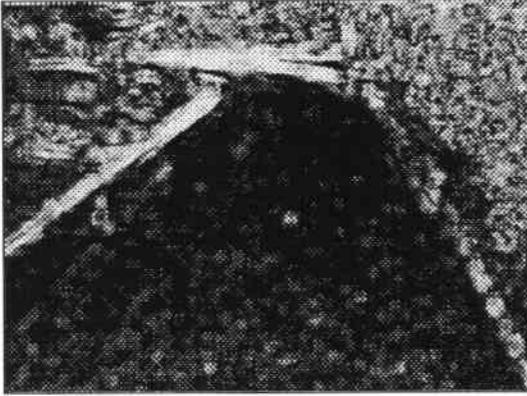
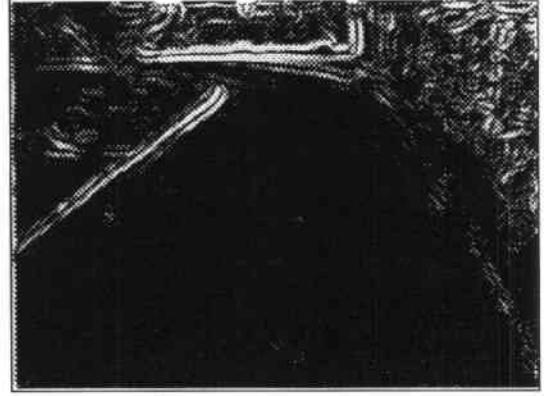
(b) $\sigma = 0.8$ und $f-g = 3 \times 3$.(c) $\sigma = 0.8$ und $f-g = 5 \times 5$.(d) $\sigma = 0.8$ und $f-g = 7 \times 7$.(e) $\sigma = 1.2$ und $f-g = 3 \times 3$.(f) $\sigma = 1.2$ und $f-g = 5 \times 5$.

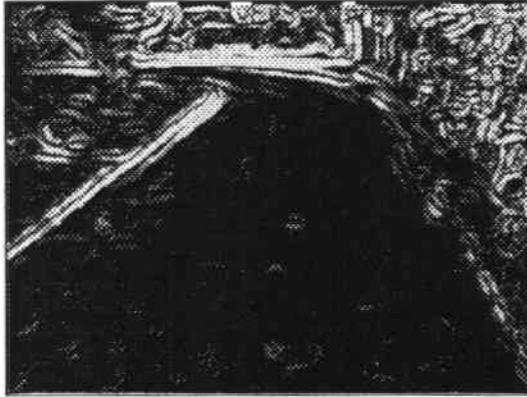
Abbildung 6.3: Texturanisotropie eines Straßenbildes mit unterschiedlichen konstanten σ und $f-g$ (Größe des Integrationsfensters) für das gesamte Bild: (a) das Grauwertbild; (b) - (j) Texturanisotropie mit $\sigma \in \{0.8, 1.2, 2.5\}$ und $f-g \in \{3 \times 3, 5 \times 5, 7 \times 7\}$; (k) Texturanisotropie mit dem optimalen σ_{opt} bzw. Fenstergröße, σ_0 beträgt 1.5; (l) Texturanisotropie mit dem optimalen σ_{opt} bzw. Fenstergröße, σ_0 beträgt 2.5.



(g) $\sigma = 1.2$ und $f-g = 7 \times 7$.



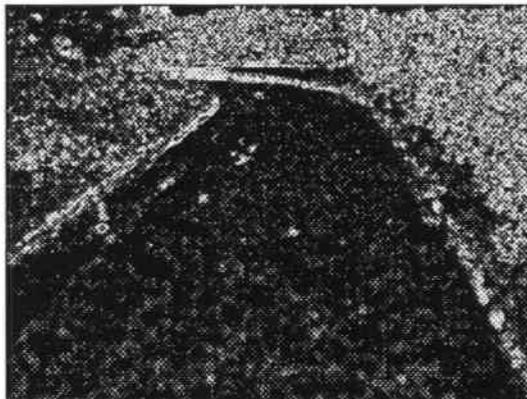
(h) $\sigma = 2.5$ und $f-g = 3 \times 3$.



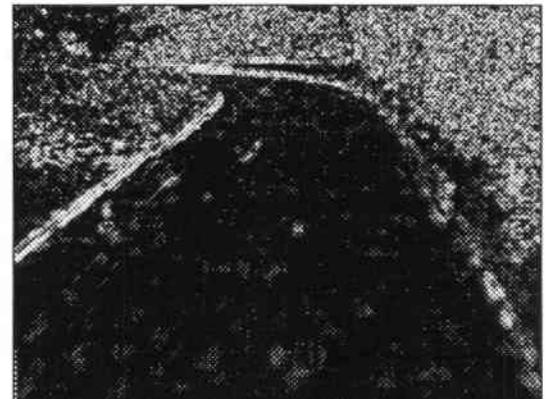
(i) $\sigma = 2.5$ und $f-g = 5 \times 5$.



(j) $\sigma = 2.5$ und $f-g = 7 \times 7$.



(k) σ_{opt} mit $\sigma_0 = 1.5$.

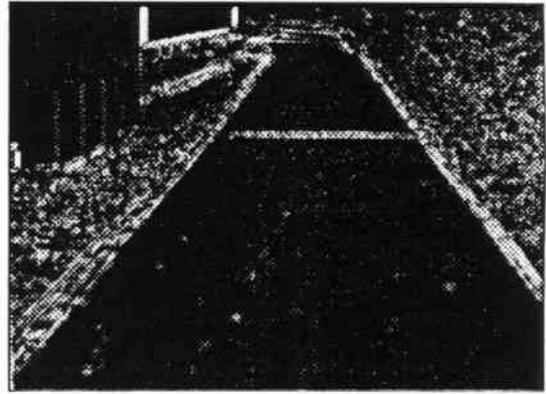


(l) σ_{opt} mit $\sigma_0 = 2.5$.

(Fortsetzung der vorherigen Seite)



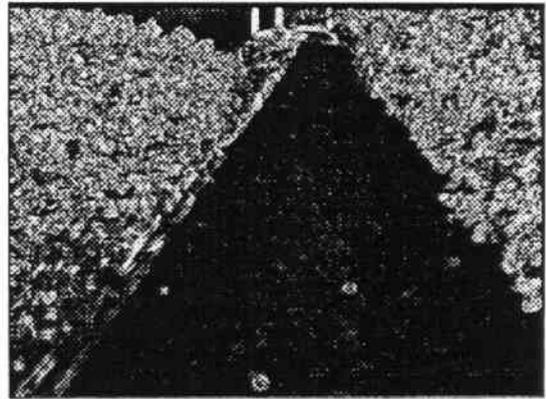
(a) Ein Straßenbild.



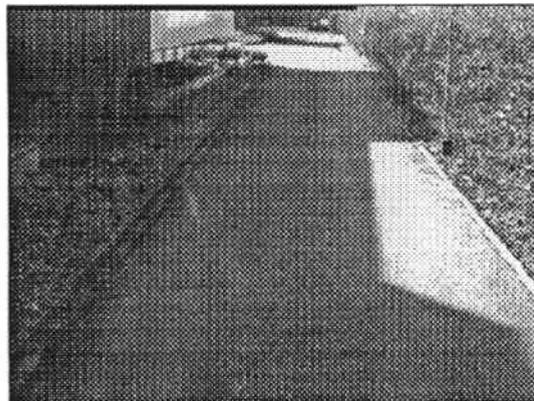
(b) Die Texturanisotropie von (a).



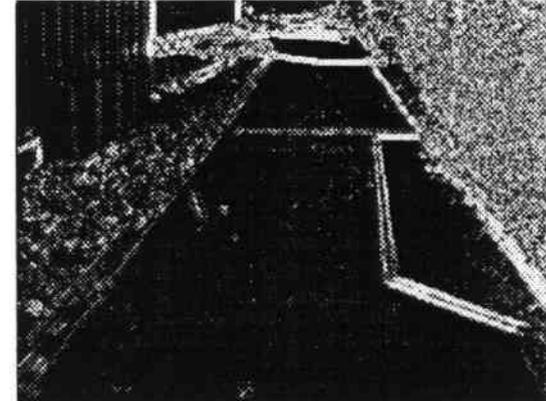
(c) Ein Straßenbild.



(d) Die Texturanisotropie von (c).



(e) Ein Straßenbild.

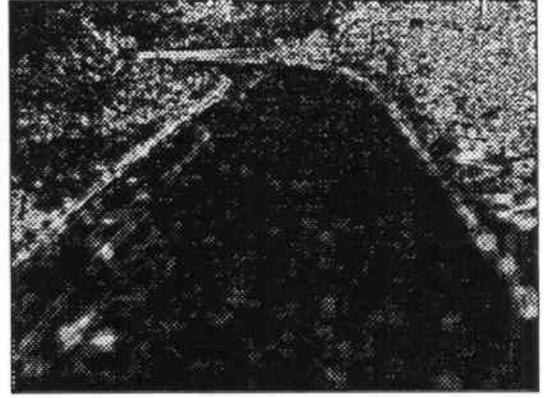


(f) Die Texturanisotropie von (e).

Abbildung 6.4: Ergebnisse der berechneten Texturanisotropie mit σ_{opt} für verschiedene Bilder einer Straßenszene. Dabei beträgt σ_0 für alle Bilder 1.5.



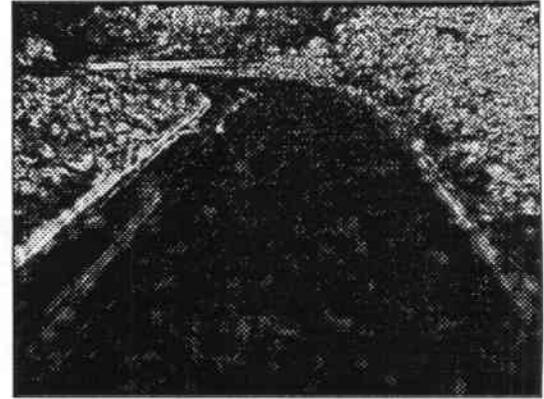
(a) Bild # 0.



(b) Die Texturanisotropie von (a).



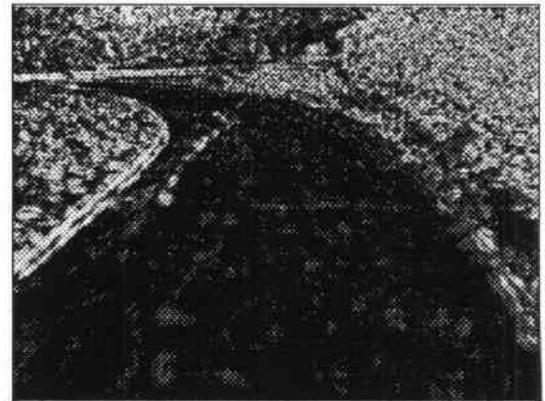
(c) Bild # 10.



(d) Die Texturanisotropie von (c).

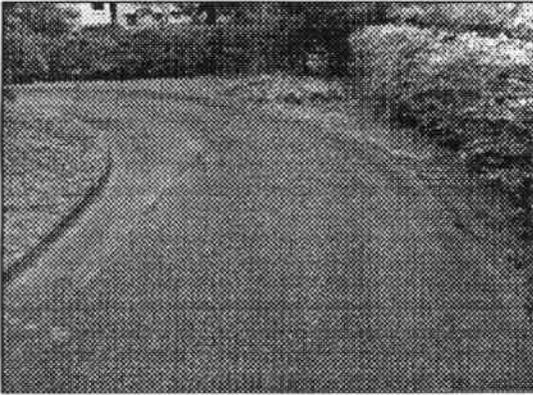


(e) Bild # 20.

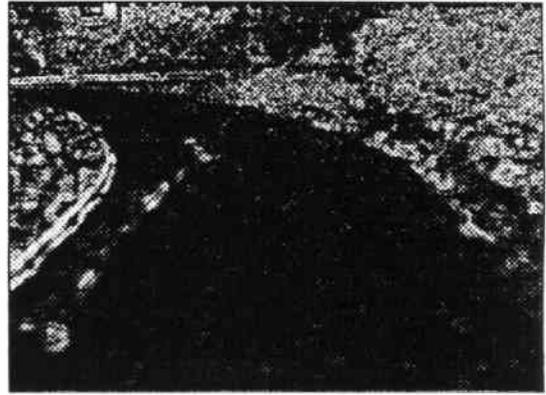


(f) Die Texturanisotropie von (e).

Abbildung 6.5: Ergebnisse der berechneten Texturanisotropie in einer im Institutsgelände (IITB) aufgenommenen Bildfolge mit σ_{opt} . Dargestellt sind das Grauwertbild und das Bild der Texturanisotropie der 0., 10., 20., 30., 40., 50., 60., 70. und 80. Aufnahme. Dabei beträgt σ_0 für alle Bilder 2.5.



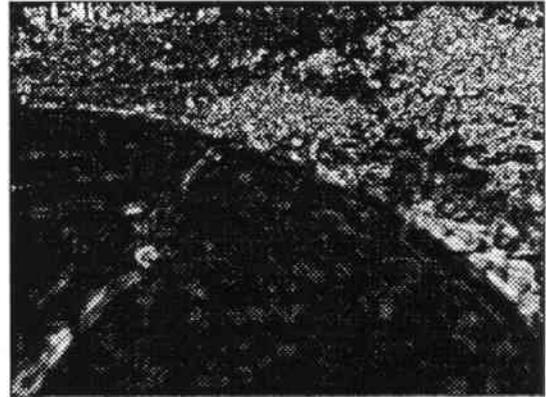
(g) Bild # 30.



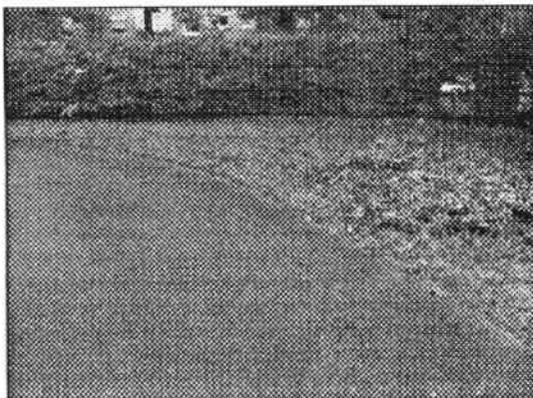
(h) Die Texturanisotropie von (g).



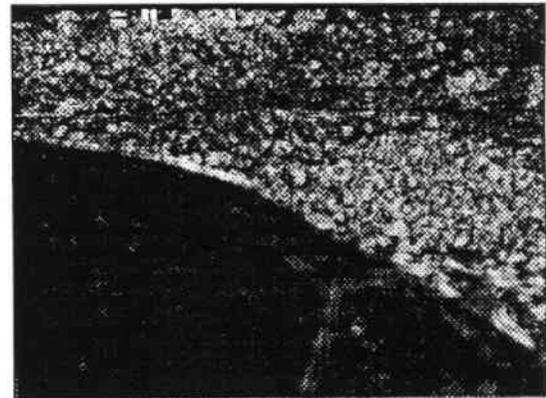
(i) Bild # 40.



(j) Die Texturanisotropie von (i).



(k) Bild # 50.

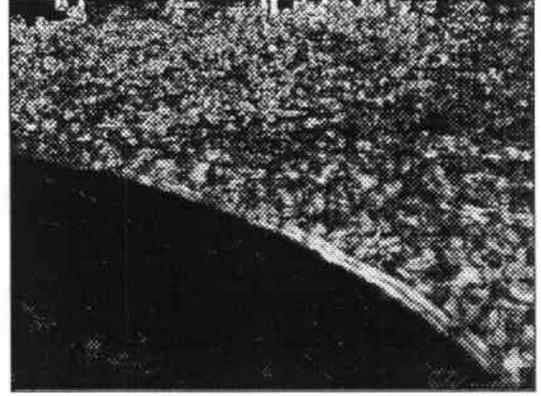


(l) Die Texturanisotropie von (k).

(Fortsetzung der vorherigen Seite)



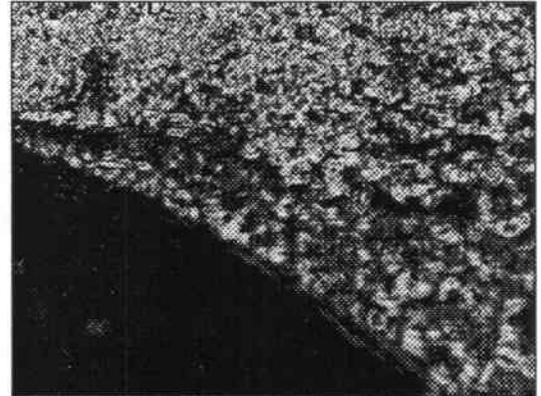
(m) Bild # 60.



(n) Die Texturanisotropie von (m).



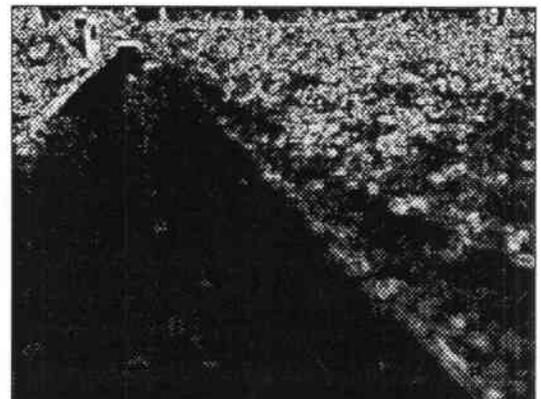
(o) Bild # 70.



(p) Die Texturanisotropie von (o).



(q) Bild # 80.



(r) Die Texturanisotropie von (q).

(Fortsetzung der vorherigen Seite)

Integrationsfensters ermittelte Texturanisotropie. In (a) wird ein Straßenbild dargestellt. Deutlich zu erkennen ist, daß der Grauwert zwischen der Straßenoberfläche und -umgebung keinen signifikanten Unterschied aufweist. In (b) - (j) wird jeweils das Ergebnis mit einem konstanten Maßstab und ebenfalls einer konstanten Größe des Integrationsfensters für das gesamte Bild dargestellt. In (k) und (l) wird die mit dem optimalen Maßstab ermittelte Texturanisotropie wiedergegeben. Dabei beträgt σ_0 1.5 für das Bild in (k) und 2.5 für das Bild in (l). Dementsprechend ist die Größe des Integrationsfensters jeweils 9×9 und 13×13 . Wie schon erwähnt wurde, beträgt die Schwelle 0 für alle Bilder.

Weitere Beispiele in Abb. 6.4 machen deutlich, daß die Texturanisotropie ein aussagekräftiges Merkmal ist. In den Grauwertbildern in Abb. 6.4 kommen unterschiedliche Störungen vor: in (a) ist auf der Straßenoberfläche eine senkrecht zur Straßenbegrenzung liegende Abflußrinne vorhanden; die Straße in (b) wird von Bäumen verdeckt; in (c) befindet sich ein Schattenwurf auf der Straßenoberfläche. Trotz solcher Störungen ist die ermittelte Anisotropie zwischen Straßenoberfläche und -umgebung eindeutig unterschiedlich.

In Abb. 6.5 wird die in einer Bildfolge mit σ_{opt} ermittelte Texturanisotropie gezeigt. Dargestellt sind das Grauwertbild und das Bild der Texturanisotropie der 0., 10., 20., 30., 40., 50., 60., 70. und 80. Aufnahme. Für die gesamte Bildfolge wird σ_0 mit 2.5 angenommen.

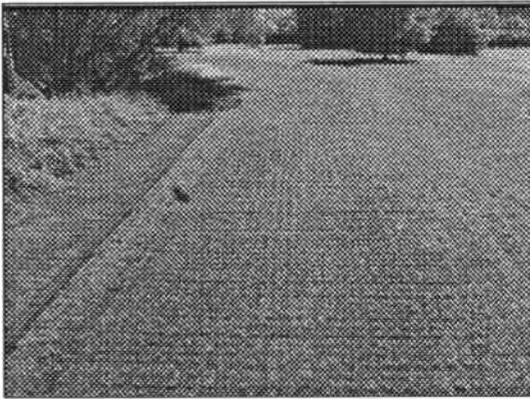
Das Straßenbild in Abb. 6.6 wurde im Institutsgelände (IITB) aufgenommen. Das Bild in (a) ist oft von Menschen auch auf den ersten Blick nicht als die Abbildung einer Straße zu erkennen. Die Straßenoberfläche besteht aus Steinen, die teilweise von Rasen verdeckt werden. Auf der linken Seite der Straße befindet sich ein Fußweg, der ebenfalls aus Steinen besteht. Auf der rechten Seite der Straße ist Rasen zu sehen. In Abb. 6.6 lassen sich die Ortsfrequenzen der Grauwertänderung zwischen der Straßenoberfläche und -umgebung nicht eindeutig voneinander unterscheiden. Die Ergebnisse der berechneten Texturanisotropie mit verschiedenen σ_0 werden in (b), (c) und (d) dargestellt. Es ist dabei zu erkennen, daß die Texturanisotropie in Straßen dieser Art kein geeignetes Merkmal erscheint. Mit diesem Beispiel wollen wir die Grenze des in dieser Arbeit neu entwickelten Verfahrens zeigen.

6.3 Klassifizierung von Straßenbildern

Zur Klassifizierung eines Bildes mit stochastischen Klassifikatoren gibt es zwei grundsätzlich verschiedene Vorgehensweisen: Klassifikation mit unüberwachten Lernverfahren oder mit überwachten Lernverfahren. Sie kommen bei der Klassifizierung von Straßenbildern jeweils in der Initialisierungs- und Verfolgungsphase zum Einsatz.

Die Klassifizierung eines Straßenbildes mit unüberwachten Lernverfahren benötigt keine Prototypen von Klassen. Die Gewinnung der Parameter erfolgt durch eine iterative Ballung der Eingabedaten. Daher ist das Verfahren geeignet für die Segmentierung der Straßenbilder in der Initialisierungsphase.

Eine Ballungsanalyse ist von Natur aus ein iteratives Verfahren. Verfahren dieser Art



(a) Ein Straßenbild.

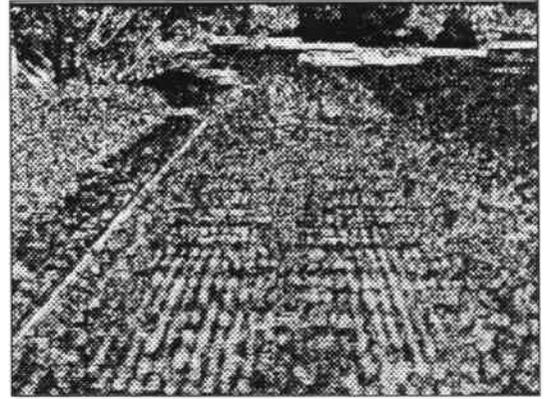
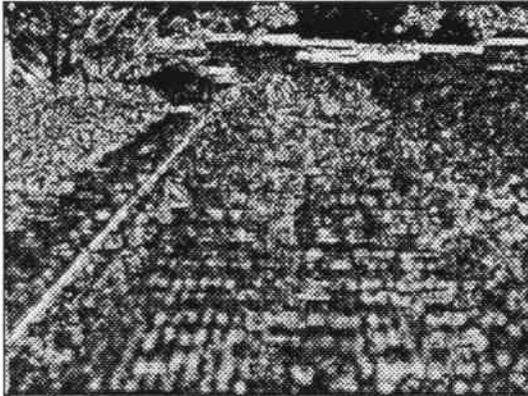
(b) Die Texturanisotropie mit σ_{opt} ($\sigma_0 = 1.5$).(c) Die Texturanisotropie mit σ_{opt} ($\sigma_0 = 2.0$).(d) Die Texturanisotropie mit σ_{opt} ($\sigma_0 = 2.5$).

Abbildung 6.6: Ergebnisse der berechneten Texturanisotropie in einem im Institutsgelände (IITB) aufgenommenen Bild mit σ_{opt} . Dabei beträgt σ_0 1.5 in (b), 2.0 in (c) und 2.5 in (d).

sind daher nicht geeignet, wenn hohe Verarbeitungsgeschwindigkeiten gebraucht werden. Die Geschwindigkeit der Segmentierung eines Bildes ist eigentlich nicht der Forschungsschwerpunkt der vorliegenden Arbeit. Trotzdem soll die Reduzierung der Verarbeitungszeit bei der Entwicklung von Algorithmen für Aufgaben wie etwa Straßenerkennung berücksichtigt werden. Um die Zeit für die Klassifikation eines Bildes zu reduzieren, verwenden wir deshalb in der Verfolgungsphase einen überwacht eingelernten Klassifikator.

6.3.1 *Isodata*-Algorithmus in der Initialisierungsphase

Als unüberwacht lernendes Verfahren haben wir den modifizierten *Isodata*-Algorithmus in Kapitel 5 bereits ausführlich diskutiert. Dieser Algorithmus wurde dort für die Klassifizierung der Orientierungsfelder gerichteter Texturen eingesetzt. Dabei wurde das Verfahren mit der Absicht entwickelt, Bilder verschiedener Art auszuwerten. Wir verwenden diesen Algorithmus hier für die Klassifizierung von Straßenbildern während der Initialisierungsphase.

Der *Isodata*-Algorithmus führt eine iterative Auswertung der Eingabedaten durch. Als Resultate ergeben sich zum einen das segmentierte Bild für die Erkennung der Straße, zum anderen die Statistiken jeder Klasse für die Segmentierung des nächsten Bildes in derselben Bildfolge.

6.3.2 Statistische Klassenmodelle

Das statistische Modell einer Klasse ist ein entscheidender Teil eines stochastischen Klassifikators. Durch den Vergleich eines Eingabevektors mit diesen Modellen wird ein Eingabevektor einer der Klassen zugeordnet. Kenngrößen der Statistik, wie etwa der Erwartungswert zur Lagebeschreibung, die Varianz als Variabilitätsmaß sowie die Kovarianz zur Quantifizierung der Abhängigkeiten zwischen Komponenten von Merkmalsvektoren geben die Eigenschaften einer Klasse an. Sie stehen deshalb im Mittelpunkt unserer Betrachtung als Modellparameter. Außerdem ist die Punktzahl einer Klasse ebenfalls ein wichtiger Parameter. Als Klassenmodelle definieren wir

$$\{\vec{\mu}_j, \Sigma_j, N_j \mid j = 1, \dots, s\}, \quad (6.15)$$

wobei $\vec{\mu}_j$ der Mittelwertvektor, Σ_j die Kovarianzmatrix, N_j die Punktzahl einer Klasse und s die Klassenanzahl sind. Der Mittelwertvektor und die Kovarianzmatrix seien definiert als:

$$\vec{\mu}_j = \frac{1}{N_j} \sum_{\vec{x}_i \in C_j} \vec{x}_i \quad (6.16)$$

und

$$\Sigma_j = \frac{1}{N_j} \sum_{\vec{x}_i \in C_j} \vec{x}_i \vec{x}_i^T - \vec{\mu}_j \vec{\mu}_j^T. \quad (6.17)$$

Hier ist C_j die Teilmenge der Bildpunkte, die zu der Klasse j gehören.

6.3.3 Bayesscher Klassifikator in der Verfolgungsphase

Nach der Klassifizierung aller Punkte eines Straßenbildes in der Initialisierungsphase ergeben sich die Statistiken jeder Klasse. Diese Statistiken sollen als Modelle der Klassen für die Klassifikation der Bilder in der Verfolgungsphase eingesetzt werden. Wegen der

örtlichen und zeitlichen Variationen der Straßenszenen führt ein überwachter eingelernter Klassifikator, der mit fest definierten Klassenmodellen arbeitet, auf Schwierigkeiten bei der Segmentierung. Daher ist es sinnvoll, die statistischen Modelle zu aktualisieren. Wir realisieren diese Idee durch die Erneuerung der Klassenstatistiken nach der Segmentierung jedes Bildes in der Verfolgungsphase. Daher nennen wir unseren überwachter eingelernten Klassifikator einen adaptiven Klassifikator.

Als überwachter eingelernter Klassifikator in der Verfolgungsphase verwenden wir hier den Bayesschen Klassifikator. Die Grundlagen dieses stochastischen Klassifikators findet man z.B. in [Duda & Hart 73], [Nagel 91a] und [Schalkoff 92]. Als Entscheidungsfunktion haben wir

$$g_i(\vec{x}) = \log p(\vec{x} | \omega_i) + \log P(\omega_i). \quad (6.18)$$

Für unsere Aufgabe nehmen wir weiter an, daß die bedingte Verteilungsdichte $p(\vec{x} | \omega_i)$ eine Normalverteilung $N(\vec{\mu}_i, \Sigma_i)$ ist. Diese Dichte ist vollständig bestimmt durch den bedingten Mittelwertvektor $\vec{\mu}_i$ und die bedingte Kovarianzmatrix Σ_i in den Gleichungen (6.16) und (6.17). Weiterhin verwenden wir das Verhältnis der Anzahl N_j der Merkmalsvektoren in der Klasse ω_j zur Anzahl N der gesamten Merkmalsvektoren als Schätzwert für die *A-Priori*-Wahrscheinlichkeit $P(\omega_j)$. Nach der zusätzlichen Eliminierung der unbedeutenden Terme erhalten wir:

$$g_i(\vec{x}) = -\frac{1}{2}(\vec{x} - \vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i) - \frac{1}{2} \log |\Sigma_i| + \log \frac{N_j}{N}. \quad (6.19)$$

Dies ist der in dieser Arbeit eingesetzte Bayessche Klassifikator.

6.3.4 Beispiele von Klassifizierungsergebnissen

Wir zeigen nun die Anwendungsbeispiele des in diesem Kapitel aufgestellten, aus zwei Klassifikatoren kombinierten Verfahrens zur Segmentierung von Straßenbildern in Bildfolgen.

Zur Gewinnung homogener Bereiche und zur Reduzierung der Verarbeitungszeit kann mit Hilfe einer Pyramide aus einem Originalbild eine Reihe von Bildern reduzierter Auflösung erzeugt werden ([Burt 81+83]). Wir verwenden diese Idee in der Segmentierung von Straßenbildern. Die Pyramide wird durch eine Mittelwertbildung aufgebaut. Die Auswahl der Ebenennummer wird in der Initialisierungsphase vor der Segmentierung durchgeführt.

In Abb. 6.7 werden die Ergebnisse der Segmentierung des in Abb. 6.4 (c) dargestellten Straßenbildes in der Initialisierungsphase auf den Pyramidenebenen 0 - 3 gezeigt. Die Klassenanzahl beträgt dabei für alle Ebenen 10. Die Straßenbegrenzung bestehend aus Bordsteinrändern wird teilweise von Bäumen verdeckt. Das Originalbild auf der Ebene 0 besteht aus 380×282 Pixeln. Trotz der Reduzierung der Bildauflösung ist deutlich zu sehen, daß die Segmentierung auf jeder Pyramidenebene die gewünschten Ergebnisse liefert.

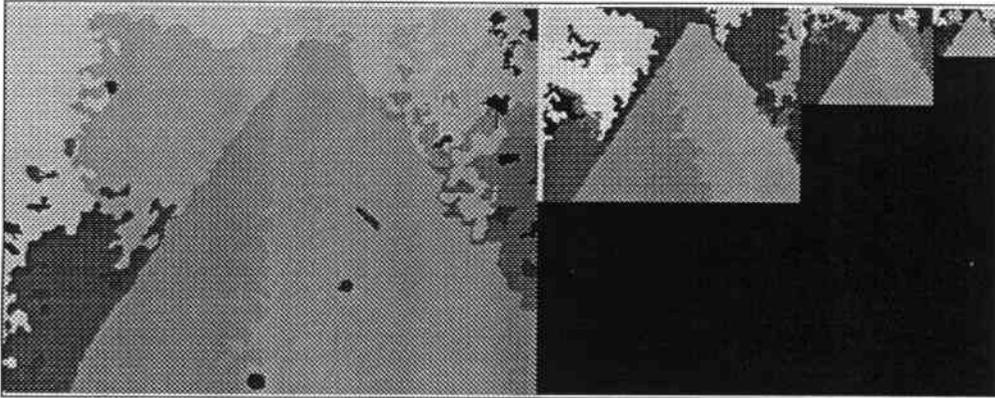


Abbildung 6.7: Die Segmentierung des in Abb. 6.4 (c) dargestellten Straßenbildes auf den Pyramidenebenen 0 - 3. Die Klassenanzahl beträgt dabei 10.

In Abb. 6.8 werden die Segmentierungsergebnisse der in Abb. 6.5 gezeigten Bildfolge dargestellt. Die Klassifizierung wird auf der Pyramidenebene 1 durchgeführt. Dabei wird die Klassenanzahl mit 3 angenommen. In der Initialisierungsphase wird die Klassifizierung der 0. Aufnahme mit dem modifizierten *Isodata*-Algorithmus als Lernverfahren erzielt. Das Ergebnis wird in Abb. 6.8 (a) dargestellt. Die dabei ermittelten Straßenmodelle werden dann in die Verfolgungsphase übertragen. In Abb. 6.8 (b) - (i) werden die Ergebnisse der Klassifizierung der 10., 20., 30., 40., 50., 60., 70. und 80. Aufnahme mit dem Bayesschen Klassifikator in der Verfolgungsphase wiedergegeben.

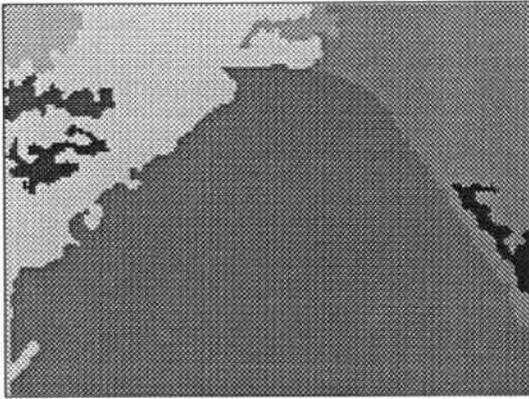
Bisher haben wir die Texturanisotropie aus Straßenbildern extrahiert und darauf aufbauend die Bildpunkte in mehrere Bereiche segmentiert. Als Bildbereichshinweise stehen die segmentierten Bilder der modellgestützten Erkennung von Straßen in Kapitel 7 zur Verfügung.

6.3.5 Beispiele von Segmentierungsergebnissen durch konturbasierte Bereichsanalyse

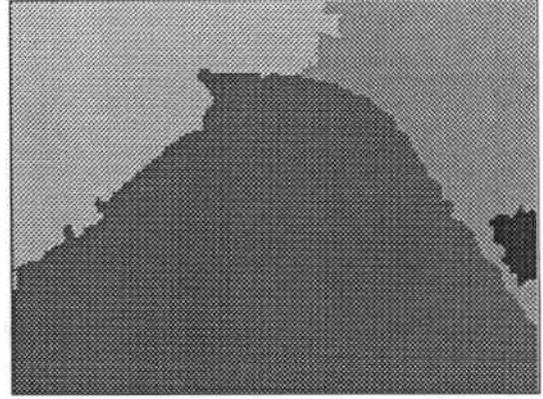
Nach der in Kapitel 2 eingeführten Untergliederung von Straßenszenen gehören die in diesem Kapitel als Beispiele verwendeten Bilder zu Straßen der vierten Klasse. Bei der Analyse von Straßen der vierten Klasse in Kapitel 2 wurde bereits darauf aufmerksam gemacht, daß in solchen Bildern die Diskontinuität in Grauwert oder Farbe zwischen der Straßenoberfläche und -umgebung nicht mehr das signifikante Merkmal ist. Daher erscheinen konturbasierte Verfahren für die Segmentierung solcher Straßenbilder nicht geeignet.

Das Verfahren zur Segmentierung von Straßenbildern durch konturbasierte Bereichsanalyse wird in Anhang E beschrieben. Die Anwendungsergebnisse auf Straßenbilder der ersten und dritten Klasse werden dort ebenfalls dargestellt.

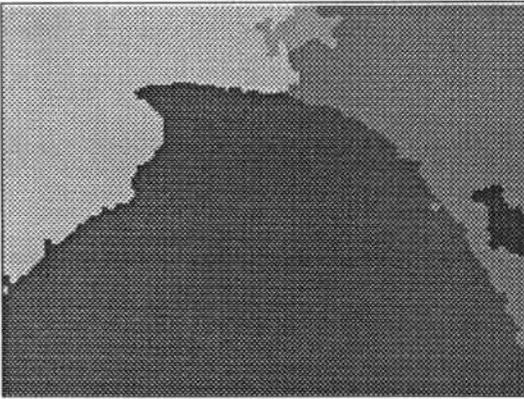
Wir verwenden das in Anhang E beschriebene Verfahren in diesem Abschnitt zur



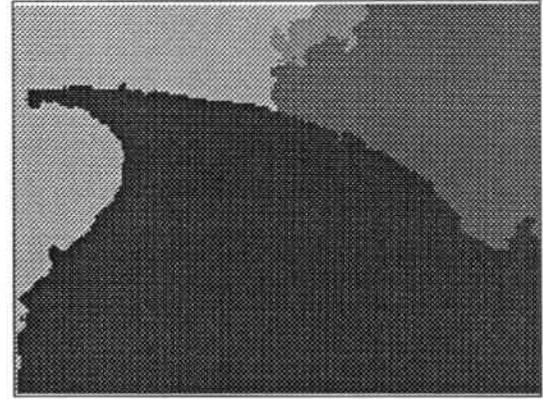
(a) Die Segmentierung von Abb. 6.5 (a).



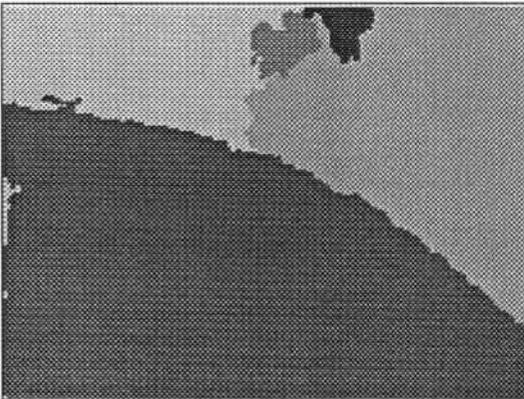
(b) Die Segmentierung von Abb. 6.5 (c).



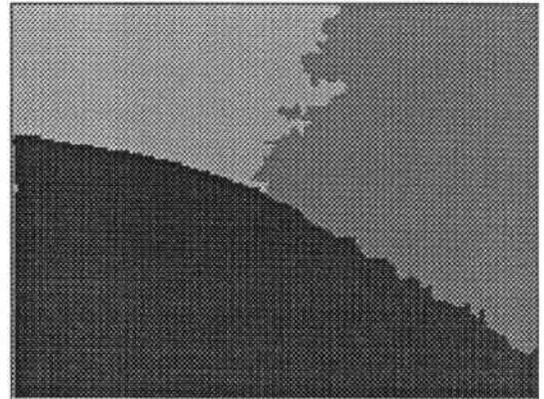
(c) Die Segmentierung von Abb. 6.5 (e).



(d) Die Segmentierung von Abb. 6.5 (g).

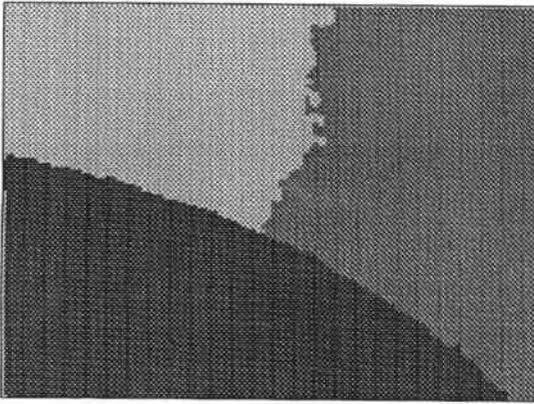


(e) Die Segmentierung von Abb. 6.5 (i).

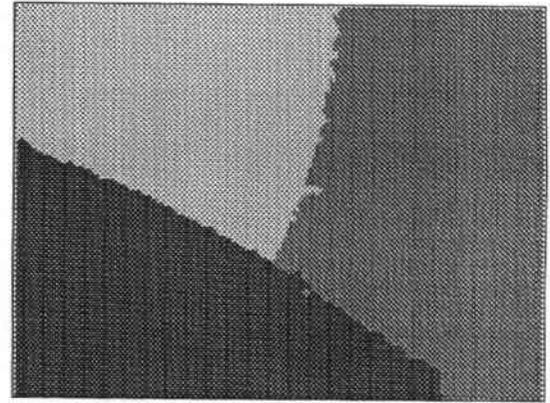


(f) Die Segmentierung von Abb. 6.5 (k).

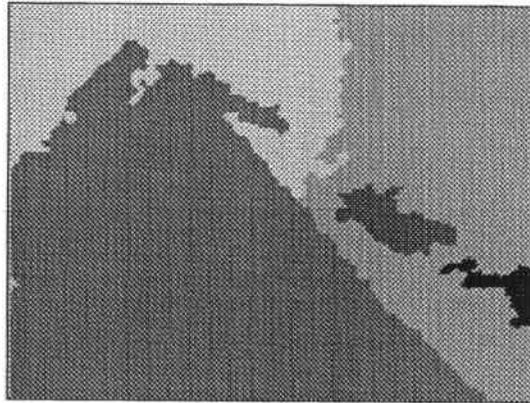
Abbildung 6.8: Ergebnisse der Bereichssegmentierung der in Abb. 6.5 dargestellten Bildfolge mit dem modifizierten Isodata-Algorithmus als Lernverfahren in (a) und dem Bayesschen Klassifikator in (b) - (i). Dargestellt sind das segmentierte Bild der 0., 10., 20., 30., 40., 50., 60., 70. und 80. Aufnahme. Dabei beträgt die Klassenanzahl für alle Bilder 3. Die Grauwerte kennzeichnen hier nicht die Klassen, sondern nur zusammenhängende Bereiche.



(g) Die Segmentierung von Abb. 6.5 (m).



(h) Die Segmentierung von Abb. 6.5 (o).



(i) Die Segmentierung von Abb. 6.5 (q).

(Fortsetzung der vorherigen Seite)

Betrachtungen. Mit diesen Merkmalen kann ein Bild sowohl durch Bereichsklassifikation als auch durch konturbasierte Bereichsanalyse segmentiert werden. Durch die folgende Anordnung von Merkmalen und Verfahren

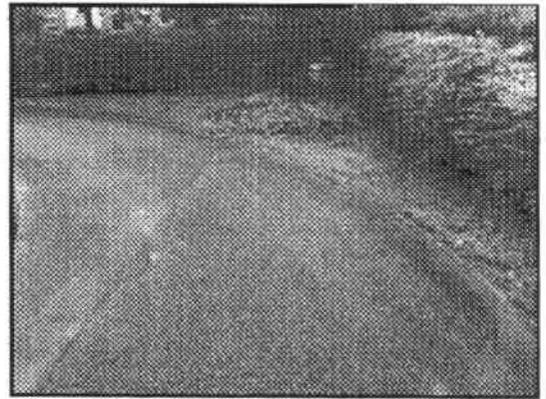
	Texturanisotropie	Grauwert
Bereichsklassifikation	●	○
Konturbasierte Bereichsanalyse	○	○

entstehen vier Möglichkeiten für die Segmentierung eines Bildes: texturbasierte Bereichsklassifikation, grauwertbasierte Bereichsklassifikation, konturbasierte Bereichsanalyse im Grauwertbild und konturbasierte Bereichsanalyse im Bild der Texturanisotropie. Die Segmentierung eines Bildes durch texturbasierte Bereichsklassifikation wurde in den vorherigen Abschnitten bereits dargestellt. Wir zeigen im folgenden die Ergebnisse der anderen drei Verfahren. Das Ziel dieser Vorgehensweise besteht darin, durch den Vergleich der Er-

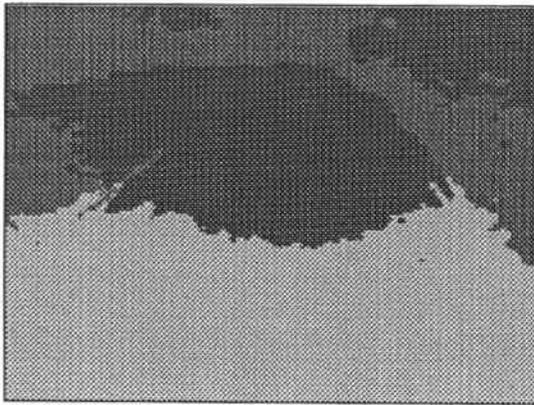
gebnisse verschiedener Verfahren zu zeigen, daß die texturbasierte Bereichsklassifikation bessere Segmentierungsergebnisse liefert.



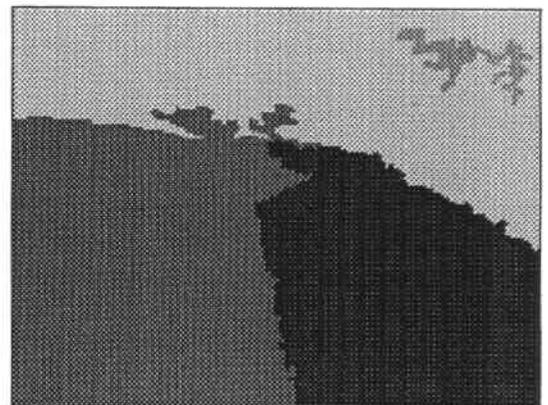
(a) Das Bild von Abb. 6.5 (c).



(b) Das Bild von Abb. 6.5 (i).



(c) Grauwertbasierte Bereichsklassifikation von (a)
(Vergleiche mit Abb. 6.8 (b)).



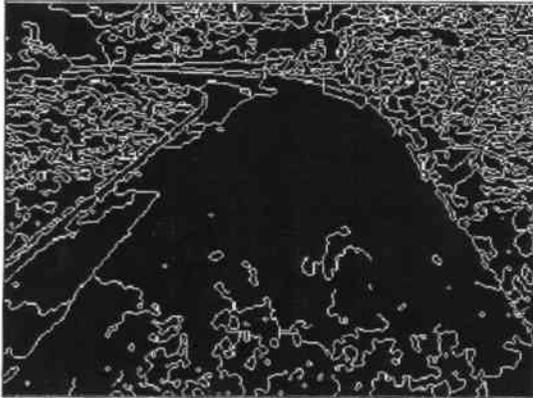
(d) Grauwertbasierte Bereichsklassifikation von (b)
(Vergleiche mit Abb. 6.8 (e)).

Abbildung 6.9: Ergebnisse der grauwertbasierten Bereichsklassifikation von Straßenbildern der vierten Klasse.

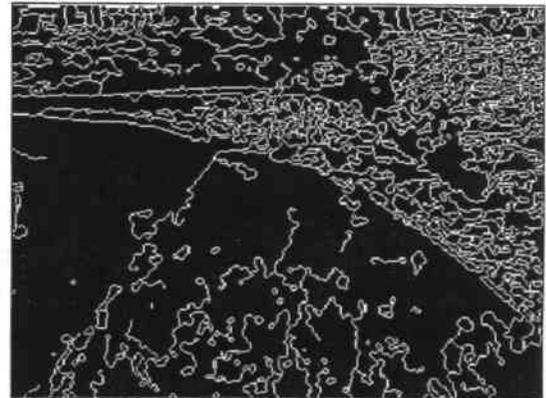
In Abb.6.9 zeigen wir die Ergebnisse der grauwertbasierten Bereichsklassifikation von Straßenbildern der vierten Klasse. Dabei werden $(\text{Grauwert}, \text{Zeile}, \text{Spalte})^T$ als Merkmal und $k = 3$ als Klassenanzahl genommen.

In Anhang E wird ein Verfahren zur Segmentierung von Straßenbildern durch konturbasierte Bereichsanalyse beschrieben. Die Anwendungsergebnisse des Verfahrens auf Straßenbilder der ersten und dritten Klasse werden dort ebenfalls dargestellt. Wir verwenden nun das Verfahren in diesem Abschnitt zur Segmentierung von Straßenbildern der vierten Klasse.

In Abb.6.10 werden die Ergebnisse der Konturverkettung und der Bereichsegmentierung von den Straßenbildern in Abb.6.9 (a) und (b) dargestellt. Wie in Anhang E wird



(a) Konturpunktketten von 6.9 (a).



(b) Konturpunktketten von 6.9 (b).

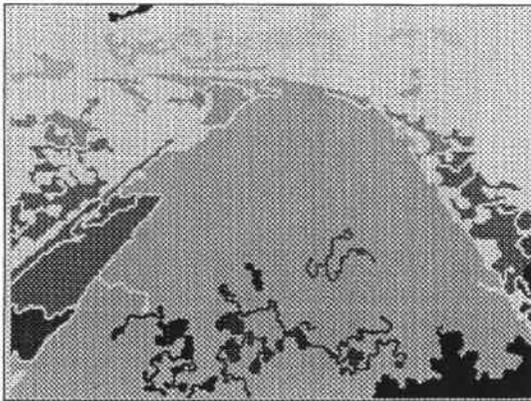
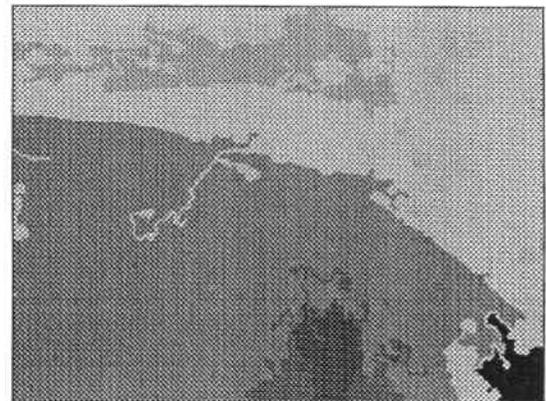
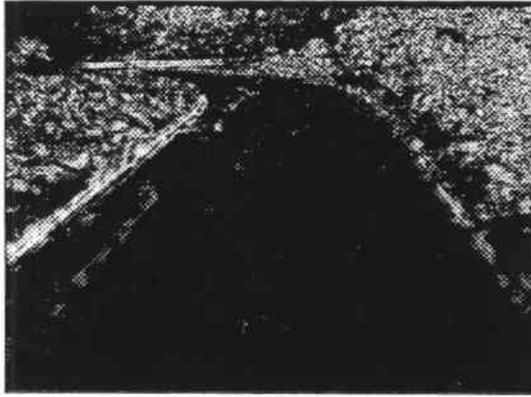
(c) Segmentierung von 6.9 (a)
(Vergleiche mit Abb. 6.8 (b)).(d) Segmentierung von 6.9 (b)
(Vergleiche mit Abb. 6.8 (e)).

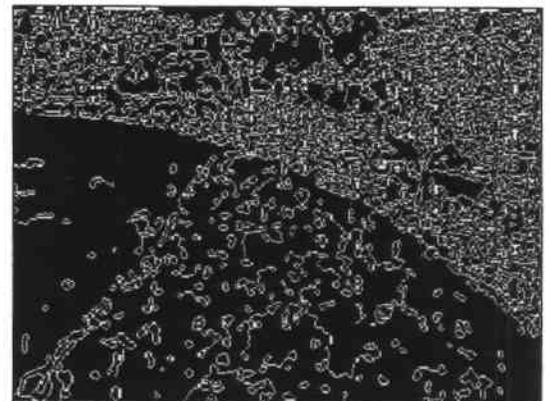
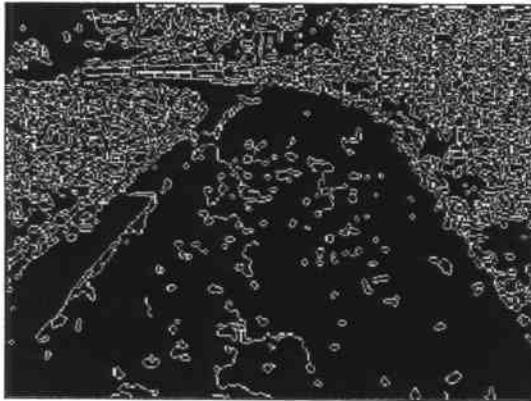
Abbildung 6.10: Ergebnisse der Konturpunktverkettung und der Segmentierung von Straßenbildern der vierten Klasse durch konturbasierte Bereichsanalyse im Grauwertbild.

hier die Schwelle für die Bestimmung der Konturpunkte auf Null gesetzt. Die Schwelle für die Selektion der Kennpunkte im ersten Schritt in Algorithmus *E.1* wird für die Bilder in (a) und (b) jeweils mit 20 und 23 angenommen. Bei der Segmentierung werden diejenigen Bereiche, deren Punktzahl weniger als 1% der gesamten Bildpunkte beträgt, mit ihren benachbarten Bereichen verschmolzen.

Wir zeigen in Abb.6.11 die Ergebnisse der konturbasierten Bereichsanalyse im Bild der Texturanisotropie. Bei der Bestimmung von Konturpunkten wird für die Schwelle Null eingesetzt. Die Schwelle für die Selektion der Kennpunkte wird mit 40 angenommen.

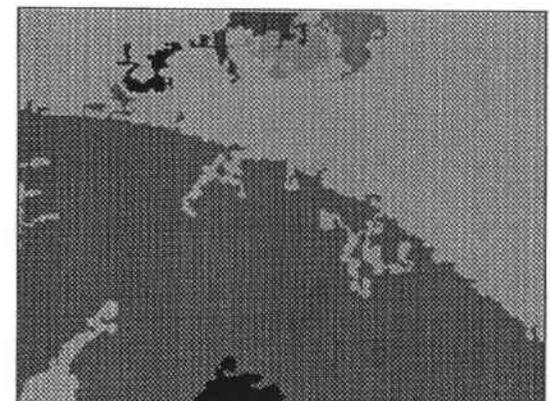
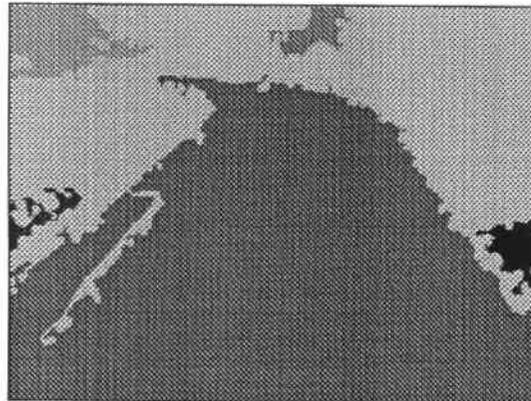


(a) Die Texturanisotropie von Abb. 6.9 (a). (b) Die Texturanisotropie von Abb. 6.9 (b).



(c) Konturpunktketten von (a).

(d) Konturpunktketten von (b).



(e) Segmentierung von (a)
(Vergleiche mit Abb. 6.8 (b)).

(f) Segmentierung von (b)
(Vergleiche mit Abb. 6.8 (e)).

Abbildung 6.11: Ergebnisse der Konturpunktverkettung und der Segmentierung von Straßenbildern der vierten Klasse durch konturbasierte Bereichsanalyse im Bild der Texturanisotropie.

zurück. In Abb. 7.1 wird das Schema dargestellt. Anhand dieses Schemas beschreiben wir im folgenden die in dieser Arbeit bei der Interpretation von Straßenbildern zu lösenden Aufgaben.

Als Bildbereichshinweise werden zunächst aus einem segmentierten Bild der Straßebereich und seine Konturpunkte bestimmt. Diese Bildbereichshinweise können mit Hilfe der bei der Kamerakalibrierung ermittelten Transformationsmatrix in Szenenbereichshinweise überführt werden. Als generische Beschreibung von Straßen wird ein geometrisches Straßenmodell definiert. Durch die Ausprägung des Modells läßt sich die Lagehypothese des Modells generieren. Die Verifikation dieser Hypothese erfolgt iterativ durch die Zuordnung zwischen Modell- und Bildmerkmalen sowie die Bestimmung einer Transformation des Modells.

Zur Lösung dieser Aufgaben für die Straßenerkennung haben wir das System *SeTa*¹ entwickelt. Der Aufbau des Systems wird schematisch in Abb. E.1 dargestellt. Es besteht aus zwei Teilen. Der erste Teil führt zur Straßenerkennung in der Initialisierungsphase. Der zweite Teil behandelt die Straßenerkennung in der Verfolgungsphase. Jede Phase umfaßt zwei wesentliche Schritte:

- Ermittlung von Klassenstatistiken und Klassifikation sowie
- Generierung bzw. Anpassung geometrischer Modellhypothesen.

In der Initialisierungsphase werden die Klassenstatistiken durch eine Ballungsanalyse iterativ berechnet. Die Hypothesen über die Position und Orientierung des geometrischen Modells werden in dieser Phase durch die Auswertung der im Bild ermittelten Daten generiert. Die Erkennung von Straßen in der Verfolgungsphase stützt sich auf die Ergebnisse der Initialisierungsphase. Sowohl die ermittelten statistischen Klassenmodelle als auch die verifizierten geometrischen Modelle werden von der Initialisierungsphase zur Verfolgungsphase sowie in der Verfolgungsphase selbst von einem Laufzyklus in den nächsten übertragen und angepaßt.

7.1 Geometrisches Modell von Straßen

Ein geometrisches Straßenmodell stellt die Erwartung für die geometrische Form der zu erkennenden Straße dar. Diese Beschreibung von Straßengeometrien bildet die Voraussetzung für die modellgestützte Erkennung von Straßen. Der Erkennungsprozeß hängt im allgemeinen stark von der gewählten Repräsentation der Modelle ab. Wir stellen im folgenden die Modellierung von Straßen im Szenenbereich und die Abbildung des verwendeten Modells im Bildbereich vor.

7.1.1 Modellierung von Straßen im Szenenbereich

In Kapitel 2 wird bei der Analyse von Straßen festgestellt, daß die Straßen der vierten Klasse im allgemeinen nicht streng nach den Vorschriften für die Straßenkonstruktion

¹*SeTa: Straßenerkennung durch Texturanalysen.*

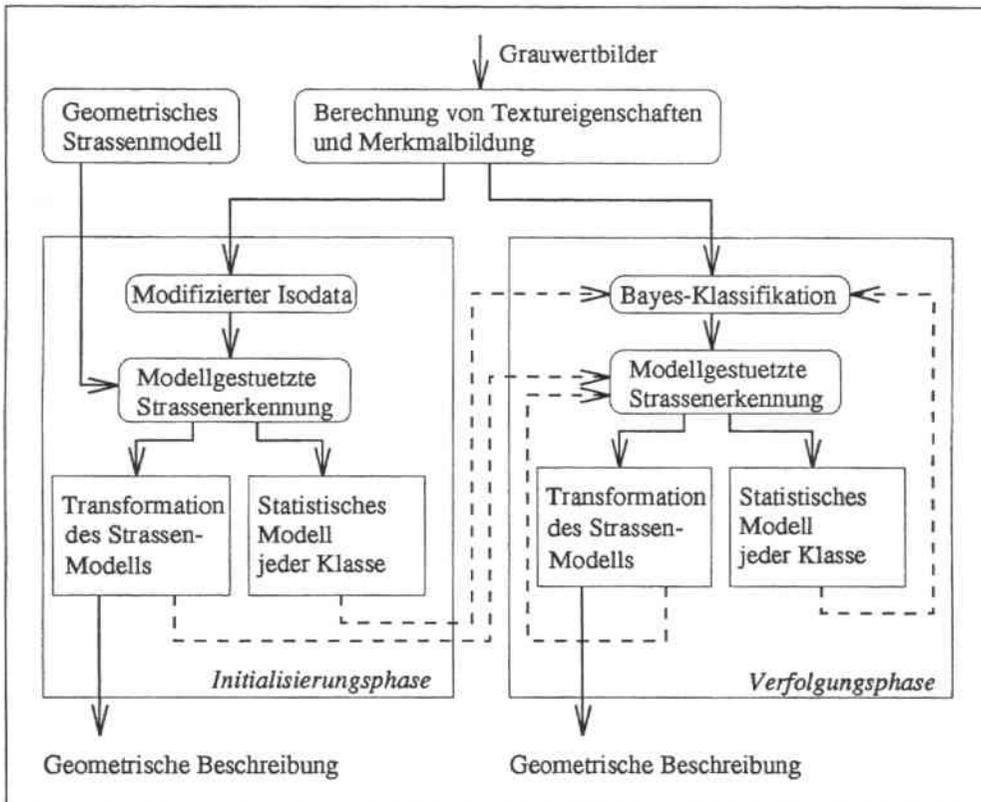


Abbildung 7.2: Aufbau des Systems SeTa zur modellgestützten Erkennung von Straßen der vierten Klasse. Die gestrichelten Linien zeigen die Übertragung der geometrischen bzw. statistischen Modelle von der Initialisierungsphase zur Verfolgungsphase sowie in der Verfolgungsphase von einem Laufzyklus in den nächsten.

gebaut wurden. Zur Modellierung solcher Straßen ist das Ebenen-Modell geeignet, das besagt, daß die betrachtete Straße auf einer Ebene liegt. Die genaue Form der Straße in dieser Ebene läßt sich durch zusätzliche Forderungen bestimmen.

Als geometrisches Straßenmodell übernehmen wir das in Kapitel 2 beschriebene generische Straßenmodell ([Ozawa & Rosenfeld 86]) und beschränken uns in diesem Modell weiter darauf, daß die betrachtete Straße auf einer Ebene liegt. Wir nennen es das GKG²-Modell von Straßen. Es besteht aus zwei Geradensegmenten und einem Kreisbogen. Die beiden Geradensegmenten treffen sich mit dem Kreisbogen jeweils in tangentialer Richtung. In Abb. 7.3 werden das GKG-Modell sowie seine Abbildung im Bildbereich dargestellt.

Die beiden Straßenbegrenzungen in Abb. 7.3 (a) lassen sich durch die folgenden Glei-

²GKG: Geradensegment-Kreisbogen-Geradensegment

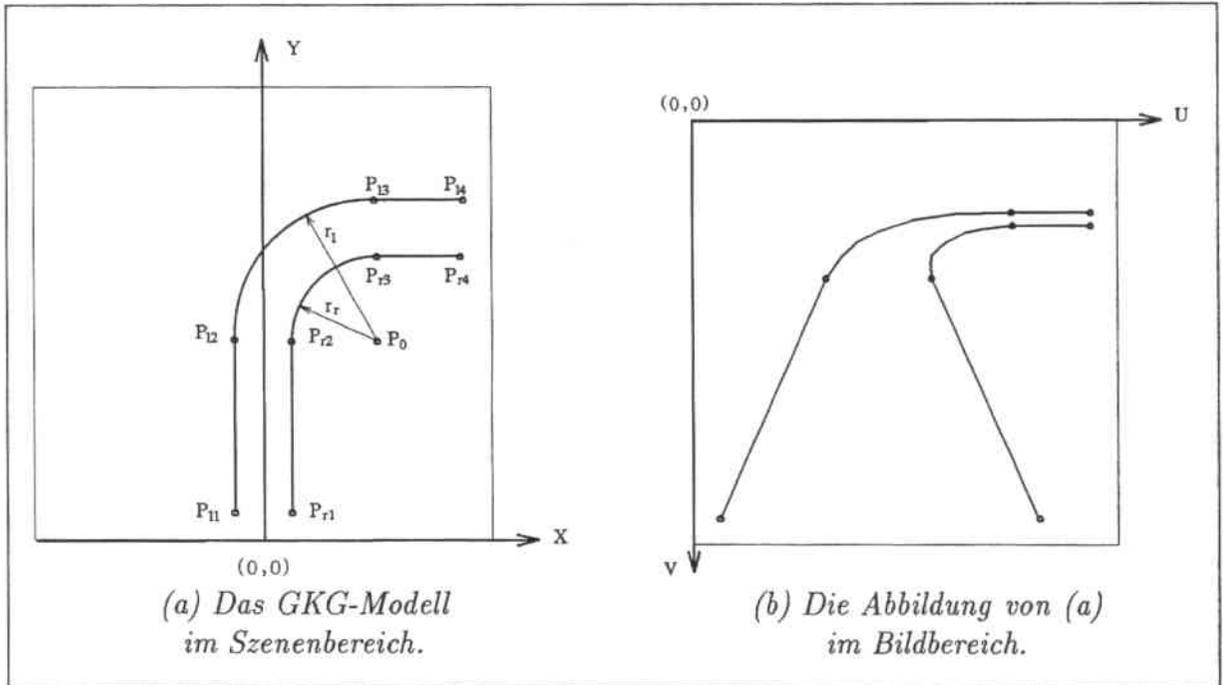


Abbildung 7.3: Darstellung des geometrischen GKG-Modells von Straßen in (a) und seine Abbildung im Bildbereich in (b). Dabei werden die Begrenzungen einer Straße im Szenenbereich als eine Kombination von zwei Geradensegmenten und einem Kreisbogen modelliert. Das GKG-Modell befindet sich in einer Ebene.

chungen (siehe auch Gleichungen (2.1) - (2.3))

$$\begin{cases} y = m_{l1}x + n_{l1} \\ (x - p)^2 + (y - q)^2 = r_l^2 \\ y = m_{l2}x + n_{l2} \end{cases} \quad (7.1)$$

und

$$\begin{cases} y = m_{r1}x + n_{r1} \\ (x - p)^2 + (y - q)^2 = r_r^2 \\ y = m_{r2}x + n_{r2} \end{cases} \quad (7.2)$$

zusammen mit den Positionsparametern $P_0, P_{l1}, P_{l2}, P_{l3}, P_{l4}, P_{r1}, P_{r2}, P_{r3}, P_{r4}$ darstellen. Durch Einschränkung freier Variablen können aus dem GKG-Modell das Fluchtpunkt-Modell und das Kreisbogen-Modell abgeleitet werden. Das GKG-Modell kann allerdings ebenfalls durch die Mittellinie und Breite von Straßen definiert werden.

7.1.2 Abbildung des GKG-Modells im Bildbereich

Die Straßenerkennung und ihre Lagebestimmung lassen sich im allgemeinen sowohl im Szenenbereich als auch im Bildbereich erledigen. In dieser Arbeit geschehen sie im Bildbereich. Daher ist es notwendig, die Abbildung des GKG-Modells im Bild zu bestimmen.

Es ist im allgemeinen bekannt, daß eine perspektivische Projektion ein Geradensegment in ein Geradensegment und einen Kegelschnitt in einen Kegelschnitt abbildet. Als Kegelschnitt kommen hier Ellipse, Parabel und Hyperbel in Betracht. Das bedeutet, daß die Projektion der Kreisbögen des GKG-Modells im Bild einen Abschnitt einer Ellipse, einer Parabel oder einer Hyperbel darstellen kann.

Zur Klassifizierung der Form des abgebildeten Kreisbogens betrachten wir die in Kapitel 6 abgeleitete Koeffizientenmatrix (siehe Gleichung (6.5)) der Abbildung eines Kreises sowie die daraus ermittelten Parameter μ , λ_1 und λ_2 (siehe Gleichungen (6.7) und (6.9)). Diese drei Parameter klassifizieren die Form des mit der Gleichung (6.5) beschriebenen Kegelschnittes in den folgenden drei Kategorien:

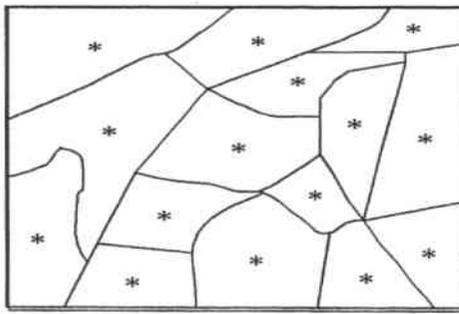
1. Falls $\mu\lambda_1 > 0$ und $\mu\lambda_2 > 0$ sind, stellt der Kegelschnitt eine Ellipse dar.
2. Falls $\lambda_1\lambda_2 < 0$ ist, ist der Kegelschnitt eine Hyperbel.
3. Falls $\lambda_1 = 0$ oder $\lambda_2 = 0$ ist, läßt sich der Kegelschnitt als eine Parabel klassifizieren.

Diese Klassifizierung läßt sich ebenfalls unter der Betrachtung der Position der Kamera relativ zum Kreis formulieren. Dazu legt wir eine Ebene parallel zum Kreisbogen durch den Augenpunkt der Kamera und projizieren das Kreisbogen parallel in dieser Ebene. Befindet sich der Augenpunkt der Kamera auf dem Kreis, außerhalb oder innerhalb des Kreises, wird dann in der Bildebene entsprechend eine Parabel, eine Ellipse oder eine Hyperbel abgebildet.

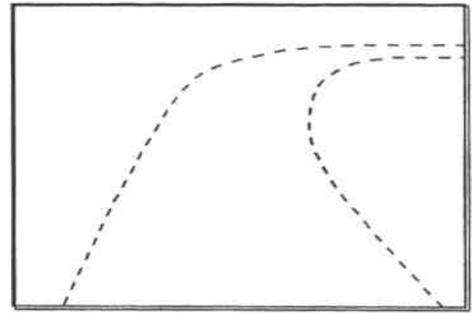
7.2 Bestimmung des Straßenbereiches

Die nächste Aufgabe im System *SeTa* besteht in der Bestimmung von Bereichen im segmentierten Bild, die zur Straße gehören. Bei der Straßenverfolgung wird vorausgesetzt, daß Straßen keine großen physikalischen bzw. geometrischen Änderungen zwischen zwei unmittelbar aufeinanderfolgenden Aufnahmen aufweisen. Dieses *A-Priori*-Wissen ist sehr hilfreich für die Bestimmung des Straßenbereiches. Mit dieser Voraussetzung läßt sich der Suchbereich für die Verfolgung der Straße auf kleine Bereiche beschränken.

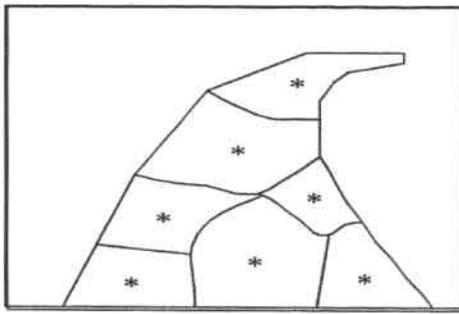
Die Vorgehensweise zur Bestimmung des Straßenbereiches aus einem segmentierten Bild wird in Abb. 7.4 graphisch mit den skizzierten Bildern gezeigt. Unsere Ausgangsdaten sind das klassifizierte Bild und die Projektion der Lagehypothese der Straße im Bildbereich. Sie werden jeweils in Abb. 7.4 (a) und (b) dargestellt. Die Lagehypothese der Straße entsteht eigentlich auf zwei verschiedene Weisen im *SeTa*. Zum einen wird sie in der Initialisierungsphase mit Hilfe von Kenntnissen über die betrachtete Szene (mit den *A-Priori*-Startwerten) definiert. Zum anderen läßt sie sich in der Verfolgungsphase direkt aus dem in der vorherigen Szene ausgeprägten Modell herleiten.



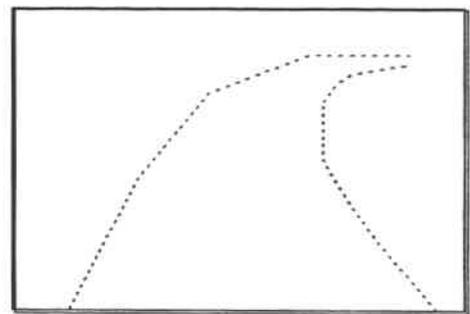
(a) Ein segmentiertes Bild.



(b) Hypothesenbereich.



(c) Selektierte Bereiche.



(d) Konturpunkte.

Abbildung 7.4: Bestimmung des Straßenbereiches im Bild: (a) das segmentierte Bild; (b) die Lagehypothese der Straße im Bildbereich; (c) die selektierten Kandidatenbereiche für die Straße und (d) die Konturpunkte des Straßenbereichs.

Im folgenden wird die in Abb. 7.4 dargestellte Bestimmung des Straßenbereiches in einzelnen beschrieben:

- Erste Stufe dabei ist, diejenigen Bereiche herauszufinden, deren Mittelpunkt innerhalb des Hypothesenbereichs liegt.
- Aufgrund der Tatsache, daß die im Merkmalsraum in einer Teilmenge klassifizierten Punkte im 2D-Bild nicht mehr zusammenhängend sein müssen, soll eine zusätzliche Prüfung der Kandidatenbereiche durchgeführt werden. Dies erfolgt durch den Vergleich des Grauwerts am Mittelpunkt jedes Kandidatenbereichs im Bild mit dem entsprechenden Mittelwert im Merkmalsraum.
- In der dritten Stufe untersuchen wir die geometrische Beziehung zwischen den Kandidatenbereichen. Liegt ein Bereich innerhalb der bisher als Straße vermuteten Straßenfläche, dann betrachten wir diesen Bereich als einen Teil der Straße. Dies setzt voraus, daß wir zunächst keine Hinderniserkennung durchführen. Ansonsten müßten noch zusätzliche Bedingungen wie z.B. Unterschiede in Grauwert und Textur zwischen den Bereichen geprüft werden, um Abbildungen von Hindernissen von

der Abbildung der Straße zu trennen.

- In der letzten Stufe werden die als Straße ausgewählten Bereiche miteinander verschmolzen.

Die Bestimmung der Konturpunkte des Straßenbereiches erfolgt durch binäre morphologische Operationen. Die Grundlagen der Morphologie zur Bildauswertung findet man in [Giordina & Dougherty 88] und [Haralick & Shapiro 92]. In Abb. 7.4 (d) werden die Konturpunkte im Bildbereich dargestellt. Als Bildmerkmale stehen die Konturpunkte der Modellanpassung zur Verfügung.

7.3 Schätzung der Parameter des GKG-Modells

Bisher werden die Konturpunkte des Straßenbereiches im Bild bestimmt. Bei der modellgestützten Erkennung von Straßen bestehen die nächsten Aufgaben in der Zuordnung der extrahierten Straßenmerkmale zu den projizierten Modellmerkmalen und in der Verifikation der Lagehypothese des Modells.

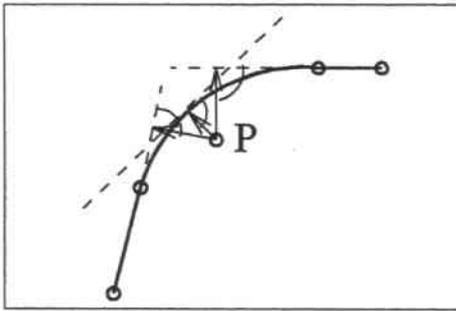
Die Merkmalszuordnung und Lagebestimmung werden in der vorliegenden Arbeit im Bildbereich durchgeführt. Der Vorteil dafür liegt darin, daß sich die projizierten Modellmerkmale direkt mit den aus dem Bild extrahierten Merkmalen vergleichen lassen und man damit eine hohe Genauigkeit erreichen kann.

7.3.1 Merkmalszuordnung

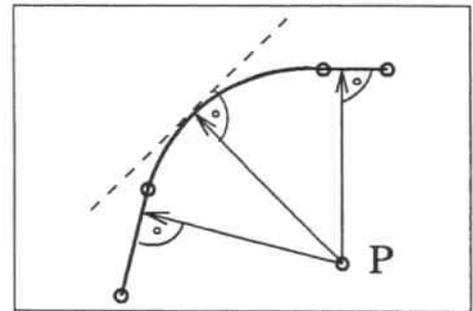
Wir haben gezeigt, daß die Abbildung der Kreisbögen des GKG-Modells im Bild eine der drei möglichen Formen (Ellipse, Parabel und Hyperbel) darstellt. Nun kann es zwar gelingen, aus den Konturpunktketten der Straßenränder Abschnitte von Kegelschnitten zu gewinnen. Jedoch führt das örtliche Rauschen der Konturpunkte zu eher unsicheren Aussagen. Abgesehen von der Ungenauigkeit der so gewonnenen Parameter kann es leicht geschehen, daß aus den Bildwerten z.B. eine Hyperbel ermittelt wird, während die Parabel der passende Kegelschnitt wäre. Aus diesem Grund bevorzugen wir eine modellgestützte Vorgehensweise und passen projizierte Hypothesen über die Gestalt der Fahrbahnränder an die gemessenen Konturpunkte an.

Bei der Ermittlung der Konturpunkte des Straßenbereiches werden diese bereits der linken bzw. rechten Straßenbegrenzung zugeordnet. Die nächste Aufgabe besteht darin, die Konturpunkte jeweils zu einem der projizierten Geradensegmenten oder dem projizierten Kreisbogen zuzuteilen.

Dazu ermitteln wir den jeweils kürzesten Abstand jedes Punktes zu allen drei Modellsegmenten, bzw. ihren Verlängerungen (siehe Abb. 7.5). Anschließend überprüfen wir die Positionen der Schnittpunkte der Linien der kürzesten Abstände mit den Modellsegmenten, ob sie innerhalb der entsprechenden Segmente liegen. Zwei Möglichkeiten lassen sich dabei unterscheiden:



(a) Eine Zuordnung nach der Position des Schnittpunktes.



(b) Eine Zuordnung nach dem euklidischen Abstand.

Abbildung 7.5: Zuordnung eines im Bild extrahierten Konturpunktes zu einem der drei Segmente des projizierten GKG-Modells nach der Position des Schnittpunktes in (a) sowie nach dem euklidischen Abstand in (b).

1. Liegt nur ein Schnittpunkt innerhalb des entsprechenden Segmentes, wird dieser Konturpunkt diesem Segment zugeordnet. Der euklidische Abstand des Punktes zum Segment wird dann berechnet.
2. Sind zwei oder alle drei Schnittpunkte innerhalb der entsprechenden Segmente, müssen in diesem Fall die euklidischen Abstände des Konturpunktes zu allen Segmenten zuerst berechnet werden. Der Punkt wird dem Segment zugeordnet, von dem er den kleinsten Abstand hat.

Bei diesem Prozeß werden nur diejenigen Abschnitte der Abbildungen des Modells herangezogen, die innerhalb des Bildes liegen.

7.3.2 Bestimmung der Modelltransformation

Die bei der Merkmalszuordnung berechneten Abstände zwischen jedem Konturpunkt und den projizierten Modellmerkmalen werden als die Positionsabweichung für die Verifikation der Lagehypothese des Modells verwendet. Wenn der Mittelwert der Abweichungen unter einer vorgegebenen Schwelle liegt, wird die Lagehypothese angenommen, ansonsten wird das Modell im Szenenbereich entsprechend transformiert und die Abweichung erneut im Bildbereich berechnet. Dieser Prozeß wird iterativ durchgeführt. Gesucht wird hier die Transformation des Modells, mit der die Positionsabweichung zwischen den Konturpunkten und den projizierten Modellmerkmalen minimal ist. So läßt sich das Problem auch als eine Minimierungsaufgabe formulieren.

Da das GKG-Modell in einer Ebene definiert ist, haben wir drei unabhängige freie Variablen für die Modelltransformation zu schätzen, nämlich die Translation t_x und t_y sowie den Rotationswinkel β . Das Minimierungsverfahren soll für diese drei Variablen die Werte finden, mit denen der Mittelwert der Abweichungen ein Minimum hat.

Für die Minimierung einer Fehlerfunktion bieten sich verschiedene Verfahren. Die Wahl eines geeigneten Verfahrens hängt im allgemeinen von der Anzahl der unabhängigen Variablen und der Komplexität der Fehlerfunktion ab. Die Anzahl der unabhängigen Variablen bestimmt die Dimension des Minimierungsverfahrens. Unter der Berücksichtigung der Komplexität der Fehlerfunktion soll dasjenige Verfahren ausgewählt werden, mit dem der Rechenaufwand zur Ermittlung der Funktionswerte bzw. auch der Funktionsgradienten klein bleibt.

Wir haben in unserer Minimierungsaufgabe drei freie Variablen. Für jede davon gibt es auch einen begrenzten Wertebereich. Für die bedingte, dreidimensionale Minimierung der Positionsabweichung des GKG-Modells verwenden wir das sogenannte *Simplex*-Verfahren. Ein *Simplex* ist eine n -dimensionale Figur bestehend aus $n+1$ n -dimensionalen Punkten. Für drei Variablen ist das *Simplex* ein Tetraeder. Für einen Anfangsvektor \vec{p}_0 lassen sich die anderen n Vektoren wie folgt bestimmen:

$$\vec{p}_i = \vec{p}_0 + \lambda \vec{e}_i, \quad (7.3)$$

wobei λ eine Konstante und \vec{e}_i einen Einheitsvektor bezeichnet. Das *Simplex*-Verfahren führt eine iterative Optimierung durch. In jedem Schritt wird der Vektor \vec{p}_i , der die Fehlerfunktion maximiert, in seinem i -ten Wert verändert. Das Optimierungsverfahren wird solange fortgesetzt, bis die Abweichung zwischen dem Vektor mit dem höchsten und dem mit dem niedrigsten Fehler unterhalb einer Schwelle liegt oder die Anzahl der Berechnungen eine Grenze übersteigt.

Wie in Gleichung (7.3) zu sehen ist, hängt die Länge der Suche nach dem Minimum stark von der Wahl des Startvektors \vec{p}_0 ab. Wir nehmen die Transformation für die Lagehypothese des Modells als diesen Startvektor. Die Wahl der Lagehypothese läßt sich im System *SeTa* in der Initialisierungs- und Verfolgungsphase auf zwei verschiedene Weisen erledigen. In der Verfolgungsphase wird die in der vorherigen Szene ermittelte Lage als die Lagehypothese übernommen. In der Initialisierungsphase läßt sie sich aus dem segmentierten Bild bestimmen.

7.3.3 Initialisierung der Modellposition

Die Schätzung der Anfangsposition des GKG-Modells in der Initialisierungsphase des Systems *SeTa* wird im Szenenbereich durchgeführt. Ein erster Schritt besteht in der Rückprojektion der Konturpunkte des Straßenbereiches in den Szenenbereich. Im Szenenbereich wollen wir dann den Übergangspunkt zwischen dem ersten Geradensegment und dem Kreisbogen des GKG-Modells durch die Auswertung der lokalen Krümmung der Konturpunkte lokalisieren. Die Lokalisierung des Übergangspunktes dient dazu, die Anfangsposition des Modells zu ermitteln und damit eine automatische Erkennung von Straßen in der Initialisierungsphase zu ermöglichen.

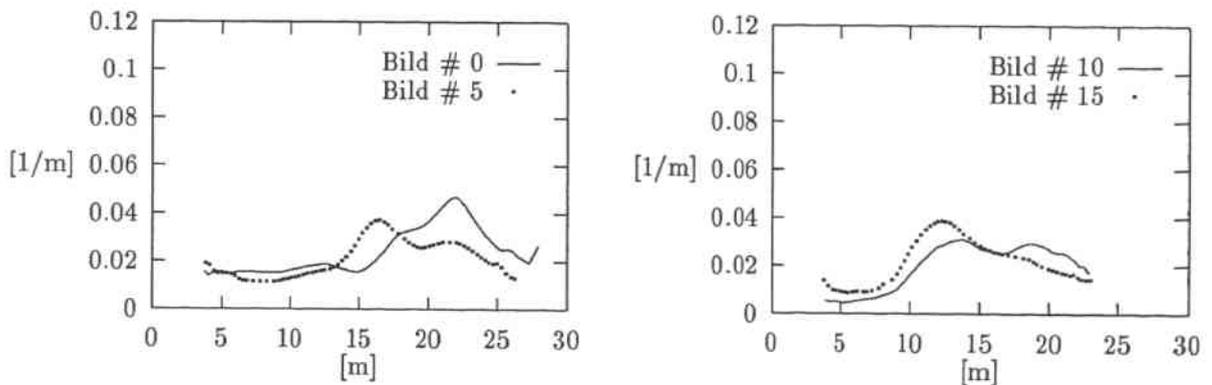
Unter der Annahme, daß sich unmittelbar vor dem kameratragenden Fahrzeug ein befahrbares, d.h. freies und im Bild sichtbares Stück der Straße befindet, bestimmen wir zunächst die Mittelpunkte zwischen den beiden Begrenzungen von nahe nach fern. Wegen

der Ungenauigkeit der Kamerakalibrierung werden die Punkte, die weiter als 30 m vom Fahrzeug entfernt sind, bei der Schätzung der Krümmung nicht berücksichtigt.

Um Rauschen zu unterdrücken, werden anschließend die X-Koordinaten der Mittelpunkte mit einer eindimensionalen Gaußfunktion geglättet. Die Berechnung der lokalen Krümmung der Konturpunkte erfolgt durch die folgende Gleichung:

$$\frac{d^2x/dy^2}{(1 + (dx/dy)^2)^{3/2}} \quad (7.4)$$

In Abb. 7.6 werden Beispiele der ermittelten Krümmungen aus vier Bildern gezeigt. Die Bilder werden während der Einfahrt in eine Kurve mit einer Kamera aufgenommen. Auf der horizontalen Achse ist die Entfernung der ermittelten Punkte der Straßen-



(a) Krümmung aus dem 0. und 5. Bild. (b) Krümmung aus dem 10. und 15. Bild.

Abbildung 7.6: Beispiele der ermittelten lokalen Krümmung der Straßenmittelpunkte aus der 0., 5., 10. und 15. Aufnahme einer Bildfolge. Auf der horizontalen Achse ist die Entfernung der ermittelten Punkte der Straßenmittellinie von der Kamera aufgetragen. Die vertikale Achse stellt die ermittelte lokale Krümmung dar.

mittellinie von der Kamera aufgetragen. Die vertikale Achse stellt die ermittelte lokale Krümmung dar. Im 0. Bild befindet sich das Maximum der lokalen Krümmung ca. 22 m entfernt vor dem Fahrzeug. In den nachfolgenden Bildern kommt das Maximum schrittweise näher auf die Kamera zu. Der Vorausschauabstand der Kamera beträgt dabei 4 m.

Zur Lokalisierung des Übergangspunktes werden die berechneten Krümmung mit einer Schwelle verglichen. Die Punkte, deren Krümmung unterhalb dieser Schwelle liegt, werden dem Geradensegment zugeteilt. Es gibt dabei zwei Spezialfälle: Liegt die Krümmung aller Punkte unterhalb der Schwelle, wird angenommen, daß die Straße vor dem Fahrzeug geradlinig ist; Ist die Krümmung der unmittelbar vor dem Fahrzeug liegenden Punkte

größer als diese Schwelle, wird davon ausgegangen, daß das Fahrzeug sich bereits in der Kurve befindet.

Aus der ermittelten Position des Übergangspunktes läßt sich die Translation des Modells bestimmen. Unter der Annahme, daß die Straßenmittellinie parallel zur Hauptachse des Fahrzeugs steht, wird die Rotation des Modells bei der Schätzung seiner Anfangsposition vernachlässigt.

7.4 Beispiele von Straßenerkennungsergebnissen

Wegen des Zeitaufwandes bei der Berechnung der Texturanisotropie sind wir zur Zeit noch nicht in der Lage, das System *SeTa* in unserem Versuchsträger für die Echtzeiterkennung von Straßen zu implementieren. Zur Erprobung des Systems für die Straßenerkennung in der Verfolgungsphase haben wir Straßenszenen im Institutsgelände (IITB) in unserem Fahrzeug aufgezeichnet und als Bildfolge digitalisiert. Anschließend wird die Texturanisotropie berechnet und der Erkennung von Straßen im System *SeTa* zur Verfügung gestellt.

Die Ausgangsdaten für die Erstellung des geometrischen Modells sind die aus dem *Travelpilot* extrahierten Beschreibungen desselben Straßenabschnittes im Institutsgelände. Die Daten im *Travelpilot* sind mit einem eigenen Format beschrieben. Mittels eines Konvertierungsprogrammes werden die Daten in einer Form gebracht, in der die Länge, Richtung und evtl. auch die Krümmung der Mittellinie des Straßenstückes zwischen zwei benachbarten Knoten gegeben wird. In Tabelle 7.1 werden die konvertierten Daten wiedergegeben. Der Verlauf der Straßenabschnitte zwischen 2. und 4. Knoten ist eigentlich

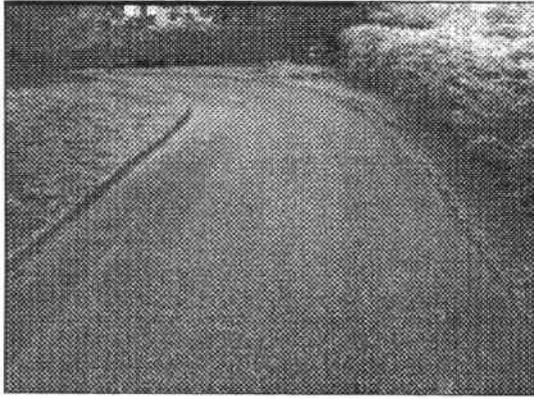
Knoten	Abstand (m)	Winkel (Grad)
1		0.0
2	25.6	-27.3
3	9.2	-39.5
4	7.5	-23.2
5	63.3	

Tabelle 7.1: Die geometrische Beschreibung der Mittellinie eines Straßenabschnittes im Institutsgelände (IITB).

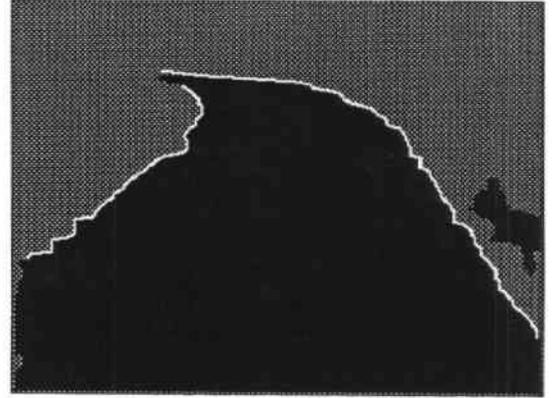
ein Kreisbogen. Wegen des geringen Radius wird er nur mit drei Punkten dargestellt.

Aus den Daten in Tabelle 7.1 lassen sich die Parameter des GKG-Modells ermitteln. Die Straßenbreite wird dabei mit 4 m angenommen. Damit stehen alle geometrischen Kenngrößen der Straße in einem Datensatz dem Erkennungsprozeß zur Verfügung.

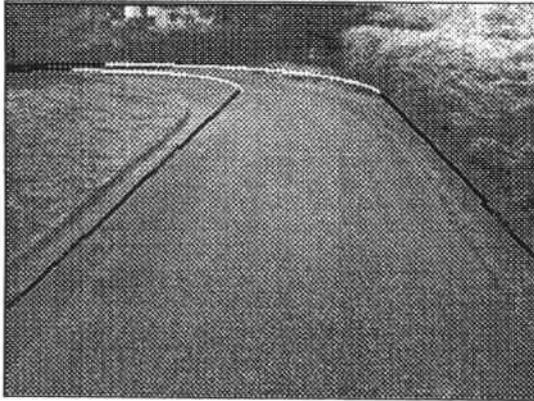
In Abb. 7.7 zeigen wir ein Straßenbild und den bestimmten Straßenbereich, die Konturpunkte des Bereiches, die Lagehypothese des geometrischen Modells, die geschätzte Position der Straßenbegrenzungen und die aktuelle Lage des Fahrzeugs im Bezug auf das geometrische Modell. Die Darstellung der Zwischenergebnisse sollte die Vorgehensweise



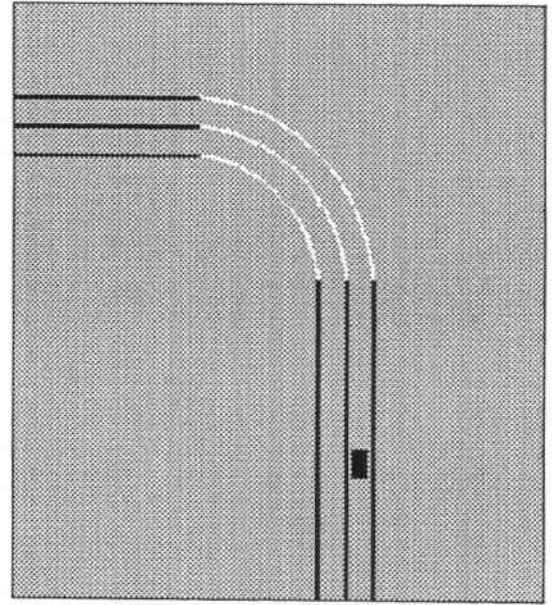
(a) Ein Straßenbild.



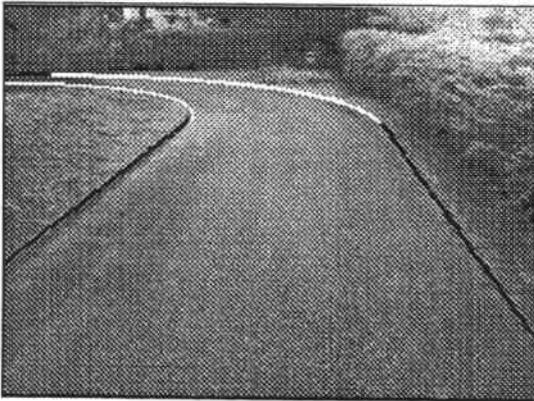
(b) Der bestimmte Straßenbereich.



(c) Die Lagehypothese.

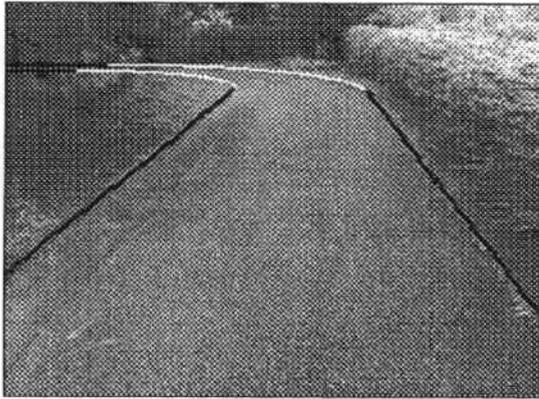


(e) Die aktuelle Position
des Fahrzeugs.

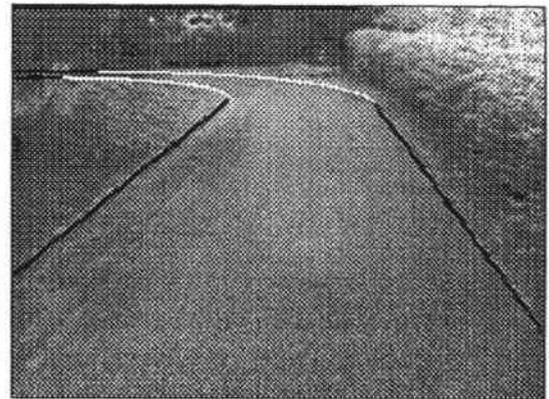


(d) Die ermittelte Straßenbegrenzung.

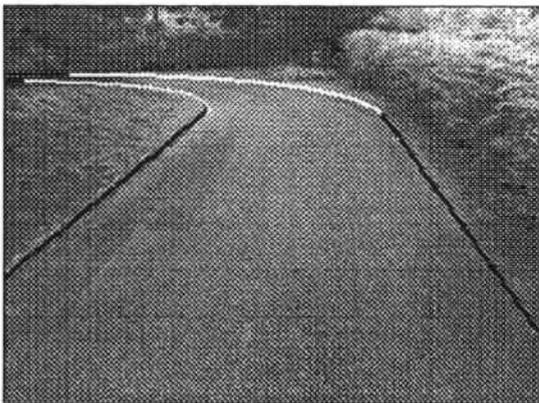
Abbildung 7.7: Darstellung der Zwischenergebnisse bei der modellgestützten Erkennung einer Straße: (a) das Grauwertbild; (b) der selektierte Straßenbereich sowie seine Konturpunkte; (c) die Lagehypothese im Bild; (d) die verifizierte Position der Straße und (e) die aktuelle Lage des Fahrzeugs.



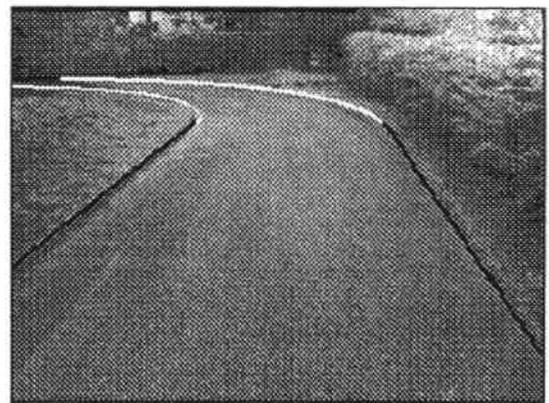
(a) Straßenerkennung im 0. Bild.



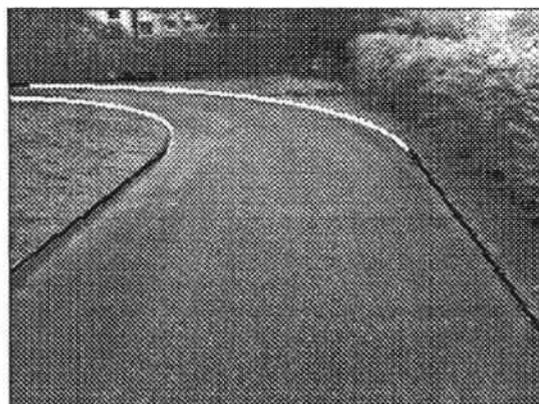
(b) Straßenerkennung im 5. Bild.



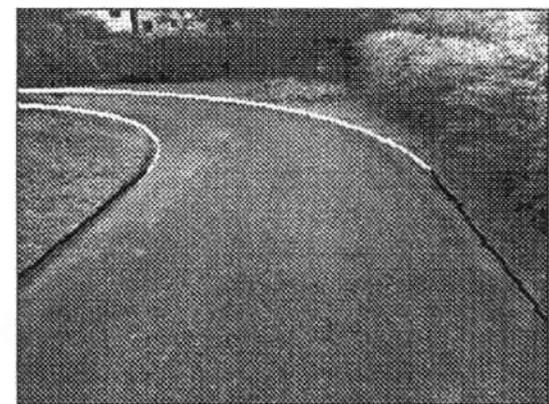
(c) Straßenerkennung im 10. Bild.



(d) Straßenerkennung im 15. Bild.

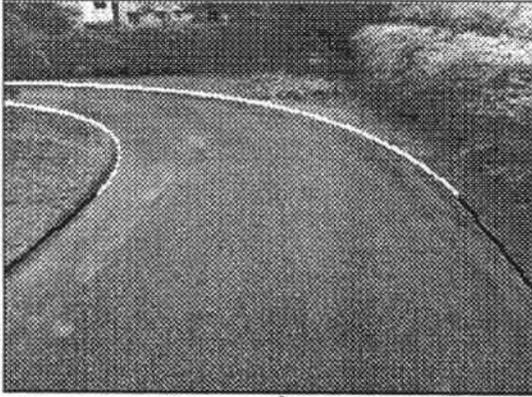


(e) Straßenerkennung im 20. Bild.

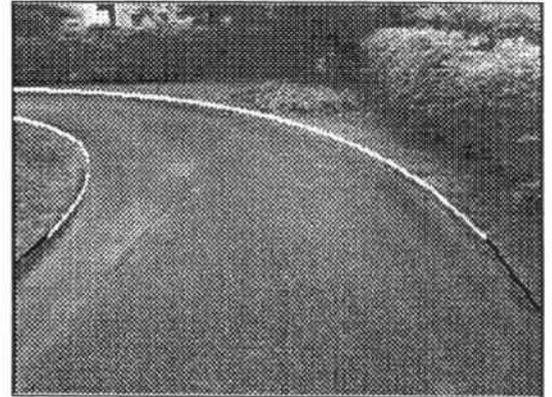


(f) Straßenerkennung im 25. Bild.

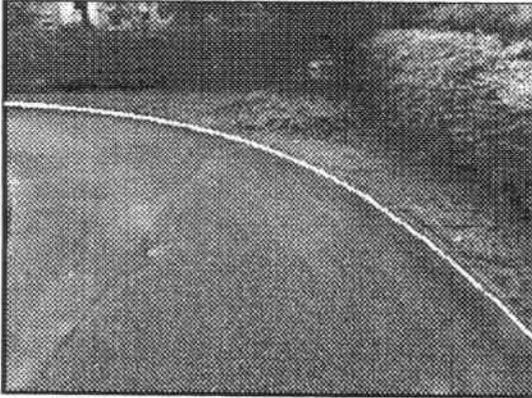
Abbildung 7.8: Ergebnisse der modellgestützten Erkennung von Straßen in der 0., 5., 10., 15., 20., 25., 30., 35., 40. und 45. Aufnahme.



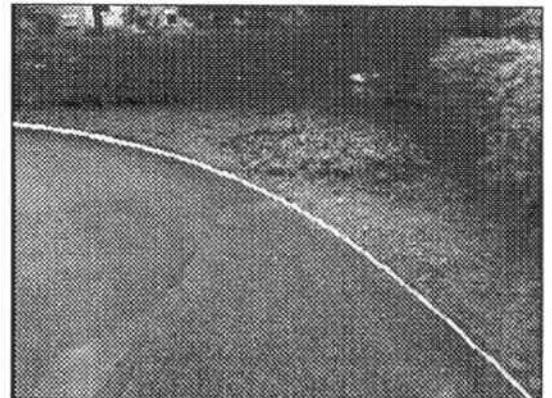
(g) Straßenerkennung im 30. Bild.



(h) Straßenerkennung im 35. Bild.



(i) Straßenerkennung im 40. Bild.



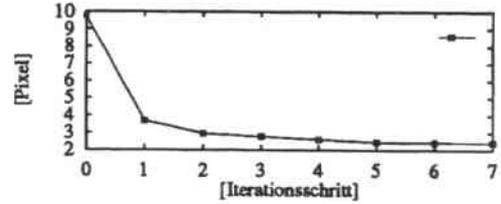
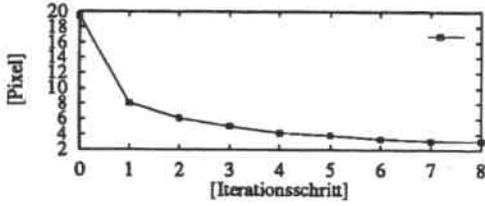
(j) Straßenerkennung im 45. Bild.

(Fortsetzung der vorherigen Seite)

des Verfahrens besser verständlich machen.

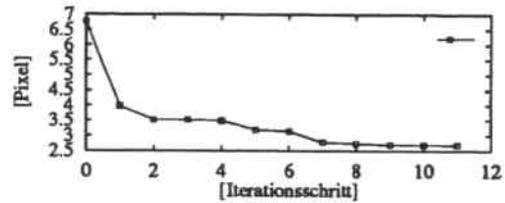
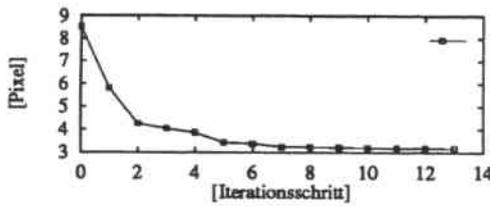
Die beiden Kreisbögen des GKG-Modells in der mit Hand ausgewählten Hypothese-
senlage werden als Ellipse in Abb. 7.7 (c) abgebildet. Nach der Verifikation dieser La-
gehypothese mit den in (b) dargestellten Konturpunkten des Straßenbereiches läßt sich
die Position der Straßenbegrenzungen durch die Transformation des Modells gewinnen.
Sie wird in (d) dargestellt. Dabei wird der Kreisbogen der linken Begrenzung zu einer
Ellipse, der Kreisbogen der rechten zu einer Hyperbel abgebildet. Wie in (d) zu sehen
ist, reicht diese mit der modellgestützten Straßenerkennung erreichte Positions-
genauigkeit für die Straßenverfolgung aus. Die Abweichung zwischen Hypothese und End-
position ist $(-0.458\text{ m}, -4.981\text{ m}, -0.126\text{ Grad})$. Entsprechend dieser geschätzten
Transformation des Modells ergibt sich die aktuelle Lage des Fahrzeugs. Das Fahrzeug
wird in (e) mit einem rechteckigen schwarzen Fleck gezeigt.

Wir zeigen in Abb. 7.8 die Ergebnisse der Straßenerkennung aus der obengenannten



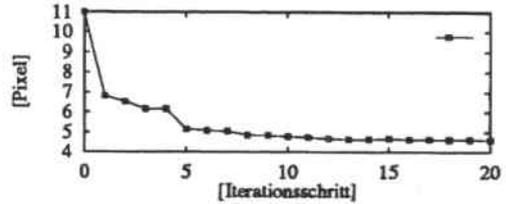
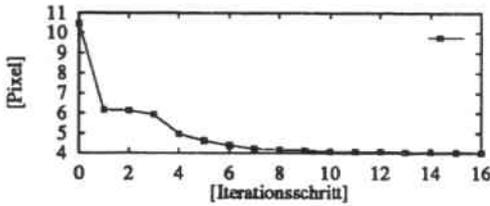
(a) Positionsabweichung in Abb. 7.8 (a).

(b) Positionsabweichung in Abb. 7.8 (b).



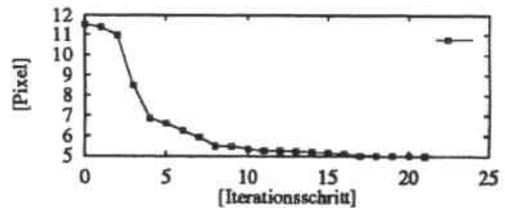
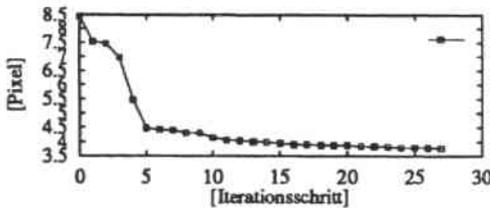
(c) Positionsabweichung in Abb. 7.8 (c).

(d) Positionsabweichung in Abb. 7.8 (d).



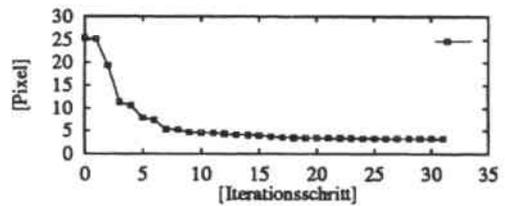
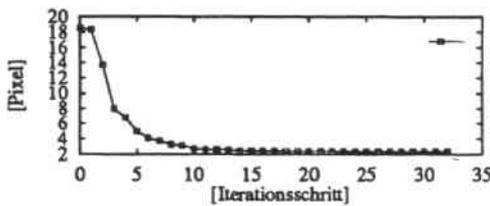
(e) Positionsabweichung in Abb. 7.8 (e).

(f) Positionsabweichung in Abb. 7.8 (f).



(g) Positionsabweichung in Abb. 7.8 (g).

(h) Positionsabweichung in Abb. 7.8 (h).



(i) Positionsabweichung in Abb. 7.8 (i).

(j) Positionsabweichung in Abb. 7.8 (j).

Abbildung 7.9: Die mittlere Positionsabweichung zwischen den im Bild bestimmten Konturpunkten und den projizierten Modellmerkmalen in jedem Iterationsschritt des Minimierungsverfahrens.

Bildfolge. Das Bild in (a) zeigt das Ergebnis der Straßenerkennung im Bild in der Initialisierungsphase. In (b) - (j) werden die Ergebnisse in der Verfolgungsphase dargestellt.

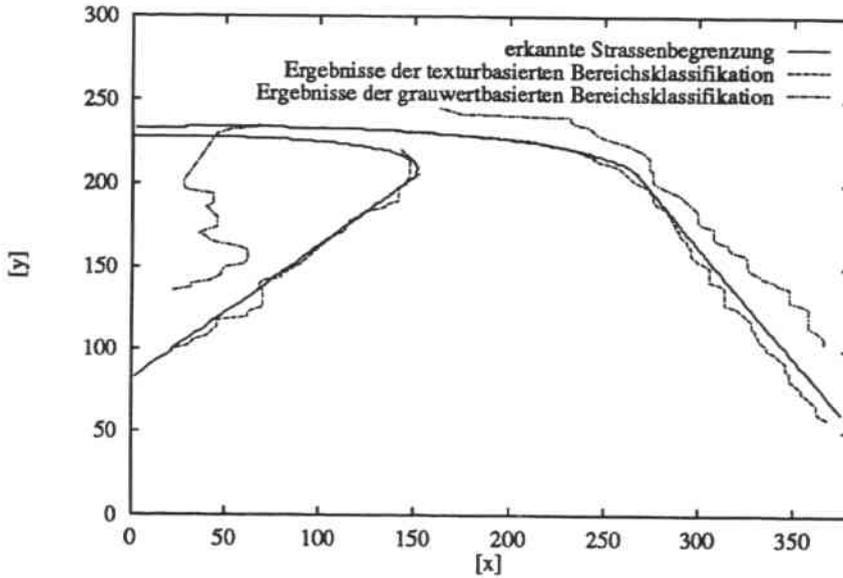
Bei dem oben dargestellten Erkennungsprozeß wurden die Abstände zwischen jedem im Bild bestimmten Konturpunkt und den projizierten Modellmerkmalen als die Positionsabweichung für die Verifikation der Lagehypothese des Modells verwendet. Mit Hilfe eines Minimierungsverfahrens wurden für die freien Variablen der Modelltransformation diejenigen Werte gefunden, mit denen die mittlere Positionsabweichung ein Minimum ist. In Abb. 7.9 zeigen wir die mittlere Positionsabweichung in jedem Iterationsschritt für die in Abb. 7.8 dargestellten Bilder. Auf der horizontalen Achse sind die Iterationsschritte des Minimierungsverfahrens aufgetragen. Dabei beträgt die Schwelle für die Abweichung zwischen dem Vektor mit dem höchsten und dem mit dem niedrigsten Positionsfehler im Minimierungsverfahren 0.005 für alle Bilder.

In Abschnitt 6.3.5 wurde gezeigt, daß ein Straßenbild durch vier verschiedene Verfahren segmentiert werden kann. Das System *SeTa* verwendete die texturbasierte Bereichsklassifikation zur Bildsegmentierung. Mit der daran anschließenden modellgestützt erkannten Straßenbegrenzung ist es nun möglich, die Segmentierungsergebnisse des *SeTa* mit den Ergebnissen der anderen drei Verfahren quantitativ zu vergleichen. Das Bewertungskriterium soll die mittlere Positionsabweichung zwischen der erkannten Straßenbegrenzung und den Konturpunkten des Straßenbereiches im Bild sein.

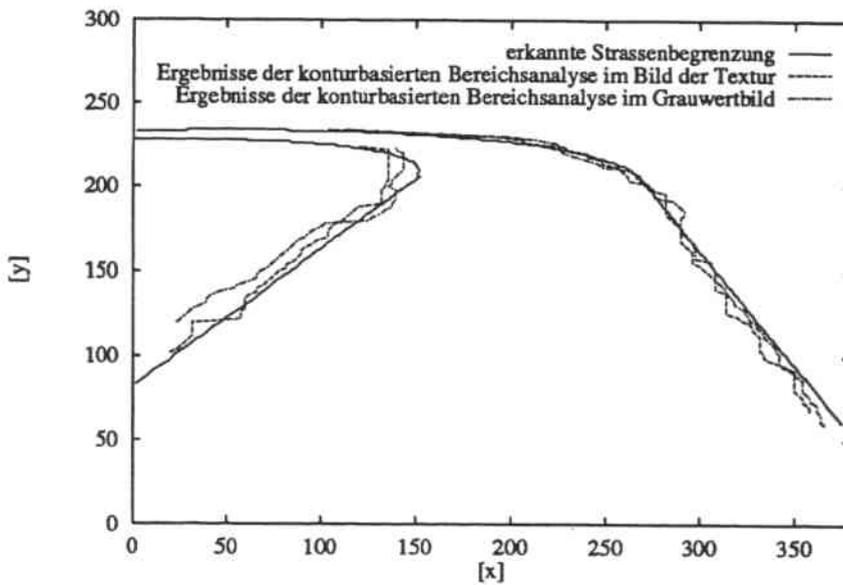
Als Beispiele dienen die in Abschnitt 6.3.5 gezeigten segmentierten Bilder des 10. und 40. Bildes in der Bildfolge. Um die Konturpunkte des Straßenbereiches in den segmentierten Bildern zu bestimmen, projizieren wir zunächst die mit dem *SeTa* erkannte Straßenbegrenzung in jedes segmentierte Bild. Mit Hilfe des in Abschnitt 7.2 beschriebenen Verfahrens werden der Straßenbereich selektiert und anschließend seine Konturpunkte bestimmt. In Abb. 7.10 und 7.11 wird die erkannte Straßenbegrenzung zusammen mit den durch alle vier Verfahren gewonnenen Konturpunkten des Straßenbereiches dargestellt. Die mittlere Positionsabweichung zeigt die folgende Tabelle.

	Bild #10	Bild #40
Texturbasierte Bereichsklassifikation	3.2 (Pixel)	2.3 (Pixel)
Grauwertbasierte Bereichsklassifikation	19.1 (Pixel)	20.5 (Pixel)
Konturbasierte Bereichsanalyse im Bild der Textur	3.9 (Pixel)	4.6 (Pixel)
Konturbasierte Bereichsanalyse im Grauwertbild	4.1 (Pixel)	5.1 (Pixel)

Es ist in dieser Tabelle zu erkennen, daß die texturbasierte Bereichsklassifikation die besten Segmentierungsergebnisse liefert. Dieses Ergebnis unterstützt die Analyse in Kapitel 2, daß die Texturanisotropie im Vergleich zum Grauwert ein aussagekräftigeres Merkmal für Straßenbilder der vierten Klasse ist.

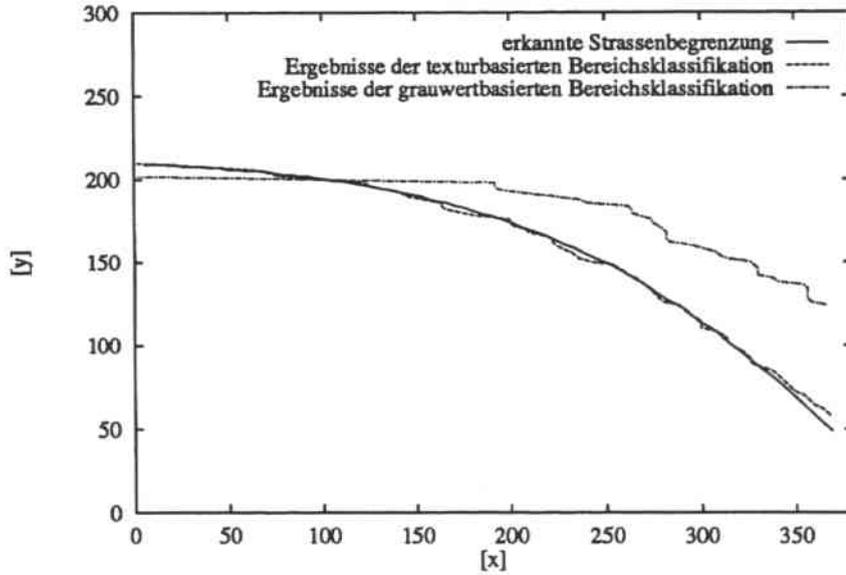


(a) Konturpunkte des Straßenbereiches durch Bereichsklassifikation.

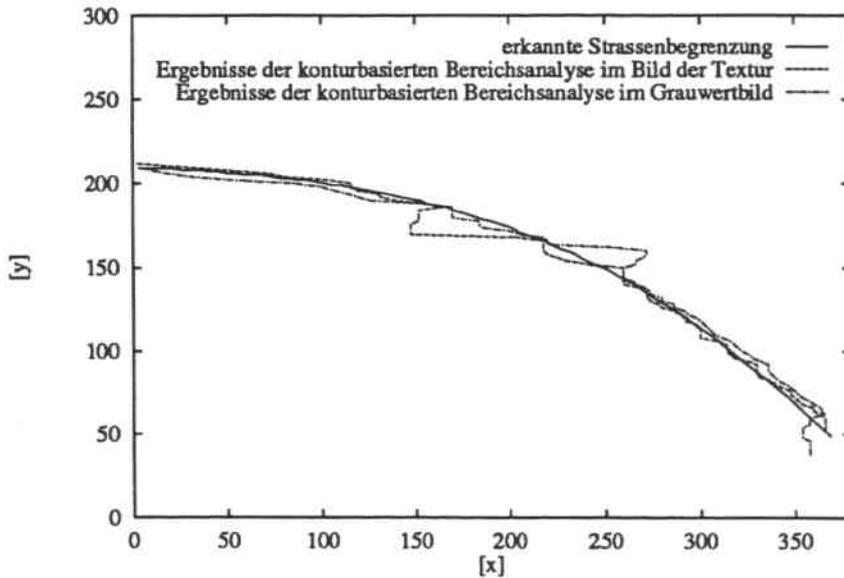


(b) Konturpunkte des Straßenbereiches durch konturbasierte Bereichsanalyse.

Abbildung 7.10: Darstellung der modellgestützt erkannten Straßenbegrenzung (Abb. 7.8 (c)) und der durch vier verschiedene Verfahren bestimmten Konturpunkte des Straßenbereichs (Abb. 6.8 (b), 6.9 (c), 6.10 (c) und 6.11 (e)) des in Abb. 6.9 (a) dargestellten Bildes.



(a) Konturpunkte des Straßenbereiches durch Bereichsklassifikation.



(b) Konturpunkte des Straßenbereiches durch konturbasierte Bereichsanalyse.

Abbildung 7.11: Darstellung der modellgestützt erkannten Straßenbegrenzung (Abb. 7.8 (i)) und der durch vier verschiedene Verfahren bestimmten Konturpunkte des Straßenbereichs (Abb. 6.8 (e), 6.9 (d), 6.10 (d) und 6.11 (f)) des in Abb. 6.9 (b) dargestellten Bildes.

Kapitel 8

Zusammenfassung und Ausblick

Die autonome Fahrzeugführung wurde in den letzten Jahren in zahlreichen Forschungsgruppen in verschiedenen Ländern intensiv untersucht. Sie hat inzwischen den Stand erreicht, Versuchsträger unter *bestimmten* Bedingungen autonom bzw. semi-autonom führen zu können. Bevor ein wirklich autonomes System am öffentlichen Straßenverkehr teilnehmen kann, müssen noch viele grundlegende Probleme gelöst werden. Vor allem ist das Sichtsystem, ein Bestandteil eines autonomen Systems, noch nicht in der Lage, bei allen Bedingungen hilfreiche und präzise Informationen über den Umgebungszustand für Quer- und Längsführung anzubieten.

Die Aufgabe eines Sichtsystems zur Fahrzeugführung besteht in der Straßenerkennung, der Hindernisdetektion und der Landmarkenerkennung. Dabei ist die zuverlässige Erkennung von Straßen die Voraussetzung für die Detektion der sich auf Straßen befindenden Hindernisse und für die Landmarkenerkennung. In der vorliegenden Arbeit wurden neue Verfahren für eine robuste Straßenerkennung vorgestellt. Im folgenden fassen wir die Beiträge dieser Arbeit zur Straßenerkennung und zur Texturanalyse zusammen. Den Abschluß bildet eine Bewertung der neu entwickelten Verfahren in Bezug auf den Einsatz im Fahrzeug.

8.1 Beiträge dieser Arbeit

Der Beitrag dieser Arbeit zur Straßenerkennung

Ausgehend von der Analyse der Komplexität verschiedener Straßen wurde in dieser Arbeit eine Untergliederung von Straßen in vier Klassen vorgenommen. Existierende Verfahren haben Schwierigkeiten bei der Erkennung von Straßen der dritten und vierten Klasse. Der Hauptbeitrag der vorliegenden Arbeit zur Straßenerkennung ist der Versuch zur Überwindung dieser Schwierigkeiten durch die Verwendung von bereichsbasierten Verfahren. Die Beiträge dieser Arbeit bestehen im einzelnen aus den folgenden Punkten:

1. Wir haben die geometrischen und physikalischen Erscheinungsformen von Straßen systematisch analysiert. Aus dieser Analyse resultierte eine Untergliederung von

Straßen in vier Klassen. Für jede Klasse wurden Vorschläge zu folgenden Gesichtspunkten gemacht:

- die aussagekräftigen Merkmale,
- die Verfahren zur Extraktion der entsprechenden Merkmale,
- die geometrischen Modelle und
- die Verfahren zur Bestimmung der Modellparameter.

Damit wurde ein einheitlicher Beschreibungsrahmen für die Beantwortung der Fragen geschaffen, was bereits gelöst wurde und welche noch zu lösenden Aufgaben verbleiben. Eine derartige Analyse von Straßen fehlte noch in der Literatur.

2. Im Gegensatz zu bisherigen Verfahren, die meistens eine Grauwert- bzw. Farbklassifikation durchführen, wurde in dieser Arbeit die Texturanisotropie als Merkmal in der Segmentierung von Straßenbildern eingesetzt. Der Unterschied in der Textur zwischen der Straßenoberfläche und ihrer Umgebung ist dabei das wichtigste Merkmal in Bildern von Straßen der vierten Klasse. Mit Hilfe der quantitativen Bewertung von Segmentierungsergebnissen wurde gezeigt, daß die texturbasierte Bereichsklassifikation auf den Straßenbildern bessere Segmentierungsergebnisse liefert als grauwertbasierte Verfahren, die in diesen Bereich Schwierigkeiten haben.
3. Zur modellgestützten Straßenerkennung wurde in dieser Arbeit das System *SeTa* aufgebaut. Wir haben dabei ein neues Konzept zur Straßenerkennung eingeführt, nämlich Straßenbilder sowohl mit unüberwacht eingelerntem Klassifikator als auch mit adaptivem überwacht eingelerntem Klassifikator in den unterschiedlichen Phasen der Straßenerkennung zu segmentieren und anschließend modellgestützt zu erkennen. Existierende bereichsbasierte Verfahren verwenden entweder einen überwacht eingelernten Klassifikator oder einen unüberwacht eingelernten Klassifikator.

Der Beitrag dieser Arbeit zur Texturanalyse

Textur wurde in dieser Arbeit als Merkmal für die Segmentierung der Bilder von Straßen der vierten Klasse verwendet. Dazu haben wir ein neues Verfahren zur Texturanalyse entwickelt. Unser Beitrag zur Texturanalyse besteht in den folgenden zwei Punkten:

1. In dieser Arbeit wurde eine neue Definition über die gerichteten Texturen eingeführt. Anschließend wurde ein Verfahren zur Ermittlung des Orientierungsfeldes, welches die gerichteten Texturen beschreibt, aus den Kovarianzmatrizen der Grauwertänderungen aufgestellt. Gerichtete Texturen bilden eine wichtige Klasse von Texturen.
2. Die durch die perspektivische Projektion verursachten Änderungen von Texturen im Bild wurde in dieser Arbeit bei der Ermittlung der Texturanisotropie berücksichtigt. Mit Hilfe der Parameter der Kamerakalibrierung wurde ein neues Verfahren zur Ermittlung der optimalen Maßstäbe aufgestellt. Die bisherigen Verfahren zur

Ermittlung eines optimalen Maßstabs verwenden im allgemeinen einen Satz von Filtern mit unterschiedlichen Maßstäben und selektieren daraus einen optimalen.

Mit den Orientierungsfeldern von Texturen lassen sich nicht nur Straßenbilder charakterisieren. Orientierungsfelder von Texturen eignen sich besonders für die Beschreibung gerichteter Texturen, wie z.B. bei Holzoberflächen oder bei handgesponnenen Baumwollstoffen.

8.2 Einsetzbarkeit der Verfahren im Fahrzeug

Bei der Anwendung eines Sichtsystems zur Fahrzeugführung besteht die Notwendigkeit darin, daß das Sichtsystem die Information so schnell wie möglich verarbeiten muß, damit das Fahrzeug weder zu langsam noch zeitweise blind fährt. Die in dieser Arbeit vorgestellten Verfahren wurden bisher mit aufgezeichneten Bildfolgen realer Straßenszenen erprobt. Im folgenden diskutieren wir die Möglichkeiten, die Verfahren zur Segmentierung von Straßenbildern und das System *SeTa* im Fahrzeug in Echtzeit zur Straßenerkennung zu implementieren.

Texturanisotropie wurde als Merkmal für die Klassifizierung eines Bildes im System *SeTa* verwendet. Die Berechnung von Texturanisotropien besteht in den folgenden Schritten:

1. der Detektion der lokalen Grauwertänderungen,
2. der Ermittlung der Varianzen und Kovarianzen sowie
3. der Hauptachsentransformation der Kovarianzmatrix.

Als lokale Grauwertänderungen im Bild können die Faltungsergebnisse aus dem Einschubmodul *NAG* im Bildauswertesystem *VISTA* verwendet werden. Um die Berechnung der Schritte 2 und 3 in Echtzeit durchführen zu können, sind zusätzliche spezifische elektronische Schaltungsmodule erforderlich.

Die Segmentierung von Straßenbildern und anschließend die Erkennung der Straßen wurden im System *SeTa* auf verschiedene Weise realisiert. Bei der Bildsegmentierung führte das System *SeTa* in der Initialisierungsphase eine Ballungsanalyse im 3D-Merkmalraum durch. Eine Ballungsanalyse ist von Natur aus iterativ und im allgemeinen nicht in Echtzeit durchführbar. Im Erkennungsprozeß wurde zur Verifikation der Lagehypothese des Modells ein Minimierungsverfahren eingesetzt. Die Berechnungszeiten der Ballungsanalyse und des Minimierungsverfahrens sind jedoch vertretbar, da sich das Fahrzeug in dieser Phase noch nicht bewegt.

In der Verfolgungsphase klassifiziert *SeTa* ein Bild mit den vorher ermittelten Klassenmodellen. Mit Hilfe von Pyramidenstrukturen und der Benutzung eines Transputernetzes kann die Verarbeitungszeit stark reduziert werden. Bei der Straßenerkennung in der Verfolgungsphase wird die in der vorherigen Szene ermittelte Straßenposition als Lagehypothese des Modells eingesetzt. Damit kann die Suche nach dem Minimum in der Verfolgungsphase beschleunigt werden.

Eine Erweiterungsmöglichkeit bei dem Erkennungsprozeß in der Verfolgungsphase besteht darin, daß die Bewegungsparameter des Fahrzeugs bei der Aufstellung der Lagehypothese berücksichtigt werden und damit der Suchraum in der Minimierungsaufgabe verkleinert wird.

Anhang A

Hauptachsentransformation

Das Hauptachsenproblem wird in der Matrizenalgebra allgemein behandelt. Die wesentlichen Grundgedanken und Lösungsschritte werden anhand des folgenden Beispiels der Hauptachsenermittlung für eine Ellipse im 2-dimensionalen Raum dargestellt.

Eine Ellipse kann mit der Gleichung

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = C \quad (\text{A.1})$$

beschrieben werden, wenn a_{ij} reell, $a_{11} \times a_{22} - a_{12}^2 > 0$ und $C > 0$ ist. Dabei sind x_1 und x_2 zwei Achsen des kartesischen Koordinatensystems. Mit

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \quad \text{und} \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

wird aus (A.1)

$$\vec{x}^T \mathbf{A} \vec{x} = C. \quad (\text{A.2})$$

Drehen wir nun das Koordinatensystem um einen Winkel θ so, daß eine Achse in die dominante Richtung der Ellipse zeigt, dann reduziert sich die Matrix \mathbf{A} in (A.2) auf eine Diagonalmatrix. Im folgenden sollen die Eigenwerte der Matrix und der Drehwinkel ermittelt werden.

Die Bestimmung der Eigenwerte kann mit der charakteristischen Gleichung

$$\|\mathbf{A} - \lambda \mathbf{I}\| = 0$$

beschrieben werden, dabei ist \mathbf{I} die Einheitsmatrix. Daraus resultiert

$$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}^2 = 0$$

mit den Lösungen:

$$\lambda_{1,2} = \frac{1}{2} \left(a_{11} + a_{22} \pm \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2} \right). \quad (\text{A.3})$$

Die beiden Lösungen werden die Eigenwerte der Matrix \mathbf{A} genannt. Die Richtung des Eigenvektors mit dem größten Eigenwert stellt die Richtung der ersten Hauptachse dar.

Wir zeigen jetzt die Ermittlung des Drehwinkels. In der transformierten Matrix stehen die Eigenwerte λ_1 und λ_2 auf den Diagonalen. Daraus berechnen wir den Drehwinkel der Hauptachse mit der Gleichung:

$$\begin{aligned} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} &= \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \\ &= \begin{pmatrix} a_{11} \cos^2 \theta + a_{22} \sin^2 \theta + a_{12} \sin 2\theta & \frac{1}{2}(a_{22} - a_{11}) \sin 2\theta + a_{12} \cos 2\theta \\ \frac{1}{2}(a_{22} - a_{11}) \sin 2\theta + a_{12} \cos 2\theta & a_{11} \cos^2 \theta + a_{22} \sin^2 \theta - a_{12} \sin 2\theta \end{pmatrix}. \end{aligned}$$

Nach einfacher Umformung ergibt sich aus den Nichtdiagonalelementen unmittelbar der Winkel der Hauptachse zur alten X-Achse:

$$\theta = \frac{1}{2} \arctan \frac{2a_{12}}{a_{11} - a_{22}}. \quad (\text{A.4})$$

Zum Schluß stellen wir noch die Ellipsengleichung (A.1) auf die neuen Hauptachsenkoordinaten y_1, y_2 um:

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = C, \quad (\text{A.5})$$

mit

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Anhang B

Die k -Mittelwert-Ballung

Die k -Mittelwert-Ballung¹ ist eines der wichtigsten unüberwachten nichtparametrischen Lernverfahren in der Mustererkennung. Dabei wird keine vorklassifizierte Stichprobe benötigt, wie für die überwacht eingelernten Klassifikatoren. Hier ist eine Menge von Mustern gegeben, für die eine natürliche Klassenverteilung gesucht wird. Deswegen wird Ballungsanalyse auch als ein unüberwachtes Lernverfahren bezeichnet. Wir werden im folgenden die k -Mittelwert-Ballung vorstellen.

Die k -Mittelwert-Ballung, wie sie in der Fachliteratur oft genannt wird, ist ein *Basic-Isodata* Algorithmus. Bei diesem Verfahren ist die Anzahl der zu bildenden Klassen vorgegeben. Dieses Verfahren basiert auf der Minimierung eines Gütekriteriums, nämlich der Summe der quadrierten Abstände aller Punkte in einem Ballungsgebiet vom Ballungsmittelpunkt. Ballungszentren werden iterativ geschätzt. Das ganze Verfahren besteht aus den folgenden Schritten:

Algorithmus B.1 (k -Mittelwert-Ballung)

Initialisierung: Gebe die Anzahl der zu bildenden Klassen k , die maximale Anzahl der Iterationsschritte L und das Abbruchkriterium $\bar{\epsilon}$ an. Wähle k Ausgangszentren:

$$\bar{z}_1(1), \bar{z}_2(1), \dots, \bar{z}_k(1),$$

dabei ist $\bar{z}_k(1)$ ein Vektor im Merkmalsraum. Gewöhnlich werden die ersten k Eingabevektoren als Ausgangszentren ausgewählt.

Beginn der Iterationen: Die folgenden Schritte sollen iterativ durchgeführt werden.

1. Im l -ten Iterationsschritt verteile die Muster $\{\bar{x}_i\}$ auf die k Ballungsgebiete mit Hilfe der Relation:

$$\bar{x}_i \in C_j(l) \quad \text{falls} \quad \|\bar{x}_i - \bar{z}_j(l)\| < \|\bar{x}_i - \bar{z}_i(l)\| \quad (\text{B.1})$$

für alle $i = 1, 2, \dots, k$ und $i \neq j$, wobei $C_j(l)$ die Teilmenge mit dem Ballungszentrum $\bar{z}_j(l)$ ist. Bei Gleichheit wird beliebig entschieden.

¹engl.: *C-Means Clustering Algorithm*

2. Aus den Ergebnissen von Schritt 1 werden die nächsten Ballungszentren $\bar{z}_j(l+1)$, $j = 1, 2, \dots, k$ so berechnet, daß die Summe der quadrierten Abstände aller Punkte aus $C_j(l)$ zum neuen Ballungszentrum minimiert wird. Das neue Ballungszentrum $\bar{z}_j(k+1)$ wird also durch Minimierung des Kompaktheitsmaßes:

$$e(C_j) = \sum_{\vec{x}_i \in C_j} \|\vec{x}_i - \bar{z}_j(l+1)\|^2 \quad (\text{B.2})$$

bestimmt. Nach einer einfachen Umrechnung erhalten wir

$$\bar{z}_j(l+1) = \frac{1}{m_j} \sum_{\vec{x}_i \in C_j} \vec{x}_i, \quad (\text{B.3})$$

dabei ist m_j die Anzahl der Muster in C_j .

3. Wenn $\|\bar{z}_j(l+1) - \bar{z}_j(l)\| \leq e$ für $j = 1, 2, \dots, k$ erfüllt ist oder L erreicht wird, wird das Verfahren beendet. Sonst gehe nach Schritt 1 zurück.

Ende der Iterationen: Gebe die klassifizierten Muster aus.

Bei der k -Mittelwert-Ballung besteht die Schwierigkeit in der Auswahl der Anzahl von Klassen, der Auswahl von Ausgangszentren und der Festlegung eines Abbruchkriteriums. Eine ausführliche Behandlung dieser Probleme ist bei [Duda & Hart 73] zu finden.

Anhang C

Isodata-Algorithmus mit dem Varianzkriterium

Mit der k -Mittelwert-Ballung werden die Merkmalsvektoren in k Klassen unterteilt. Gegeben ist dabei eine Menge M von Bildpunkten. Jeder davon ist beschreibbar durch einen s -dimensionalen Vektor in der Form

$$\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{is})^T.$$

Nach der in Anhang B beschriebenen k -Mittelwert-Ballung sollen alle Bildpunkte in k Teilmengen unter den Bedingungen

$$C_1 \cup C_2 \cup \dots \cup C_k = M$$

und

$$C_i \cup C_j = \phi \quad (i \neq j)$$

eingeteilt werden.

Als Zielfunktion der k -Mittelwert-Ballung wird die Minimierung der Summation der quadratischen euklidischen Abstände zwischen den Eingabedaten und Klassenzentren gewählt. Mathematisch kann dieses Gütekriterium für jede einzelne Teilmenge repräsentiert werden als

$$\min_{\vec{z}_j} e(C_j) = \min_{\vec{z}_j} \sum_{i \in C_j} \|\vec{x}_i - \vec{z}_j\|^2. \quad (\text{C.1})$$

Dabei ist

$$\vec{z}_j = (z_{j1}, z_{j2}, \dots, z_{js})^T$$

das Zentrum der Teilmenge C_j . Die komplette k -Mittelwert-Ballung wird in Anhang B beschrieben. Dabei wird die Klassenanzahl festgelegt.

Für die Frage, welche Klassenanzahl optimal ist, gibt es noch keine analytische Lösung. Sie ist im allgemeinen nur durch Probieren beantwortbar. Der *Isodata*-Algorithmus versucht, das Problem mit zusätzlichen Bedingungen durch iterative Verbesserung der mit

der k -Mittelwert-Ballung erzielten Klassifikationen zu lösen. Als Gütekriterien für die Qualität der Klassifikationsergebnisse nach der k -Mittelwert-Ballung werden die folgenden beiden Bedingungen gestellt:

- Der Abstand $\| \bar{z}_i - \bar{z}_j \|$ zweier Klassenzentren soll nie kleiner als ein vorgegebener Wert δ ($\delta > 0$) sein.
- Die Heterogenität $\frac{1}{n-1} \sum_{x_1 \in C_i} \| \bar{x}_1 - \bar{x}_{C_i} \|^2$ jeder Klasse soll nie größer als eine Konstante η ($\eta > 0$) sein.

Mit diesen Bedingungen besteht der *Isodata*-Algorithmus aus den folgenden Schritten:

Algorithmus C.1 (*Isodata*-Algorithmus)

1. Man wählt zunächst k typische Punkte aus dem Merkmalsraum, so daß deren gegenseitiger Abstand nicht kleiner als eine Konstante θ ($\theta > 0$) ist. Diese Punkte werden als die Zentren der ersten k Klassen angenommen.
2. Mit einer gegebenen Anzahl der zu bildenden Klassen wird die im Anhang B beschriebene k -Mittelwert-Ballung durchgeführt.
3. Falls die Heterogenität einer Klasse zu groß ist, spalte man sie in zwei Unterklassen auf.
4. Man iteriere die Schritte 2 und 3, bis kein Klassenzentrum mehr heterogen ist.
5. Man fusioniere alle Klassen, deren Abstand kleiner als δ ist.
6. Man iteriere die Schritte 2 bis 5 solange, bis eine stabile Gruppierung erreicht ist.
7. Man eliminiere die Klassen mit zu wenig Punkten und wende die Schritte 2 bis 6 auf die Restmenge an.
8. Iteration der Schritte 2 bis 7 für schrittweise verkleinerte Objektmenge.

Anhang D

Der Mahalanobis-Abstand

Der *Mahalanobis*-Abstand ist definiert als

$$r_{ij}^2 = (\vec{x}_i - \vec{z}_j)^T \Sigma^{-1} (\vec{x}_i - \vec{z}_j). \quad (\text{D.1})$$

Dabei ist Σ eine $s \times s$ Kovarianzmatrix, \vec{x}_i ein Eingabevektor und \vec{z}_j ein Klassenzentrum.

Der *Mahalanobis*-Abstand kann eigentlich auch als eine Verallgemeinerung des Euklid'schen Abstandes angesehen werden. Dies kann man wie folgt zeigen: Seien $\lambda_1 \geq \lambda_2 \geq \dots > 0$ die Eigenwerte der Matrix Σ und $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_s \in R^s$ die zugehörigen Eigenvektoren. Mit den Matrizen $V := (\vec{v}_1, \dots, \vec{v}_s)$ und $\Lambda := \text{diag}(\lambda_1, \dots, \lambda_s)$ kann die Matrix Σ zerlegt werden als

$$\begin{cases} \Sigma = V \Lambda V^T \\ \Sigma^{-1} = V \Lambda^{-1} V^T, \end{cases} \quad (\text{D.2})$$

wobei gilt $V^T = V^{-1}$. Damit kann man den *Mahalanobis*-Abstand folgendermaßen zerlegen:

$$\begin{aligned} r_{ij}^2 &= (\vec{x}_i - \vec{z}_j)^T \Sigma^{-1} (\vec{x}_i - \vec{z}_j) \\ &= (\vec{x}_i - \vec{z}_j)^T V \Lambda^{-\frac{1}{2}} \Lambda^{-\frac{1}{2}} V^T (\vec{x}_i - \vec{z}_j) \\ &= \underbrace{(\Lambda^{-\frac{1}{2}} V^T \vec{x}_i)}_{\vec{y}_i} - \underbrace{\Lambda^{-\frac{1}{2}} V^T \vec{z}_j}_{\vec{y}_j} \quad (\Lambda^{-\frac{1}{2}} V^T \vec{x}_i - \Lambda^{-\frac{1}{2}} V^T \vec{z}_j) \\ &= \|\vec{y}_i - \vec{y}_j\|^2. \end{aligned} \quad (\text{D.3})$$

Es gibt noch den erwähnenswerten Aspekt aus der obigen Betrachtung, daß die Form einer Klasse im Merkmalsraum prinzipiell mit den Verhältnissen der Eigenwerte $\lambda_1 \geq \lambda_2 \geq \dots > 0$ der Matrix Σ bestimmt werden kann. Unterschiedliche Eigenwerte erlauben die Entstehung von ellipsenförmigen Unterbereichen im Bild.

Die Dekorrelation von Merkmalen wird im allgemeinen realisiert durch die Rotation um eine Achse im multidimensionalen Merkmalsraum, wie z.B. der *Karhunen-Loeve*-Transformation. Die transformierten Merkmale sind nicht mehr korreliert. Bei der Verwendung des *Mahalanobis*-Abstandes kann die Dekorrelation von Merkmalen eingespart

werden. Wir wollen diese Behauptung mathematisch herleiten. Dazu berechnen wir die Kovarianzmatrix der Komponenten der Vektoren $\vec{y}_1, \vec{y}_2, \dots, \vec{y}_N$ in (D.3), wobei mit N die Anzahl der Eingabemuster bezeichnet wird. Wir erhalten

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N (\vec{y}_i - \vec{\bar{y}})(\vec{y}_i - \vec{\bar{y}})^T &= \Lambda^{-\frac{1}{2}} V^T \cdot \left[\frac{1}{N} \sum_{i=1}^N (\vec{x}_i - \vec{z})(\vec{x}_i - \vec{z})^T \right] \cdot V \Lambda^{-\frac{1}{2}} \\ &= \Lambda^{-\frac{1}{2}} V^T \Sigma V \Lambda^{-\frac{1}{2}} = I. \end{aligned} \quad (D.4)$$

Dabei ist \vec{z} das Zentrum aller Eingabemuster. Daraus folgt, daß die Komponenten der Vektoren $\vec{y}_1, \vec{y}_2, \dots, \vec{y}_N$ unkorreliert sind. Nach (D.3) kann es auch so interpretiert werden, daß bei der Verwendung des *Mahalanobis*-Abstandes die Merkmale $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N$ dekorreliert werden.

Bei der Berechnung hat der *Mahalanobis*-Abstand zweifellos den Nachteil, daß die Anzahl der Multiplikationen größer ist als beim Euklid'schen Abstand. Zur Reduzierung des Rechenaufwandes gibt es verschiedene Möglichkeiten. Eine davon besteht darin, Gleichung (D.1) zu vereinfachen ([Venkateswarlu & Raju 92]). Dazu schreiben wir den *Mahalanobis*-Abstand um:

$$r_{ij}^2 = \vec{x}_i^T \Sigma^{-1} \vec{x}_i - 2\vec{x}_i^T \Sigma^{-1} \vec{z}_j + \vec{z}_j^T \Sigma^{-1} \vec{z}_j. \quad (D.5)$$

In Gleichung (D.5) ist der erste Term auf der rechten Seite unabhängig von den Klassenzentren. Dieser Term kann bei der Berechnung vernachlässigt werden. Ersetzen wir $2\Sigma^{-1}\vec{z}_j$ durch einen Vektor \vec{w}_j , dann erhalten wir:

$$r_{ij}^2 = \vec{z}_j^T \Sigma^{-1} \vec{z}_j - \vec{x}_i^T \vec{w}_j. \quad (D.6)$$

Der erste Term und der Vektor \vec{w}_j in (D.6) sind für jede Klasse konstant und müssen nur einmal in einer Schleife berechnet werden. Mit dieser Vorgehensweise besteht der Zeitaufwand hauptsächlich in der Multiplikation zweier Vektoren: \vec{x}_i und \vec{w}_j .

Anhang E

Segmentierung von Straßenbildern durch konturbasierte Bereichsanalyse

Obwohl der Schwerpunkt dieser Arbeit auf der texturbasierten Erkennung von Straßen liegt, soll an dieser Stelle eine konturbasierte Bereichssegmentierung vorgestellt werden, mit der besonders Straßen der ersten und dritten Klasse gut zu erkennen sind. Das Verfahren besteht aus den folgenden drei Verarbeitungsschritten:

1. Bestimmung von Konturpunkten mit einem gradientenbasierten Verfahren,
2. Verkettung der Konturpunkte zu Bereichskonturen und
3. Bereichsmarkierung und Berechnung von Bereichsattributen.

Als Bildbereichshinweise steht das segmentierte Bild zusammen mit den Attributen aller Bereiche der Interpretation des Bildes zur Verfügung.

E.1 Bestimmung von Konturpunkten

Der erste Schritt eines konturbasierten Verfahrens ist die Detektion von Kantenelementen. Die Kantendetektion läßt sich im allgemeinen auf zwei verschiedene Weisen erledigen: Die Suche nach einem Maximum in der ersten Ableitung der Bildfunktion oder nach einem Nulldurchgang in ihrer zweiten Ableitung.

Für die Auswahl der Art bzw. auch der Größe und Orientierung der Ableitungsoperatoren existieren schon viele verschiedene Ansätze (siehe z.B. [Marr & Hildreth 80], [Witkin 83], [Haralick 84], [Grimson & Hildreth 85], [Canny 86], [Korn 88] und [Perona 91]). Je höher die Ordnung eines Faltungsoperators ist, desto empfindlicher reagiert er auf Rauschstörungen. Aus diesem Grund ziehen wir die Gradientenoperatoren zur Kantendetektion vor. Außerdem liefert ein gradientenbasiertes Verfahren bei der Kantendetektion noch die Richtungsinformation, die wir bei der Konturpunktverkettung verwenden können.

Zur Kantendetektion verwenden wir hier das in [Korn 88] beschriebene Verfahren. Das Verfahren beruht auf einer zweifachen Filterung eines Grauwertbildes mit normierten ersten Ableitungen der Gaußfunktion und anschließender Berechnung von Beträgen und Richtungen des Gradienten an jedem Ort des Bildes. Die beiden Ableitungsmasken sind in Abb. 5.1 dargestellt. Die Bestimmung der Konturpunkte erfolgt durch die Feststellung eines Maximums des Gradientenbetrages in Richtung des Gradienten. Am Ende der Konturbestimmung in [Korn 88] stehen vier Bilder,

- das Bild der Gradientenbeträge,
- das Bild der Gradientenrichtungen,
- das Bild der lokalen Maxima des Gradientenbetrages und
- das Bild der Gradientenrichtungen an den lokalen Maxima des Gradientenbetrages,

für die weiteren Verarbeitungsschritte zur Verfügung. Als Beispiel findet man in [Müller 91] eine Anwendung des Korn'schen Verfahrens zur Konturdetektion in Bildern polyedrischer Szenen.

E.2 Verkettung von Konturpunkten

Kantendetektoren liefern häufig nur eine unvollständige Objektkontur. Die Konturlinie ist in einzelne Konturelemente aufgebrochen: Lücken, Doppelkonturen und isolierte Fragmente kennzeichnen das typische Ergebnis einer Kantendetektion. Die Gründe dafür sind vielfältig, wie z.B. die unangepasste Größe und Orientierung der Kantenoperatoren an die betrachteten Grauwertverläufe im Bild.

Bei der Kantendetektion in Straßenbildern besteht zusätzlich noch ein Problem darin, daß die Straßenbegrenzungen infolge von Störungen nicht mehr das signifikante Merkmal sein könnten. Daher werden einige Konturpunkte durch eine Schwelle unterdrückt. Ein rein konturbasiertes Verfahren zur Straßenerkennung, in dem nach der Bestimmung der Konturpunkte die Zusammenfassung solcher Punkte zu einer symbolischen Beschreibung stattfindet, führt hierbei zu einer unvollständigen Beschreibung der Straßenbegrenzung. Die resultierende Beschreibung der Straßenbegrenzung umfaßt Lücken und isolierte Liniensegmente. Zur Überwindung dieses Problems ist es nötig, eine symbolische Gruppierung der extrahierten Liniensegmente durchzuführen. Verfahren zur Extraktion globaler Bildstrukturen aus einem Bild durch eine symbolische Gruppierung der Unterstrukturen findet man z.B. in [Lowe 85] und [Boldt *et al.* 89].

Um die Bereichsinformation über Straßen auch bei der Erkennung von Straßen ausnutzen zu können, stellen wir eine konturbasierte Bereichsanalyse zur Segmentierung von Straßenbildern vor. Hier bedarf es der Verkettung der Konturpunkte, der Bereichsmarkierung und der Analyse der Bereichsattribute.

Die Konturpunktverkettung verbindet die einzelnen Konturelemente und beseitigt die überflüssigen Konturpunktketten. Mit der Überflüssigkeit einer Konturpunkt-kette ist

gemeint, daß sie entweder nicht geschlossen oder kürzer als eine vorgegebene Schwelle ist. Da wir bei der Bestimmung von Konturpunkten keine Schwelle eingesetzt haben, werden beim Schließen unterbrochener Konturen alle Konturpunkte gefunden. Dieser Prozeß liefert geschlossene Konturlinien. Zu den geschlossenen Ketten gehören auch diejenigen Ketten, die zusammen mit der Bildrändern einen geschlossenen Bereich bilden.

Das Verkettungsverfahren beginnt in der linken oberen Ecke eines Bildes und arbeitet dann zeilenweise. Im folgenden erläutern wir die einzelnen Schritte des Verkettungsverfahrens.

Algorithmus E.1 (Konturpunktverkettung-Algorithmus)

1. *Im ersten Schritt markiert man diejenigen Punkte im Bild der lokalen Maxima des Gradientenbetrages, deren Wert größer als eine vorgegebene Schwelle ist, als Kennpunkte für die Verkettung. Ausgehend von diesen Initialpunkten wird die Konturverkettung sequentiell durchgeführt.*
2. *Für jeden Kennpunkt werden seine 8 Nachbarpunkte im Bild der lokalen Maxima des Gradientenbetrages überprüft, um die Beziehung zwischen ihm und seinen Nachbarn festzustellen.*
 - (a) *Ist der Kennpunkt ein isolierter Punkt, so wird er gelöscht. Der Prozeß beginnt erneut bei 2 und überprüft den nächsten Kennpunkt.*
 - (b) *Ist dieser Kennpunkt ein Endpunkt einer Konturkette, geht der Prozeß zu Schritt 3, wo die Suche nach seinem Nachfolger stattfindet.*
 - (c) *Ist der Kennpunkt in einer Kette, läßt man den Kennpunkt unverändert. Der Prozeß beginnt erneut bei 2.*
 - (d) *Ansonsten geht der Prozeß zu Schritt 4.*
3. *In diesem Schritt soll die Suche nach der Fortsetzung einer Konturpunktkette im Bild der Gradientenbeträge durchgeführt werden. Entsprechend der jeweiligen Gradientenrichtung eines Kennpunktes werden drei Nachbarpunkte als Kandidaten ausgewählt. Daraus wird der Punkt mit dem größten Gradientenbetrag bestimmt. Ist der Gradientenbetrag an diesem Punkt größer als eine vorgegebene Schwelle, so wird dieser Punkt als eine Fortsetzung der Kette betrachtet. Mit diesem als neuem Kennpunkt wird der Verkettungsprozeß erneut bei 2 weiter durchgeführt. Ist die Suche nach neuen Konturpunkten in sämtlichen 8 Richtungen nicht erfolgreich, so bricht der Prozeß ab und geht zu Schritt 4 über.*
4. *Bildet die gefundene Punktkette einen geschlossenen Bereich oder ist ihre Länge größer als eine vorgegebene Schwelle, so wird die Kette abgespeichert, ansonsten wird sie eliminiert. Anschließend beginnt der Prozeß erneut bei 2.*

Um den oben beschriebenen Algorithmus besser verstehen zu können, wird im folgenden auf jeden einzelnen Schritt des Verfahrens detailliert eingegangen.

Die Verwendung der Schwelle im ersten Schritt führt zu einer starken Reduzierung der Kennpunkte. Bei der Bestimmung von Konturpunkten im Verfahren nach [Korn 88] haben wir für die Schwelle des Gradientenbetrages Null eingesetzt. Daher werden die Konturpunkte mit schwachem Kontrast nicht durch die Schwelle unterdrückt. Sie werden auch an der Konturverkettung teilnehmen. Ausgehend von den Kennpunkten werden alle zu einer Objektkontur gehörenden Punkte in die Verkettung miteinbezogen.

Eine der wichtigsten Aufgaben bei der Konturverkettung besteht in der Feststellung der Beziehung eines Kennpunktes zu seinen Nachbarn. Für jeden Kennpunkt werden seine 8 Nachbarn betrachtet. Diese Nachbarschaft wird in Abb. E.1 (a) dargestellt. Die Beziehungen eines Kennpunktes zu seinen Nachbarn können eigentlich in 3 Klassen

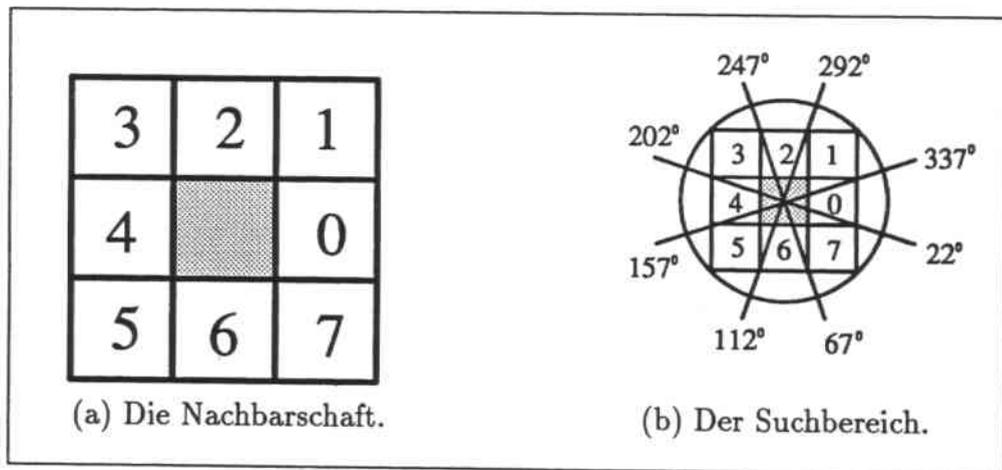


Abbildung E.1: Die Definition der Nachbarschaft eines Konturpunktes in (a) und der Suchbereich in Bezug auf die Gradientenrichtung für die Konturverkettung im Bild des Gradientenbetrages in (b).

eingeteilt werden:

1. Der Kennpunkt ist ein *isolierter* Punkt, wenn kein Konturpunkt in seiner Nachbarschaft gefunden wird.
2. Der Kennpunkt ist ein Endpunkt einer Kette. Dafür ergeben sich folgende zwei Möglichkeiten:
 - (a) Es existiert nur ein Konturpunkt in der Nachbarschaft.
 - (b) Zwei Konturpunkte werden gefunden und sind untereinander benachbart.
3. Der Kennpunkt befindet sich in einer Kette, wenn zwei nicht untereinander benachbarte Kennpunkte oder mehr als zwei gefunden werden.

Da die Breite der Kontur nach [Korn 88] im allgemeinen mehr als ein Pixel beträgt, kann eine Verdünnung der Kontur in unserem Verfahren nicht entfallen. Dies wird durch die Eliminierung der überflüssigen Punkte realisiert.

In der jüngsten Literatur wird die Einbeziehung der Gradientenrichtung bei der Konturverkettung als eine vielversprechende Vorgehensweise angesehen. Wir realisieren diese Idee bei der Suche nach der Fortsetzung einer Konturkette im Bild des Gradientenbetrages.

Die Unterbrechung einer Kontur kann verschiedene Gründe haben. An einer solchen Stelle weist die Gradientenrichtung oft eine Unzuverlässigkeit auf. Zur Überwindung des Problems beziehen wir noch zusätzlich zwei Nachbarn des Kandidatenpunktes ein. Wie in Abb. E.1 (b) dargestellt, wird zuerst nach der Gradientenrichtung des Kennpunktes ein entsprechender Punkt i als ein erster Kandidat bestimmt. Nachfolgend nehmen wir auch die Punkte $i - 1$ und $i + 1$ zur Entscheidung hinzu. Die Verkettung wird zu dem Punkt hin fortgesetzt, der den größten Gradientenbetrag besitzt.

Bevor der Algorithmus weiter nach nächsten Konturpunkten sucht, ist es wichtig, den gefundenen Konturpunkt auf Null zu setzen oder auf eine andere Weise als „gefunden“ zu kennzeichnen, um eine eventuelle zweite Verarbeitung zu vermeiden.

E.3 Bereichsmarkierung

Der nächste Verarbeitungsschritt nach der Konturverkettung richtet sich nach dem Zweck der Bildauswertung. Für manche Anwendungen werden die jeweils gefundenen Konturpunkte einer linearen Regression unterzogen, damit wird eine symbolische Beschreibung der Grauwertänderungen als Geradensegment gewonnen. Für die konturbasierte Bereichsanalyse soll vor der Gewinnung symbolischer Beschreibungen noch eine Analyse der durch die Konturpunkte geschlossenen Bereiche erfolgen.

Zur Berechnung der Bereichsattribute soll zuerst die Markierung aller Bildpunkte zu einzelnen Bereichen durchgeführt werden. Da die geschlossenen Konturen schon gefunden wurden, ist das Markierungsverfahren nicht kompliziert. Ausgehend von einem Punkt wird eine Suche nach dem Nachbarn gestartet. Dieser Prozeß wird solange fortgesetzt, bis ein Konturpunkt oder der Bildrahmen erreicht wird. Es ist dabei keinerlei *a priori* Wissen über das Bild erforderlich.

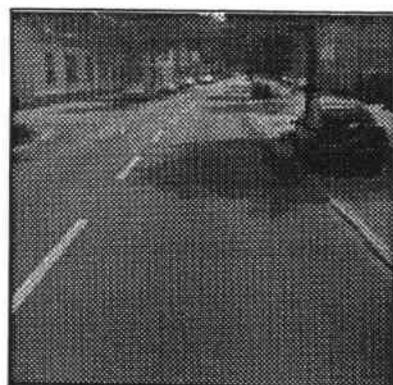
Nach der Markierung aller Bildpunkte können anschließend Attribute aller Bereiche berechnet werden. Die folgenden Attribute für einen Bereich werden berechnet:

- die Anzahl der Punkte,
- der mittleren Grauwert,
- der mittleren Kontrast und
- die Koordinaten des Mittelpunktes.

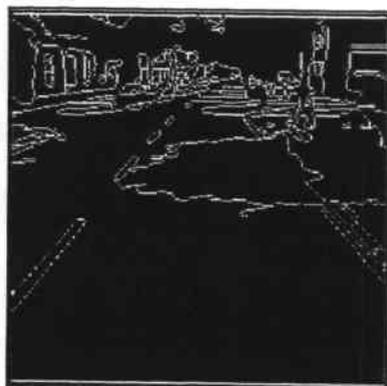
Einige Attribute davon, z.B. die Anzahl der Bildpunkte und die Koordinaten des Mittelpunktes in jedem Bereich, werden schon bei der Markierung der Bildpunkte ermittelt. Der mittleren Kontrast eines Bereiches läßt sich aus den Gradientenbeträgen berechnen.

E.4 Beispiele von Segmentierungsergebnissen

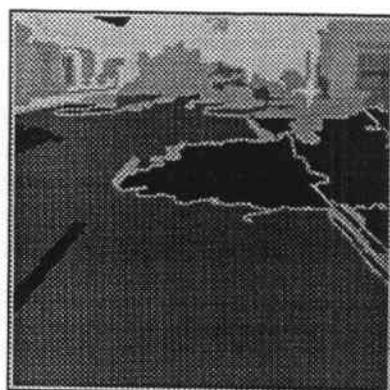
In Abb. E.2 wird die Segmentierung zweier gestörter Straßenbilder dargestellt. In (a) befinden sich Schattenwürfe auf der Straßenoberfläche. In (d) bestehen die Straßenbeläge aus Materialien verschiedener Farben. Auf der Straßenoberfläche ist zusätzlich ein Kanaldeckel vorhanden. Die Segmentierungsergebnisse werden in (c) und (f) gezeigt.



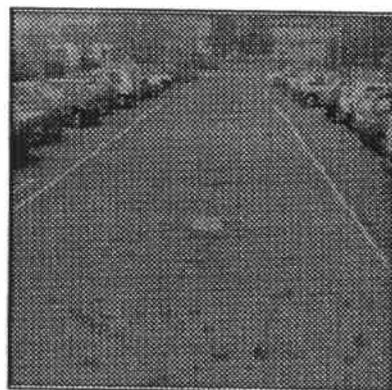
(a) Ein Straßenbild.



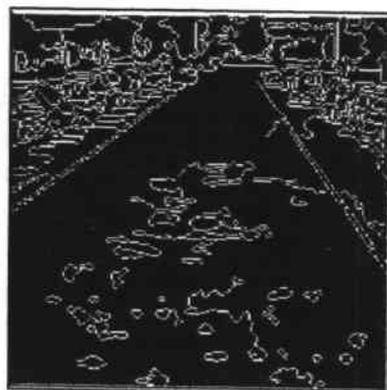
(b) Konturpunktketten von (a).



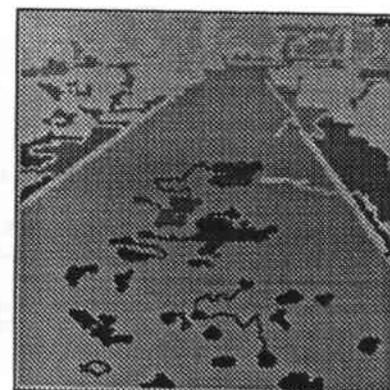
(c) Segmentierung von (a).



(d) Ein Straßenbild.



(e) Konturpunktketten von (d).



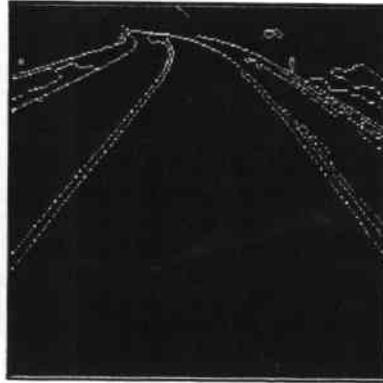
(f) Segmentierung von (d).

Abbildung E.2: Ergebnisse der Konturpunktverkettung und der Segmentierung von zwei gestörten Straßenbildern. Bei der Bestimmung von Konturpunkten wird für die Schwelle Null eingesetzt. Die Schwelle für die Selektion der Kennpunkte wird mit 20 angenommen.

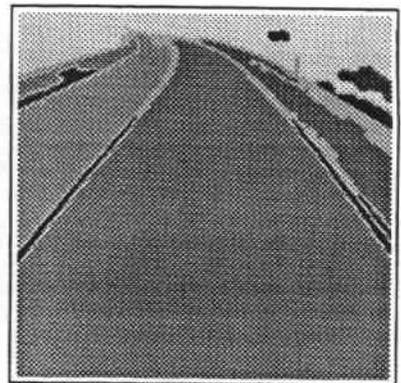
In Abb. E.3 zeigen wir die Ergebnisse der Segmentierung von Straßenbildern in einer Bildfolge. Die Bildfolge wird auf einer Landstraße aufgenommen. Dabei treten verschiedene Störungen auf. Die Fahrbahnmarkierungen sind an manchen Stellen nicht sichtbar



(a1) Bild # 0.



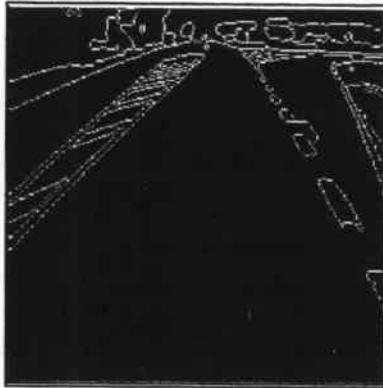
(a2) Konturpunktketten.



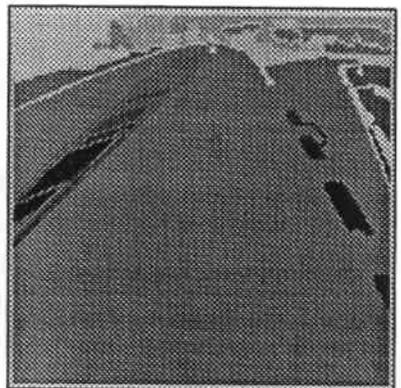
(a3) Segmentierung.



(b1) Bild # 45.



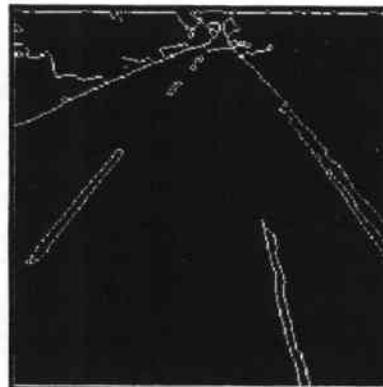
(b2) Konturpunktketten.



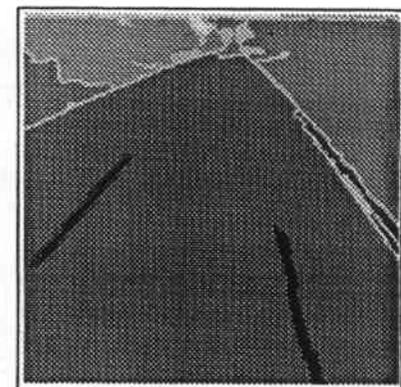
(b3) Segmentierung.



(c1) Bild # 75.



(c2) Konturpunktketten.



(c3) Segmentierung.

Abbildung E.3: Ergebnisse der Konturpunktverketzung und der Segmentierung der 0., 45. und 75. Aufnahme in einer Bildfolge. Bei der Bestimmung von Konturpunkten wird für die Schwelle Null eingesetzt. Die Schwelle für die Selektion der Kennpunkte wird mit 20 angenommen.

geworden. Verkehrszeichen wie z.B. die zur Trennung verschiedener Spuren verwendeten Markierungsbänder sind im Straßenbereich zu sehen.

In der linken Spalte in Abb. *E.2* und *E.3* werden Grauwertbilder dargestellt. Nach wie vor wird bei unseren Versuchen die Schwelle für die Bestimmung der Konturpunkte im Verfahren nach [Korn 88] auf Null gesetzt. Die extrahierten Konturpunktketten werden in der mittleren Spalte gezeigt. Die Schwelle für die Selektion der Kennpunkte im ersten Schritt in Algorithmus E.1 wird für alle Bilder in Abb. *E.2* und *E.3* mit 20 angenommen. Die Konturpunkte werden dabei mit Intensitätswert 255 kodiert und dargestellt. Die Bilder in der rechten Spalte sind die Ergebnisse der Segmentierung. Dabei werden diejenigen Bereiche, deren Punktzahl weniger als 1% der gesamten Bildpunkte beträgt, mit ihren benachbarten Bereichen verschmolzen. Die Beispiele zeigen eine gute Übereinstimmung zwischen den segmentierten Bildern und unserer Erwartung.

Literaturverzeichnis

- [Ahuja & Rosenfeld 81] N. Ahuja and A. Rosenfeld, Mosaic Models for Textures, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-3:1 (1981) 1-11.
- [Aloimonos & Swain 85] J. Aloimonos and M.J. Swain, Shape from Texture, in *Proc. Int. Joint. Conf. on Artif. Intel.*, Los Angeles, CA, August 18-23, 1985, pp. 926-931.
- [Amadasun & King 89] M. Amadasun and R. King, Textural Features Corresponding to Textural Properties, *IEEE Trans. Systems, Man, and Cybernetics* SMC-19:5 (1989) 1264-1275.
- [Aono 89] S. Aono, Technology for the Intelligent Car of the Future, in *Proc. JSK Int. Symp. on Technological Innovations for Tomorrow's Automobile Traffic and Driving Information Systems*, Tokyo, Japan, Nov. 1-2, 1989, pp. 110-119.
- [Arkin 87] R.C. Arkin, *Towards Cosmopolitan Robots: Intelligent Navigation in Extended Man-Made Environments*, PhD thesis, University of Massachusetts, Amherst, MA, Sept. 1987.
- [Arkin et al. 87] R.C. Arkin, E.M. Riseman, A. Hanson, Visual Strategies for Mobile Robot Navigation, in *Proc. IEEE Workshop on Computer Vision*, Miami Beach/FL, Nov. 30 - Dec. 2, 1987, pp. 176-181.
- [Ball & Hall 65] G.H. Ball and D.J. Hall, *ISODATA: A Novel Technique for Data Analysis and Pattern Classification*, Tech. Report, Stanford Research Institute, Menlo Park, CA, 1965.
- [Bares et al. 89] J. Bares, M. Hebert, T. Kanade, E. Krotkov, T. Mitchell, R. Simmons, W. Whittaker, Ambler: An Autonomous Rover for Planetary Exploration, *Computer* 22:6 (1989) 18-26.
- [Beck 83] J. Beck, Textural Segmentation, Second-Order Statistics, and Textural Elements, *Biological Cybernetics* 48 (1983) 125-130.
- [Behringer et al. 92] R. Behringer, V.v. Holt, D. Dickmanns, Road and Relative Ego-State Recognition, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 385-390.
- [Bergman & Cowan 86] A. Bergman and C.K. Cowan, Noise-Tolerant Range Analysis for Autonomous Navigation, in *Proc. 5th National Conf. on AI*, Philadelphia, PA, Aug. 11-15, 1986, pp. 1122-1126.
- [Blake & Marinos 90] A. Blake and C. Marinos, Shape from Texture: Estimation, Isotropy, Moments, *Artificial Intelligence* 45 (1990) 324-380.

- [Blidberg 89] D.R. Blidberg, Autonomous Underwater Vehicles: Current Activities and Research Opportunities, in *Intelligent Autonomous System (IAS-2), An International Conference*, Amsterdam, The Netherlands, Dec. 11-14, 1989, pp. 51-64.
- [Blostein & Ahuja 89a] D. Blostein and N. Ahuja, A Multi-Scale Region Detector, *Computer Vision, Graphics, and Image Processing* **45** (1989) 22-41.
- [Blostein & Ahuja 89b] D. Blostein and N. Ahuja, Shape from Texture: Integrating Texture-Element Extraction and Surface Estimation, *IEEE Trans. Pattern Analysis and Machine Intelligence* **PAMI-11:12** (1989) 1233-1251.
- [Boldt et al. 89] M. Boldt, R. Weiss, E.M. Riseman, Token-Based Extraction of Straight Lines, *IEEE Trans. Systems, Man, and Cybernetics* **SMC-19:6** (1989) 1581-1594.
- [Bovik et al. 90] A.C. Bovik, M. Clark, V.S. Geisler, Multichannel Texture Analysis Using Localized Spatial Filters, *IEEE Trans. Pattern Analysis and Machine Intelligence* **PAMI-12:1** (1990) 55-73.
- [Brodatz 66] P. Brodatz, *Texture: A Photographic Album for Artists and Designers*, Dover Publications, New York, NY, 1966.
- [Bruyelle & Postaire 92] J.-L. Bruyelle and J.-G. Postaire, Disparity Analysis for Real Time Obstacle Detection by Linear Stereovision, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 51-56.
- [Bruyelle & Postaire 93] J.-L. Bruyelle and J.-G. Postaire, Direct Range Measurement by Linear Stereovision for Real-Time Obstacle Detection in Road Traffic, in *Intelligent Autonomous System (IAS-3), An International Conference*, Pittsburgh, Pennsylvania, February 15-18, 1993, pp. 15-23.
- [Büker et al. 93] U. Büker, H. Austermeier, G. Hartmann, B. Mertsching, Verkehrsszenenanalyse in Hierarchisch Codierten Bildern, in S.J. Pöppel and H. Handels (Hrsg.), *DAGM-Symposium Mustererkennung 1993*, Lübeck, Informatik Aktuell, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1993, pp. 694-701.
- [Burt 81] P.J. Burt, Fast Filter Transformations for Image Processing, *Computer Graphics and Image Processing* **16** (1981) 20-51.
- [Burt 83] P.J. Burt, Fast Algorithms for Estimating Local Image Properties, *Computer Vision, Graphics, and Image Processing* **21** (1983) 368-382.
- [Campbell & Thomas 92] N.W. Campbell and B.T. Thomas, Lane Boundary Tracking for an Autonomous Road Vehicle, in *Proc. of the British Machine Vision Conference*, Leeds, Sept. 22-24, 1992, pp. 157-166.
- [Canny 86] J. Canny, A Computational Approach to Edge Detection, *IEEE Trans. Pattern Analysis and Machine Intelligence* **PAMI-8:6** (1986) 679-698.
- [Català et al. 92] A. Català, A. Grau, B. Morcego, J.M. Fuertes, A Neural Network Texture Segmentation System for Open Road Vehicle Guidance, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 247-252.

- [Chan & Cheung 92] K.P. Chan and Y.S. Cheung, Clustering of Clusters, *Pattern Recognition* 25:2 (1992) 211-217.
- [Chetverikov 91] D. Chetverikov, Generating Contrast Curves for Texture Regularity Analysis, *Pattern Recognition Letters* 12 (1991) 437-444.
- [Coggins & Jain 85] J.M. Coggins and A.K. Jain, A Spatial Filtering Approach to Texture Analysis, *Pattern Recognition Letters* 3 (1985) 195-203.
- [Coleman & Andrew 79] G.B. Coleman and H.C. Andrew, Image Segmentation by Clustering, *Proceedings of the IEEE* 76:5 (1979) 773-785.
- [Conners & Harlow 80a] R.W. Conners and C.A. Harlow, A Theoretical Comparison of Texture Algorithms, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-2:3 (1980) 204-222.
- [Conners & Harlow 80b] R.W. Conners and C.A. Harlow, Toward a Structural Textural Analyzer Based on Statistical Methods, *Computer Vision, Graphics, and Image Processing* 12 (1980) 224-256.
- [Cox & Wilfong 90] I.J. Cox and G.T. Wilfong (eds.), *Autonomous Robot Vehicles*, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1990.
- [Crisman & Thorpe 88] J.D. Crisman and C.E. Thorpe, Color Vision for Road Following, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Mobile Robots III*, Cambridge, MA, November 10-11, 1988, pp. 175-184.
- [Crisman & Thorpe 91] J.D. Crisman and C.E. Thorpe, UNSCARF: A Color Vision System for the Detection of Unstructured Roads, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Sacramento, California, April 9-11, 1991, pp. 2496-2501.
- [Crisman & Webb 91] J.D. Crisman and J.A. Webb, The Warp Machine on Navlab, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-13:5 (1991) 451-465.
- [Cross & Jain 83] G.R. Cross and A.K. Jain, Markov Random Field Texture Models, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-5:1 (1983) 25-39.
- [Daily et al. 88] M. Daily, J. Harris, D. Keirse, K. Olin, D. Payton, K. Reiser, J. Rosenblatt, D. Tseng, V. Wong, Autonomous Cross-Country Navigation with the ALV, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Philadelphia, PA, April 24-29, 1988, pp. 718-726.
- [Davis et al. 79] L.S. Davis, S.A. Johns, J.K. Aggarwal, Texture Analysis Using Generalized Co-Occurrence Matrices, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-1:3 (1979) 251-259.
- [Davis et al. 86] L.S. Davis, T.R. Kushner, J.J. LeMoigne, A.M. Waxman, Road Boundary Detection for Autonomous Vehicle Navigation, *Optical Engineering* 25:3 (1986) 409-414.
- [Davis et al. 92] L.S. Davis, D. DeMenthon, S. Dickinson, P.A. Veatch, Algorithms for Road Navigation, in I. Masaki (ed.), *Vision-based Vehicle Guidance*, Proc. of Roundtable Discussion on Vision-Based Vehicle Guidance '90, Tokyo, Japan, July 2, 1990, Springer-Verlag, Berlin, Heidelberg, New York, Paris, Tokyo, Hong Kong, 1992, pp. 83-110.

- [DeMenthon 87] D. DeMenthon, A Zero-Bank Algorithm for Inverse Perspective of a Road from a Single Image, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Raleigh, North Carolina, March 31 - April 2, 1987, pp. 1444-1449.
- [DeMenthon & Davis 90] D. DeMenthon and L.S. Davis, Reconstruction of a Road by Local Image Matches and Global 3D Optimization, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Cincinnati, Ohio, May 13-18, 1990, pp. 1337-1342.
- [Dickinson & Davis 90] S.J. Dickinson and L.S. Davis, A Flexible Tool for Prototyping ALV Road Following Algorithms, *IEEE Trans. on Robotics and Automation* **RA-6:2** (1990) 232-242.
- [Dickmanns & Zapp 86] E.D. Dickmanns and A. Zapp, A Curvature-Based Scheme for Improving Road Vehicle Guidance by Computer Vision, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Volume 727, Mobile Robots*, Cambridge, Massachusetts, Oct. 30-31, 1986, pp. 161-168.
- [Dickmanns & Graefe 88a] E.D. Dickmanns and V. Graefe, Dynamic Monocular Machine Vision, *Machine Vision and Applications* **1** (1988) 223-240.
- [Dickmanns & Graefe 88b] E.D. Dickmanns and V. Graefe, Applications of Dynamic Monocular Machine Vision, *Machine Vision and Applications* **1** (1988) 241-261.
- [Dickmanns & Christians 89] E.D. Dickmanns and T. Christians, Relative 3D-State Estimation for Autonomous Visual Guidance of Road Vehicles, in *Intelligent Autonomous System (IAS-2), An International Conference*, Amsterdam, The Netherlands, Dec. 11-14, 1989, pp. 683-693.
- [Dickmanns & Mysliwetz 92] E.D. Dickmanns and B. Mysliwetz, Recursive 3D Road and Relative Ego-State Recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence* **PAMI-14:2** (1992) 199-213.
- [Dickmanns et al. 93] E.D. Dickmanns, R. Behringer, C. Brüdigam, D. Dickmanns, F. Thomanek, V.v. Holt, An All-Transputer Visual Autobahn-Autopilot/Copilot, in *Proc. Int. Conf. on Computer Vision, ICCV'93*, Berlin, Germany, May 11-14, 1993, pp. 608-615.
- [Dreyer et al. 90] W. Dreyer, U. Jacob, P. Hoppe, J. Marezke, Intelligent Computer Aided Driving, *Automobil-Industrie* **2** (1990) 147-151.
- [Duda & Hart 73] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, London, Sydney, Toronto, 1973.
- [Dunlay & Morgenthaler 86] R.T. Dunlay and D.G. Morgenthaler, Obstacle Avoidance on Roadway Using Range Data, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Volume 727, Mobile Robots*, Cambridge, Massachusetts, Oct. 30-31, 1986, pp. 110-116.
- [Dunlay 88] R.T. Dunlay, Obstacle Avoidance Perception for Autonomous Land Vehicle, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Philadelphia, PA, April 24-29, 1988, pp. 912-917.

- [Durbin & Willshaw 87] R. Durbin and D. Willshaw, An Analogue Approach to the Travelling Salesman Problem Using an Elastic Net Method, *Nature* **326** (1987) 689-691.
- [Efenberger *et al.* 92] W. Efenberger, Q.-H. Ta, L. Tsinas, V. Graefe, Automatic Recognition of Vehicles Approaching From Behind, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 57-62.
- [Enkelmann 90] W. Enkelmann, Obstacle Detection by Evaluation of Optical Flow Fields from Image Sequences, in *Proc. First European Conference on Computer Vision*, Antibes, France, Apr. 23-26, 1990, O. Faugeras (ed.), Lecture Notes in Computer Science 427, Springer-Verlag, Berlin, Heidelberg, New York etc., 1990, pp. 134-138.
- [Enkelmann 91] W. Enkelmann, Obstacle Detection by Evaluation of Optical Flow Fields from Image Sequences, *Image and Vision Computing* **9:3** (1991) 160-168.
- [Faugeras 78] O.D. Faugeras, Texture Analysis and Classification Using a Human Visual Model, in *Proc. Int. Conf. on Pattern Recognition*, Kyoto, Japan, Nov. 7-10, 1978, pp. 549-559.
- [Fernandez-Maloigne *et al.* 92] C. Fernandez-Maloigne, D. Laugier, A. Bekkhoucha, Texture Analysis for Road Detection, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 219-224.
- [Fisher 85] J.J. Fisher, Applying Robots in Nuclear Applications, in *ROBOTS-9 Conf. Proc., Volume 1, Advancing Applications*, Detroit, Michigan, June 2-6, 1985, pp. 8.1-8.20.
- [Franke 89] U. Franke, PROMETHEUS- wissensbasierte Systeme eröffnen neue Perspektiven im Straßenverkehr, in W. Brauer and C. Freksa (Hrsg.), *Wissensbasierte Systeme*, Informatik-Fachberichte 227, Springer-Verlag, Berlin, Heidelberg, New York, 1989, pp. 363-376.
- [Franke 91] U. Franke, Real Time 3D-Road Modeling for Autonomous Vehicle Guidance, in *Proc. of the 7th Scandinavian Conf. on Image Analysis*, Kopenhagen, August 13-18, 1991, pp. 316-323.
- [Franke *et al.* 91] U. Franke, H. Fritz, S. Mehring, Long Distance Driving With the Daimler-Benz Autonomous Vehicle VITA, in *Proc. of Prometheus PRO-ART Workshop on Intelligent Co-Pilot*, Grenoble, France, Dec. 12-13, 1991, pp. 239-246.
- [Fu & Mui 81] K.S. Fu and J.K. Mui, A Survey on Image Segmentation, *Pattern Recognition* **13** (1981) 3-16.
- [Galloway 75] M. M. Galloway, Texture Classification Using Gray Level Run Length, *Computer Graphics and Image Processing* **4** (1975) 172-179.
- [Geisler *et al.* 90] J. Geisler, H.-H. Nagel, G. Siegle, G. Struck, Schritte zum autarken Fahren auf privaten Verkehrswegen, in *6. Fachgespräch Autonom Mobile Systeme*, U. Rembold, R. Dillmann und P. Levi (Hrsg.), Karlsruhe, November 26-27, 1990, pp. 233-239.
- [Gilmore & Semeco 86] J.F. Gilmore and A.C. Semeco, Knowledge-Based Approach Toward Developing an Autonomous Helicopter System, *Optical Engineering* **25:3** (1986) 415-427.

- [Giordina & Dougherty 88] C.R. Giordina and E.R. Dougherty, *Morphological Methods in Image and Signal Processing*, Prentice Hall, Englewood Cliffs, New Jersey, 1988.
- [Gool *et al.* 85] L. Van Gool, P. Dewable, A. Oosterlinck, SURVEY : Texture Analysis Anno 1983, *Computer Vision, Graphics, and Image Processing* **29** (1985) 336–357.
- [Gotlieb & Kreyszig 90] C.C. Gotlieb and H.E. Kreyszig, Texture Descriptors Based on Co-Occurrence Matrices, *Computer Vision, Graphics, and Image Processing* **51** (1990) 70–86.
- [Goto & Stentz 87] Y. Goto and A. Stentz, The CMU System for Mobile Robot Navigation, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Raleigh, North Carolina, March 31 - April 2, 1987, pp. 99–105.
- [Graefe *et al.* 88] V. Graefe, U. Regensburger, U. Solder, Visuelle Entdeckung und Vermessung von Objekten in der Bahn eines autonom mobilen Systems, in H. Bunke, O. Kübler, P. Stucki (Hrsg.), *DAGM-Symposium Mustererkennung 1988*, Zürich, Informatik-Fachbericht **180**, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1988, pp. 312–318.
- [Graefe 90] V. Graefe, The BVV-Family of Robot Vision Systems, in *Proc. IEEE Workshop on Intelligent Motion Control*, Istanbul, Turkey, Aug. 20-22, 1990, pp. IP55–IP65.
- [Grimson & Hildreth 85] W.E.L. Grimson and E. Hildreth, Comments on: Digital Step Edges from Zero Crossings of Second Directional Derivatives, *IEEE Trans. Pattern Analysis and Machine Intelligence* **PAMI-7:1** (1985) 121–129.
- [Grimson 90] W.E.L. Grimson, *Object Recognition by Computer: The Role of Geometric Constraints*, The MIT Press, Cambridge, MA, London, UK, 1990.
- [Groen *et al.* 93] F.C.A. Groen, S. Hirose, C.E. Thorpe (eds.), *Intelligent Autonomous Systems (IAS-3)*, An International Conference, Pittsburgh, Pennsylvania, February 15-18, 1993, IOS Press, 1993.
- [Haddon *et al.* 93] J.F. Haddon, J.F. Boyce, S. Protheroe, S. Hesketh, Neural Networks for the Texture Classification of Segmented Regions of Forward Looking Infrared Images, in *Proc. of the British Machine Vision Conference*, Uni. of Surrey, Guildford, Sept. 21-23, 1993, pp. 197–206.
- [Hallset 92] J.O. Hallset, A Vision System for an Autonomous Underwater Vehicle, in *Proc. Int. Conf. on Pattern Recognition, Volume I, Computer Vision and Applications*, The Hague, The Netherlands, Aug. 30 - Sept. 3, 1992, p. 320–323.
- [Hanson & Riseman 78] A.R. Hanson and E.M. Riseman, VISIONS: A Computer System for Interpreting Scenes, in A.R. Hanson and E.M. Riseman (eds.), *Computer Vision Systems*, Academic Press, New York, San Francisco, London, 1978, pp. 303–333.
- [Haralick *et al.* 73] R.M. Haralick, K. Shanmugam, I. Dinstein, Texture Features for Image Classification, *IEEE Trans. Systems, Man, and Cybernetics* **SMC-3:6** (1973) 610–621.
- [Haralick 79] R.M. Haralick, Statistical and Structural Approaches to Texture, *Proceedings of the IEEE* **67:5** (1979) 796–804.

- [Haralick 84] R.M. Haralick, Digital Step Edges from Zero Crossing of Second Directional Derivatives, *IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-6:1* (1984) 58-68.
- [Haralick & Shapiro 85] R.M. Haralick and L.G. Shapiro, Image Segmentation Techniques, *Computer Vision, Graphics, and Image Processing* 29 (1985) 100-132.
- [Haralick & Shapiro 91] R.M. Haralick and L.G. Shapiro, Glossary of Computer Vision Terms, *Pattern Recognition* 24:1 (1991) 69-93.
- [Haralick & Shapiro 92] R.M. Haralick and L.G. Shapiro, *Computer and Robot Vision: Volumes 1 and 2*, Addison-Wesley Publishing Company, Amsterdam, Bonn, Don Mills, Madrid, New York, Paris, Singapore, Tokyo, 1992.
- [Hartmann & Mertsching 92] G. Hartmann and B. Mertsching, A Hierarchical Vision System, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 18-23.
- [Hashimoto *et al.* 92] K. Hashimoto, S. Nakayama, T. Saito, S. Ishida, K. Unoura, J. Ishii, N. Oono, Y. Okada, An Image Processing Architecture and a Motion Control Method for an Autonomous Vehicle, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 213-218.
- [Hattori *et al.* 92] A. Hattori, A. Hosaka, M. Taniguchi, E. Nakano, Driving Control System for an Autonomous Vehicle Using Multiple Observed Point Information, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 207-212.
- [He & Wang 91] D.-C. He and L. Wang, Textural Filters Based on the Texture Spectrum, *Pattern Recognition* 24:12 (1991) 1187-1195.
- [Hebert & Kanade 86] M. Hebert and T. Kanade, Outdoor Scene Analysis Using Range Data, in *Proc. IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, April 7-10, 1986, pp. 1426-1432.
- [Hebert 89] M. Hebert, Building and Navigating Maps of Road Scenes Using an Active Sensor, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Scottsdale, Arizona, May 14-19, 1989, pp. 1136-1142.
- [Hertz *et al.* 91] J. Hertz, A. Krogh, R.G. Palmer, *Introduction to the Theory of Neural Computation*, Addison-Wesley Publishing Company, Amsterdam, Boon, Don Mills, Madrid, New York, Paris, Singapore, Tokyo, 1991.
- [Horn 86] B.K.P. Horn, *Robot Vision*, MIT Press, Cambridge, MA, 1986.
- [Hsiao & Sawchuk 89] J.Y. Hsiao and A.A. Sawchuk, Supervised Textured Image Segmentation Using Feature Smoothing and Probabilistic Relaxation Techniques, *IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-11:12* (1989) 1279-1292.
- [Hu 62] M.-K. Hu, Visual Pattern Recognition by Moment Invariants, *IRE Trans. on Inform. Theory* IT-8 (1962) 179-187.

- [Huntsberger *et al.* 85] T.L. Huntsberger, C.L. Jacobs, R.L. Cannon, Iterative Fuzzy Image Segmentation, *Pattern Recognition* **18:2** (1985) 131-138.
- [Ito *et al.* 92] T. Ito, T. Sakagami, S. Kawakatsu, A Real Time Distance Headway Measurement Method Using Stereo and Optical Flow, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 230-235.
- [Jain & Binford 91] R.C. Jain and T.O. Binford, Dialogue: Ignorance, Myopia, and Naiveté in Computer Vision Systems, *Computer Vision, Graphics, and Image Processing: Image Understanding* **51:1** (1991) 112-117.
- [Jain & Farrokhnia 91] A.K. Jain and F. Farrokhnia, Unsupervised Texture Segmentation Using Gabor Filters, *Pattern Recognition* **24:12** (1991) 1167-1186.
- [Jochem *et al.* 93] T.M. Jochem, D.A. Pomerleau, C.E. Thorpe, MANIAC: A Next Generation Neurally Based Autonomous Road Follower, in *Intelligent Autonomous System (IAS-3), An International Conference*, Pittsburgh, Pennsylvania, February 15-18, 1993, pp. 592-599.
- [Jolion *et al.* 91] J.-M. Jolion, P. Meer, S. Bataouche, Robust Clustering with Applications in Computer Vision, *IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-13:8* (1991) 791-802.
- [Julesz 75] B. Julesz, Experiments in the Visual Perception of Texture, *Scientific American* **232** (1975) 34-43.
- [Julesz & Bergen 83] B. Julesz and J.R. Bergen, Textons, The Fundamental Elements in Preattentive Vision and Perception of Textures, *The Bell System Technical Journal* **62:6** (1983) 1619-1645.
- [Julesz 86] B. Julesz, Texton Gradients: The Texton Theory Revisited, *Biological Cybernetics* **54** (1986) 245-251.
- [Kanade 78] T. Kanade, Region Segmentation: Signal versus Semantics, in *Proc. Int. Conf. on Pattern Recognition*, Kyoto, Japan, Nov. 7-10, 1978, pp. 95-105.
- [Kanade *et al.* 89] T. Kanade, F.C.A. Groen, L.O. Hertzberger (eds.), *Intelligent Autonomous Systems (IAS-2)*, An International Conference, Amsterdam, The Netherlands, Dec. 11-14, 1989, IOS Press, 1989.
- [Kass & Witkin 87] M. Kass and A.P. Witkin, Analyzing Oriented Patterns, *Computer Vision, Graphics, and Image Processing* **37** (1987) 362-385.
- [Kenue 89a] S.K. Kenue, Lanelok: Detection of Lane Boundaries and Vehicle Tracking Using Image-Processing Techniques - Part I: Hough-Transform, Region-Tracking and Correlation Algorithms, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Volume 1195, Mobile Robots IV*, Philadelphia, PA, Nov. 6-7, 1989, pp. 221-233.
- [Kenue 89b] S.K. Kenue, Lanelok: Detection of Lane Boundaries and Vehicle Tracking Using Image-Processing Techniques - Part II: Template Matching Algorithms, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Volume 1195, Mobile Robots IV*, Philadelphia, PA, Nov. 6-7, 1989, pp. 234-245.

- [Khotanzad & Chen 89] A. Khotanzad and J.-Y. Chen, Unsupervised Segmentation of Textured Images by Edge Detection in Multidimensional Features, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-11:4 (1989) 414-421.
- [Kluge & Thorpe 92] K. Kluge and C.E. Thorpe, Representation and Recovery of Road Geometry in YARF, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 114-119.
- [Kluge & Thorpe 93] K. Kluge and C.E. Thorpe, Intersection Detection in the YARF Road Following System, in *Intelligent Autonomous System (IAS-3), An International Conference*, Pittsburgh, Pennsylvania, February 15-18, 1993, pp. 145-154.
- [Korn 88] A. Korn, Toward a Symbolic Representation of Intensity Changes in Images, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-10 (1988) 610-625.
- [Kristen & Munkelt 92] H. Kristen and O. Munkelt, Markov-Feld-basierte Bildinterpretation mit automatisch generierten Datenbasen, in S. Fuchs and R. Hoffmann (Hrsg.), *DAGM-Symposium Mustererkennung 1992*, Dresden, Informatik Aktuell, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1992, pp. 50-57.
- [Kuan et al. 86] D. Kuan, G. Phipps, A.C. Hsueh, A Real-Time Road Following and Road Junction Detection Vision System for Autonomous Vehicles, in *Proc. 5th National Conf. on AI*, Philadelphia, PA, Aug. 11-15, 1986, pp. 1127-1132.
- [Kuan & Sharma 87] D. Kuan and U.K. Sharma, Model-Based Geometric Reasoning for Autonomous Road Following, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Raleigh, North Carolina, March 31 - April 2, 1987, pp. 416-423.
- [Kuan et al. 88] D. Kuan, G. Phipps, A.C. Hsueh, Autonomous Robotic Vehicle Road Following, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-10:5 (1988) 648-658.
- [Kuhnert 86] K.-D. Kuhnert, A Vision System for Real Time Road and Object Recognition for Vehicle Guidance, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Volume 727, Mobile Robots*, Cambridge, Massachusetts, Oct. 30-31, 1986, pp. 267-272.
- [Kurita et al. 92] T. Kurita, N. Otsu, N. Abdelmalik, Maximum likelihood thresholding based on population mixture models, *Pattern Recognition* 25 (1992) 1231-1240.
- [Larkin & Burt 83] L.I. Larkin and P.J. Burt, Multi-Resolution Texture Energy Measures, in *IEEE Conf. on Computer Vision and Pattern Recognition*, Washington, D.C., June 19-23, 1983, pp. 519-520.
- [Laws 80] K. Laws, Rapid Texture Identification, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Volume 238, Image Processing for Missile Guidance*, San Diego, CA, July 29 - Aug. 1, 1980, pp. 367-380.
- [Lawton et al. 86] D.T. Lawton, T.S. Levitt, C. McConnell, J. Glicksman, Terrain Models for an Autonomous Land Vehicle, in *Proc. IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, April 7-10, 1986, pp. 2043-2051.

- [LeMoigne *et al.* 86] J. LeMoigne, A.M. Waxman, L.S. Davis, Knowledge-Based Reasoning for Road Following, in *Proc. IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, April 7-10, 1986, pp. 244-247.
- [Leu & Wee 85] J.-G. Leu and W.G. Wee, Detecting the Spatial Structure of Natural Textures Based on Shape Analysis, *Computer Vision, Graphics, and Image Processing* **31** (1985) 67-88.
- [Levitt & Lawton 89] T.S. Levitt and D.T. Lawton, Visual Re-Acquisition of Geographic Locations, in *DARPA Image Understanding Workshop*, Palo Alto, CA, May 23-26, 1989, pp. 950-956.
- [Levitt & Lawton 90] T.S. Levitt and D.T. Lawton, Qualitative Navigation for Mobile Robots, *Artificial Intelligence* **44** (1990) 305-360.
- [Li 92] Y. Li, Reforming the Theory of Invariant Moments for Pattern Recognition, *Pattern Recognition* **25:7** (1992) 723-730.
- [Lim & Malik 81] J.S. Lim and N.A. Malik, A New Algorithm for Two-Dimensional Maximum Entropy Power Spectral Estimation, *IEEE Trans. Acoustics, Speech, and Signal Processing ASSP-29:3* (1981) 401-413.
- [Liou & Jain 87] S.P. Liou and R.C. Jain, Road Following Using Vanishing Points, *Computer Vision, Graphics, and Image Processing* **39** (1987) 116-130.
- [Liu & Jernigan 90] S.-S. Liu and M.E. Jernigan, Texture Analysis and Discrimination in Additive Noise, *Computer Vision, Graphics, and Image Processing* **49** (1990) 52-67.
- [Lohmann 93] G. Lohmann, Texturklassifikation mit Polynomialverteilungen, in S.J. Pöpl and H. Handels (Hrsg.), *DAGM-Symposium Mustererkennung 1993*, Lübeck, Informatik Aktuell, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1993, pp. 576-583.
- [Lowe 85] D.G. Lowe, *Perceptual Organization and Visual Recognition*, Kluwer Academic Publishers, Boston, Dordrecht, Lancaster, 1985.
- [Lu & Fu 78] S.Y. Lu and K.S. Fu, A Syntactic Approach to Texture analysis, *Computer Graphics and Image Processing* **7** (1978) 303-330.
- [Marr 82] D. Marr, *Vision*, W.H. Freeman, San Francisco, CA, 1982.
- [Marr & Hildreth 80] D. Marr and E. Hildreth, Theory of Edge Detection, *Proc. Royal Society of London B* **207** (1980) 187-217.
- [Masaki 92] I. Masaki (ed.), *Vision-based Vehicle Guidance*, Proc. of Roundtable Discussion on Vision-Based Vehicle Guidance '90, Tokyo, Japan, July 2, 1990, Springer-Verlag, Berlin, Heidelberg, New York, Paris, Tokyo, Hong Kong, 1992.
- [Monga 87] O. Monga, An Optimal Region Growing Algorithm for Image Segmentation, *International Journal of Pattern Recognition and Artificial Intelligence* **1** (1987) 351-375.

- [Moravec 83] H.P. Moravec, The Stanford Cart and the CMU Rover, *Proceedings of the IEEE* 71:7 (1983) 872-884.
- [Morgan *et al.* 90] A.D. Morgan, E.L. Dagless, D.J. Milford, B.T. Thomas, Road Edge Tracking for Robot Road Following : A Real-Time Implementation, *Image and Vision Computing* 8:3 (1990) 233-240.
- [Morgenthaler *et al.* 90] D.G. Morgenthaler, S.J. Hennessy, D. DeMenthon, Range-Video Fusion and Comparison of Inverse Perspective Algorithms in Static Images, *IEEE Trans. Systems, Man, and Cybernetics* SMC-20:6 (1990) 1301-1312.
- [Mori 90] H. Mori, Active Sensing in Stereotyped Motion, in *Proc. IEEE Workshop on Intelligent Motion Control*, Istanbul, Turkey, Aug. 20-22, 1990, pp. IP11-IP19.
- [Mori 92] H. Mori, Recent Progress in Mobile Robot Harunobu-4, in I. Masaki (ed.), *Vision-based Vehicle Guidance*, Proc. of Roundtable Discussion on Vision-Based Vehicle Guidance '90, Tokyo, Japan, July 2, 1990, Springer-Verlag, Berlin, Heidelberg, New York, Paris, Tokyo, Hong Kong, 1992, pp. 204-221.
- [Müller 91] C. Müller, *Verwendung von Bildauswertungsmethoden zur Erkennung und Lagebestimmung von generischen polyedrischen Objekten im Raum*, Dissertation, Fakultät für Informatik der Universität Karlsruhe, durchgeführt am Fraunhofer-Institut für Informations- und Datenverarbeitung (IITB), Nov. 1991.
- [Münkel & Welz 91] H. Münkel and K. Welz, Verfolgung des Straßenverlaufes in einer Farbbildfolge, in B. Radig (Hrsg.), *DAGM-Symposium Mustererkennung 1991*, München, Informatik-Fachbericht 290, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1991, pp. 515-520.
- [Nagel 83] H.-H. Nagel, Displacement Vectors Derived from Second-Order Intensity Variations in Image Sequences, *Computer Vision, Graphics, and Image Processing* 21 (1983) 85-117.
- [Nagel 87] H.-H. Nagel, Principles of (Low-Level) Computer Vision, in J.P. Haton (ed.), *Fundamentals in Computer Understanding: Speech, Vision, and Natural Language*, Cambridge Univ. Press, Cambridge, UK, 1987, pp. 113-139.
- [Nagel 88] H.-H. Nagel, Wodurch zeichnet sich ein "umsichtiges" Automobil aus? oder Zur Frage der impliziten und expliziten Repräsentation von Wissen in maschinellen Ko-Piloten für den Straßenverkehr, in *Tagungsband VDI-Fachtagung "Elektronik im Kraftfahrzeug"*, Baden-Baden, 8.-9. September 1988, VDI Berichte 687, VDI Verlag Düsseldorf, 1988, pp. 25-40.
- [Nagel 91a] H.-H. Nagel, *Kognitive Systeme*, Vorlesungsskript für SS 1990 und SS 1991, Fakultät für Informatik der Universität Karlsruhe, September 1991.
- [Nagel 91b] H.-H. Nagel, Wissenbasierte Systeme für Anwendungen im Verkehr, in *Entwicklungslinien in Kraftfahrzeugtechnik und Straßenverkehr: Forschungsbilanz 1991*, 14. Statuseminar des Bundesministeriums für Forschung und Technologie, 13.-15. Mai 1991, Dresden, Bundesrepublik Deutschland, Projektbegleitung TÜV Rheinland e.V., Köln, 1991, pp. 457-466.

- [Nasr & Bhanu 88] H. Nasr and B. Bhanu, Landmark Recognition for Autonomous Mobile Robots, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Philadelphia, PA, April 24-29, 1988, pp. 1218-1223.
- [Negahdaripour & Jain 91] S. Negahdaripour and A.K. Jain (eds.), *Final Report of the NSF Workshop on the Challenges in Computer Vision Research: Future Directions of Research*, Maui, Hawaii, June 7-8, 1991.
- [Negahdaripour & Jain 92] S. Negahdaripour and A.K. Jain, Challenges in Computer Vision: Future Research Directions, in *IEEE Conf. on Computer Vision and Pattern Recognition*, Champaign, Illinois, June 15-18, 1992, pp. 189-199.
- [Okuno et al. 92] A. Okuno, K. Fujita, A. Kutami, Visual Navigation of Autonomous On-Road Vehicle - Autonomous Cruising on Highways -, in I. Masaki (ed.), *Vision-based Vehicle Guidance*, Proc. of Roundtable Discussion on Vision-Based Vehicle Guidance '90, Tokyo, Japan, July 2, 1990, Springer-Verlag, Berlin, Heidelberg, New York, Paris, Tokyo, Hong Kong, 1992, pp. 222-236.
- [Olin & Tseng 91] K.E. Olin and D.Y. Tseng, Autonomous Cross-Country Navigation, *IEEE Expert* 6:4 (1991) 16-30.
- [Orban & Nagel 92] G.A. Orban and H.-H. Nagel (eds.), *Artificial and Biological Vision Systems*, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1992.
- [Otsu 78] N. Otsu, Discriminant and Least Squares Threshold Selection, in *Proc. Int. Conf. on Pattern Recognition*, Kyoto, Japan, Nov. 7-10, 1978, pp. 592-596.
- [Ozaki et al. 89] T. Ozaki, M. Ohzora, K. Kurahashi, An Image Processing System for Autonomous Vehicle, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Volume 1195, Mobile Robots IV*, Philadelphia, PA, Nov. 6-7, 1989, pp. 256-266.
- [Ozawa & Rosenfeld 86] S. Ozawa and A. Rosenfeld, Synthesis of a Road Image as Seen from a Vehicle, *Pattern Recognition* 19:2 (1986) 123-145.
- [Pappas 92] T.N. Pappas, An Adaptive Clustering Algorithm for Image Segmentation, *IEEE Trans. on Signal Processing* 40:4 (1992) 901-914.
- [Paul et al. 88] D. Paul, W. Hättich, W. Nill, S. Tatari, G. Winkler, VISTA : Visual Interpretation System for Technical Applications - Architecture and Use, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-10:3 (1988) 399-407.
- [Paulus et al. 93] D.W.R. Paulus, H. Niemann, C. Lenz, L. Demling, C. Ell, Fraktale Dimension der Kontur endoskopisch ermittelter Farbbilder von Geschwüren des Magens, in S.J. Pöppel and H. Handels (Hrsg.), *DAGM-Symposium Mustererkennung 1993*, Lübeck, Informatik Aktuell, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1993, pp. 484-491.
- [Pavlidis 77] T. Pavlidis, *Structural Pattern Recognition*, Springer-Verlag, Berlin, Heidelberg, New York, 1977.

- [Pavlidis & Liow 90] T. Pavlidis and Y.-T. Liow, Integrating Region Growing and Edge Detection, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-12:3 (1990) 225-233.
- [Pavlidis 92] T. Pavlidis, Why Progress in Machine Vision is so Slow, *Pattern Recognition Letters* 13 (1992) 221-225.
- [Peckinpaugh 91] S.H. Peckinpaugh, An Improved Method for Computing Gray-Level Cooccurrence Matrix Based Texture Measures, *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing* 53:6 (1991) 574-580.
- [Pentland 84] A. Pentland, Fractal-Based Description of Natural Scenes, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-6:6 (1984) 661-674.
- [Perona 91] P. Perona, Deformable kernels for early vision, in *IEEE Conf. on Computer Vision and Pattern Recognition*, Lahaina, Maui, Hawaii, June 3-6, 1991, pp. 222-227.
- [Perry & Lowe 89] A. Perry and D.G. Lowe, Segmentation of Textured Images, in *IEEE Conf. on Computer Vision and Pattern Recognition*, San Diego, CA, June 4-8, 1989, pp. 319-325.
- [Petkovic & Wilder 91] D. Petkovic and J. Wilder, Machine Vision in the 1990s: Applications and How to Get There, *Machine Vision and Applications* 4 (1991) 113-126.
- [Phillips & Rosenfeld 88] T.-Y. Phillips and A. Rosenfeld, An Isodata Algorithm for Straight Line Fitting, *Pattern Recognition Letters* 7 (1988) 291-297.
- [Pietikäinen & Rosenfeld 82] M. Pietikäinen and A. Rosenfeld, Edge-Based Texture Measures, in *Proc. Int. Conf. on Pattern Recognition*, Munich, Germany, Oct. 19-22, 1982, pp. 298-300.
- [Pietikäinen *et al.* 82] M. Pietikäinen, A. Rosenfeld, L.S. Davis, Texture Classification Using Averages of Local Pattern Matches, in *Proc. Int. Conf. on Pattern Recognition*, Munich, Germany, Oct. 19-22, 1982, pp. 301-303.
- [Polk & Jain 92] A. Polk and R. Jain, A Parallel Architecture for Curvature-Based Road Scene Classification, in I. Masaki (ed.), *Vision-based Vehicle Guidance*, Proc. of Roundtable Discussion on Vision-Based Vehicle Guidance '90, Tokyo, Japan, July 2, 1990, Springer-Verlag, Berlin, Heidelberg, New York, Paris, Tokyo, Hong Kong, 1992, pp. 284-299.
- [Pomerleau 89] D.A. Pomerleau, ALVINN: An Autonomous Land Vehicle In a Neural Network, in D.S. Touretzky (ed.), *Advances in Neural Information Processing Systems 1*, Morgan Kaufmann Publishers, San Mateo, CA, USA, 1989, pp. 305-313.
- [Pomerleau 92] D.A. Pomerleau, Progress in Neural Network-based Vision for Autonomous Robot Driving, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 391-396.
- [Pomerleau & Touretzky 93] D.A. Pomerleau and D.S. Touretzky, Analysis of Feature Detectors Learned by a Neural Network Autonomous Driving System, in *Intelligent Autonomous System (IAS-3)*, An International Conference, Pittsburgh, Pennsylvania, February 15-18, 1993, pp. 572-581.

- [Pong *et al.* 92] T.-C. Pong, S. Madarasmi, Y.-L. Chu, Texture Segmentation Using Topographic Labels, in L. Shapiro and A. Rosenfeld (eds.), *Computer Vision and Image Processing*, Academic Press, Inc., Boston, San Diego, New York, London, Tokyo, Toronto, 1992, pp. 537-554.
- [Porat & Zeevi 88] M. Porat and Y.Y. Zeevi, The Generalized Gabor Scheme of Image Representation in Biological and Machine Vision, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-10:4 (1988) 452-467.
- [Rao & Schunck 89] A.R. Rao and B.G. Schunck, Computing Oriented Texture Fields, in *IEEE Conf. on Computer Vision and Pattern Recognition*, San Diego, CA, June 4-8, 1989, pp. 61-68.
- [Rao 90] A.R. Rao, *A Taxonomy for Texture Description and Identification*, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1990.
- [Ras 91] Special Issue: Intelligent Autonomous Systems, *Robotics and Autonomous Systems* 7:2-3 (1991).
- [RAS-L-1 84] *Richtlinien für die Anlage von Straßen (RAS), Teil: Linienführung (RAS-L), Abschnitt 1: Elemente der Linienführung (RAS-L-1)*, Kirschbaum Verlag, Bonn - Bad Godesberg, 1984.
- [Reed & Buf 93] T.R. Reed and J.M.H. Du Buf, A Review of Recent Texture Segmentation and Feature Extraction Techniques, *Computer Vision, Graphics, and Image Processing: Image Understanding* 57:3 (1993) 359-372.
- [Rembold 88] U. Rembold, The Karlsruhe Autonomous Mobile Assembly Robot, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Philadelphia, PA, April 24-29, 1988, pp. 598-603.
- [Ritter 92] W. Ritter, Traffic Sign Recognition in Color Image Sequences, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 12-17.
- [Rueff 89] M. Rueff, Scale Space Filtering and the Scaling Regions of Fractals, in J.C. Simon (ed.), *From Pixels to Features*, Elsevier Science Publishers B.V. (North-Holland), Amsterdam, New York, Oxford, Tokyo, 1989, pp. 49-60.
- [Sahoo *et al.* 88] P.K. Sahoo, S. Soltani, A.K. Wong, Y.C. Chen, SURVEY: A Survey of Thresholding Techniques, *Computer Vision, Graphics, and Image Processing* 41 (1988) 233-260.
- [Sakurai *et al.* 87] K. Sakurai, H. Zen, H. Ohta, Y. Ushioda, S. Ozawa, Analysis of a Road Image as Seen from a Vehicle, in *Proc. Int. Conf. on Computer Vision, ICCV'87*, London, UK, June 8-11, 1987, pp. 651-656.
- [Schaaser & Thomas 92] L.T. Schaaser and B.T. Thomas, Finding Road Lane Boundaries for Vision Guided Vehicle Navigation, in I. Masaki (ed.), *Vision-based Vehicle Guidance*, Proc. of Roundtable Discussion on Vision-Based Vehicle Guidance '90, Tokyo, Japan, July 2, 1990, Springer-Verlag, Berlin, Heidelberg, New York, Paris, Tokyo, Hong Kong, 1992, pp. 238-254.
- [Schalkoff 92] R.J. Schalkoff, *Pattern Recognition*, John Wiley & Sons, New York, NY, 1992.

- [Schwarzinger *et al.* 92a] M. Schwarzinger, T. Zielke, D. Noll, M. Brauckmann, W. von Seelen, Vision-Based Car-Following: Detection, Tracking, and Identification, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 24-29.
- [Schwarzinger *et al.* 92b] M. Schwarzinger, D. Noll, W. von Seelen, Object Recognition with Deformable Models Using Constrained Elastic Nets, in S. Fuchs and R. Hoffmann (Hrsg.), *DAGM-Symposium Mustererkennung 1992*, Dresden, Informatik Aktuell, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1992, pp. 96-104.
- [Seitz *et al.* 91] P. Seitz, G.K. Lang, B. Gilliard, J.C. Pandazis, The Robust Recognition of Traffic Signs from a Moving Car, in B. Radig (Hrsg.), *DAGM-Symposium Mustererkennung 1991*, München, Informatik-Fachbericht 290, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1991, pp. 287-294.
- [Sekine *et al.* 92] M. Sekine, T. Senoo, I. Morita, H. Endo, Design Method for an Automotive Laser Radar System and Future Prospects for Laser Radar, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 120-125.
- [Sharma & Davis 86] U.K. Sharma and L.S. Davis, Road Following by an Autonomous Vehicle Using Range Data, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Volume 727, Mobile Robots*, Cambridge, Massachusetts, Oct. 30-31, 1986, pp. 169-179.
- [Sher & Rosenfeld 89] C.A. Sher and A. Rosenfeld, Detecting and Extracting Compact Textured Regions Using Pyramids, *Image and Vision Computing* 7:2 (1989) 129-134.
- [Shigematu *et al.* 92] T. Shigematu, Y. Hashimoto, Y. Watanabe, Development of Automatic Driving System on Rough Road - Automatic Steering Control by Fuzzy Algorithm -, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 154-159.
- [Siegle *et al.* 92] G. Siegle, J. Geisler, F. Laubenstein, H.-H. Nagel, G. Struck, Autonomous Driving on a Road Network, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 403-408.
- [Singh & Keller 91] S. Singh and P. Keller, Obstacle Detection for High Speed Autonomous Navigation, in *Proc. IEEE Int. Conf. on Robotics and Automation*, Sacramento, California, April 9-11, 1991, pp. 2798-2805.
- [Solder & Graefe 90] U. Solder and V. Graefe, Object Detection in Real Time, in *Proc. of SPIE - The Inter. Soc. for Optical Engineering, Volume 1388, Mobile Robots V*, Boston, Nov. 8-9, 1990, pp. 104-111.
- [Struck *et al.* 93] G. Struck, J. Geisler, F. Laubenstein, H.-H. Nagel, G. Siegle, Interaction Between Digital Road Map Systems and Trinocular Autonomous Driving, in *Proc. Symposium on Intelligent Vehicles '93*, Tokyo, Japan, July 14 - 16, 1993, pp. 461-466.
- [Thorpe *et al.* 85] C.E. Thorpe, L. Mathies, H. Moravec, Experiments and Thoughts on Visual Navigation, in *Proc. IEEE Int. Conf. on Robotics and Automation*, St. Louis, Missouri, March 25-28, 1985, pp. 830-835.

- [Thorpe *et al.* 88] C.E. Thorpe, M. Hebert, T. Kanade, S.A. Shafer, Vision and Navigation for the Carnegie-Mellon Navlab, *IEEE Trans. Pattern Analysis and Machine Intelligence* 10:3 (1988) 362-373.
- [Thorpe 90] C.E. Thorpe (ed.), *Vision and Navigation - The Carnegie Mellon Navlab*, Kluwer Academic Publishers, Boston, Dordrecht, London, 1990.
- [Thorpe 91] C.E. Thorpe, Mobile Robots, *International Journal of Pattern Recognition and Artificial Intelligence* 5:3 (1991) 383-397.
- [Thorpe *et al.* 91] C.E. Thorpe, M. Hebert, T. Kanade, S.A. Shafer, Toward Autonomous Driving: The CMU Navlab, *IEEE Expert* 6:4 (1991) 31-52.
- [Thorpe *et al.* 92] C.E. Thorpe, M. Hebert, T. Kanade, S. Shafer, The New Generation System for the CMU Navlab, in I. Masaki (ed.), *Vision-based Vehicle Guidance*, Proc. of Roundtable Discussion on Vision-Based Vehicle Guidance '90, Tokyo, Japan, July 2, 1990, Springer-Verlag, Berlin, Heidelberg, New York, Paris, Tokyo, Hong Kong, 1992, pp. 30-82.
- [Tomita *et al.* 82] F. Tomita, Y. Shirai, S. Tsuji, Description of Textures by a Structural Analysis, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-4:2 (1982) 183-191.
- [Tomita & Tsuji 90] F. Tomita and S. Tsuji, *Computer Analysis of Visual Textures*, Kluwer Academic Publishers, Boston, Dordrecht, London, 1990.
- [Tsugawa *et al.* 92] S. Tsugawa, T. Saito, A. Hosaka, Super Smart Vehicle System: AVCS Related Systems for the Future, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 132-137.
- [Tuceryan 92] M. Tuceryan, Moment Based Texture Segmentation, in *Proc. Int. Conf. on Pattern Recognition, Volume III, Image, Speech, and Signal Analysis*, The Hague, The Netherlands, Aug. 30 - Sept. 3, 1992, pp. 45-48.
- [Turk & Marra 86] M.A. Turk and M. Marra, Color Road Segmentation and Video Obstacle Detection, in *Proc. of SPIE - Soc. of Photo-Optical Instrumentation Engineers, Volume 727, Mobile Robots*, Cambridge, Massachusetts, Oct. 30-31, 1986, pp. 136-142.
- [Turk *et al.* 88] M.A. Turk, D.G. Morgenthaler, K.D. Gremban, M. Marra, VITS - A Vision System for Autonomous Land Vehicle Navigation, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-10:3 (1988) 342-361.
- [Ulmer 92] B. Ulmer, VITA - An Autonomous Road Vehicle (ARV) for Collision Avoidance in Traffic, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 36-41.
- [Unser 86] M. Unser, Sum and Difference Histograms for Texture Classification, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-8:1 (1986) 118-125.
- [Unser & Eden 89] M. Unser and M. Eden, Multiresolution Feature Extraction and Selection for Texture Segmentation, *IEEE Trans. Pattern Analysis and Machine Intelligence* PAMI-11:7 (1989) 717-728.

- [Unser & Eden 90] M. Unser and M. Eden, Nonlinear Operators for Improving Texture Segmentation Based on Features Extracted by Spatial Filtering, *IEEE Trans. Systems, Man, and Cybernetics SMC-20:4* (1990) 804-815.
- [Veatch & Davis 90] P.A. Veatch and L.S. Davis, Efficient Algorithms for Obstacle Detection Using Range Data, *Computer Vision, Graphics, and Image Processing* **50** (1990) 50-74.
- [Venkateswarlu & Raju 92] N.B. Venkateswarlu and P.S.V.S.K. Raju, Fast Isodata Clustering Algorithms, *Pattern Recognition* **25:3** (1992) 335-342.
- [Vilnrotter *et al.* 86] F.M. Vilnrotter, R. Nevatia, K.E. Price, Structural Analysis of Natural Textures, *IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-8:1* (1986) 76-89.
- [Vistnes 89] R. Vistnes, Texture Models and Image Measures for Texture Discrimination, *International Journal of Computer Vision* **3** (1989) 313-336.
- [Voorhees & Poggio 87] H. Voorhees and T. Poggio, Detecting Textons and Texture Boundaries in Natural Images, in *Proc. Int. Conf. on Computer Vision, ICCV'87*, London, UK, June 8-11, 1987, pp. 250-258.
- [Waibel *et al.* 89] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, K. Lang, Phoneme Recognition Using Time-Delay Neural Networks, *IEEE Trans. Acoustics, Speech, and Signal Processing ASSP-37* (1989) 328-339.
- [Wallace *et al.* 86] R. Wallace, K. Matsuzaki, Y. Goto, J.D. Crisman, J.A. Webb, T. Kanade, Progress in Robot Road-Following, in *Proc. IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, April 7-10, 1986, pp. 1615-1621.
- [Waxman *et al.* 85] A.M. Waxman, J. LeMoigne, B. Srinivasan, Visual Navigation of Roadways, in *Proc. IEEE Int. Conf. on Robotics and Automation*, St. Louis, Missouri, March 25-28, 1985, pp. 862-867.
- [Wechsler 80] H. Wechsler, Texture Analysis - A Survey, *Signal Processing* **2** (1980) 271-282.
- [Wen & Fryer 91] W. Wen and R.J. Fryer, Multiscale Texture Element Detection, *Pattern Recognition Letters* **12** (1991) 793-804.
- [Weszka *et al.* 76] J.S. Weszka, C.R. Dyer, A. Rosenfeld, A Comparative Study of Texture Measures for Terrain Classification, *IEEE Trans. Systems, Man, and Cybernetics SMC-6* (1976) 269-285.
- [Weszka & Rosenfeld 79] J.S. Weszka and A. Rosenfeld, Histogram Modification for Threshold Selection, *IEEE Trans. Systems, Man, and Cybernetics SMC-9:1* (1979) 38-52.
- [Williams & Wechsler 92] D. Williams and H. Wechsler, Filters Versus Textons in Human and Machine Texture Discrimination, in H. Wechsler (ed.), *Neural Networks for Perception: Volume 1, Human and Machine Perception*, ACADEMIC PRESS, INC. Harcourt Brace Jovanovich, Publishers, Boston, San Diego, New York, London, Sydney, Tokyo, Toronto, 1992, pp. 145-175.

- [Wilm 90] T. Wilm, Bildverarbeitung im Kraftfahrzeug - nur ein Forschungsansatz? Automotive image processing - only an initial stage of research?, in *Elektronik im Kraftfahrzeug*, Baden-Baden, 5.-6. September 1990, VDI Berichte 819, VDI Verlag Düsseldorf, 1990, pp. 613-629.
- [Witkin 81] A.P. Witkin, Recovering Surface Shape and Orientation from Texture, *Artificial Intelligence* 17 (1981) 17-46.
- [Witkin 83] A.P. Witkin, Scale-Space Filtering, in *Proc. Int. Joint. Conf. on Artif. Intel.*, Karlsruhe, Germany, Aug. 8-12, 1983, pp. 1019-1022.
- [Xie et al. 93] M. Xie, L. Trassoudaine, J. Alizon, M. Thonnat, J. Gallice, Active and Intelligent Sensing of Road Obstacles: Application to The European Eureka-PROMETHEUS Project, in *Proc. Int. Conf. on Computer Vision, ICCV'93*, Berlin, Germany, May 11-14, 1993, pp. 616-623.
- [Yu et al. 91] X. Yu, Y.-J. Juha, B. Yuan, A New Algorithm for Texture Segmentation Based on Edge Detection, *Pattern Recognition* 24:11 (1991) 1105-1112.
- [Yu et al. 92] X. Yu, S. Beucher, M. Bilodeau, Road Tracking, Lane Segmentation and Obstacle Recognition by Mathematical Morphology, in *Proc. Symposium on Intelligent Vehicles '92*, Detroit, USA, June 29 - July 1, 1992, pp. 166-172.
- [Zahlten et al. 93] C. Zahlten, C.J.G. Evertsz, H.-O. Peitgen, I. Zuna, S. Delorme, G. van Kaick, Fraktale in der Analyse von Ultraschallbildern der Leber, in S.J. Pöppel and H. Handels (Hrsg.), *DAGM-Symposium Mustererkennung 1993*, Lübeck, Informatik Aktuell, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1993, pp. 150-157.
- [Zapp 88] A. Zapp, *Automatische Straßenfahrzeugführung durch Rechnersehen*, Dissertation, Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München, Neubiberg, September 1988.
- [Zhang & Wu 87] J. Zhang and J. Wu, A Fuzzy Clustering Algorithm for Image Noises Smoothing, *ACTA Automation Sinica* 13:5 (1987) 352-359.
- [Zhang 91] J. Zhang, *Erkennen von Straßenbegrenzungen durch Auswertung der Richtungen des örtlichen Grauwertgradienten im Videobild*, Studienarbeit (Aufbaustudium), Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe, durchgeführt am Fraunhofer-Institut für Informations- und Datenverarbeitung (ITB), Februar 1991.
- [Zimdahl et al. 86] W. Zimdahl, I. Rackow, T. Wilm, OPTOPILOT - Ein Forschungsansatz zur Spurerkennung und Spurführung bei Straßenfahrzeugen, in *Fahrzeugtechnik*, Baden-Baden, 9.-10. Oktober 1986, VDI Berichte 612, VDI Verlag Düsseldorf, 1986, pp. 49-60.
- [Zucker 76a] S. Zucker, Toward a Model of Textur, *Computer Graphics and Image Processing* 5 (1976) 190-202.
- [Zucker 76b] S. Zucker, Region Growing: Childhood and Adolescence, *Computer Graphics and Image Processing* 5 (1976) 382-399.

Autorenverzeichnis

- Abdelmalik, N., 91
Aggarwal, J.K., 61, 65, 69
Ahuja, N., 68, 70
Alizon, J., 45
Aloimonos, J., 54
Amadasun, M., 64
Andrew, H.C., 92, 95
Aono, S., 33
Arkin, R.C., 43, 46
Austermeier, H., 48
- Ball, G.H., 92
Bares, J., 1
Bataouche, S., 92
Beck, J., 56
Behringer, R., 32
Bekkhoucha, A., 41
Bergen, J.R., 53, 56, 69
Bergman, A., 38
Beucher, S., 46
Bhanu, B., 47
Bilodeau, M., 46
Binford, T.O., 49
Blake, A., 54
Blidberg, D.R., 1
Blostein, D., 70
Boldt, M., 150
Bovik, A.C., 72
Boyce, J.F., 61
Brauckmann, M., 45
Brodatz, P., 54, 64, 84
Bruyelle, J.-L., 43
Brüdigam, C., 32
Burt, P.J., 66, 118
Büker, U., 48
- Campbell, N.W., 35
Cannon, R.L., 92
Canny, J., 78, 149
- Català, A., 41
Chan, K.P., 92
Chen, J.-Y., 58
Chen, Y.C., 69, 91
Chetverikov, D., 61
Cheung, Y.S., 92
Christians, T., 45
Chu, Y.-L., 61
Clark, M., 72
Coggins, J.M., 71
Coleman, G.B., 92, 95
Connors, R.W., 62, 63, 71
Cowan, C.K., 38
Cox, I.J., 3
Crisman, J.D., 31, 39, 41-43
Cross, G.R., 58
- Dagless, E.L., 18, 35
Daily, M., 44
Davis, L.S., 19, 36, 37, 44, 46, 61, 65, 69
Delorme, S., 59
DeMenthon, D., 18, 19, 46
Demling, L., 59
Dewable, P., 54, 57
Dickinson, S.J., 46
Dickmanns, D., 32
Dickmanns, E.D., 32, 35, 45
Dinstein, I., 60, 61
Dougherty, E.R., 129
Dreyer, W., 32
Du Buf, J.M.H., 54
Duda, R.O., 118, 144
Dunlay, R.T., 44
Durbin, R., 45
Dyer, C.R., 60, 62
- Eden, M., 67
Efenberger, W., 45
Ell, C., 59

- Endo, H., 33
 Enkelmann, W., 43
 Evertsz, C.J.G., 59

 Farrokhnia, F., 72
 Faugeras, O.D., 71
 Fernandez-Maloigne, C., 41
 Fisher, J.J., 1
 Franke, U., 31, 32, 35
 Fritz, H., 32
 Fryer, R.J., 70
 Fu, K.S., 71, 89
 Fuertes, J.M., 41
 Fujita, K., 33

 Gallice, J., 45
 Galloway, M.M., 41, 62
 Geisler, J., 33, 48
 Geisler, V.S., 72
 Gilliard, B., 47
 Gilmore, J.F., 1
 Giordina, C.R., 129
 Glicksman, J., 46
 Gotlieb, C.C., 61
 Goto, Y., 31
 Graefe, V., 32, 35, 45
 Grau, A., 41
 Gremban, K.D., 17, 19, 30, 39, 41
 Grimson, W.E.L., 11, 149
 Groen, F.C.A., 3

 Haddon, J.F., 61
 Hall, D.J., 92
 Hallset, J.O., 1
 Hanazawa, T., 40
 Hanson, A.R., 46
 Haralick, R.M., 53, 54, 57, 60, 61, 63, 89,
 129, 149
 Harlow, C.A., 62, 63, 71
 Harris, J., 44
 Hart, P.E., 118, 144
 Hartmann, G., 48
 Hashimoto, K., 33
 Hashimoto, Y., 32
 Hattori, A., 33
 He, D.-C., 63
 Hebert, M., 1, 19, 31, 36, 39, 44

 Hennessy, S.J., 19
 Hertz, J., 40
 Hertzberger, L.O., 3
 Hesketh, S., 61
 Hildreth, E., 149
 Hinton, G., 40
 Hirose, S., 3
 Holt, V.v., 32
 Hoppe, P., 32
 Horn, B.K.P., 5
 Hosaka, A., 32, 33
 Hsiao, J.Y., 65
 Hsueh, A.C., 31, 39, 41
 Hu, M.-K., 93
 Huntsberger, T.L., 92
 Hättich, W., 33

 Ishida, S., 33
 Ishii, J., 33
 Ito, T., 33

 Jacob, U., 32
 Jacobs, C.L., 92
 Jain, A.K., 49, 58, 71, 72
 Jain, R.C., 17, 37, 49
 Jernigan, M.E., 63
 Jochem, T.M., 41
 Johns, S.A., 61, 65, 69
 Jolion, J.-M., 92
 Juha, Y.-J., 65
 Julesz, B., 53, 55-57, 69

 Kanade, T., 1, 3, 19, 31, 36, 39, 44, 123
 Kass, M., 75, 76
 Kawakatsu, S., 33
 Keirse, D., 44
 Keller, P., 44
 Kenue, S.K., 37, 46
 Khotanzad, A., 58
 King, R., 64
 Kluge, K., 36, 48
 Korn, A., 78, 149, 150, 152, 153, 156
 Kreyszig, H.E., 61
 Kristen, H., 47
 Krogh, A., 40
 Krotkov, E., 1
 Kuan, D., 31, 39, 41

- Kuhnert, K.-D., 32, 35
Kurahashi, K., 33, 43
Kurita, T., 91
Kushner, T.R., 36
Kutami, A., 33
- Lang, G.K., 47
Lang, K., 40
Larkin, L.I., 66
Laubenstein, F., 33, 48
Laugier, D., 41
Laws, K., 65, 66, 72
Lawton, D.T., 46
LeMoigne, J., 36
Lenz, C., 59
Leu, J.-G., 69
Levitt, T.S., 46
Li, Y., 93
Lim, J.S., 63
Liou, S.P., 17, 37
Liow, Y.-T., 90
Liu, S.-S., 63
Lohmann, G., 62
Lowe, D.G., 72, 150
Lu, S.Y., 71
- Madarasmi, S., 61
Malik, N.A., 63
Maretzke, J., 32
Marinos, C., 54
Marr, D., 56, 149
Marra, M., 17, 19, 30, 39, 41, 44
Masaki, I., 3
Mathies, L., 43
Matsuzaki, K., 31
McConnell, C., 46
Meer, P., 92
Mehring, S., 32
Mertsching, B., 48
Milford, D.J., 18, 35
Mitchell, T., 1
Monga, O., 90
Moravec, H.P., 42, 43
Morcego, B., 41
Morgan, A.D., 18, 35
Morgenthaler, D.G., 17, 19, 30, 39, 41, 44
Mori, H., 47
- Morita, I., 33
Mui, J.K., 89
Munkelt, O., 47
Mysliwetz, B., 45
Müller, C., 150
Münkel, H., 36
- Nagel, H.-H., 1, 5, 31, 33, 48, 75, 118, 123
Nakano, E., 33
Nakayama, S., 33
Nasr, H., 47
Negahdaripour, S., 49
Nevatia, R., 69
Niemann, H., 59
Nill, W., 33
Noll, D., 45
- Ohta, H., 12
Ohzora, M., 33, 43
Okada, Y., 33
Okuno, A., 33
Olin, K., 44
Oono, N., 33
Oosterlinck, A., 54, 57
Orban, G.A., 5
Otsu, N., 90, 91
Ozaki, T., 33, 43
Ozawa, S., 12, 18
- Palmer, R.G., 40
Pandazis, J.C., 47
Pappas, T.N., 92
Paul, D., 33
Paulus, D.W.R., 59
Pavlidis, T., 49, 89, 90
Payton, D., 44
Peckinpaugh, S.H., 62
Peitgen, H.-O., 59
Pentland, A., 59
Perona, P., 149
Perry, A., 72
Petkovic, D., 49
Phillips, T.-Y., 92
Phipps, G., 31, 39, 41
Pietikäinen, M., 64, 65
Poggio, T., 69, 70
Polk, A., 37

- Pomerleau, D.A., 40, 41
Pong, T.-C., 61
Porat, M., 72
Postaire, J.-G., 43
Price, K.E., 69
Protheroe, S., 61
- Rackow, I., 36
Raju, P.S.V.S.K., 148
Rao, A.R., 54, 76
Reed, T.R., 54
Regensburger, U., 45
Reiser, K., 44
Rembold, U., 1
Riseman, E.M., 46, 150
Ritter, W., 48
Rosenblatt, J., 44
Rosenfeld, A., 12, 18, 60, 62, 64-66, 68, 90, 92
Rueff, M., 59
- Sahoo, P.K., 69, 91
Saito, T., 32, 33
Sakagami, T., 33
Sakurai, K., 12
Sawchuk, A.A., 65
Schaaser, L.T., 35
Schalkoff, R.J., 118
Schunck, B.G., 76
Schwarzinger, M., 45
Seitz, P., 47
Sekine, M., 33
Semeco, A.C., 1
Senoo, T., 33
Shafer, S.A., 19, 31, 36, 39
Shanmugam, K., 60, 61
Shapiro, L.G., 53, 89, 129
Sharma, U.K., 37, 39
Sher, C.A., 66
Shigematu, T., 32
Shikano, K., 40
Shirai, Y., 69
Siegler, G., 33, 48
Simmons, R., 1
Singh, S., 44
Solder, U., 45
Soltani, S., 69, 91
- Srinivasan, B., 36
Stentz, A., 31
Struck, G., 33, 48
Swain, M.J., 54
- Ta, Q.-H., 45
Taniguchi, M., 33
Tatari, S., 33
Thomanek, F., 32
Thomas, B.T., 18, 35
Thonnat, M., 45
Thorpe, C.E., 3, 19, 31, 36, 39, 41-43, 48
Tomita, F., 53, 54, 69
Touretzky, D.S., 40
Trassoudaine, L., 45
Tseng, D., 44
Tsinas, L., 45
Tsugawa, S., 32
Tsuji, S., 53, 54, 69
Tuceryan, M., 65
Turk, M.A., 17, 19, 30, 39, 41, 44
- Ulmer, B., 32
Unoura, K., 33
Unser, M., 62, 67
Ushioda, Y., 12
- Van Gool, L., 54, 57
van Kaick, G., 59
Veatch, P.A., 44, 46
Venkateswarlu, N.B., 148
Vilrotter, F.M., 69
Vistnes, R., 65
von Seelen, W., 45
Voorhees, H., 69, 70
- Waibel, A., 40
Wallace, R., 31
Wang, L., 63
Watanabe, Y., 32
Waxman, A.M., 36
Webb, J.A., 31, 43
Wechsler, H., 54, 57
Wee, W.G., 69
Weiss, R., 150
Welz, K., 36
Wen, W., 70

- Weszka, J.S., 60, 62, 90
Whittaker, M., 1
Wilder, J., 49
Wilfong, G.T., 3
Williams, D., 57
Willshaw, D., 45
Wilm, T., 32, 36, 40
Winkler, G., 33
Witkin, A.P., 54, 75, 76, 149
Wong, A.K., 69, 91
Wong, V., 44
Wu, J., 92
- Xie, M., 45
- Yu, X., 46, 65
Yuan, B., 65
- Zahlten, C., 59
Zapp, A., 18, 35
Zeevi, Y.Y., 72
Zen, H., 12
Zhang, J., 90, 92
Zielke, T., 45
Zimdahl, W., 36
Zucker, S., 68, 71, 90
Zuna, I., 59