

Ermittlung von Verkehrsgeschehen durch Bildfolgenauswertung

Henner Kollnig



Henner Kollnig
Universität Karlsruhe
Institut für Algorithmen und Kognitive Systeme
Am Fasanengarten 5
D-76128 Karlsruhe

Zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
der Fakultät für Informatik
der Universität Karlsruhe (Technische Hochschule)
genehmigte Dissertation

Erster Gutachter: Prof. Dr. H.-H. Nagel
Zweiter Gutachter: Prof. Dr. A. Waibel

Tag der mündlichen Prüfung: 15. Februar 1995

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Kollnig, Henner:

Ermittlung von Verkehrsgeschehen durch Bildfolgenauswertung

/ Henner Kollnig. – Sankt Augustin : Infix, 1995

(Dissertationen zur künstlichen Intelligenz ; Bd. 88)

Zugl.: Karlsruhe, Univ., Diss., 1995

ISBN 3-929037-88-2

NE: GT

© 1995 Dr. Ekkehard Hundt, „infix“, Ringstr. 32, 53757 Sankt Augustin

Das Werk ist in allen seinen Teilen urheberrechtlich geschützt. Jede Verwertung ohne ausdrückliche Zustimmung des Verlages ist unzulässig. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung in und Verarbeitung durch elektronische Systeme.

Reproduziert von einer Druckvorlage des Autors
Umschlaggestaltung: Art und Media, Bonn
Druck und Verarbeitung: Hundt Druck GmbH, Köln
Printed in Germany

ISSN 0941-5769

ISBN 3-929037-88-2

Zum Geleit

Diese von H. Kollnig vorgelegte Dissertation setzt eine Serie von Untersuchungen an monokularen Bildfolgen von Verkehrsszenen fort, die sich systematisch mit der Detektion von bewegten Objekten, deren modellgestützter Initialisierung und Verfolgung sowie mit einer nicht nur geometrischen, sondern auch begrifflichen Beschreibung geschätzter Bewegungen befaßt, die sich bis zur Abstraktionsebene natürlichsprachlicher Ausdrücke erstreckt.

Charakteristisch für die inzwischen erreichte Stoßrichtung dieser Untersuchungen sind vier Eigenschaften:

- Es werden nicht mehr isolierte Einzelfragestellungen untersucht, sondern ein durchgängiger Systemansatz.
- Die Konzeption und Realisierung eines Gesamtsystems erfordert die Berücksichtigung zahlreicher Einzelaspekte, ohne die das System entweder überhaupt nicht oder nur bedingt einsatzfähig ist.
- Die Leistungsfähigkeit einzelner Komponenten wird nicht isoliert untersucht, sondern durch Beurteilung ihres Beitrages zur Leistungsfähigkeit des Gesamtsystems. Dies führt auf der einen Seite zu einer realistischen Einschätzung der wahren Bedeutung erwogener 'Verbesserungen', auf der anderen Seite werden Grundlagen für eine *systematische* – im Gegensatz zu einer heuristisch gesteuerten – Optimierung sichtbar.
- Die ständige Verbesserung des Preis-/Leistungsverhältnisses von Arbeitsplatzstationen ermöglicht es, als Schwachstellen erkannte Komponenten durch leistungsfähigere Ansätze auch dann zu ersetzen, wenn dies mit nicht unbeträchtlichem Mehraufwand an Rechenzeit verbunden ist.

Mit anderen Worten: die Notwendigkeit, 'wilde' Heuristiken nur aus Rechenzeitgründen einzusetzen, geht zurück. Die Einführung aufwendigerer, aber dafür oft besser begründbarer – und damit *besser analysierbarer* – Ansätze ermöglicht es, nach und nach eine zeitraubende 'blinde' Suche nach Möglichkeiten zur Steigerung der Gesamtsystemleistung durch eine zielstrebige, auf einer begründeten Auswahl unter denkbaren Verfahrensalternativen aufbauende Vorgehensweise zu ersetzen.

Somit kann man diese Arbeit auch als Illustration einer weiteren Zwischenstufe in der langwierigen Transformation von der Grundlagenforschung zur ingenieurmäßigen Konstruktion von Bildauswertungssystemen für den Einsatz in Anwendungsfällen ansehen.

Danksagung

Herrn Prof. Dr. H.-H. Nagel bin ich für seine intensive Betreuung meiner Arbeit zu großem Dank verpflichtet. Seine sehr rasche und gründliche Durchsicht meiner Entwürfe, seine zahlreichen Hinweise und Anregungen und die konstruktiven Diskussionen mit ihm – insbesondere bei der gemeinsamen Analyse der Systemausgaben in den Abendstunden – haben nicht nur mein wissenschaftliches Arbeiten stark geprägt, sondern darüber hinaus ermöglicht, die Dissertation in weniger als drei Jahren abzuschließen.

Herrn Prof. Dr. A. Waibel danke ich für die freundliche Bereitschaft zur Übernahme des Korreferats und seine Kommentare zur Arbeit.

Bedanken möchte ich mich auch bei meinen Kollegen Dres. Konstantinos Daniilidis, Dieter Koller und Michael Otte sowie Karl Schäfer, Stefan Wachter und Martin Tonko. Insbesondere Dr. Konstantinos Daniilidis hat mir bei geometrischen Fragestellungen immer wertvolle Ratschläge geben können.

Mein Dank gilt auch Annette Storkebaum, Karl Schäfer, Markus Maier, Bernd Kaufmann, Harald Damm, Gudrun Socher, Till Osswald, Valdimar Bragason, Michael Haag, Klaus Mück, Holger Leuck, Nazanin Baumgärtner, Holger Böhnke, Ralf Gerber und Thorsten Früauf, die im Rahmen von Studien- und Diplomarbeiten für zahlreiche Anregungen sorgten. Besonders ergiebig war die Zusammenarbeit mit Holger Leuck, Michael Haag, Harald Damm und Klaus Mück. Für Kommentare zur vorliegenden Arbeit bedanke ich mich bei Michael Haag, Stefan Wachter, Klaus Mück und Holger Leuck.

Dr. Wilfried Enkelmann, Dr. Wolfgang Krüger und Norbert Heinze danke ich für ihre Unterstützung bei der Klärung von Fragen zu früheren Arbeiten und bei der Bereitstellung von Daten.

Dank auch an die Saarbrückener Forschungsgruppe um Prof. Dr. W. Wahlster für die Zusammenarbeit im Rahmen des DFG-Sonderforschungsbereiches SFB 314 'Künstliche Intelligenz – Wissensbasierte Systeme'.

Markus Weber und Bernd Kaufmann gilt mein Dank für ihre Hilfe beim Kampf mit den Idiosynkrasien der Rechnerumgebung.

Dank an Herrn Dr.-Ing. Pfundt, Büro für Verkehrssicherheit, Köln, für die Bereitstellung der Videobänder von Straßenkreuzungen in Frankfurt sowie seine Bereitschaft, die Nutzung dieser Daten für Forschungszwecke zu ermöglichen.

Bei meinen Eltern bedanke ich mich für ihre Unterstützung und für die Durchsicht der Arbeit.

Schließlich gilt mein Dank meiner Frau Silke für ihre Hilfe, ihr Verständnis und ihre Geduld.

Kurzfassung

Es ist Ziel der Bildfolgenauswertung, nicht nur geometrische, sondern auch begriffliche Beschreibungen für Objekte und ihre Veränderungen in einer dreidimensionalen Szene aus Daten zu berechnen, die eine Kamera als Folge zweidimensionaler Grauwertbilder liefert. Die vorliegende Arbeit zeigt an Aufnahmen von Straßenkreuzungen, wie Verkehrsgeschehen mittels modellgestützter Bildfolgenauswertung zuverlässig ermittelt werden können. Durch Detektion und Verfolgung bewegter Fahrzeuge sowie durch die Anpassung eines Fahrbahnmodells an den statischen Hintergrund wertet das System sowohl dynamische als auch statische Aspekte der in der Bildfolge aufgezeichneten Szene aus.

Bekannte Verfahren zur Schätzung der Lage der Kamera relativ zur Fahrbahnebene benötigen Kalibrierkörper oder Katasterpläne. Das in dieser Arbeit vorgestellte Verfahren schätzt dagegen mit Hilfe allgemeinerer Annahmen über den Straßenverlauf die Lage der Kamera relativ zur Szene sowie die Einfallsrichtung des Sonnenlichtes aus dem aufgenommenen Bildmaterial.

Unter Verwendung von Wissen über die einzelnen Fahrspuren erhält man mit einer modellbasierten Segmentierung eines dichten optischen Flußfeldes Hinweise auf Abbilder bewegter Fahrzeuge. Die gezielte Ausnutzung von Modellwissen gestattet es, auf Suchverfahren zu verzichten, und liefert im Gegensatz zu rein datengetriebenen Segmentierungsverfahren deutlich bessere Startwerte für die Verfolgung von Fahrzeugen.

Ein Schätzverfahren wurde dahingehend verallgemeinert, daß einzustellende Parameter zur Schätzung der Lage von Fahrzeugen und Fahrbahn aus einem gemeinsamen Parametersatz entwickelt werden können. Damit liegt ein systematischer Weg vor, der sowohl die Zahl der einzustellenden Parameter verringert als auch bisher eingesetzte Heuristiken überflüssig macht.

Viele Verfahren zur Objekterkennung mittels Bildauswertung basieren auf einer Merkmalsextraktion. Dabei erweist sich die Zuordnung von Kantensegmenten zu Modelldaten in einem exponentiell wachsenden Suchraum als schwierig und aufwendig. In dem neuen Verfahren wird die Lage eines Fahrzeugs durch die direkte Anpassung eines Flächenmodells an die Bilddaten geschätzt. Somit kann auf eine Merkmalsextraktion verzichtet werden. Das neue Verfahren ist bei der Verfolgung von Fahrzeugen deutlich robuster als bisherige Ansätze, die die Bildinformation auf Kantensegmente beschränken.

Das vorgestellte System liefert erstmals nicht nur retrospektiv, sondern auch inkrementell im Zeitraster der Bildfolge begriffliche Beschreibungen von Verkehrsgeschehen in Form von deutschen und englischen Bewegungsverbren. Hierfür notwendige Wegsegmente werden durch eine Anpassung eines Fahrbahnmodells an Bilddaten gewonnen, das aus einer begrifflichen Spezifikation entwickelt wird. Damit ist es nun möglich, aus Bildfolgen gewonnene Auswertungsergebnisse auch Datenbanksystemen sowie sprachverarbeitenden Systemen zugänglich zu machen.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation und Aufgabenstellung	1
1.2	Algorithmische Auswertung von Bildfolgen	2
1.2.1	Initialisierungsphase	2
1.2.2	Aktualisierungsphase	3
1.3	Ziel und Beitrag dieser Arbeit	4
1.4	Vorgehensweise und Gliederung	7
2	Methoden und Ansätze in der Literatur	9
2.1	Detektion und 2D-Verfolgung von Objekten	9
2.1.1	Änderungsdetektion	9
2.1.2	Schätzung eines Hintergrundbildes	10
2.1.3	Optischer Fluß	11
2.1.4	Hindernisdetektion	14
2.1.5	Suchverfahren	15
2.2	3D-Verfolgung	15
2.3	Schätzung von Bewegungs- und Strukturparametern	17
2.4	Systeme aus dem Verkehrswesen	17
2.5	Diskussion der Literatur	18
3	Einsatz leistungsfähiger signalnaher Komponenten	23
3.1	System von [Koller 92]	23
3.2	Entwicklung leistungsfähigerer signalnaher Komponenten	23
3.3	Segmentierung eines optischen Flußfeldes	25
3.4	Datensegmentextraktion	30
3.5	Ergebnisse der Trajektorienberechnung	31

4	Kalibrierung mit Hilfe von Fluchtpunkten	35
4.1	Übergang vom Bild- in den Szenenbereich	35
4.2	Interne und externe Kameraparameter	35
4.3	Übersicht über die zu schätzenden Parameter	37
4.4	Verfahren zur Fluchtpunktberechnung	38
4.4.1	Verfahren zur Fluchtpunkt detektion	38
4.4.2	Schätzung der Koordinaten von Fluchtpunkten und Kalibrierung	41
4.4.3	Diskussion der Literatur	43
4.5	Kalibrierung am Beispiel einer Innenstadtszene	45
4.5.1	Beschreibung der Kalibrierung	45
4.5.2	Zusammenstellung der zur Kalibrierung benötigten Eingangsdaten	46
4.5.3	Bestimmung der Rotationsparameter	46
4.5.4	Verfahren zur Berechnung der Brennweite	48
4.5.5	Iterative Verbesserung und Prädiktion von Rotationsparametern und Brennweite	49
4.5.6	Translationsparameter	50
4.5.7	Berechnung der Einfallsrichtung des Lichtes	51
5	Modellbasierte Bildfolgenauswertung	55
5.1	Systematische Erweiterung des verwendeten Modellwissens	55
5.2	Entwurf und Anpassung von Modellen innerstädtischer Straßenkreuzungen an Bildfolgen	58
5.2.1	Probleme bei rein datengetriebener Segmentierung	58
5.2.2	Einfaches geometrisches Fahrbahnmodell	60
5.2.3	Erweiterung zu schätzender Lageparameter	60
5.2.4	Änderung der Abstandsfunktion zur Anpassung eines Fahrbahnmodells an Datenkantensegmente	61
5.2.5	Berücksichtigung der endlichen Breite von Fahrbahnmarkierungen	65
5.2.6	Berücksichtigung der Richtung von Datengrauwertübergängen bei der Zuordnung von Modell- zu Datenkantensegmenten	65
5.2.7	Anpassung eines Fahrbahnmodells an Bildfolgen	66
5.2.8	Weitere Modellverfeinerungen	70
5.2.9	Modellgestützte Segmentierung	71

5.3	Verfeinerung der Fahrzeugmodelle	73
5.3.1	Entwurf eines Busmodells	73
5.3.2	Modellierung der Frontleuchten eines Fahrzeugs	75
6	Direkte Anpassung von Flächenmodellen an Grauwertgradienten	79
6.1	Problemstellung	79
6.2	Untersuchung der Mahalanobisdistanz an einem Beispiel in der Ettl- linger-Tor-Szene	80
6.3	Synthetische Gradientennorm	85
6.4	MAP-Schätzung	88
6.5	Diskussion der Abstimmungsparameter des Verfahrens	89
6.5.1	Modellgradientennorm	89
6.5.2	Grauwertgradient	90
6.5.3	Kalmanfilter	90
6.6	Ergebnisse experimenteller Untersuchungen	90
6.6.1	Ettlinger Tor	90
6.6.2	Frankfurter Kreuzung	96
6.6.3	Durlacher Tor	96
6.7	Fahrbahn	101
7	Begriffliche Beschreibungen	103
8	Diskussion und Ausblick	109
8.1	Zusammenfassung und Diskussion	109
8.2	Ausblick	112
A	Ballung optischer Flußvektoren	115
B	Kalibrierung	117
B.1	Projektive Geometrie	117
B.1.1	Homogene Koordinaten von Punkten	117
B.1.2	Homogene Koordinaten von Geraden	118
B.2	Fluchtpunkte	119
B.3	Schätzung von Fluchtpunktkoordinaten	120
B.3.1	Fluchtpunktschätzung mit inhomogenen Bildkoordinaten	120
B.3.2	Fluchtpunktschätzung mit homogenen Koordinaten	121

B.3.3	Schätzung des Azimutwinkels der 3D-Orientierung einer Fluchtpunktrichtung	122
B.4	Fluchtpunkte und Kalibrierung	124
B.4.1	Fluchtpunkt und Brennweite	124
B.4.2	Horizont und Brennweite	125
B.4.3	Zwei konjugierte Fluchtpunkte	125
B.4.4	Zwei Fluchtpunkte und Brennweite bzw. Schnittwinkel	127
B.4.5	Drei Fluchtpunkte	128
B.5	Rückprojektion von Geradenorientierungen von der Bild- in eine Raumebene	133
C	Abstand zwischen Modell- und Datenkantensegmenten	135
C.1	Schätzung von σ_1^2 aus den Bilddaten	135
C.2	Unsicherheit in der Orientierung eines Kantensegments	137
C.3	Diskussion der Mahalanobisdistanz zwischen Modell- und Datensegmenten	139
C.3.1	Definition der Mahalanobisdistanz	139
C.3.2	Mahalanobisdistanz zweier Zufallsvariablen	140
C.3.3	Verallgemeinerte Mahalanobisdistanz zweier Zufallsvariablen	141
C.3.4	Transformation der verallgemeinerten Mahalanobisdistanz zweier Zufallsvariablen bei Verkettung ihrer definierenden Funktionen	142
C.3.5	Untersuchung der Mahalanobisdistanz von [Deriche & Faugeras 90] bzw. [Koller 92]	143
C.3.6	Auswirkungen der Korrekturaddition auf die Berechnung des Abstands zweier Kantensegmente	147
C.3.7	Untersuchung anderer Abstandsmaße	148
D	Berechnung von Jakobimatrizen	155
D.1	Vorbemerkungen	155
D.2	Datenassoziation im Beispiel der Fahrbahnmodell Anpassung an Datenkantensegmente	156
D.3	Anpassung eines Fahrbahnmodells an Kantensegmente	157
D.4	Jakobimatrix der Meßfunktion bei direkter Anpassung an die Grauwertgradienten	159
	Literaturverzeichnis	163
	Stichwortverzeichnis	183

Kapitel 1

Einleitung

1.1 Motivation und Aufgabenstellung

Ein menschlicher Beobachter benutzt zur Beschreibung der in Bildfolgen eingefangenen Abläufe seine natürliche Sprache. So liegt es als ein Ziel der maschinellen Bildfolgenauswertung nahe, Ergebnisse nicht nur *geometrisch* auf niedrigen, signalnahen Abstraktionsebenen, sondern auch *begrifflich* in Form von natürlicher Sprache auf höheren, signalfernen Abstraktionsebenen auszugeben. Einen Vorschlag zur Verbindung von Bildfolgenauswertung mit natürlicher Sprache machte bereits [Nagel 77]. Begriffliche Beschreibungen liefern nicht nur eine geeignete Mensch-Maschine-Schnittstelle [Stopp *et al.* 94], sondern gestatten es auch, die vom System gelieferten Auswertungsergebnisse darunterliegender Abstraktionsebenen zu kontrollieren.

Bei dem Bemühen, mit Hilfe maschineller Bildfolgenauswertung geometrische Eingangsdaten zur Berechnung von reichhaltigen begrifflichen Beschreibungen für in Bildfolgen erfaßte Abläufe bereitzustellen, stellt sich seit Jahren immer wieder heraus, daß zur Zeit der Engpaß noch in der Ermittlung der geometrischen Beschreibungen besteht. Die Bildfolgenauswertung sollte dazu so robust sein, daß jedes von der Kamera aufgenommene Objekt detektiert und solange verfolgt werden kann, bis es das Gesichtsfeld der Kamera verläßt. Sofern Parameter auf die auszuwertenden Bildfolgen abgestimmt werden müssen, sollten sie aus den Bilddaten geschätzt werden. In Ermangelung eines solchen robusten Systems greifen bisherige Ansätze zur Extraktion von begrifflichen Beschreibungen für in Bildfolgen erfaßte Abläufe meist auf synthetische Daten zurück, ohne ein einziges Bildsignal auszuwerten [Neumann & Novak 86; Mohnhaupt 90; Mohnhaupt & Neumann 90; Mohnhaupt 92; Walter 89], oder ergänzen fehlende automatische Auswertungsergebnisse interaktiv [Schirra *et al.* 87; Herzog *et al.* 89; Retz-Schmidt 91; Retz-Schmidt 92; Herzog 92; Howarth & Buxton 92; Schirra & Stopp 93; Schirra 94]. Andere Autoren beschränken sich auf die Auswertung von Einzelbildern [Birnbaum *et al.* 93; Cooper *et al.* 94] oder auf einen einfachen Wortschatz [Toal & Buxton 92; Huang *et al.* 94]. Die Aufgabe, ein komplexes System vorzustellen, in dem alle Auswertungsschritte vom Videosignal bis zur Berechnung begrifflicher Beschreibungen realisiert sind, ist bis heute ungelöst.



Bild Nr. 1

Bild Nr. 25

Bild Nr. 50

Abbildung 1.1: Drei Aufnahmen einer Bildfolge, die das Verkehrsgeschehen am Ettlinger Tor in Karlsruhe zeigen. Die Kamera stand auf einem Hochhaus der Badenwerk AG.

Als Diskursbereich für die vorliegende Untersuchung wurde der Verkehr an einer innerstädtischen Straßenkreuzung gewählt (siehe Abb. 1.1). Der Einsatz von Videokameras im Bereich der Verkehrsanalyse bietet sich hier an, um beispielsweise ohne bauliche Maßnahmen eine Verkehrszählung durchzuführen [Wegmann & Spring 93], wodurch sich Ampelanlagen je nach Verkehrsaufkommen flexibel steuern lassen [Kilger 93a]. Mit herkömmlichen Induktionsschleifen lassen sich nur *punktueller* Messungen erzielen, was für *Streckenbeobachtungen* nicht ausreicht, um z.B. Wartezeiten an Ampeln, Anfahr- und Abbiegeverhalten an Kreuzungen zu analysieren. Mit Videokameras lassen sich die hierzu notwendigen Messungen durchführen, die einer Bildauswertung zur automatischen Verkehrsanalyse zugeführt werden.

Auch im Bereich der Fahrerassistenzsysteme bietet sich der Einsatz von Videokameras an [Nagel 94a; Nagel 94b], um den Fahrer bei einem auch in Zukunft weiter steigenden Individualverkehr durch einen sichtsystemgestützten Automaten zu unterstützen.

1.2 Algorithmische Auswertung von Bildfolgen

Im Gegensatz zur *Bildverarbeitung* hat die *Bildauswertung* zum Ziel, rechnergestützt die Bedeutung des Bildinhaltes zu ermitteln. Entsprechend versucht man bei einer *Bildfolgenauswertung*, von der Kamera erfaßte Veränderungen in der Szene zu erkennen und zu interpretieren. Als Eingabedaten zur Bildauswertung dient eine Matrix diskreter Grauwerte, die durch *Digitisierung* aus einem Kamerabild gewonnen werden. Für eine modellgestützte Bildfolgenauswertung unterscheiden wir zwei Phasen: eine *Initialisierungsphase* und eine *Aktualisierungsphase*. Dabei legen wir den Interpretationszyklus zur Auswertung von Bildern nach [Kanade 78; Kanade 80] mit Modifikationen nach [Nagel 79; Nagel 89; Nagel 95b] zugrunde (vgl. Abb. 1.2).

1.2.1 Initialisierungsphase

Mit signalnahen Verfahren werden aus den Bilddaten zunächst *Bildbereichshinweise* berechnet, beispielsweise optische Flußvektoren oder Kantensegmente. Eine Be-

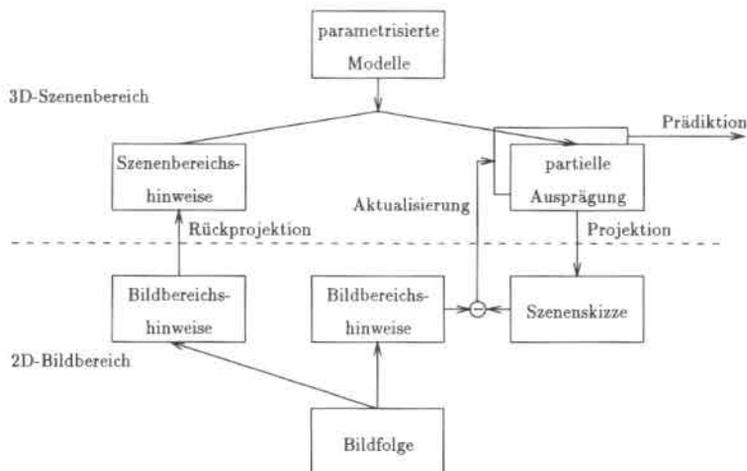


Abbildung 1.2: Interpretationszyklus für eine modellgestützte Bildfolgenauswertung nach [Kanade 78; Kanade 80] mit Modifikationen nach [Nagel 79; Nagel 95b]

schränkung der Bildauswertung auf Verfahren, die nur im 2D-Bildbereich arbeiten, reicht zur Interpretation der Bilddaten meist nicht aus. Nun ist die *perspektivische Projektion*, die man als Modell für eine Bildaufnahme verwendet, weder linear noch invertierbar. Doch kann man untersuchen, inwieweit aus 2D-Bildbereichshinweisen 3D-Informationen berechnet werden können (vgl. [Faugeras 93; Haralick & Shapiro 93]). Zur *Rückprojektion* von Bildbereichshinweisen aus dem 2D-Bildbereich in die 3D-Szene wird zum einen Wissen über den Abbildungsvorgang bei der Bildaufnahme herangezogen, insbesondere ein Kameramodell sowie externe und interne Kameraparameter. Zum anderen gehen in eine Rückprojektion noch weitere Annahmen über die Lage der Bildbereichshinweise in der 3D-Szene ein, zum Beispiel die Annahme, daß die Urbilder der Bildbereichshinweise alle in einer bekannten Ebene in der 3D-Szene liegen.

Die nach einer Rückprojektion erhaltenen *Szenenbereichshinweise* dienen zur Initialisierung eines modellgestützten Anpaßverfahrens. Mit Hilfe der Szenenbereichshinweise lassen sich *parametrisierte Modelle*, z.B. von Fahrzeugen und der Fahrbahn, ausprägen, wobei als freie Parameter in erster Linie *Orientierungs- und Positionsparameter*, d.h. *Lageparameter*, aber auch *Formparameter* in Frage kommen. Initialisierte Modellausprägungen liefern eine erste rechnerinterne Darstellung zur Beschreibung (eines Teils) der abgebildeten Szene.

1.2.2 Aktualisierungsphase

Zur Überprüfung, später auch zur iterativen Verbesserung von Hypothesen über eine Modellausprägung, die – wie gerade beschrieben – datengetrieben initialisiert

wurde, wird in einem weiteren Schritt durch *Projektion* der Modellausprägung in den 2D-Bildbereich eine *Szenenskizze* erzeugt. Ein Vergleich der Szenenskizze mit Bildbereichshinweisen läßt sich zur *Aktualisierung* der Modellausprägung ausnutzen, wobei bei guter Initialisierung eine Linearisierung der *Projektions-* oder auch *Meßfunktion* zur Aktualisierung der zu schätzenden Parameter zulässig ist [Koller 92; Freeman 93; Lowe 92b; Schick & Dickmanns 91].

Bei der Auswertung einer Bildfolge kann man die aktualisierte Modellausprägung nutzen, um – unter Berücksichtigung eines Bewegungsmodells – sofort eine Modellausprägung für das nächste Bild der Folge zu erzeugen. Diese *Prädiktion* der Modellausprägung ersetzt den nur am Anfang der Bildfolge notwendigen datengetriebenen Initialisierungsschritt. Die prädizierte Modellausprägung wird wiederum aktualisiert, dann für die darauf folgende Aufnahme prädiziert.

1.3 Ziel und Beitrag dieser Arbeit

Aufgabe der vorliegenden Arbeit ist es, ein Verfahren vorzustellen, das zuverlässig und robust Trajektorien liefert, ohne daß viele Parameter interaktiv eingestellt werden müssen. Bisherige Ansätze versagen an dieser Stelle. Die vorliegende Untersuchung zeigt – aufbauend auf den guten Ergebnissen von [Koller 92; Koller *et al.* 93a] –, daß

- eine systematische Verbesserung von Auswertungskomponenten,
- ein konsequenter Einsatz von Modellwissen und
- eine Ausnutzung von mehr Bildinformationen

es gestatten, auf eine Vielzahl von bisher eingesetzten Heuristiken zu verzichten und die Zahl interaktiv einzustellender Parameter zu verringern. Dadurch ist es nun möglich, Bildfolgen in einem weit größeren Maße als bisher automatisch auszuwerten.

Das Gesamtsystem zeichnet sich durch die folgenden Punkte aus:

- **Bildfolgenauswertung vom Signal zum Begriff:**
Im vorgestellten System ist am Beispiel von Bildfolgen des Straßenverkehrs die ganze Kette der Auswertungsstufen von *sensorisch erfaßten* Daten über *geometrische* bis hin zu *begrifflichen* Beschreibungen realisiert. Im Gegensatz hierzu beschränken sich andere Systeme meist auf die Berechnung geometrischer Beschreibungen. Ein Gesamtsystem bietet zudem die Möglichkeit, Schwachstellen in einzelnen Bausteinen zu erkennen und zu analysieren, um Engpässe im gesamten Auswertungsprozeß systematisch auszuräumen.
- **Fahrbahn und Fahrzeuge:**
Durch Detektion und Verfolgung bewegter Fahrzeuge sowie der Anpassung eines Fahrbahnmodells an den statischen Hintergrund wertet das System sowohl

dynamische als auch statische Aspekte der in der Bildfolge aufgezeichneten Szene aus. Andere Systeme konzentrieren sich auf die Detektion und Verfolgung von Fahrzeugen und berücksichtigen während ihrer Bildauswertung nicht einmal Teilaspekte des Hintergrundes.

Ein Schwerpunkt des hier vorgestellten Systems bildet dabei eine *automatische Initialisierung*: Die Initialisierungen für die Anpassung eines Fahrbahnmodells werden durch im Bild detektierte Fluchtpunkte und für die Anpassung von Fahrzeugmodellen durch Segmentierung eines optischen Flußfeldes bereitgestellt. Sowohl für Fahrzeuge als auch für die Fahrbahn wird eine Modellanpassung zur Schätzung freier Lageparameter durchgeführt (vgl. Abb. 1.3). Die vorliegende Arbeit zeigt, daß gute Initialisierungen eine notwendige Voraussetzung für ein robustes modellgestütztes Anpaß- und Verfolgungsverfahren sind. Die beiden Teilsysteme zur Auswertung statischer bzw. zur Auswertung dynamischer Aspekte der in der Bildfolge aufgezeichneten Szene sind nicht isoliert zu sehen: ein signifikanter Fortschritt bei der automatischen Bildfolgenauswertung wurde gerade durch eine Interaktion beider Programmteile auf den verschiedenen Abstraktionsstufen erreicht. Zudem werden sowohl für die Fahrbahn als auch für die Bewegung der Fahrzeuge begriffliche Beschreibungen ermittelt.

- **Einsatz leistungsfähiger signalnaher Komponenten:**

Die heute zur Verfügung stehende Rechnergeneration erlaubt es, gerade auf signalnaher Ebene aufwendigere, dafür aber leistungsfähigere Verfahren einzusetzen, die mehr Bildinformation ausnutzen. Dadurch können bisher eingesetzte Heuristiken vermieden werden, zu denen man gezwungen war, um Unzulänglichkeiten der signalnahen Komponenten zu überbrücken.

- **Automatische Kalibrierung :**

Bekanntes Verfahren zur Schätzung der Lage der Kamera relativ zur Fahrbahnebene benötigten Kalibrierkörper oder Katasterpläne. Das in dieser Arbeit vorgestellte Verfahren schätzt dagegen mit Hilfe von allgemeineren Annahmen über den Straßenverlauf die Kameralage sowie die Einfallsrichtung des Sonnenlichtes aus dem aufgenommenen Bildmaterial. Hierzu werden zunächst Fluchtpunkte im Bildmaterial detektiert. Damit können die Orientierung und die Brennweite der Kamera berechnet werden. Die Schätzung des Skalierungsparameters erfolgt mit Hilfe von Annahmen über die Breite von Fahrbahnmarkierungen.

- **Einsatz von Modellwissen anstelle von Suchverfahren:**

In diesem System wird schon auf unteren Ebenen Modellwissen eingesetzt. Beispielsweise wird ein optisches Flußfeld modellbasiert segmentiert, was im Vergleich mit rein datengetriebenen Verfahren zu deutlich besseren Initialisierungen führte. Dieses gezielte Einsetzen von Modellwissen erlaubt es, auf Suchverfahren zu verzichten.

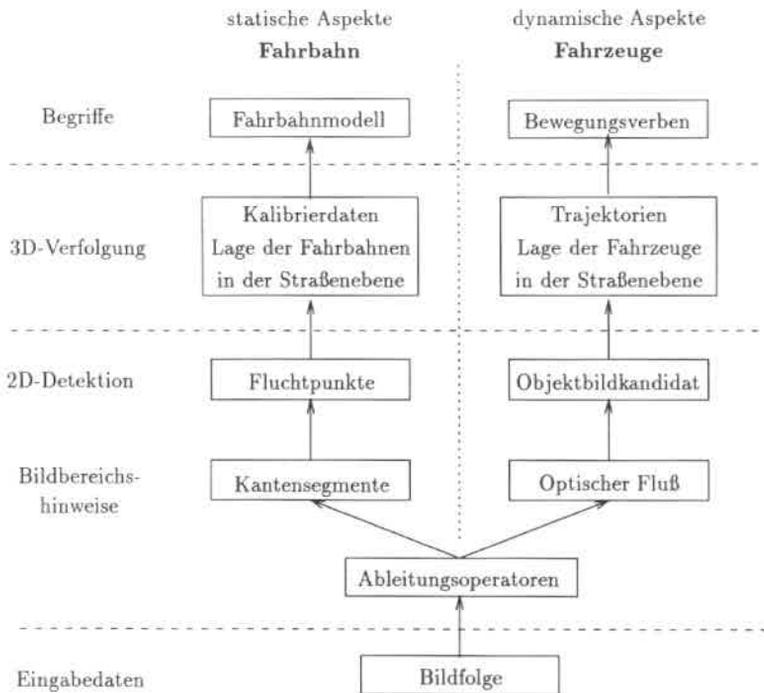


Abbildung 1.3: Das System wertet sowohl statische als auch dynamische Aspekte der in der Bildfolge aufgezeichneten Szene aus. Dabei werden beide Programmkomponenten automatisch initialisiert, und die Auswertung wird bis zu einer begrifflichen Ebene durchgeführt.

- **Flächenanpassung:**

Viele Verfahren zur Objekterkennung mittels Bildauswertung basieren auf einer Merkmalsextraktion, z.B. von Kantensegmenten. Dabei erweist sich die Zuordnung von Kantensegmenten zu Modelldaten in einem exponentiell wachsenden Suchraum als schwierig und aufwendig. In einem im Rahmen dieser Arbeit neu entwickelten Verfahren wird die Lage eines Fahrzeugs durch die direkte Anpassung eines Flächenmodells an die Bilddaten geschätzt. Somit kann auf eine Merkmalsextraktion und auf die damit verbundenen Heuristiken verzichtet werden, wie z.B. Kantenelemente zu Kantensegmenten aggregiert werden sollen [Henricsson & Heitger 94]. Das neue Verfahren ist bei der Verfolgung von Fahrzeugen sogar deutlich robuster als bisherige Ansätze, die die aus dem Bild extrahierte Information auf Kantensegmente beschränken. In einem wei-

teren Schritt gelang es, auch die Anpassung eines Fahrbahnmodells direkt an die Bilddaten durchzuführen und auch hier auf die Extraktion von Kantensegmenten zu verzichten. Damit kann auf alle zur Kantensegmentberechnung notwendigen, interaktiv ermittelten Parameter verzichtet werden. Somit liegt ein systematischer Weg vor, der es gestattet, die Lage von Fahrzeugen und Fahrbahn relativ zur Kamera zu schätzen.

1.4 Vorgehensweise und Gliederung

Kapitel 2 enthält eine Literaturübersicht mit Publikationen, in denen Fragestellungen aus dem Umfeld der vorliegenden Arbeit untersucht werden. Anschließend zeigen wir in Kapitel 3, welche Fortschritte bei der Detektion und Verfolgung von Fahrzeugen durch Verwendung leistungsfähiger signalnaher Komponenten zu erzielen sind. Ein auf Fluchtpunkten aufbauendes Kalibrierverfahren in Kapitel 4 liefert den Übergang vom 2D-Bild- in den 3D-Szenenbereich. Kapitel 5 stellt die Fortschritte bei der Bildfolgenauswertung dar, die durch systematisches, schrittweise vergrößertes 3D-Modellwissen erreicht werden. In Kapitel 6 wird das im Rahmen dieser Arbeit neu entwickelte Verfahren zur Lageschätzung von Fahrzeugen und Fahrbahnen durch eine direkte Anpassung von Flächenmodellen an Grauwertgradienten dargestellt. Kapitel 7 stellt eine Berechnung von begrifflichen Beschreibungen für in Bildfolgen erfaßte Abläufe vor.

Kapitel 2

Methoden und Ansätze in der Literatur

Wir stellen zunächst einige Methoden und Ansätze aus der Literatur vor, in denen Fragestellungen aus dem Umfeld unserer Arbeit untersucht werden und bei denen Aufnahmen von Straßenverkehrsszenen ausgewertet werden. Wir beschränken uns dabei auf Publikationen, in denen nicht nur Einzelbilder, sondern Bildfolgen analysiert werden. Ein Literaturüberblick zu früheren Arbeiten findet sich bei [Nagel 81; Nagel 85; Koller 92; Cédras & Shah 94], weshalb wir uns vor allem auf neuere Arbeiten konzentrieren können.

2.1 Detektion und 2D-Verfolgung von Objekten

Veränderungen der Grauwertstruktur von einer Aufnahme zur nächsten können Hinweise auf eine relative Bewegung zwischen Kamera und Objekten in der Szene geben [Hsu *et al.* 84]. Daher ist es naheliegend, zur Detektion bewegter Objekte in Bildfolgen Differenzbilder zu nutzen. Dies führt auf die klassische *Änderungsdetektion*.

2.1.1 Änderungsdetektion

[Dubuisson & Jain 93] detektieren Abbilder bewegter Fahrzeuge — wir nennen sie *Objektbildkandidaten* — in Farbbildern. Dabei werden bewegte Bereiche im Bild auf der Basis von Differenzbildern erkannt. Weiterhin segmentieren sie das Bild in Bereiche gleicher Farbwerte mit einer Teile- und Verschmelze-Methode. Die Zusammenfassung der Ergebnisse ihrer Bewegungsdetektion und Segmentierung liefert Konturen, die die Abbilder von Fahrzeugen umschliessen. Desweiteren stellen sie einen einfachen Ansatz zur Klassifikation von Fahrzeugen vor. In ihren Aufnahmen wurde meist nur ein Fahrzeug aufgenommen, das große Teile des Bildes umfaßt. Ihr Verfahren wurde mit 73 Aufnahmen getestet. [Dubuisson & Jain 94] verbessern ihr Segmentierungsverfahren durch Verwendung eines generischen 2D-Fahrzeugmodells, um bei in enger Nachbarschaft fahrenden Fahrzeugen zumindest die Kontur des vorderen

Fahrzeuges zu detektieren. [Dubuisson & Jain 94] setzen allerdings voraus, daß die Fahrzeuge von der Seite aufgenommen wurden. Detektierte Fahrzeugabbilder werden gespeichert, um sie in Aufnahmen, die eine in einiger Entfernung aufgestellte Kamera liefert, wieder dem Abbild desselben Fahrzeugs zuzuordnen. Hierbei verwenden sie – wie auch [Huttenlocher *et al.* 93] – eine Hausdorff-Distanz.

[Li-Qun *et al.* 92] benutzen Differenzbilder und einen Sobel-Operator zur Detektion von Abbildern bewegter Fahrzeuge in Bildfolgen. Ein Objektbildkandidat wird durch ein Rechteck repräsentiert, dessen Seiten parallel zum Bildkoordinatensystem verlaufen. Die Rechtecke werden im Bildbereich verfolgt. Daraus ermittelte Trajektorien liefern Hinweise auf Fahrspuren.

2.1.2 Schätzung eines Hintergrundbildes

Einige Ansätze für den Fall einer festen Kamera bauen eine rechnerinterne Darstellung für den statischen Hintergrund auf. In den Arbeiten von [Karmann & von Brandt 90; Kilger 93b; Koller *et al.* 94; Baumberg & Hogg 94] wird ein Grauwertbild für den statischen Hintergrund ermittelt, das mit jeder Aufnahme aktualisiert wird. Der Betrag der Differenz der Grauwerte eines neuen Bildes der Folge und des geschätzten Hintergrundbildes liefert Hinweise auf Abbilder bewegter Objekte. Ein Schwellwertoperator binarisiert das absolute Differenzbild.

[Kilger 93b] stellt ein System zur Auswertung von Straßenverkehrs-Bildfolgen vor. Die Breite des detektierten Abbildes eines Fahrzeugs wird genutzt, um detektierte Fahrzeuge einer der drei Klassen – LKW oder Bus, PKW sowie Motorräder oder Fahrräder – zuzuteilen, wobei in [Kilger 93c] für diese Breite der Schattenwurf des Fahrzeuges auf die Fahrbahn berücksichtigt wird. Hierzu müssen Tageszeit und Datum eingegeben werden. Das Zuordnungsmerkmal für eine auf einem Kalmanfilter basierende Verfolgung ist die Mitte der unteren Kante eines Rechtecks, das das Abbild eines Fahrzeugs umschreibt. Dieses System hat als zwingende Voraussetzung, daß die Fahrzeuge sich in Richtung der optischen Achse der Kamera bewegen. Schwellwerte werden adaptiv an verschiedene Bilddaten angepaßt, um unterschiedlichen Beleuchtungsverhältnissen Rechnung zu tragen. Das System wurde auch mit Bilddaten getestet, die bei Nebel aufgenommen wurden [Kilger 93a].

In dem von [Koller *et al.* 94] dargestellten Segmentieralgorithmus geht nicht nur das Differenzbild zwischen einem neuen Bild der Folge und dem Hintergrundbild ein, sondern auch Grauwertgradienten. In [Koller *et al.* 93b] nehmen die Autoren dazu allerdings an, die Abbilder bewegter Fahrzeuge bestehen aus Kanten, die im Bild horizontal oder vertikal verlaufen. Nach der Detektion eines bewegten Bereiches wird die umschließende Kontur erst polygonal, dann mit stückweise kubischen Kurven approximiert. Zur rekursiven Schätzung der Form der Kontur (12 Stützpunkte) und der Bewegungsparameter werden zwei lineare Kalmanfilter eingesetzt. Ihr 2D-Verfolgungsalgorithmus berücksichtigt auch partielle Verdeckungen der Fahrzeuge, wobei ihr Verfahren dabei nochmals die spezielle Lage der Kamera ausnutzt, so daß Verdeckungen auf der Basis vertikaler Bildkoordinaten der Objekte erkannt werden können.

[Baumberg & Hogg 94] stellen einen 2D-Ansatz zur Detektion von Fußgängern vor. Nach der Detektion eines Fußgängers wird der Umriß seines Abbildes mit stückweise kubischen Kurven approximiert. Aus der Verteilung der Stützstellen der Kurven wird ein Modell für das menschliche Profil generiert. Es ist geplant, ihr Verfahren um eine 2D-Verfolgungskomponente zu erweitern.

[Niyogi & Adelson 94] schätzen ein Hintergrundbild mittels eines Median-Filters. Dann detektieren sie bewegte Fußgänger, indem sie einen Schwellwertoperator auf die Differenz von einem neuen Bild und dem Hintergrundbild anwenden. Sie setzen voraus, daß sich die Fußgänger im Bild aufrecht und parallel zur Bildebene bewegen. Außerdem sollte die Höhe der Knie und des Kopfes über dem Boden bekannt sein. Sie untersuchen und verfolgen vor allem die zwei beim Gehen von den Beinen eines Fußgängers im xt -Raum¹ erzeugten verflochtenen Bänder.

Mit einer Gaußschen Auflösungsrampe schätzen [Irani *et al.* 94a] 2D-Bewegungsparameter für die „dominierende“ Bewegung aus zwei Aufnahmen. Damit läßt sich das Folgebild so transformieren, daß die Bildbereiche statisch bleiben, die zu dieser Bewegung gehören. Mittels zeitlicher Integration bauen die Autoren eine rechnerinterne Datenstruktur auf, in denen die zur dominierenden Bewegung gehörenden Bildbereiche scharf, andere Bildbereiche unscharf sind. Diese Datenstruktur wird jeweils mit einem neuen Bild der Folge aktualisiert, aus dem die dominierende Bewegung schon eliminiert wurde. Da die zeitliche Integration von [Irani *et al.* 94a] aber im Gegensatz zu den Verfahren [Karmann & von Brandt 90; Kilger 93c; Koller *et al.* 94; Baumberg & Hogg 94] Bewegungsparameter mit einbezieht, arbeitet das Verfahren von [Irani *et al.* 94a] anders als die zitierten Arbeiten auch mit Bilddaten, die mit einer bewegten Kamera aufgezeichnet wurden. Indem der dominierenden Bewegung zugehörige Bereiche ausgeschlossen werden, können weitere Objekte gefunden werden. Das Verfahren funktioniert auch bei Verdeckungen und transparenten Objekten. [Irani *et al.* 94b] benutzen ihren Ansatz aus [Irani *et al.* 94a], um 3D-Bewegungsparameter der bildgebenden Kamera relativ zur Szene zu schätzen.

2.1.3 Optischer Fluß

Im Vergleich mit der Änderungsdetektion liefern Verfahren, die den *optischen Fluß* schätzen und auswerten, bessere Ergebnisse bei der Detektion bewegter Objekte². Der optische Fluß bezeichnet die scheinbare Verschiebungsgeschwindigkeit von Grauwertstrukturen in der Bildebene. Dieses aus den Bilddaten meßbare Feld muß nicht immer mit der *Verschiebungsrage* – auch *Bewegungsfeld* genannt – übereinstimmen, die die Projektion der Geschwindigkeitsvektoren von Szenenpunkten in die Bildebene angibt [Nagel 85; Horn 86]. Eine Übersicht zu Schätzverfahren für den optischen Fluß findet sich in [Nagel 85; Nagel 87; Nagel 92; Barron *et al.* 94]; für die Schätzung der Ableitungen des optischen Flusses siehe auch [Otte 94].

Zum einen gibt es Verfahren, die *merkmalsbasiert* Verschiebungsvektoren schätzen,

¹Dabei bezeichnet x horizontale Bildkoordinaten, t steht für die Zeitachse.

²Wir diskutieren dies noch genauer im Abschnitt 2.5.

und zum anderen *differentielle* Ansätze zur Schätzung des optischen Flusses.

[Sung & Zimmermann 86; Sung 88] bestimmen Verschiebungsvektoren durch Zuordnung lokaler Bildmerkmale, die mit Hilfe des Monotonieoperators aus einem Bild nach einem Verfahren von [Kories & Zimmermann 84; Kories & Zimmermann 86] extrahiert werden. Eine datengetriebene Ballungsanalyse liefert Objektbildkandidaten, die im Bildbereich verfolgt werden. [Sung & Zimmermann 86] detektieren Fahrspuren aus den Verschiebungsvektoren, da bei fester Kamera vor allem Abbilder bewegter Fahrzeuge Verschiebungen von Bildmerkmalen liefern. [Sung 88] wertet eine Bildfolge mit einer Dauer von 132 Sekunden aus. Obgleich die Trajektoriendaten auch nur mit einer Abtastfrequenz von 1,25 Hz ermittelt wurden, war dies im Hinblick auf die damals zur Verfügung stehende Rechnerleistung ein beachtliches Ergebnis. Im Gegensatz zu [Sung & Zimmermann 86] geben bei [Sung 88] nicht Verschiebungsvektoren selbst, sondern – wie auch vier Jahre später von [Li-Qun *et al.* 92] vorgestellt – aus detektierten Objektbildkandidaten ermittelte Fahrzeugtrajektorien Hinweise auf Fahrbahnabbildungen und sogar auf Ampelphasen. Dabei werden die Trajektorien mittels Kalibrierdaten in die 3D-Szene zurückprojiziert. Kumulierte Trajektorien wie bei [Sung 88; Li-Qun *et al.* 92] findet man auch bei [Mohnhaupt 92], der jedoch nur mit synthetischen Daten arbeitet.

[Schirra *et al.* 87] nutzen die Daten aus [Sung & Zimmermann 86], um mit dem System Citytour retrospektiv natürlichsprachliche Anfragen zu räumlichen und orientierungsabhängigen Relationen zwischen statischen sowie bewegten Objekten in der Bildfolge zu beantworten. Der statische Hintergrund wird manuell vorgegeben, da der Schwerpunkt der Bildauswertung bei der Detektion und Verfolgung von Fahrzeugen lag. Doch wurde bei [Schirra *et al.* 87] damit erstmals die Kette vom Bildsignal zur begrifflichen Beschreibung geschlossen.

Der Ansatz von [Sung & Zimmermann 86; Sung 88] liefert nicht nur Hinweise auf Abbilder bewegter Fahrzeuge, sondern wurde auch zur Detektion und 2D-Verfolgung von Fußballspielern getestet. Die Daten dienen einem natürlichsprachlichen System als Eingabe [Herzog *et al.* 89]. Die Trajektoriendaten waren zwar teilweise unterbrochen, so daß sie interaktiv mit einem graphischen Trajektorien-Editor ergänzt und korrigiert werden mußten, trotzdem war auch hier die Kette vom Bildsignal zur begrifflichen Beschreibung geschlossen worden.

[Koller *et al.* 91] bauen auf den Arbeiten von Sung auf und verfolgen vor allem solche Fahrzeuge wesentlich robuster, die sich nicht nur rein translatorisch bewegen. [Gong & Buxton 93] benutzen Verschiebungsvektoren, die in ähnlicher Weise wie bei [Sung & Zimmermann 86; Koller *et al.* 91] berechnet werden, segmentieren das Vektorfeld aber unter Einbeziehung von mehr Wissen u.a. über die Größe der Fahrzeuge. Dabei muß allerdings die maximale Anzahl zu erwartender Fahrzeuge vorgegeben werden.

[Zhang 93a] erhält durch Zuordnung korrespondierender Eckmerkmale mittels eines Hopfield-Netzes 2D-Trajektorien für drei Fahrzeuge. Zuordnung von Eckmerkmalen findet man auch bei [Roberts & Charnley 93] bei Aufnahmen, die aus einem fahrenden Fahrzeug aufgenommen wurden. Bei beiden Autoren wird allerdings keine Ballung durchgeführt, d.h. es werden keine Hypothesen über Objektbildkandidaten

aufgestellt. Auch wird von den Autoren nicht versucht, „Struktur aus Bewegung“ zu schätzen.

[Bouthemy & François 93] detektieren mit Hilfe eines stochastischen Ansatzes eigenständig bewegte Objekte in einer Bildfolge, die auch mit bewegter Kamera aufgenommen worden sein kann. Im Gegensatz zu einer früher publizierten Methode (vgl. [Bouthemy & Santillana Rivero 87]) muß bei [Bouthemy & François 93] ein Flußfeld nicht mehr explizit geschätzt werden. Mit einem Relaxationsverfahren wird das Bild in Bereiche unterschiedlicher Bewegungen segmentiert. Außerdem werden Bereiche verkettet, die in aufeinanderfolgenden Aufnahmen korrespondieren. Ihr rein im Bildbereich arbeitendes Verfahren unterscheidet weiterhin, wie sich die Objekte relativ zur Kamera bewegen: ob sie sich der Kamera nähern, sich entfernen oder sich zur Seite bewegen. Sie testen ihr Verfahren an zahlreichen Bildfolgen. Aufbauend auf dem Segmentieralgorithmus von [Bouthemy & François 93] stellen [Meyer & Bouthemy 94] einen bereichsbasierten Verfolgungsalgorithmus vor, der auch im Falle partieller Verdeckung zu verfolgender Fahrzeuge Ergebnisse liefert. [Meyer & Bouthemy 94] benutzen Polygone zur Beschreibung von Abbildern bewegter Objekte, deren Form und 2D-Bewegungsparameter mit einem Kalmanfilter geschätzt werden.

Um zu zählen, wie viele Personen in einen Bus ein- bzw. aussteigen, segmentieren [Del Bimbo & Nesi 93] ein optisches Flußfeld aus Bilddaten einer Kamera, die sich oberhalb einer Bustür befindet und nach unten gerichtet ist. Aufgrund dieser Anordnung weisen von ein- oder aussteigenden Personen erzeugte Flußfelder fast nur Komponenten in vertikaler Bildrichtung auf. Flußfelder werden in Richtung horizontaler Bildkoordinaten integriert, dann zeitlich geglättet. Mit Hilfe eines positiven Schwellwertes erhalten [Del Bimbo & Nesi 93] Hinweise auf einsteigende, mit Hilfe eines negativen Schwellwertes Hinweise auf aussteigende Personen. Sie stellen damit ein Zählverfahren vor, das nur auf ein eindimensionales Merkmal reduzierte Bildinformation auswertet.

Wissen über den Szenenbereich läßt sich zur Segmentierung von Flußfeldern heranziehen: Mit der Annahme, daß die Kamera sich parallel zu einer ebenen Fahrbahn bewegt, berechnet [Enkelmann 91] unter Verwendung von Kalibrierdaten sowie eines Meßwertes für die Geschwindigkeit der Kamera relativ zur Straße ein Bewegungsfeld, das die Projektion der Geschwindigkeitsvektoren relativ zur Kamera von auf der Fahrbahn liegenden Szenenpunkten angibt. Diskrepanzen zwischen dem Bewegungsfeld und dem geschätzten optischen Fluß geben Hinweise auf eigenständig bewegte Objekte und Hindernisse.

[Nagel 95a] detektiert Diskontinuitäten im optischen Flußfeld mit Hilfe von stochastischen Tests. Er läßt die bei bisherigen lokalen differentiellen Verfahren zur Schätzung des optischen Flusses getroffene, im allgemeinen jedoch nicht zutreffende Annahme fallen, die Schätzwerte für die Ableitungen der gemessenen Grauwertfunktion seien stochastisch unabhängig. Ergebnisse einer ersten Implementierung finden sich in [Nagel *et al.* 94; Socher 93]. Dort werden Abbilder bewegter Fahrzeuge vom Hintergrund getrennt, ohne das a-priori Wissen über eine feste Kamera in die Segmentierung des Flußfeldes mit einzubeziehen.

Diskontinuitäten im Fluffeld nutzen auch [Proesmans *et al.* 94] zur Detektion von Abbildern bewegter Objekte. Sie detektieren Bildbereiche, die sichtbar werden, sowie Bildbereiche, die gerade verdeckt werden. Daraus ergibt sich dann der Rand eines bewegten Objektbildes. Ihr Verfahren liefert sowohl bei synthetischen als auch bei realen Bilddaten – der Durlacher-Tor-Folge aus [Koller 92] – gute Ergebnisse.

2.1.4 Hindernisdetektion

Im speziellen, aber natürlich wichtigen Fall, in dem die Kamera hinter der Windschutzscheibe eines Fahrzeugs nach vorne gerichtet montiert ist, um im Rahmen eines Fahrerassistenzsystems den Fahrer zu unterstützen, reduziert sich die Objekterkennung zunächst auf eine einfache *Hindernisdetektion*. Bei dieser Kameraanordnung verläuft die der Kamera zugewandte Fläche eines vorausfahrenden Fahrzeuges parallel zur Bildebene, das Abbild eines solchen Hindernisses erscheint folglich meist rechteckförmig im Bildbereich. Die Echtzeitrandbedingung der hier vorliegenden Aufgabenstellung zwingt oft zu starken Vereinfachungen der Modellierung und zu starken Einschränkungen der Verfahren auf die Auswertung von Bildfolgen, die eine Kamera hinter der Windschutzscheibe eines auf der Autobahn fahrenden Fahrzeuges aufgenommen hat.

Zum automatischen Kolonnenfahren und zur Abstandshaltung bei einer Autobahnfahrt nutzen [Zielke *et al.* 93] zur Detektion vorausfahrender Fahrzeuge aus, daß die PKW-Rückansicht im Bild achsensymmetrisch zu einer vertikalen Achse erscheint. Symmetrien und Differenzbilder nutzt auch das Verfahren von [Bernasch & Koutny 93], wobei von ihnen detektierte Bereiche meist nur Teile eines Fahrzeuges erfassen und selten das ganze Abbild eines Fahrzeuges überdecken. Ihre Objektbildkandidaten sind achsenparallele Quadrate. Die zeitliche Stabilisierung ihres 2D-Verfahrens erfolgt mittels eines Kalmanfilters.

[Solder 92; Solder & Graefe 93] entdecken Objekte in großer Entfernung beim Betrieb eines autonomen Fahrzeuges. Bei einer weiten Vorausschau auf der Autobahn unterscheiden sich vorausfahrende Fahrzeuge von Flecken auf der Fahrbahn dadurch, daß ein Fleck relativ lang sein müßte, um im Bild die Größe des Abbildes eines Fahrzeuges aufzuweisen. Ein Fahrzeug, das auf derselben Spur vorausfährt, erscheint im Bild als Rechteck. Da die obere Begrenzungslinie – je nach Hintergrund – nicht so zuverlässig zu detektieren ist wie die untere und die seitlichen Begrenzungslinien des Rechtecks, suchen die Autoren in einem mit einer ternären Maske (Prewitt-Operator) erzeugten Grauwertgradientenbild nach U-förmigen Merkmalen zur Detektion von Fahrzeugabbildern. Ihr heuristisches, sehr einfaches Verfahren ermöglicht eine kurze Zykluszeit für den Fahrzeugdetektionsprozeß. Die Hinderniskandidaten von [Solder 92] werden bei [Regensburger 93] in Falschalarme und tatsächliche Hindernisse klassifiziert und verfolgt, wobei Kanten- und Eckmerkmale wiederum mittels einfacher ternärer Masken berechnet werden. Zur Hindernisidentifikation setzt er sowohl für den Nah- als auch für den Fernbereich ein generisches 2D-Fahrzeugmodell für eine PKW-Rückansicht ein.

Detektion und Erkennung von Fahrzeugen sind nicht nur mit einer Videokamera,

sondern auch mit anderen Sensoren möglich. Einen Vergleich zwischen der Verwendung einer Videokamera als bildgebendem Sensor und anderen Sensoren, die auf Radar-, Ultraschall- oder Lasertechniken beruhen, findet sich in [Soldner 92]. Beim Einsatz mehrerer Sensoren muß auch geklärt werden, wie die Daten der Sensoren zusammengefaßt werden und wie die Sensoren zu kalibrieren sind [Regensburger 93].

2.1.5 Suchverfahren

Darüber hinaus gibt es Publikationen, in denen *Suchverfahren* zur Detektion von Objekten eingesetzt werden.

[Kinzel 94] detektiert aufrecht gehende Fußgänger in Bildfolgen, die eine bewegte Kamera aufgenommen hat. Das Längen- und Breitenverhältnis des Abbildes eines Fußgängers müssen ungefähr bekannt sein. Kantensegmente werden berechnet und unter Verwendung von Heuristiken und Suchverfahren paarweise zugeordnet. [Kinzel 94] formuliert jedoch keine Hypothesen, um einen modellbasierten Verfolgungsansatz zu initialisieren.

2.2 3D-Verfolgung

Wir fassen zunächst eine Reihe von Arbeiten zusammen, die an der Universität Reading in England durchgeführt wurden:

[Zhang *et al.* 93] schlagen für eine 3D-Objekterkennung ein relationales Modell vor, das Objektwissen enthält, das unabhängig von der Betrachtungsrichtung ist. Dabei wird das Wissen über ein Objekt in einer Graphenstruktur rechnerintern dargestellt. Die Knoten des Graphen repräsentieren Modellmerkmale. Die Kanten werden sowohl mit Gewichten versehen, die für zwei Modellmerkmale angeben, ob sie gemeinsam sichtbar sind, als auch mit Zwangsbedingungen, die ansichtsunabhängige Relationen zwischen Modellmerkmalen erfassen, z.B. Parallelität oder Größenverhältnisse. Aus einem CAD-Drahtmodell konstruiert ihr System ein ansichtsunabhängiges, relationales Modell, mit dessen Hilfe man das Objekt auch ohne Wissen über die Lage der Kamera erkennen kann.

[Tan *et al.* 92b] stellen einen nichtstatistischen, linearen Algorithmus zur Lageschätzung von Objekten vor, wobei die Zuordnungen zwischen Daten- und Modellkantensegmenten interaktiv vorgegeben werden bzw. in [Tan *et al.* 93; Tan *et al.* 94a] mit einer verallgemeinerten Hough-Transformation algorithmisch ermittelt werden. Dieses Verfahren wird im Rahmen ihres Bildauswertungssystems für Straßenverkehrsszenen benutzt, um Daten mehrerer Kameras [Tan *et al.* 94b] gleichzeitig auszuwerten sowie gegliederte Objekte zu verfolgen, z.B. einen LKW mit Anhänger [Worrall *et al.* 93].

Im Gegensatz zu Ansätzen, in denen mögliche Zuordnungen zwischen Modell- und Datenmerkmalen mit Hilfe eines Interpretationsbaumes verwaltet werden [Grimson 90] und der dann nötige Suchaufwand mit geeigneten Algorithmen verringert werden muß [Lowe 92a], führen [Du *et al.* 93] eine 3D-Gruppierung von Kantensegmenten durch, indem eine Funktion monoton verbessert wird, die die Modellanpassung an Datenkantensegmente bewertet. Experimentell zeigen sie, daß ihr Ansatz bei mit

Störungen behafteten Kantensegmenten robuster als Lowes Verfahren ist.

Unter der Voraussetzung, Punktmerkmale seien bereits aus den Bilddaten berechnet und korrekt zugeordnet, ermitteln [Tan *et al.* 92a] Struktur (relative Tiefe) und Bewegungsparameter unter der Annahme, daß die Objekte nur zwei translatorische und einen rotatorischen Parameter haben. Mit interaktiver Hilfe lassen sie die erhaltenen 3D-Punkte in ein vollständiges polyhedrales Objektmodell überführen. [Tan *et al.* 92c] erweitern ihr Verfahren auf mehr als zwei Aufnahmen als Eingabedaten. Weitere Untersuchungen finden sich in [Tan *et al.* 94d].

Unter der Annahme, ein interaktiv ausgewählter Bildausschnitt enthält höchstens ein Fahrzeug, schätzen [Tan *et al.* 94c] die Orientierung eines Fahrzeuges relativ zur Fahrbahnnormalen durch eine Suche nach einem Maximum in einem Histogramm akkumulierter Bildgradienten. Mit der weiteren Annahme, das Fahrzeug sei nahezu rechtwinklig, werden die drei Hauptrichtungen des Fahrzeugmodells in die Bildebene projiziert. Dazu modellieren die Autoren die Bildaufnahme als schwache perspektivische Abbildung. Für jede der projizierten Hauptrichtungen wird ein 1D-Modell-Profil erzeugt, indem die Szenenskizze einer in das Bild projizierten Fahrzeugausprägung senkrecht auf die jeweilige projizierte Hauptrichtung abgebildet wird. Maxima einer Korrelation eines 1D-Modell-Profiles mit einem aus den Bildgradienten gewonnenen 1D-Daten-Profil geben Hinweise auf (eine der beiden) Komponenten der Bildposition des Fahrzeugmodells. Diese Komponenten werden mit einer Ausgleichsrechnung zu einer Bildposition des Fahrzeugmodells zusammengefaßt.

Bei gegebener Initialisierung projizieren [Worrall *et al.* 94] ein polyedrisches Fahrzeugmodell in die Bildebene. Senkrecht zu den projizierten Modellkanten werden Richtungsableitungen aus dem Bildmaterial geschätzt, ohne Datenkantensegmente aus dem Bild zu extrahieren. Es handelt sich aber immer noch um ein Suchverfahren um die interaktiv vorgegebene richtige Position. Ihr Verfahren oszilliert oft und braucht viele Iterationsschritte.

Die von dieser Gruppe benutzten Fahrzeugmodelle werden mit Hilfe eines interaktiven Werkzeugs von [Attwood *et al.* 93] erstellt: nach Vorgabe eines Flächenmodells markiert der Benutzer in einer Fahrzeugaufnahme eine Menge von Punkten, die symmetrisch zu der Symmetrieebene des Fahrzeugs liegen. Das System berechnet aus den 2D-Bild- die zugehörigen 3D-Modellkoordinaten.

Mit Ausnahme älterer Arbeiten dieser Gruppe — z.B. [Worrall *et al.* 91] sowie der mittels Kalmanfilter verbesserte Ansatz von [Marslin *et al.* 91; Sullivan 92]— findet in den neueren Arbeiten dieser Forschungsgruppe nur die Auswertung einzelner Bilder statt, wobei die Stetigkeit von Bewegungen in Bildfolgen nicht ausgenutzt wird. Hinweise zur Verfolgung finden sich in [Worrall *et al.* 94], wobei Trajektorien allerdings nicht angegeben werden.

Ein 3D-Polyedermodell wie in [Koller 92] verwenden auch [Schick & Dickmanns 91; von Holt 94]. Mit Hilfe einiger Abstandsheuristiken — und nicht mit einer Mahalanobis-Distanz wie [Deriche & Faugeras 90] — werden Modellsegmente Datensegmenten zugeordnet. In den Arbeiten von [Schick & Dickmanns 91; von Holt 94] werden allerdings keine realen Bilddaten ausgewertet.

Eine Reihe von Autoren befaßt sich mit der Detektion und Verfolgung nicht starrer Körper. [Kellner 93] verwendet das Schätzverfahren aus [Koller 92], um die Position und Orientierung eines Fußgängers zu schätzen. Arme, Beine, Rumpf und Kopf werden durch 14 elliptische Zylinder und 13 Gelenke modelliert. [Schürmann 94] führte die Arbeit von [Kellner 93] fort, indem er die Gelenkwinkeinstellungen in den Schätzprozeß integrierte und eine zeitliche Fortschreibung einführte.

[Rehg & Kanade 94] verfolgen mit einer Frequenz von 10 Hz eine menschliche Hand, die mit 27 Freiheitsgraden repräsentiert wurde. Sie benutzen jedoch eine orthographische Projektion. Ihre Minimierung ist dem Verfahren von [Lowe 92a] ähnlich.

2.3 Schätzung von Bewegungs- und Strukturparametern

Zahlreiche Arbeiten befassen sich vor allem mit der Schätzung von Bewegungs- und Strukturparametern. Dabei werden – im Gegensatz zu dichten Flußfeldern – meist wenige, unter Umständen interaktiv vorgegebene Merkmalskorrespondenzen verwendet. Einen Überblick geben [Faugeras 93; Haralick & Shapiro 93; Kanatani 93a; Daniilidis 92; Maybank 92]. Neben den oben schon zitierten Arbeiten von [Tan *et al.* 92a; Tan *et al.* 92c; Irani *et al.* 94b] gehen wir kurz noch auf weitere ein.

[Liu & Huang 93] verwenden ein Bewegungsmodell, das eine 3D-Bewegung als eine Rotation um die Fahrzeugachse mit anschließender Translation repräsentiert. Sie führen zusätzliche Zwangsbedingungen für die Rotation und Translation ein und erhalten verschiedene Bewegungstypen. Doch ihr Ansatz zur Schätzung der Fahrzeugbewegung wurde nur mit fünf Aufnahmen getestet, die jeweils nur ein einziges Fahrzeug enthalten.

Mehrere Aufnahmen – z.B. von einem Haus – und viele Punkte benutzen [Tomasi & Kanade 92], die unter der Annahme einer orthographischen Projektion mit einer Faktorisierungsmethode, die auf einer Singulärwertzerlegung beruht, Form und Bewegung aus einem Bildstrom berechnen. Ihr Verfahren arbeitet damit retrospektiv und nicht inkrementell. Eine Verallgemeinerung der Methode findet man bei [Poelman & Kanade 94], die die perspektivische Projektion nicht mehr durch eine orthographische, sondern durch eine sog. paraperspektivische Projektion approximieren.

Einen Bildstrom werten auch [Szeliski & Kang 94] aus. Mittels – wiederum nach [Kories & Zimmermann 86] – gewonnenen Punkt- sowie Geradensegmentmerkmalen berechnen sie mit dem Gauß-Newton-Verfahren 3D-Form- und Bewegungsparameter aus synthetischen sowie aus realen Daten, wobei keine orthographische, sondern eine perspektivische Projektion zugrundegelegt wird.

2.4 Systeme aus dem Verkehrswesen

Bei videobasierten Systemen aus dem Bereich des Verkehrswesens stehen vor allem photogrammetrische Aspekte im Vordergrund, weniger eine automatische Bildfol-

genauswertung. So versuchen [Herslund *et al.* 93], aus mit einem Stereokamerasystem aufgenommenen Bilddaten interaktiv (mit der Maus) eingegebene Bildkoordinaten in einen Plan zurückzuprojizieren, um Konflikte und Koordinationsverhalten im Straßenverkehr zu studieren. Die Entzerrung perspektivischer Aufnahmen behandeln [Wegmann & Spring 93], wobei Fahrspuren interaktiv ins Bild eingezeichnet werden. Wenn auch Fahrzeuge interaktiv in jeder Aufnahme „punktiert“ werden, d.h. der Benutzer gibt die Stützpunkte von Fahrzeugtrajektorien ein, so gestattet es der Einsatz von Videokameras, auf sonst zur Verkehrsanalyse notwendige bauliche Maßnahmen, wie Induktionsschleifen, zu verzichten [Wegmann & Spring 93]. Beim System von [Hupfer 93] werden dagegen in Video-Bilddaten erfaßte Fahrzeuge automatisch gezählt, wobei Fahrspuren jedoch interaktiv vorgegeben werden müssen. Aber auch bei [Hupfer 93] steht die Umrechnung von Bild- in Weltkoordinaten im Vordergrund. Die Kalibrierung geschieht dabei über die interaktive Eingabe von Weltkoordinaten ausgewählter Punkte.

Die vorliegende Arbeit hat dagegen gerade die Bildfolgenauswertung als Schwerpunkt. Darüber hinaus liefert das in Kapitel 4 vorgestellte Kalibrierverfahren die Möglichkeit einer automatisch initialisierten perspektivischen Bildentzerrung, wobei interaktive Eingaben durch allgemeiner gültige Annahmen ersetzt werden.

2.5 Diskussion der Literatur

Wir vergleichen unser System mit den zitierten Arbeiten:

- **Komplexität des Bildmaterials:**

Im Vergleich zu den zitierten Arbeiten ist das von uns verwendete Bildmaterial reichhaltiger: Die in dieser Arbeit vorgestellten Verfahren werden nicht nur an einem Einzelbild, sondern an Bildfolgen getestet. Desweiteren werden verschiedene Bildfolgen untersucht, die von unterschiedlichen Kameras, Röhren- sowie Halbleiterkameras, an mehreren Kreuzungen in verschiedenen Städten aufgenommen wurden. Wir verzichten bewußt auf die Auswertung synthetischer Daten sowie auf die Auswertung aufgebauter Spielzeugwelten (im Gegensatz zu [Worrall *et al.* 93]) und konzentrieren uns auf die Auswertung realer Daten. Die im Rahmen der vorliegenden Arbeit untersuchten Kreuzungen zeigen komplexes Verkehrsgeschehen mit vielen Fahrzeugen. In den Bilddaten von [Bouthemy & François 93] sowie von [Zhang 93a] finden sich beispielsweise maximal drei Fahrzeuge, bei [Irani *et al.* 94a] ein Fahrzeug und ein Fußgänger, bei [Liu & Huang 93] ein Fahrzeug. [Solder 92; Regensburger 93; Dubuisson & Jain 93; Dubuisson & Jain 94] detektieren meist nur jeweils ein Fahrzeug pro Aufnahme, wenn auch viele Aufnahmen ausgewertet werden. Wenige Fahrzeuge findet man in den Aufnahmen von [Karmann & von Brandt 90; Gong & Buxton 93] sowie bei [Sullivan 92] und anderen Publikationen dieser Gruppe. Deutlich mehr Fahrzeugabbilder werten [Sung 88; Li-Qun *et al.* 92; Kilger 93c; Koller *et al.* 94] aus.

- **Randbedingungen bei der Bildaufnahme:**

Manche Verfahren (z.B. [Kilger 93c; Dubuisson & Jain 93; Dubuisson & Jain 94; Koller *et al.* 94]) sind so entwickelt, daß sie eine feste Kamera als notwendige Voraussetzung haben. In unserer Arbeit nutzen wir diese Voraussetzung auch, z.B. bei der Detektion bewegter Fahrzeugabbilder. Vor allem Verfahren, die nur im Bildbereich arbeiten, sind oft nur bei einer eng eingegrenzten Lage der Kamera einsetzbar: die Methoden von [Zielke *et al.* 93; Bernasch & Koutny 93; Solder 92; Regensburger 93] benötigen eine Kamera im Fahrzeug, bei [Kilger 93c; Koller *et al.* 94] sollte die Kamera am besten auf einer Autobahnbrücke stehen, etwas nach unten geneigt und in Richtung des Straßenverlaufs orientiert sein, bei [Dubuisson & Jain 94] orthogonal zur Straße. Personen werden von [Kinzel 94; Niyogi & Adelson 94] auch nur detektiert, wenn sie im Bild aufrecht abgebildet werden. Unser Ansatz schließt eine allgemeine Kameralage nicht aus.

- **Verwendung anderer Sensoren:**

In unserem System versuchen wir, nach Möglichkeit alle Informationen aus dem Bildmaterial zu schätzen. Andere Verfahren, beispielsweise im Bereich der automatischen Fahrzeugführung, wo die Regelungstechnik und Echtzeitanforderungen im Vordergrund stehen [Dickmanns & Mysliwetz 92; Solder 92; Regensburger 93], basieren stark auf Eingabedaten, die von anderen Sensoren geliefert werden.

- **Detektion bewegter Objekte mit optischem Fluß sowie mit einer Änderungsdetektion:**

Die *Änderungsdetektion* gehört zu den ersten Verfahren, mit denen man in der Bildauswertung bewegte Objekte detektierte. Die Probleme bei Verwendung einer Änderungsdetektion werden beispielsweise von [Nägel 85] und [Hsu *et al.* 84] diskutiert. Bei Einsatz einer Änderungsdetektion erweist sich insbesondere die Wahl des Schwellwertes als schwierig, mit dessen Hilfe man Grauwertänderungen, die durch Veränderungen der relativen Lage von Kamera und Szene hervorgerufen werden, von denen trennt, die durch Rauschen entstehen [Nägel 85]. Bei bewegter Kamera sind Grauwertänderungen an einer festen Bildposition in aufeinanderfolgenden Aufnahmen nicht mehr die Ausnahme, sondern die Regel. Damit läßt sich eine Änderungsdetektion – wie z.B. von [Kilger 93c; Koller *et al.* 94; Dubuisson & Jain 93; Dubuisson & Jain 94; Niyogi & Adelson 94] eingesetzt – in diesem Fall nicht mehr zur Detektion bewegter Objekte verwenden.

Die Berechnung von optischen Flußvektoren ist aufwendiger als eine einfache Änderungsdetektion, dafür enthält sie aber Bewegungs- und Richtungsinformation, wodurch wir im Gegensatz zu z.B. den Arbeiten von [Kilger 93c; Dubuisson & Jain 93; Dubuisson & Jain 94; Niyogi & Adelson 94] auch eine Initialisierung für die Geschwindigkeit berechnen können. Eine optische Flußschätzung liefert im Gegensatz zur primitiven Änderungsdetektion auch bei bewegter Kamera Hinweise auf Objektbildkandidaten, auch wenn die Segmentierung des Flußfeldes schwieriger ist als bei fester Kamera. Ansätze zur

Detektion von Objektbildkandidaten bei bewegter Kamera finden sich u.a. bei [Bouthemy & François 93; Irani *et al.* 94a], aber auch bei [Wachter 93].

[Dubuisson & Jain 94] weisen darauf hin, daß Fahrzeuge, die sich begegnen, dann gut zu trennen wären, wenn man ihre Geschwindigkeiten kennt. Im Gegensatz zu [Dubuisson & Jain 94] können wir durch Ausnutzung von Schätzwerten für den optischen Fluß in diesem Fall Abbilder zweier sich begegnender Fahrzeuge gut trennen. Den anderen Fall, bei dem sich Fahrzeuge in enger Nachbarschaft mit gleicher Geschwindigkeit bewegen, was auch bei [Sung & Zimmermann 86] zu Problemen führte, lösen wir durch eine modellbasierte Segmentierung, die im Gegensatz zur Segmentierung von [Dubuisson & Jain 94] mit allgemeiner gültigen Annahmen über die 3D-Szene arbeitet.

Manche Autoren halten optische Flußfelder für ungeeignet zur Detektion von Objekten. [Koller *et al.* 93b] verzichten beispielsweise auf den Einsatz optischer Flußfelder zur Detektion von Abbildern bewegter Fahrzeuge, da sie aus den Bilddaten einer auf einer Autobahnbrücke montierten Kamera als Folge von Vibrationen keine guten Flußfelder erhalten. Die Güte der Flußvektoren hängt stark von der Wahl des Verfahrens zur Schätzung der Vektoren ab. Einen Vergleich von Verfahren findet sich in [Otte 94; Barron *et al.* 94]

Aufgrund von Fortschritten bei der Auswertung von Halbbildern können wir erstmals optische Flußfelder mit 50Hz segmentieren [Kollnig *et al.* 94], ohne wie [Enkelmann 87] Flußvektoren über vier aufeinanderfolgende Bildpaare konkatenieren zu müssen. Auch entfällt die bei [Koller *et al.* 91; Koller 92] notwendige Stabilisation der Verschiebungsvektorfelder, die als Bildabstanzvektoren aus über mehreren Aufnahmen verfolgten Bildmerkmalen ermittelt wurden.

- **Zeitliche Faltung:**

Im Gegensatz zu räumlichen Operatoren, wo Nachteile von Faltungsmasken gerader Maskenbreite schon früh erkannt wurden, benutzen viele Autoren für zeitliche Operatoren Faltungsmasken mit gerader Maskenbreite, meistens der Breite zwei. [Irani *et al.* 94a] initialisieren mit zwei Aufnahmen, wenn auch später zeitlich integriert wird. [Koller *et al.* 94] berechnen die initiale Geschwindigkeit als Differenz zweier Aufnahmen. Die von uns verwendeten optischen Flußvektoren werden mit Faltungsmasken berechnet, die auf das laufende Bild zentriert sind. Die Geschwindigkeit ist dann eine Momentangeschwindigkeit für den Zeitpunkt eines Bildes, nicht für einen Zeitpunkt zwischen zwei Aufnahmen.

- **Initialisierung:**

Während vor allem bei Arbeiten der Gruppe an der Universität in Reading Startwerte für eine Verfolgung interaktiv vorgegeben oder mit Suchverfahren [Tan *et al.* 94a; Tan *et al.* 94b] gefunden werden, werden diese z.B. bei [Baumberg & Hogg 94; Koller *et al.* 94] automatisch berechnet. Durch Ausnutzung der im optischen Fluß enthaltenen Information erhalten wir gute Initialisierungen, so daß wir im Gegensatz zu [Tan *et al.* 94a; Tan *et al.* 94b] vermeiden,

alle möglichen Positionen und Orientierungen des Modells durchzuprobieren.

- **Abtastfrequenz:**

Unsere Trajektorien werden zwar nicht in Echtzeit berechnet, dafür sind sie dichter abgetastet: mit 50 Hz im Vergleich zu 1,25 Hz [Sung 88], 0,48 Hz [Tan *et al.* 92b], 3 Hz [Li-Qun *et al.* 92], 3-5 Hz [Kilger 93c], 5 Hz [Tan *et al.* 93; Worrall *et al.* 93], 10 Hz [Rehg & Kanade 94]. Eine höhere Abtaststrategie rechtfertigt eine Linearisierung von Systemfunktionen und stabilisiert damit die Verfolgung von Fahrzeugen.

- **Klassifikation:**

[Kilger 93c] und [Dubuisson & Jain 93] stellen auch in ihren Arbeiten erste Ansätze zur Klassifikation von Fahrzeugen vor. Wir haben dies in unserer Arbeit zunächst zurückgestellt und den Schwerpunkt der Arbeit auf eine robuste Verfolgung gesetzt.

- **Lichteinfall:**

Im Gegensatz zur Arbeit von [Kilger 93c], der Beleuchtungsverhältnisse, Schattwurf und Lichteinfallrichtung berücksichtigt, müssen wir nicht Ort, Tag und Uhrzeit notiert haben.

- **Verfolgung:**

- Einige Verfahren zur Verfolgung arbeiten nur im *2D-Bildbereich*: [Li-Qun *et al.* 92; Karmann & von Brandt 90; Kilger 93c; Koller *et al.* 94; Sung & Zimmermann 86; Sung 88; Koller *et al.* 91; Meyer & Bouthemy 94; Solder 92; Regensburger 93; Bernasch & Koutny 93; Dubuisson & Jain 93; Blake *et al.* 93; Baumberg & Hogg 94]. Unser Verfolgungsalgorithmus arbeitet im *3D-Szenenbereich*.
- *Detektion und Verfolgung* findet man bei [Rehg & Kanade 94; Kilger 93c; Koller *et al.* 94; Li-Qun *et al.* 92], *nur Detektion, keine Verfolgung* bieten die Publikationen von [Baumberg & Hogg 94; Dubuisson & Jain 93]. Häufig wird zur Verfolgung von Objekten ein *Kalmanfilter* eingesetzt: [Bernasch & Koutny 93; Karmann & von Brandt 90; Kilger 93c; Koller *et al.* 94; Meyer & Bouthemy 94; Blake *et al.* 93].
- Ein generisches, jedoch *2D-Fahrzeugmodell* findet sich bei [Regensburger 93; Dubuisson & Jain 94], *3D-Modelle* benutzen die Gruppe aus Reading und [Schick & Dickmanns 91; von Holt 94]. Unsere Arbeit baut auf den *3D-Fahrzeugmodellen* von [Koller 92] auf, sie werden allerdings verfeinert und ergänzt.
- *Detektion gegliederter Objekte* findet man bei [Baumberg & Hogg 94; Kinzel 94] sowie ihre Verfolgung bei [Rehg & Kanade 94; Kellner 93; Schürmann 94], ansatzweise bei einem LKW mit Anhänger auch bei [Sullivan 92; Worrall *et al.* 93]. Abbilder von *Fahrzeugen* werden z.B. von [Dubuisson & Jain 93; Dubuisson & Jain 94; Li-Qun *et al.* 92; Karmann & von Brandt 90; Kilger 93c; Koller *et al.* 94; Irani *et al.* 94a; Sung &

Zimmermann 86; Sung 88; Koller *et al.* 91; Gong & Buxton 93; Zhang 93a; Roberts & Charnley 93; Bouthemy & François 93; Meyer & Bouthemy 94; Enkelmann 91; Nagel *et al.* 94; Proesmans *et al.* 94; Zielke *et al.* 93; Bernasch & Koutny 93; Solder 92; Regensburger 93; Schick & Dickmanns 91] und in den Arbeiten an der Universität in Reading, z.B. [Sullivan 92; Worrall *et al.* 93; Worrall *et al.* 94], ausgewertet. [Sung & Zimmermann 86; Sung 88] detektieren Abbilder einer bewegten *Straßenbahn*, wenn auch das Gelenk nicht modelliert wurde. Selten findet man Aufnahmen von *Radfahrern*, eine Ausnahme bilden [Bouthemy & François 93], die den Radfahrer aber nicht explizit modellieren. Aufnahmen von *Fußgängern* werden bei [Herzog *et al.* 89; Baumberg & Hogg 94; Niyogi & Adelson 94; Bouthemy & François 93; Del Bimbo & Nesi 93; Kinzel 94; Kellner 93; Schürmann 94], ansatzweise auch von [Regensburger 93] ausgewertet. Einen Überblick über Verfahren zur *Fahrbahnerkennung* gibt [Zhang 94].

- **Berechnung eines Grauwertbildes für den Hintergrund:**

[Karmann & von Brandt 90; Kilger 93b; Koller *et al.* 94; Baumberg & Hogg 94; Niyogi & Adelson 94; Irani *et al.* 94a] schätzen ein Grauwertbild für den Hintergrund. Allerdings sind dies alles 2D-Ansätze. Wir benutzen für den Hintergrund ein Fahrbahnmodell und passen es mit einem 3D-Verfahren an die Bilddaten an.

- **Explizite Angabe von Abstimmparametern:**

Selten werden interaktiv einzustellende Parameter explizit angegeben. Damit wird es schwierig, bereits publizierte Verfahren mit neuen Verfahren auch experimentell zu vergleichen.

- **Gesamtsystem**

Desweiteren beschränken wir uns nicht auf ein einzelnes Detail der Bildfolgenauswertung, sondern stellen ein System vor, das alle Auswertungsschritte von den Grauwerten bis zu den begrifflichen Beschreibungen umfaßt. Bisher publizierte Verfahren im Bereich der Detektion und Verfolgung von Fahrzeugen konzentrieren sich meist auf die Lösung von Detailfragen, beispielsweise die Arbeiten von Baker, Sullivan und Mitarbeitern aus Reading. Fragestellungen zu dem Verhalten eines ganzen Systems wurden kaum diskutiert. Doch zeigte es sich – auch im Rahmen der vorliegenden Untersuchung –, daß Verfahren, die isoliert gut funktionieren, sich als Schwachpunkte herausstellen können, wenn sie in einen Systemkontext eingebunden werden. In dieser Arbeit zeigen wir, daß ganze Systemkomponenten – beispielsweise die Extraktion von Kantensegmenten – weggelassen und durch leistungsfähigere ersetzt werden müssen, um das System insgesamt robuster zu machen.

Kapitel 3

Einsatz leistungsfähiger signalnaher Komponenten

3.1 System von [Koller 92]

Ausgangspunkt für die vorliegenden Untersuchungen bildet ein Bildfolgenauswertungssystem von [Koller 92], das bewegte Fahrzeuge in Videoaufnahmen des Straßenverkehrs detektiert und verfolgt (vgl. Abb. 3.1). Mit einem merkmalsbasierten Verfahren bestimmt [Koller 92] ein *Verschiebungsvektorfeld*. Eine Segmentierung dieses Feldes liefert Bereiche, die als Kandidaten für Abbilder bewegter Objekte – sog. *Objektbildkandidaten* – dienen. Mittels Kalibrierdaten, die nach einem Verfahren von [Tsai 87] ermittelt wurden, werden die Objektbildkandidaten in die Szene zurückprojiziert. Man erhält dadurch *Objektkandidaten*. Damit wird ein modellbasierter Verfolgungsalgorithmus initialisiert. Durch Projektion einer Ausprägung eines Fahrzeugmodells in die Bildebene erhält man unter Berücksichtigung verdeckender Flächen Modellkantensegmente, die aus dem Bild extrahierten Datenkantensegmenten zugeordnet werden. Damit läßt sich die Ausprägung mit einer MAP-Schätzung aktualisieren. Eine Kalmanfilter-basierte Prädiktion nutzt ein Bewegungsmodell.

3.2 Entwicklung leistungsfähigerer signalnaher Komponenten

Aufbauend auf den Arbeiten, Werkzeugen und Erfahrungen von [Koller 92] gelang es, durch Entwicklung leistungsfähigerer *datengetriebener* Programmkomponenten im Vergleich zum System von [Koller 92] folgende Verbesserungen zu erzielen:

- **Auswertung von Halbbildern:**

Durch das in der Videotechnik übliche Zeilensprungverfahren bei der Aufzeichnung von Bildern werden im ersten Halbbild die ungeraden, im zweiten Halbbild die geraden Zeilen mit einem zeitlichen Versatz von 20 ms aufgenommen.

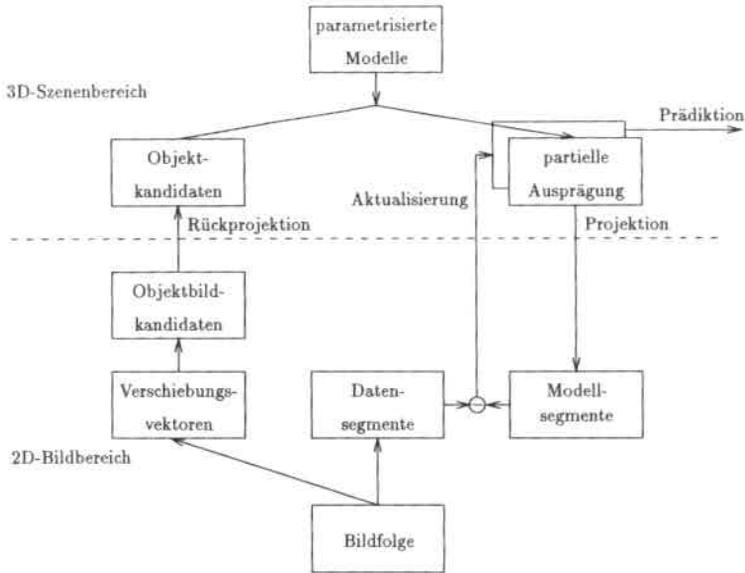


Abbildung 3.1: System von [Koller 92], entwickelt nach Abb. 1.2: Mit einem merkmalsbasierten Verfahren werden aus den Bilddaten *Verschiebungsvektoren* ermittelt. Eine Segmentierung der Verschiebungsvektoren liefert Hinweise auf Abbilder bewegter Fahrzeuge (*Objektbildkandidaten*). Die nach einer Rückprojektion in den 3D-Szenenbereich gewonnenen *Objektkandidaten* dienen zur Initialisierung eines modellgestützten Verfolgungsverfahrens: Die Projektion einer Ausprägung eines Fahrzeugmodells in die Bildebene liefert unter Berücksichtigung verdeckender Flächen *Modellkantensegmente*, die aus dem Bild extrahierten *Datenkantensegmente* zugeordnet werden. Die Differenz korrespondierender Segmente dient zur *Aktualisierung* der Modellausprägung, die anschließend zeitlich fortgeschrieben wird. In der Folgeaufnahme entfällt der Initialisierungsschritt, nur die präzidierte Modellausprägung wird aktualisiert.

men [Mäusel 91]. In dem durch Verschränkung der Halbbilder entstandenen Vollbild treten bei der Abbildung sich in horizontaler Richtung bewegender Objekte Verzerrungen auf. Dies äußert sich in Inkonsistenzen und Artefakten bei der Ermittlung von Bildbereichshinweisen [Koller 92]. Durch Fortschritte bei der Berechnung von Ableitungsoperatoren [Otte 94] wurden Filter bereitgestellt, die es gestatten, jedes Halbbild der Bildfolge auszuwerten. Damit konnte im Vergleich zu [Koller 92] die Abtastfrequenz von 25 Hz auf 50 Hz verdoppelt werden. Die Ableitungsoperatoren werden zur optischen Flußschätzung, zur Kantenberechnung und – in Kapitel 6.1 – zur Flächenanpassung eingesetzt. Damit basiert die Bildfolgenauswertung auf einem einheitlichen Satz signalnaher Verfahren, was die Systematik des gesamten Systems erhöht. Bis einschließlich des Programmteils zur Berechnung von begrifflichen Beschreibungen arbeitet jede Abstraktionsebene mit dieser Abtastfrequenz.

- **Austausch signalnaher Systemkomponenten:**

Das Programmodul zur Schätzung von *Verschiebungsvektoren*, in dem mit Hilfe des *Monotonieoperators* aus einem Bild extrahierte Merkmale für eine heuristische Zuordnungsstrategie eingesetzt werden [Kories & Zimmermann 84; Kories & Zimmermann 86], wurde durch ein auf Ableitungsoperatoren aufbauendes differentielles Verfahren zur Schätzung des optischen Flußfeldes ersetzt.

- **Verbesserung des datengetriebenen Ballungsverfahrens:**

Das jetzt zur Verfügung stehende, dichtere Feld erzwang eine Überarbeitung der Verfahren zur Ballung von Vektoren, um durch eine Segmentierung des Vektorfeldes Objektbildkandidaten zu erhalten. Erste Ergebnisse findet man in [Kollnig *et al.* 94].

- **Datensegmentextraktion:**

Außerdem wurde das von Koller benutzte Modul zur Berechnung von Datenkantensegmenten durch ein von [Otte 94; Otte & Nagel 92] entwickeltes Verfahren ersetzt, das bessere Kantensegmente aus den Bilddaten berechnet, was ein Einengen der Toleranzen zur Verfolgung der Fahrzeuge gestattet.

- **Einzustellende Parameter:**

Bei [Koller 92] muß für jedes Fahrzeug ein eigener Parametersatz für den Verfolgungsalgorithmus eingestellt werden, wohingegen bereits in [Kollnig *et al.* 94] eine einzige Parametereinstellung zur Verfolgung mehrerer Fahrzeuge gefunden wurde. Darüber hinaus gelang es, die Anzahl der einzustellenden Parameter zu verringern.

Abb. 3.2 faßt die dargestellten Änderungen zusammen.

3.3 Segmentierung eines optischen Flußfeldes

Eine robuste Detektion bewegter Objektbildkandidaten durch einen datengetriebenen Ansatz ist eine wesentliche Voraussetzung für den darauf aufbauenden modell-

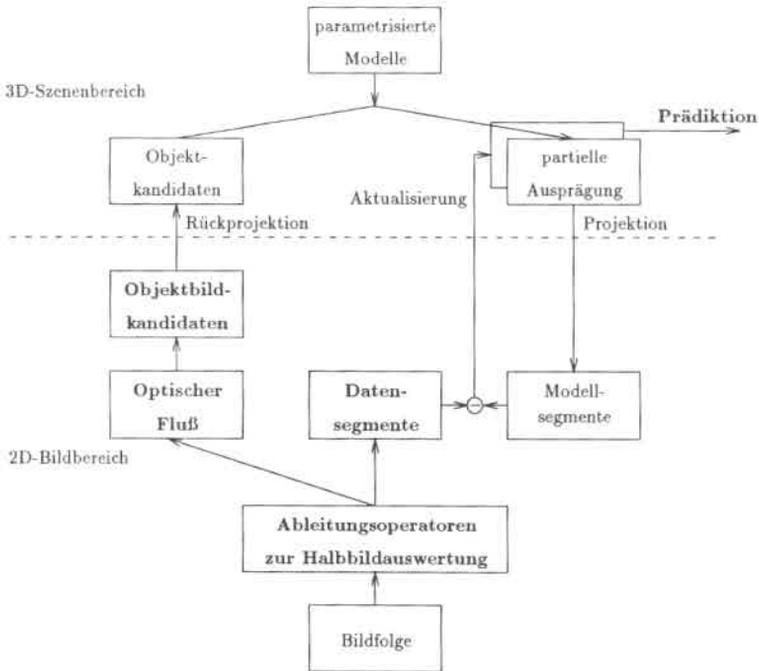


Abbildung 3.2: Vergleich der Systemkomponenten, die Fahrzeuge detektieren und verfolgen, mit den entsprechenden Komponenten des Systems von [Koller 92] aus Abb. 3.1. Die Systemverbesserungen durch Einsatz leistungsfähigerer signalnaher Komponenten sind **fett** markiert: Durch Einsatz neuer Ableitungsoperatoren [Otte 94] konnte auf die Auswertung von Halbbildern übergegangen werden. Das fleckbasierte Verschiebungsvektorfeld wurde durch ein *dichtes optisches* Flußfeld nach [Otte 94] ersetzt. Außerdem wurde das von Koller benutzte Modul zur Berechnung von Datenkantensegmenten durch ein von [Otte 94; Otte & Nagel 92] entwickeltes Verfahren ersetzt.

basierten Ansatz: dessen Toleranzen können enger gezogen werden, da sie nicht mehr den bei zu grober Initialisierung notwendigen großen Fangbereich erfassen müssen.

In vielen Experimenten mit dem System von [Koller 92] zeigte sich, daß der merkmalsbasierte Segmentierungsalgorithmus nicht immer genaue Initialisierungen liefert. Auch wurde nicht jedes Objekt detektiert. Es war daher manchmal notwendig, die initiale Lageschätzung interaktiv zu korrigieren, um ein Fahrzeug verfolgen zu können. Probleme bei der Verwendung einer ähnlichen Komponente zur Schätzung von Verschiebungsvektoren lösen [Gong & Buxton 93] durch Einsatz von mehr Wissen beim Segmentieren und 2D-Verfolgungsalgorithmus, beispielsweise über die Breite der Fahrzeuge. Wir dagegen ersetzen zum einen das spärlich besetzte Vektorfeld durch ein dichtes und setzen zum anderen im nächsten Kapitel ebenfalls für die Segmentierungskomponente mehr Wissen über die Situation in der Szene ein, allerdings über die Lage der Fahrspuren in der Szene.

Das Programmmodul zur Schätzung von Verschiebungsvektoren wurde durch ein auf Ableitungsoperatoren aufbauendes differentielles Verfahren nach [Campani & Verri 92] zur Schätzung des optischen Flusses ersetzt, das von [Otte 94] erweitert und implementiert wurde. In Abb. 3.3 wird das nach [Koller 92] berechnete Verschiebungsvektorfeld dem nach [Otte 94] berechneten optischen Flußfeld am Beispiel der Ettlinger-Tor-Szene (siehe Abb. 1.1) gegenübergestellt. Das optische Flußfeld wurde dabei auf Vektoren beschränkt, die sowohl eine bestimmte Länge – was eine feste Kamera voraussetzt – als auch eine Güteschwelle überschreiten. Als Güteschwelle wurde der kleinste Singulärwert der Koeffizientenmatrix benutzt, der angibt, wie stark die Grauwertvariation in der Umgebung des Bildpunktes ist, an dessen Stelle der Flußvektor geschätzt wird. Der in Abb. 3.4 dargestellte vergrößerte Ausschnitt von Abb. 3.3 (d) zeigt, wie gut z.B. das Fahrzeug Nr. 9 von Flußvektoren überdeckt wird.

- **Rechenaufwand:**

Das neue Verfahren benötigt im Vergleich zum alten größere Rechenkapazitäten, die früher nicht zur Verfügung standen, so daß bis vor kurzem Experimente im hier benötigten Umfang mit dem neuen Verfahren aufgrund technischer Grenzen praktisch nicht durchführbar waren. Das alte Verfahren benötigt 13 Sekunden pro Aufnahme auf einer SPARCstation 20¹. Das neue Verfahren benötigt zur Schätzung der Ableitungen der Grauwertfunktion und zur Schätzung des optischen Flusses jeweils 110 Sekunden pro Halbbild.

- **Dichtes Feld:**

Das nun vorliegende optische Flußfeld ist wesentlich *dichter* als das Verschiebungsvektorfeld (vgl. Tab. 3.1), *überdeckt* die Abbilder bewegter Objekte besser (Abb. 3.3) und wird darüber hinaus mit weniger Heuristiken berechnet. Dadurch wird es erst mathematisch analysierbar, und seine Leistungsgrenzen werden überschaubarer, so daß sich das Verfahren gezielter verbessern und robuster machen läßt (vgl. [Nagel *et al.* 94]).

¹SPARCstation 20 ist ein eingetragenes Warenzeichen der Firma Sun Microsystems, Inc.

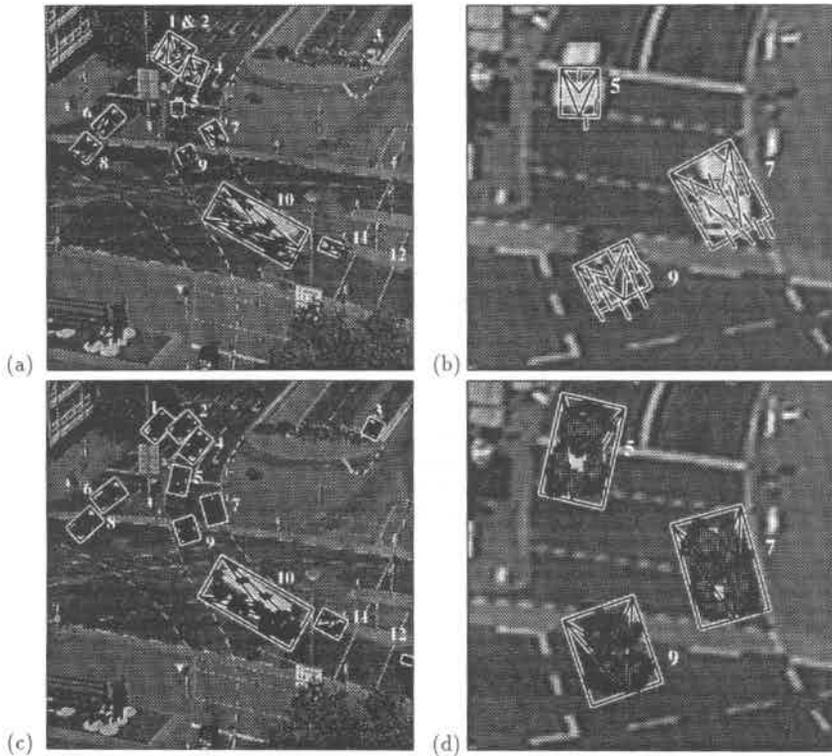


Abbildung 3.3: Verbesserungen der Initialisierungen am Beispiel der Ettlinger-Tor-Szene aus Abb. 1.1: Die erste Zeile zeigt die Ergebnisse der Schätzung von Verschiebungsvektoren nach [Koller *et al.* 91; Koller 92] und die zu einer Ballung zusammengefaßten Vektoren; jede Ballung wird durch ein umschreibendes Rechteck visualisiert. (b) zeigt einen Ausschnitt des oberen linken Quadranten von (a). Die Verschiebungsvektoren sind das Ergebnis der Zuordnungen von Fleckmerkmalen über vier aufeinanderfolgende Aufnahmen. Die zweite Zeile zeigt das Ergebnis des Ballungsverfahrens bei Verwendung von optischen Flußvektoren, die nach einem Verfahren von [Campani & Verri 92] berechnet wurden, das von [Otte 94] erweitert und implementiert wurde. Die optischen Flußvektoren geben die Verschiebung über ein Halbbild an, deshalb sind sie kürzer als die Verschiebungsvektoren. Eine Vergrößerung des unteren Fahrzeuges Nr. 9 aus (d) zeigt Abb. 3.4.

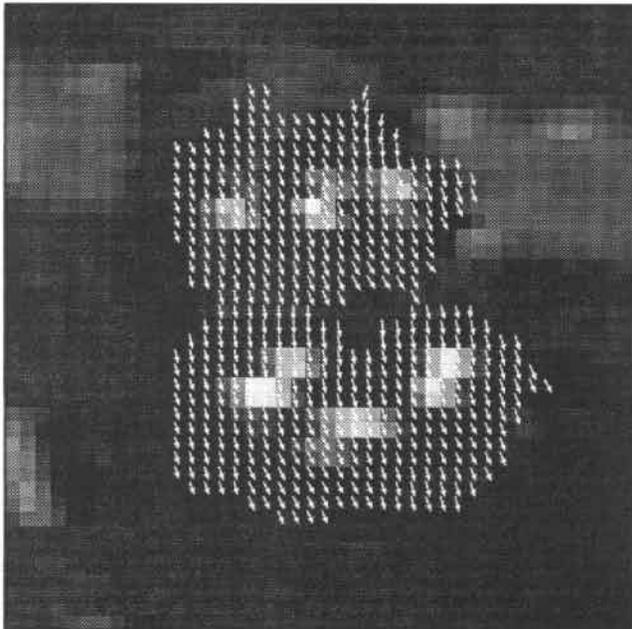


Abbildung 3.4: Optische Flußvektoren für das untere Fahrzeug Nr. 9 aus Abb. 3.3 (d). Diese Abbildung zeigt, daß auch Fahrzeugabbilder in kontrastarmen Bildbereichen detektiert werden.

- **Ballungsanalyse:**

Das jetzt zur Verfügung stehende dichtere Feld erzwang eine Überarbeitung des Verfahrens zur Ballung der Vektoren, um durch eine Segmentierung Objektbildkandidaten zu erhalten. Erste Untersuchungen findet man in [Kaufmann 93]. Einzelheiten zur – jetzt vereinfachten – Ballungsanalyse, die auf die Arbeit von [Sung 88] zurückgeht und von [Koller 92] verbessert wurde, findet man im Anhang A. Dabei werden benachbarte Vektoren mit ungefähr gleicher Länge und Orientierung zu Objektbildkandidaten zusammengefaßt.

- **Genauere Initialisierungen:**

Die Startwerte für eine Fahrzeugmodellanpassung können aus einem einzigen Flußfeld gewonnen werden. Eine über mehrere Aufnahmen durchzuführende Mittelung der Schätzwerte für die Fahrzeugpositionen zur Gewinnung eines Schätzwertes für die Geschwindigkeit entfällt: das optische Flußfeld wird nicht nur zur Schätzung initialer Werte für Position und Orientierung eines Fahrzeuges benutzt, sondern auch zur Schätzung des Betrags der Geschwindigkeit. Deshalb ist der nun vorliegende Initialisierungsschritt homogener als der von

Fahrzeug Nr.	Verschiebungsvektoren	optische Flußvektoren
1	22	583
2		550
3	—	274
4	17	574
5	5	567
6	25	623
7	19	697
8	23	707
9	14	532
10	84	3068
11	7	369
12	—	79

Tabelle 3.1: Für die in Abb. 3.3 detektierten Fahrzeuge wird die Anzahl der in die Ballungsanalyse eingehenden Verschiebungsvektoren der Anzahl optischer Flußvektoren gegenübergestellt. Das optische Flußfeld ist um ein bis zwei Größenordnungen dichter als das Verschiebungsvektorfeld.

[Koller 92], der das Verschiebungsvektorfeld nur zur Trennung bewegter Bereiche vom statischen Hintergrund benutzen konnte.

3.4 Datensegmentextraktion

[Koller 92] benutzte Kantensegmente, die mit einer linearen Regression aus solchen Kantenelementen gewonnen wurden, die als lokale Maxima des Grauwertgradientenbetrags in Gradientenrichtung einen festen Schwellwert überschritten. In Bildbereichen mit geringem Kontrast kann diese Schwellwertoperation nicht nur Rauschen, sondern auch Kantenelemente unterdrücken, die Teil einer wichtigen Bildstruktur sind. Dadurch können Kantensegmente aufbrechen oder ganz verlorengehen. Im Gegensatz zu herkömmlichen bildpunktorientierten Gradientenbetragschwellen schlagen [Otte 94; Otte & Nagel 92] vor, Kantenelemente ohne Anwendung eines Schwellwertoperators zu Kantenelementketten und Vertizes zu verketten. Die Auswertung von Eigenschaften einer Kette, wie durchschnittlicher Gradientenbetrag, Kettenlänge und zweite Momente der Gradientenrichtungsänderung, werden genutzt, um entweder eine Kette als Rauschen zu verwerfen oder als detektierte

Bildstruktur aufzufassen. Kantenelementketten umfassen globalere Bildinformation als die Bildinformation, die ein einzelnes Kantenelement hat. Daher wurde die Berechnung von Kantensegmenten verbessert.

Anstatt die Unsicherheiten der Datensegmente interaktiv vorzugeben (wie zum Beispiel [Tan *et al.* 93; Deriche & Faugeras 90; Koller 92]), schätzen wir die Unsicherheit σ_{\perp} senkrecht zum Datensegment aus den Bilddaten und verringern damit die Menge der freien Parameter. Mit der Mittelpunkt-Repräsentation für Kantensegmente wie in [Deriche & Faugeras 90] berechnen wir den kleineren Eigenwert beim Eigenvektor-Anpaßprozeß eines Kantensegmentes an eine Menge von Kantenelementen [Duda & Hart 73; Pavlidis & Horowitz 74], um σ_{\perp} zu schätzen (vgl. Anhang C.1).

3.5 Ergebnisse der Trajektorienberechnung

Das folgende Experiment wurde mit einer Bildfolge durchgeführt, die am Ettlinger-Tor-Platz in Karlsruhe aufgezeichnet wurde. Die Größe der Abbilder der Fahrzeuge schwankt von 25×30 bis 30×35 (mit der Ausnahme des Busses: 110×110) Bildpunkten in einer Aufnahme. Die Abbilder der kleinsten Fahrzeuge in diese Bildfolge sind kleiner als die 20×40 Bildpunkte umfassenden Abbilder der Fahrzeuge in den Aufnahmen des Durlacher-Tor-Platzes aus [Koller 92].

Die mit dem neuen Verfahren ermittelten Objektbildkandidaten lassen sich mit Kenntnis der internen sowie externen Kameraparameter in die Szene zurückprojizieren und dienen dann zur Initialisierung einer modellgestützten Objektverfolgung. Die verbesserten Initialisierungen erlauben um Faktor 100 verringerte Abstimmungsparameter für den Kalmanfilter.

Die gerade beschriebenen Änderungen führten dazu, daß erstmals mehrere Fahrzeuge (Nr. 2, 4 bis 8) mit einem einheitlichen Parametersatz verfolgt werden konnten, obgleich noch viele Zufälligkeiten im Spiel waren. Die Ergebnisse der Trajektorienberechnung finden sich in Abb. 3.5 sowie in Abb. 3.6. Fahrzeug Nr. 1 konnte nur über 40 Halbbilder verfolgt werden, da es teilweise durch ein Straßenschild verdeckt wird. Für Fahrzeug Nr. 3, das gerade aus einer Unterführung herausfährt, stimmen die Kalibrierdaten nicht, in deren Ermittlung die Annahme einging, die gesamte Fahrbahn läge in einer Ebene. Fahrzeug Nr. 11 ist problematisch, da der detektierte bewegte Bildbereich aufgrund des Laternenpostens das Fahrzeug nicht ganz überdeckt. Beim Bus waren wir gezwungen, den Schwellwert für die maximale Mahalanobisdistanz zwischen Modell- und Datenkantensegment zu erhöhen, da Radkästen und Türen nicht im Modell aufgenommen sind und deshalb nur verhältnismäßig lange Kantensegmente erwartet werden (vgl. Abb. 3.7).

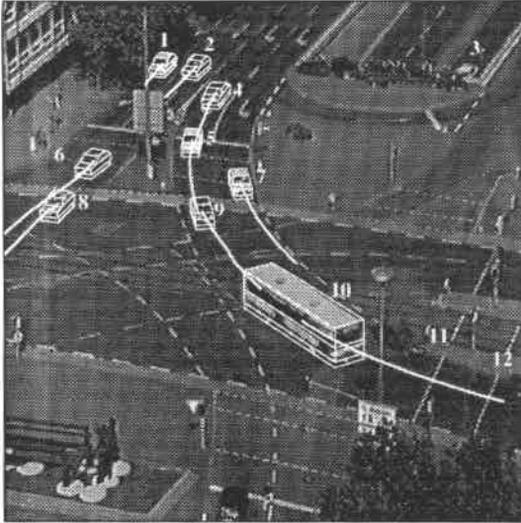


Abbildung 3.5: Die geschätzten Trajektoriendaten für die Ettlinger-Tor-Bildfolge.

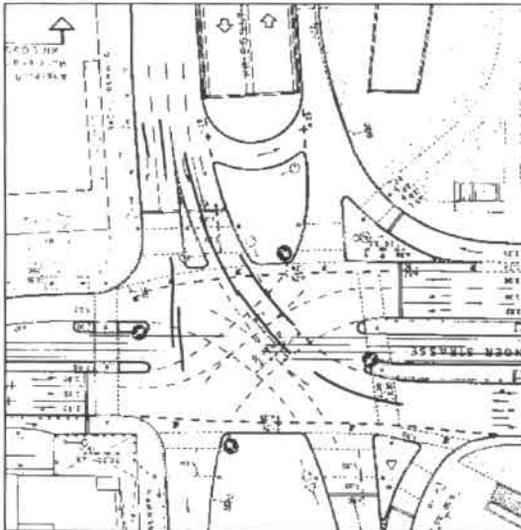


Abbildung 3.6: Eine Projektion der geschätzten Trajektoriendaten aus Abb. 3.5 in die Fahrbahnebene, wobei eine digitalisierte Karte für diese Straßenkreuzung unterlegt wurde.

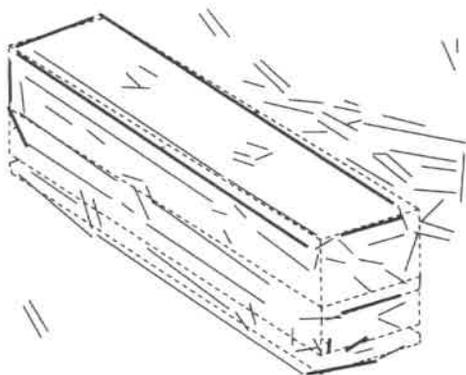


Abbildung 3.7: Ergebnis einer Anpassung eines einfachen Busmodells an Datenkantensegmente: Modellkantensegmente sind gestrichelt, zugeordnete Datenkantensegmente fett eingezeichnet. Radkästen und Türen sind nicht im Modell aufgenommen, stattdessen wurde z.B. die Busunterkante als ununterbrochene lange Kante modelliert, für die dann kein Zuordnungskandidat gefunden werden konnte. Desweiteren mußten an der Busfront Modellkantensegmente von der Zuordnung ausgeschlossen werden, da das Modell den Bus an dieser Stelle nicht richtig erfaßt. Nur so konnte der Bus überhaupt verfolgt werden.

Kapitel 4

Kalibrierung mit Hilfe von Fluchtpunkten

4.1 Übergang vom Bild- in den Szenenbereich

In Innenstadtszenen gibt es auf der Fahrbahn, an Häusern und Fenstern, an Masten der Straßenbeleuchtung und Verkehrsschilder viele Scharen paralleler Geraden. Die Bilder von im Raum paralleler, jedoch nicht notwendigerweise komplanarer Geraden schneiden sich bei perspektivischer Projektion in der Bildebene. Dieser gemeinsame, eventuell im Unendlichen liegende Schnittpunkt heißt *Fluchtpunkt*. In diesem Kapitel wird untersucht, wie im Bild detektierte Kantensegmente und Fluchtpunkte die Rekonstruktion von 3D-Informationen aus einer einzelnen Aufnahme erlauben.

Annahmen über den Straßenverlauf ermöglichen mit Hilfe von im Bild detektierten Kantensegmenten über eine Berechnung von Fluchtpunkten eine initiale Kalibrierung der Kamera, ohne daß ein Markierungs- oder Katasterplan vorhanden zu sein braucht. Die Kalibrierung schafft die Voraussetzung, ein Fahrbahnmodell an ein Bild anzupassen. Die berechnete Ausprägung des Fahrbahnmodells wird zur Korrektur der Kalibrierungsparameter verwendet.

Auswertung der Abbilder von Ampelmasten können die Kalibrierung deutlich verbessern. Desweiteren wird der Schatten, den die Masten bei Sonnenlicht auf die Fahrbahn werfen, zur Berechnung der Einfallsrichtung des Lichtes herangezogen.

4.2 Interne und externe Kameraparameter

Im folgenden Abschnitt bildet die Fahrbahnebene die XY -Ebene des Weltkoordinatensystems (WKS), die Z -Achse des WKS zeigt nach oben (siehe Abb. 4.1), wobei Punkte und Koordinaten in der 3D-Szene mit großen, im 2D-Bildbereich dagegen mit kleinen Buchstaben bezeichnet werden. Das Kamerakoordinatensystem (KKS) ist mit dem Ursprung im *Augenpunkt* oder *optischen Zentrum* so angeordnet, daß die Z -Achse in Richtung der optischen Achse zeigt. Die Bildebene verläuft im Ab-

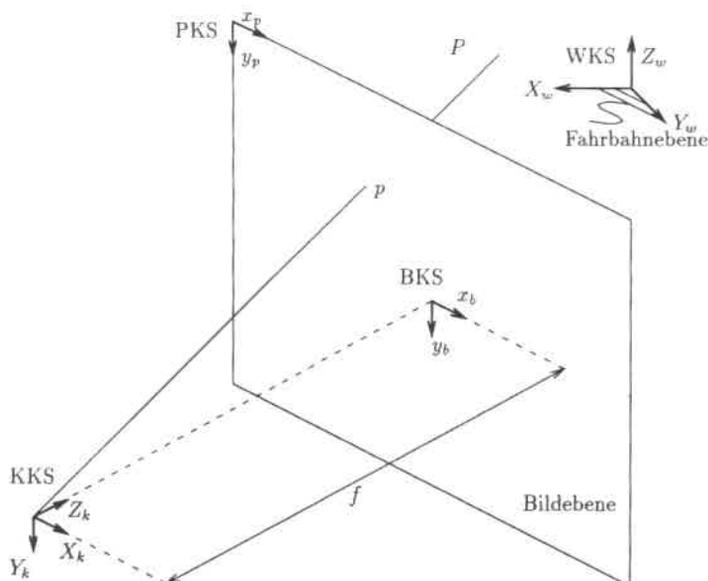


Abbildung 4.1: Die Geometrie der Abbildung eines Szenenpunktes bei perspektivischer Projektion nach [Krein 90]. Im Gegensatz zu [Koller 92] haben hier die positiven x -Achsen bzw. y -Achsen im Kamerakoordinatensystem (KKS) und Bildkoordinatensystem (BKS) die gleiche Orientierung. Weltkoordinatensystem (WKS) und Pixelkoordinatensystem (PKS) werden im Text erklärt.

stand f , der *Brennweite*, vom optischen Zentrum senkrecht zur optischen Achse. Der Schnittpunkt der optischen Achse mit der Bildebene heißt *Hauptpunkt*.

Für die Abbildung eines Punktes $(X_k, Y_k, Z_k)^T$ im KKS erhält man die Bildkoordinaten $(x_b, y_b)^T$ aus:

$$\begin{aligned} x_b &= f \frac{X_k}{Z_k}, \\ y_b &= f \frac{Y_k}{Z_k} \end{aligned} \quad (4.1)$$

bzw. seine Pixelkoordinaten $(x_p, y_p)^T$ aus:

$$\begin{aligned} x_p &= \frac{x_b}{s_x} + x_0, \\ y_p &= \frac{y_b}{s_y} + y_0. \end{aligned} \quad (4.2)$$

Der Hauptpunkt $(x_0, y_0)^T$, die Brennweite f , den horizontalen und vertikalen Skalierungsfaktor s_x und s_y bzw. das Pixelseitenverhältnis $s = s_x : s_y$ nennt man die *internen Kameraparameter*.

[Penna 91] bestimmt die internen Kameraparameter aus dem Abbild eines Kreises, [Lenz & Tsai 88] schlagen zur Berechnung von horizontalem Skalierungsfaktor und Hauptpunkt drei Methoden vor: eine direkte optische Methode mittels Laserstrahl, einen auf Variation der Brennweite basierenden Ansatz und ein auf Kalibrierplatte und Kalibrierpunkten aufbauendes Verfahren (siehe auch [Tsai 87; Lenz 87]). Interne Kameraparameter lassen sich — wie wir später noch diskutieren werden — auch mittels im Bild detektierter Fluchtpunkte bestimmen [Caprile & Torre 90; Kanatani 91; Beardsley & Murray 92; Lavest *et al.* 93].

Die Koordinatentransformation von Welt- nach Kamerakoordinaten wird durch eine Rotationsmatrix \mathcal{R} und einen Translationsvektor $\mathbf{T} = (T_x, T_y, T_z)^T$ — die *externen Kameraparameter* — beschrieben:

$$\begin{pmatrix} X_k \\ Y_k \\ Z_k \end{pmatrix} = \mathcal{R} \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix}. \quad (4.3)$$

4.3 Übersicht über die zu schätzenden Parameter

Von den *internen Kameraparametern* schätzen wir nur die *Brennweite* f , den Hauptpunkt nehmen wir in der Bildmitte an, das Pixelseitenverhältnis zu Eins¹. Linsenverzerrungen werden vernachlässigt.

Die Orientierung der Kamera relativ zur Fahrbahnebene wird durch den Normalenvektor der Fahrbahnebene im KKS festgelegt, d.h. durch *zwei Rotationsparameter*. Ein dritter Rotationsparameter legt die Orientierung der Straßenzüge in der Fahrbahnebene fest. Der Normalenvektor der Fahrbahnebene erlaubt — zusammen mit der Brennweite — eine Rückprojektion von Geradenorientierungen (vgl. Lemma B.46) sowie von Geradenschnittwinkeln (vgl. Lemma B.48) aus der Bild- in die Fahrbahnebene, da es dazu nicht von Bedeutung ist, wie das WKS um den Normalenvektor der Fahrbahnebene gedreht wird.

Da wir den einzelnen Bildpunkten keine festen Weltkoordinaten zuordnen, können wir das Weltkoordinatensystem so wählen, daß sein Ursprung der Durchstoßpunkt der optischen Achse mit der Fahrbahnebene ist. Hiermit reduzieren wir die Zahl der Translationsparameter von drei auf einen: den Abstand T_z des optischen Zentrums vom Durchstoßpunkt. Damit ist $\mathbf{T} = (0, 0, T_z)^T$ der *Translationsparameter*. T_z ist gleichzeitig der *Skalierungsfaktor* bzw. die *Tiefe*.

¹[Koller 92] gab bei dem von ihm eingesetzten Verfahren von [Tsai 87; Lenz 87] auch Hauptpunkt und Pixelseitenverhältnis vor.

Als weitere zu schätzende Parameter ergeben sich *Azimutwinkel* ϕ_s und *Polarwinkel* θ_s (im WKS) der Einfallrichtung der Lichtes. Wir fassen die zu schätzenden Parameter zusammen:

- Brennweite f ,
- zwei Rotationsparameter,
- ein Translationsparameter: Tiefe T_z ,
- Azimutwinkel ϕ_s und Polarwinkel θ_s der Einfallrichtung der Lichtes.

4.4 Verfahren zur Fluchtpunktberechnung

Ein geeignetes mathematisches Modell für den durch eine Kameraaufnahme mittels *perspektivischer* Projektion entstandenen zweidimensionalen Bildbereich ist die *projektive Ebene*. Ein Vorteil des projektiven Modells besteht darin, daß man nicht zwischen im Endlichen und im Unendlichen liegenden Fluchtpunkten unterscheiden muß. Eine kurze Einführung in die zweidimensionale projektive Geometrie findet man im Anhang B.1. Dort werden *homogene* Koordinaten und *N-Vektoren* von Punkten und Geraden sowie der Übergang von einer Bildebene zur *Bildsphäre* wie üblich definiert.

Fluchtpunkte werden in Anhang B.2 eingeführt. Im Anhang B.3 diskutieren wir, wie man aus im Bild detektierten Kantensegmenten Fluchtpunktkoordinaten schätzen kann. Anhang B.4 enthält eine Erörterung der Frage, wie Fluchtpunkte und Kalibrierungsparameter zusammenhängen. Tab. B.1 faßt verschiedene Möglichkeiten zur Berechnung von Kalibrierungsparametern aus Fluchtpunkten zusammen.

4.4.1 Verfahren zur Fluchtpunkt detektion

Die Berechnung von Fluchtpunkten aus Datenkantensegmenten gliedert sich nach [Collins & Weiss 90] in zwei Teilaufgaben: zunächst sind diejenigen Teilmengen von Datenkantensegmenten zu bestimmen, die sich jeweils in einem Fluchtpunkt schneiden, anschließend muß für jede Teilmenge der Fluchtpunkt bestimmt werden. Um die erste Aufgabe zu lösen, nutzen [Barnard 83; Magee & Aggarwal 84; Quan & Mohr 89; Li-Qun *et al.* 92; Schuster *et al.* 93] eine Hough-Transformation, mittels der zu Geraden verlängerte Kantensegmente vom Bildbereich auf Großkreise einer Einheitskugel (Gaußsche Sphäre) projiziert werden, deren Mittelpunkt sich im optischen Zentrum befindet, indem die Projektionsebene jeder Geraden (Def. B.8) mit der Einheitskugel zum Schnitt gebracht wird. Die Gaußsche Sphäre ist eine beschränkte Menge, daher rechnerintern, beispielsweise als ein zweidimensionales Feld mit Azimut- und Polarwinkeln als Parameter, besser darstellbar als eine unbeschränkte Ebene [Barnard 83]. In unseren Untersuchungen (vgl. Abb. 4.2) zeigte

es sich, daß solche Zählverfahren zu spröde sind, wenn man sie auf Kantensegmente anwendet, die aus unseren komplexen Innenstadtszenen ermittelt wurden. Diese Zählverfahren können gut funktionieren, wenn statt Innenstadtszenen die Aufnahme einer Kalibrierplatte mit regelmäßigen Mustern verwendet wird. Hinzu kommt, daß die Wahl eines Schätzwertes für die in unserem Fall unbekannte Brennweite die Ergebnisse empfindlich beeinflußt.

Als Ergebnis weiterer Messungen [Lebègue & Aggarwal 93] können auch Schätzwerte für die Kameraorientierung relativ zur Welt ermittelt werden. Dieses Wissen kann genutzt werden, um mögliche Positionen der Fluchtpunkte einzuschränken [Liou & Jain 87; Schuster *et al.* 93] (vgl. Satz B.14). Im klassischen Beispiel mit der Kamera im Fahrzeug (siehe Abb. 4.3) liegt ein Hauptfluchtpunkt (Def. B.24) in der Bildmitte, die anderen beiden im Unendlichen [Shigang *et al.* 90; Straforini *et al.* 93]. Auf der Basis eines bekannten Neig- und Kippwinkels der Kamera kombinieren auch [Liou & Jain 87] die Berechnung von Kantensegmenten der rechten und linken Fahrbahnbegrenzung und die Detektion des zugehörigen Fluchtpunktes.

Bei [Shigang *et al.* 90] ist die Kamera so montiert, daß nur ihr Schwenkwinkel als Freiheitsgrad bleibt. Sie berechnen den verbleibenden Rotationsparameter nach einer Klassifizierung der Kantensegmente in horizontale — d.h. in parallel zur Fahrbahnebene liegende — und in nicht horizontale Segmente. Dabei nutzen sie aus, daß bei einer nicht rein rotatorischen Bewegung der Kamera ein Fluchtpunkt eines horizontalen Kantensegmentes eine andere Verschiebungsrate hat als der Schnittpunkt eines nicht horizontalen Kantensegments mit dem Horizont.

[Straforini *et al.* 93] nehmen an, der Neigwinkel der Kamera verschwinde. Sie transformieren die Bildgeraden in Zylinderkoordinaten in einen sog. Polar-Raum. Eine Schar von Geraden schneidet sich genau dann in einem Punkt, wenn die in den Polar-Raum transformierten Bilder der Geraden auf einer Sinuskurve liegen. Damit vereinfacht sich der Gruppierungsprozeß von Geraden. Bei dem von ihnen vorgestellten Verfahren zur Fluchtpunktdetektion handelt es sich um ein Invertieren des Hough-Transformations-Verfahrens aus [Duda & Hart 73] zur Detektion von Kantensegmenten aus Kantenelementen.

Die Ausnutzung von Orientierungs- und Bewegungsinformation zur Gruppierung der Kantensegmente nach Hauptfluchtpunkten bietet sich geradezu an: Im Gegensatz zu den Verfahren von [Straforini *et al.* 93; Schuster *et al.* 93], bei denen der Neigwinkel bzw. der Kippwinkel der Kamera verschwindet, und im Gegensatz zu [Shigang *et al.* 90; Liou & Jain 87], in deren Ansätze sogar Neig- und Kippwinkel der Kamera bekannt sein müssen und nur der Schwenkwinkel als Rotationsfreiheitsgrad bleibt, lassen wir eine allgemeine Kameraorientierung zu: insbesondere ergeben sich damit im Gegensatz zu den zitierten Arbeiten gerade Neig- und Kippwinkel als Freiheitsgrade.

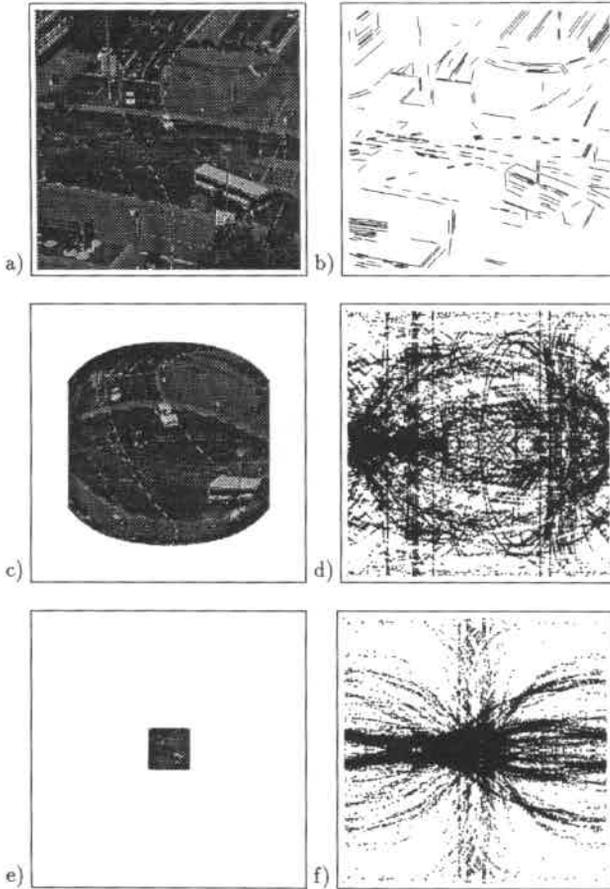


Abbildung 4.2: (a) Aufnahme des Ettliger-Tor-Platzes (vgl. Abb. 1.1); (b) detektierte Kantensegmente, die länger als 15 Bildpunkte sind. Mit quadratischem Aufwand lassen sich die homogenen Koordinatenvektoren der Schnittpunkte von je zwei Bildgeraden als Kreuzprodukt der Normalenvektoren der Projektionsebenen der Bildgeraden berechnen [Magee & Aggarwal 84; Schuster *et al.* 93]. Dargestellt sind (c) das auf die Einheits(halb)kugel im Abstand $f = 100$ von der Bildebene projizierte Bild und (d) die auf die Einheits(halb)kugel projizierten homogenen Koordinatenvektoren der Schnittpunkte der zu Geraden verlängerten Kantensegmente aus (b), wobei die Oberfläche der Einheits(halb)kugel durch Azimut- (Abszisse) und Polarwinkel (Ordinate) parametrisiert wurde. (e) und (f) zeigen das Ergebnis bei $f = 1000$. Häufungspunkte sind nicht genau zu erkennen, doch die Richtung, in der ein Fluchtpunkt zu suchen ist, wird in (f) deutlich. Die Bilder (d) und (f) verdeutlichen, daß dieses Verfahren stark von der i.a. a-priori nicht bekannten Brennweite f abhängt.



Abbildung 4.3: Ein Hauptfluchtpunkt liegt in der Nähe der Bildmitte, die vertikalen und horizontalen Bildkantensegmente verlaufen fast parallel. (Dieses Bild wurde von Dr. W. Enkelmann, Fraunhofer-Institut für Informations- und Datenverarbeitung (IITB), Karlsruhe, aufgenommen.)

4.4.2 Schätzung der Koordinaten von Fluchtpunkten und Kalibrierung

Laut [Collins & Weiss 90] lassen sich Histogramm-Ansätze zur Zusammenfassung von Geraden heranziehen, die sich in einem Fluchtpunkt schneiden. Die genaue Lokalisierung des Fluchtpunktes soll jedoch auf den Kantensegmenten selbst und nicht auf der Histogramm-Datenstruktur beruhen, die in Zellen willkürlicher Breite eingeteilt ist. [Collins & Weiss 90] gehen in ihrem Ansatz davon aus, die Teilmengen von Kantensegmenten, deren Verlängerungen zu Geraden sich in einem Fluchtpunkt schneiden, seien bereits gefunden. Sie fassen die N -Vektoren (siehe Def. B.13) der zu Geraden verlängerten Kantensegmente als Stichproben einer Äquator-Verteilung auf der Einheitskugel auf, wobei der Kameraöffnungswinkel bekannt sein muß. Daraus bestimmen sie einen Schätzwert und ein Konfidenzintervall für die 3D-Orientierung der Geraden, indem sie die Polarachse (vgl. Korollar B.12, Lemma B.25) schätzen. Die Autoren zeigen, daß ihr Ansatz das gleiche Ergebnis liefert wie bei Anwendung der Methode der kleinsten Quadrate (siehe Anhang B.3).

[Caprile & Torre 90] stellen eine Methode zur Kalibrierung von zwei oder mehr Kameras vor: In einem ersten Schritt wird das Pixelseitenverhältnis aus dem Abbild eines quadratischen Kalibrierungsmusters berechnet. In einem zweiten Schritt werden für jede Kamera Brennweite (nach Satz B.33) und Hauptpunkt durch das Abbild eines Würfels bestimmt. Zur Bestimmung der Rotationsparameter zwischen den Kameras werden zwei komplanare Fluchtpunkte (Lemma B.36), die aus Aufnahmen einer ebenen Kalibrierplatte berechnet werden, in beiden Aufnahmen zugeordnet. Die Translationsparameter werden aus einer Triangulation bestimmt, wobei ein Bild von fünf parallelen Kantensegmenten verwendet wird, deren Länge und deren gegenseitigen Abstand man kennt.

[Echigo 90] ermittelt die Rotationsparameter nach Korollar B.45 aus den Abbildern dreier Scharen paralleler Geraden auf einer mit einem gitterförmigen Muster versehenen Pyramide. Er berechnet für jede der drei Geradenscharen mit Hilfe der Methode der kleinsten Quadrate die Gerade durch den Haupt- und Fluchtpunkt. Die Lage dieser Gerade ist unabhängig von der Brennweite (Satz B.39). Bei einer Änderung der Brennweite wandert der Fluchtpunkt auf dieser Geraden (vgl. Abb. B.1). Die Translationsparameter und die Brennweite erhält er, wenn die Weltkoordinaten von mindestens zwei Punkten bekannt sind. Das Pixelseitenverhältnis berechnet er aus den Datenblättern von Kamera und von A/D-Wandler. Der Hauptpunkt wird mit einem Laserstrahl ermittelt.

Ausgehend vom Abbild eines Würfels berechnen [Wang & Tsai 90] den Hauptpunkt aus den drei Hauptfluchtpunkten im Bildbereich. Sie zeigen, daß die drei Rotationsparameter nur von den Steigungen der drei Fluchtlinien durch die Hauptpunkte im Bild abhängen und aus ihnen zu berechnen sind. Die Brennweite ergibt sich aus den Rotationsparametern und dem Flächeninhalt des Dreiecks der drei Hauptfluchtpunkte. Mit Hilfe von bekannten Koordinaten zweier Punkte berechnen sie die Translationsparameter. Ihr Verfahren setzt voraus, daß man von allen drei Hauptfluchtpunkten die inhomogenen Bildkoordinaten bestimmen kann, was nur bei eingeschränkter Kameraorientierung der Fall ist.

Eine Reihe von Arbeiten behandelt die Fragestellung, wie aus dem Horizont (vgl. Anhang B.4.2) 3D-Information zu schätzen ist: In manchen Aufnahmen ist es möglich, den Horizont selbst als Kantensegment im Bild zu detektieren. [Wang & Tsai 91; Weiss *et al.* 90] schlagen vor, eine quadratische Anpassung von mehr als zwei Fluchtpunkten an den Horizont im Bildbereich zu bestimmen, obwohl eine Anpassung mittels homogener Koordinaten auf der Einheitskugel robustere Ergebnisse brächte, wenn der Horizont weit außerhalb des sichtbaren Bereiches läge. Dasselbe gilt für Verfahren, wie z. B. von [Wang & Tsai 91; Wang & Tsai 90; Weiss *et al.* 90], die auf den inhomogenen Koordinaten des Horizontes aufbauen. [Weiss *et al.* 90] schätzen den Normalenvektor einer Ebene und seinen Fehler zum einen mit Hilfe von inhomogenen Koordinaten des Horizontes und zum zweiten mittels homogener Koordinaten zweier Fluchtpunkte, wobei im zweiten Ansatz vorausgesetzt wird, daß der Fehler der homogenen Koordinaten eines Fluchtpunktes auf der Einheitskugel isotrop ist.

[Wang & Tsai 91] stellen ein auf der Berechnung des Horizontes beruhendes Verfahren zur Kalibrierung vor. Ausgehend vom Abbild eines Sechsecks — allgemeiner ist jede ebene Kalibrierplatte zulässig, die drei Scharen paralleler Geraden hat, d.h. auch ein Schachbrettmuster, bei dem man die Diagonalen als dritte Schar mit einbeziehen kann — wird der Horizont durch Mittelung der Koordinaten der drei Fluchtpunkte berechnet. Mit Hilfe der Weltkoordinaten zweier Punkte werden die Translationsparameter berechnet. Das Verfahren ist anwendbar, wenn die Kamera im Fahrzeug montiert ist, nicht aber, wenn die Kamera auf einem Haus steht und nach unten geneigt ist, da dann die Berechnung des Horizontes zu ungenau ist.

[Beardsley *et al.* 92] zeigen, wie die bei einer rotierenden Kalibrierplatte sich auf einer Quadrik bewegenden Fluchtpunkte mit der Fluchtlinie als ihre Einhüllende zur Berechnung von Pixelseitenverhältnis, Hauptpunkt und Brennweite einer stati-

onären Kamera herangezogen werden können. [Beardsley & Murray 92] bestimmen Hauptpunkt, Brennweite und Pixelseitenverhältnis mittels Fluchtpunktmethode. Sie benötigen eine Kalibrierplatte und die Zuordnung von vier Punkten der Platte zu den korrespondierenden Punkten in zwei bzw. — wenn auch das Pixelseitenverhältnis zu schätzen ist — in drei Bildaufnahmen. [Lavest *et al.* 93] schätzen aus einer durch Variation der Brennweite entstandenen Bildfolge den Hauptpunkt.

[Kanatani & Onodera 91; Kanatani 92] schätzen zunächst die Brennweite nach Lemma B.34 aus dem Abbild einer ebenen Kalibrierplatte, auf der ein großes Quadrat aus vier kleineren Quadraten zusammengesetzt ist. Darüber hinaus zeigen sie, wie man aus mehreren Schätzwerten für die Brennweite aus verschiedenen Aufnahmen, die man durch Variation der Orientierung der Kalibrierplatte relativ zur Kamera erhält, einen optimalen Schätzwert für die Brennweite ermitteln kann. [Kanatani & Onodera 91] erhalten die Rotationsparameter aus den Fluchtpunkten der Begrenzungslinien der Quadrate mit Lemma B.36. Die Rotationsparameter werden unter Berücksichtigung der zu den Diagonalen der Quadrate gehörenden Fluchtpunkte noch verbessert. Die Translationsparameter ergeben sich bei Vorgabe der Weltkoordinaten eines Punktes.

4.4.3 Diskussion der Literatur

Wir vergleichen unser Verfahren mit den zitierten Arbeiten.

Kalibrierobjekte und exakte Größen: Viele Autoren (beispielsweise [Echigo 90; Caprile & Torre 90; Chen & Jiang 91; Kanatani & Onodera 91; Kanatani 92; Beardsley & Murray 92]) verwenden zur Kamerakalibrierung eine Kalibrierplatte oder einen Kalibrierkörper mit bekannten, regelmäßigen Mustern und bekannter Größe. Darüber hinaus nutzen [Caprile & Torre 90; Kanatani & Onodera 91; Kanatani 92] die Möglichkeit, den Kalibrierkörper so zu positionieren, daß die Kameraparameter gut zu schätzen sind, wohingegen bei uns keine derartige Positionierung möglich ist. Mit der Ausnahme von [Collins & Weiss 90; Weiss *et al.* 90] findet man Bilder von Innenstadtszenen nur bei Arbeiten zur Detektion von Fluchtpunkten, beispielsweise bei [Barnard 83; Magee & Aggarwal 84; Quan & Mohr 89; Straforini *et al.* 93; Tai *et al.* 93; Palmer & Tai 93]. Dagegen werten Arbeiten, die sich mit Fragen der Kamerakalibrierung oder Orientierungsberechnung beschäftigen, meist nur Bilder von Kalibrierobjekten aus, was das Problem der Fluchtpunkt detektion wesentlich vereinfacht. In den uns vorliegenden Bildfolgen von Straßenverkehrsszenen wurden während der Bildaufnahme weder Kalibrierkörper noch Marken, wie etwa bei [Worrall *et al.* 93], positioniert. Erschwerend kommt hinzu, daß die Größe der in diesen Videoaufnahmen erfaßten Objekte uns nicht genau vermessen vorliegt. Auch schneiden sich Straßenzüge nicht unbedingt senkrecht an einer Kreuzung.

Keine weiteren Messungen oder Datenblätter: In unserem Fall ist der Kameratyp unbekannt, so daß Datenblätter keine Hinweise auf interne Kameraparameter

geben können. Es liegen keine Messungen von Kameraparametern vor, auch nicht der externen, wie bei [Lebègue & Aggarwal 93]. Im Gegensatz zu z.B. [Haralick 89; Liu *et al.* 90; Schuster *et al.* 93] können wir die Brennweite nicht als bekannt voraussetzen.

Kein Ausschluß uneigentlicher Fluchtpunkte: Es ist bei den uns vorliegenden Bildaufnahmen der Normalfall, daß zumindest ein, wenn nicht sogar zwei der drei Hauptfluchtpunkte sehr weit außerhalb des sichtbaren Bildbereiches liegen, d.h. oft verläuft eine Koordinatenachse des Weltkoordinatensystems fast parallel zur Bildebene. Die Verfahren von z.B. [Duda & Hart 73; Haralick 80; Caprile & Torre 90; Wang & Tsai 90; Wang & Tsai 91; Chen & Jiang 91; Tai *et al.* 93; Straforini *et al.* 93; Schuster *et al.* 93] schließen diesen Fall aus, da bei ihren Verfahren inhomogene Bildkoordinaten verwendet werden. Diese Ansätze scheiden daher für uns aus. Im Gegensatz zu diesen Arbeiten liefert unser Verfahren, wie auch die Verfahren von [Collins & Weiss 90; Kanatani & Onodera 91; Kanatani 92], auch im Fall, daß ein Fluchtpunkt im Unendlichen liegt, gute Schätzergebnisse für die Kameraparameter.

Entkopplung der Parameter:

- *Rotationsparameter und Brennweite:* [Haralick 80; Haralick 89; Batista *et al.* 93; Caprile & Torre 90; Collins & Weiss 90; Wang & Tsai 91; Kanatani & Onodera 91; Schuster *et al.* 93] benötigen einen Schätzwert für die Brennweite zur Bestimmung von Orientierungs- oder Rotationsparametern. [Echigo 90; Wang & Tsai 90] haben bereits gezeigt, daß die Rotationsparameter der Kamera auch ohne Kenntnis der Brennweite aus im Bild detektierten Kantensegmenten zu schätzen sind. Die Rotationsparameter der Kamera sind genauer aus den Bilddaten zu schätzen als die Brennweite. Daher sollte man vermeiden, durch einen schlechten Schätzwert für die Brennweite die Schätzung der Rotationsparameter zu verfälschen.
- *Rotationsparameter und Translationsparameter:* Mit der Ausnahme von [Lavest *et al.* 93] erfolgt bei auf Kantensegmenten aufbauenden Verfahren zur Kalibrierung die Berechnung der Rotationsparameter unabhängig von den Translationsparametern. Bei auf Punktmerkmalen aufbauenden Verfahren zur Kalibrierung von z.B. [Tsai 87; Lenz 87] werden die Rotationsparameter gleichzeitig mit den Translationsparametern T_x, T_y geschätzt.
- *Brennweite und Translationsparameter:* Im Gegensatz zu z.B. [Echigo 90], bei dem wie auch bei [Tsai 87; Lenz 87] die Brennweite der Kamera und der Translationsparameter T_x in einem gemeinsamen Schätzprozeß ermittelt werden, berechnen wir die Brennweite wie auch schon [Wang & Tsai 91; Caprile & Torre 90; Kanatani & Onodera 91; Kanatani 92] unabhängig von dem Translationsparameter.

In unserem Ansatz erfolgt die Schätzung der Kameraparameter unabhängiger voneinander als in allen anderen zitierten Verfahren: die Berechnung der Rotationspara-

meter erfolgt unabhängig von dem Translationsparameter und der Brennweite, die Berechnung der Brennweite unabhängig von dem Translationsparameter.

Brennweite aus den Bildkoordinaten eines einzigen Fluchtpunktes: Im Gegensatz zu [Duda & Hart 73; Haralick 80; Caprile & Torre 90; Kanatani & Onodera 91; Kanatani 92; Collins & Beveridge 93], die zur Berechnung der Brennweite die inhomogenen Koordinaten zweier Fluchtpunkte heranziehen, und im Gegensatz zu [Wang & Tsai 90] und auch zu [Caprile & Torre 90], für deren Verfahren sogar die inhomogenen Koordinaten dreier Fluchtpunkte benötigt werden, reichen uns die inhomogenen Koordinaten von nur *einem* Fluchtpunkt. Da nicht zugleich alle drei Hauptfluchtpunkte im Unendlichen liegen können, sind die inhomogenen Bildkoordinaten eines Hauptfluchtpunktes immer zu schätzen.

3D-Körper statt 2D-Ebene: [Wang & Tsai 91; Kanatani & Onodera 91; Kanatani 92; Chen & Jiang 91; Collins & Beveridge 93], teilweise auch [Caprile & Torre 90] benutzen nur die Aufnahme eines ebenen Körpers. Wie auch [Caprile & Torre 90; Wang & Tsai 90; Echigo 90] verwenden wir Abbilder von 3D-Körpern, indem wir nicht nur Fahrbahnkantensegmente, sondern auch die im Bild erfaßten Ampelmasten auswerten, wodurch die Schätzung von Kameraparametern robuster wird.

Berechnung der Einfallsrichtung des Lichtes: Hat man bei der Aufnahme Ort, Datum und Uhrzeit notiert, könnte man dies zur Berechnung der Einfallsrichtung des Sonnenlichts ausnutzen. In keiner der zitierten Arbeiten wird die Einfallsrichtung der Beleuchtung geschätzt, noch gehen Beleuchtungsverhältnisse in irgendeiner Weise in den zitierten Verfahren ein. Wir stellen ein Verfahren vor, das es gestattet, die Einfallsrichtung des Lichtes mit Hilfe der Rotationsparameter ohne Kenntnis der Brennweite, des Skalierungsfaktors und der Höhe der Masten aus im Bild detektierten Ampelmasten und ihrem Schattenwurf auf die Fahrbahnebene zu schätzen.

4.5 Kalibrierung am Beispiel einer Innenstadtszene

4.5.1 Beschreibung der Kalibrierung

Wir werden die Kalibrierung folgendermaßen durchführen:

1. Interaktive grobe Klassifizierung der im Bild detektierten Kantensegmente nach Hauptfluchtpunkten, wenn kein a-priori Wissen über die Orientierung der Kamera relativ zur Szene vorhanden ist.
2. Bestimmung der drei Ursprungsgeraden durch die Hauptfluchtpunkte.
3. Berechnung der Rotationsparameter.

4. Bestimmung einer zuverlässigen Schätzung der Koordinaten eines Hauptfluchtpunktes zur Berechnung der Brennweite.
5. Mit Hilfe der bereits geschätzten Parameter iterativ eine bessere Klassifizierung der Kantensegmente nach Hauptfluchtpunkten erreichen. Zurück zu Schritt 2.
6. Bestimmung des Translationsparameters.
7. Bestimmung von Azimut- und Polarwinkel der Einfallrichtung des Lichts. Dies ist auch schon nach Schritt 3 möglich, da hierfür nur die Rotationsparameter der Kamera gebraucht werden.

4.5.2 Zusammenstellung der zur Kalibrierung benötigten Eingangsdaten

Zur Kalibrierung brauchen wir als Eingangsdaten

- 2D-Kantensegmente aus dem Bild,
- die drei daraus berechneten Ursprungsgeraden durch die Hauptfluchtpunkte zur Schätzung der Rotationsparameter,
- die zuverlässige Schätzung der Bildkoordinaten eines Hauptfluchtpunktes zur Schätzung der Brennweite,
- zwei 2D-Bildpunkte der Fahrbahnebene und die Länge des 3D-Kantensegments ihres Urbilds, die die einzige, nicht aus einem Bild zu schätzende Größe ist,
- eine weitere Ursprungsgerade durch einen Fluchtpunkt von Schattenkantensegmenten senkrechter Masten zur Berechnung des Azimutwinkels und
- die Schätzung der Ursprungsgerade durch den Fluchtpunkt „virtueller“ Lichtstrahlkantensegmente (vgl. Def. 4.2) zur Berechnung des Polarwinkels der Einfallrichtung der Lichtes.

Abb. 4.4 zeigt die Abhängigkeit der zur Kalibrierung benötigten Eingangsdaten. Dabei wird deutlich, daß der Translationsparameter am schwierigsten zu schätzen ist, da er von vielen Größen abhängt.

4.5.3 Bestimmung der Rotationsparameter

In einem ersten Ansatz werden interaktiv Winkelbereiche angegeben, in denen die Kantensegmente liegen, die sich in einem Fluchtpunkt schneiden (vgl. Abb. 4.5). Eine bessere Klassifizierung ist erst nach Kenntnis der Rotationsparameter und der Brennweite mit Lemma B.26 möglich (siehe unten im Abschnitt 4.5.5).

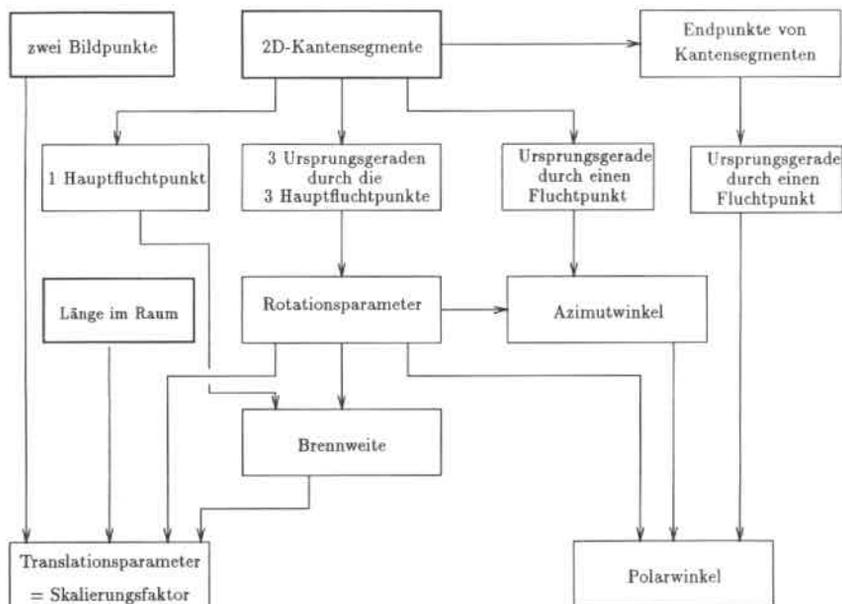


Abbildung 4.4: Eingangsdaten (fette Kästchen) und Abhängigkeiten (Pfeile) der zu bestimmenden Parameter werden aufgezeigt. Nur die „Länge im Raum“ ist eine nicht im Bildmaterial enthaltene Größe und muß interaktiv eingegeben werden.

Wie schon diskutiert, werden in den Verfahren von z.B. [Haralick 80; Caprile & Torre 90; Wang & Tsai 90; Wang & Tsai 91; Chen & Jiang 91; Tai *et al.* 93; Straforini *et al.* 93; Schuster *et al.* 93] Fluchtpunkte auf der Basis inhomogener Bildkoordinaten bestimmt. Dabei muß der Fall ausgeschlossen werden, daß der Fluchtpunkt weit entfernt vom Bildzentrum liegt, um eine stark fehlerbehaftete Schätzung der Fluchtpunktkoordinaten zu vermeiden (siehe Anhang B.3.1). Doch bei unseren Bildmaterial ist das keine seltene Ausnahme, sondern eher die Regel: zum Beispiel liegen in Abb. 4.5 alle drei Hauptfluchtpunkte weit entfernt vom Bildzentrum. Eine Lösung dieses Problems besteht im Übergang zu homogenen Bildkoordinaten, siehe z.B. [Collins & Weiss 90; Kanatani 93a]. Dies setzt einen Schätzwert für die Brennweite voraus, der bei uns nicht verfügbar ist. Doch beeinflußt die Brennweite nur den Azimut- und nicht den Polarwinkel (bzgl. des KKS) der 3D-Richtung des Verbindungsvektors vom Augen- zum Fluchtpunkt. Deshalb bestimmen wir diese Azimutwinkel mit einer Ausgleichsrechnung für jede der drei Scharen von Kantensegmenten aus Abb. 4.5, Details siehe in Anhang B.3.3. Eine ähnliche Anpassung beschreibt [Echigo 90]. Er kommt – ohne Angabe der Herleitung – zu einem anderen Ergebnis².

²Versuche, das Ergebnis von [Echigo 90] herzuleiten, führten zu keinem uns befriedigenden Vorgehen.

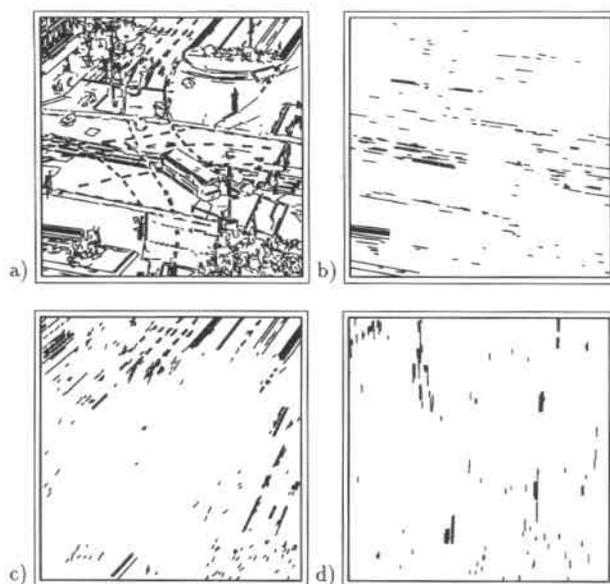


Abbildung 4.5: (a) Detektierte Kantensegmente in der dritten Aufnahme einer Folge des Ettlinger-Tor-Platzes aus Abb. 1.1; (b) (c) (d): Kantensegmente gruppiert nach X_w -, Y_w - bzw. Z_w -Richtung des Weltkoordinatensystems.

Dieser Azimutwinkel legt gleichzeitig die Bild-Ursprungsgerade durch den zur 3D-Richtung gehörenden Fluchtpunkt bzw. die Bildpunktverhältnisse der Fluchtpunkt-bildkoordinaten fest (vgl. Anhang B.4.5).

Im Anhang B.4.5 ist dargestellt, wie aus den Ursprungsgeraden durch drei Hauptfluchtpunkte die Rotationsparameter der Kamera relativ zur Welt zu ermitteln sind. Insbesondere ist – im Gegensatz zu [Echigo 90] – angegeben, welche zusätzlichen Voraussetzungen hinreichend sind.

4.5.4 Verfahren zur Berechnung der Brennweite

Hat man die Rotationsparameter der Kamera relativ zur Fahrbahnebene bestimmt, so kann man die Brennweite aus den Bildkoordinaten $(x_\infty, y_\infty)^T$ eines Hauptfluchtpunktes ermitteln. Dieses eine Koordinatenpaar ist immer vorhanden, da nicht alle drei Hauptfluchtpunkte gleichzeitig im Unendlichen liegen können. Mit der 3D-Orientierung $(U, V, W)^T$ der Achse des WKSs, die zu jenem detektierten Hauptfluchtpunkt gehört, dessen Bildkoordinaten $(x_\infty, y_\infty)^T$ geschätzt wurden, erhält man f aus der Tatsache, daß $(U, V, W)^T$ und $(x_\infty, y_\infty, f)^T$ kollinear sein müssen (Satz B.14).

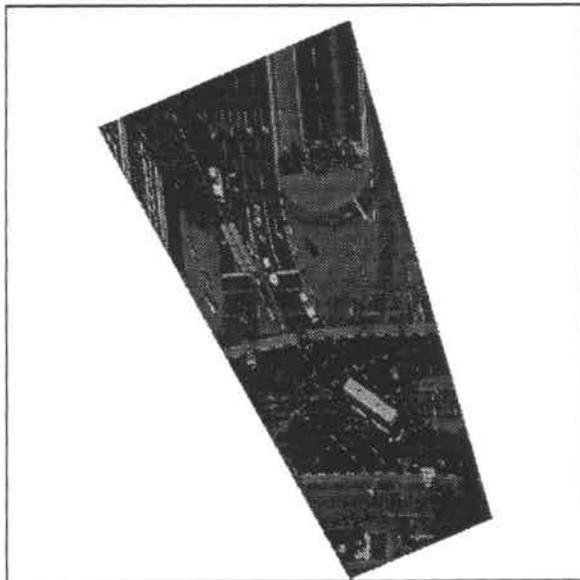


Abbildung 4.6: Projektive Bildtransformation der ersten Aufnahme der Ettlinger-Tor-Folge (vgl. Abb. 1.1), so daß die Fahrbahn von oben betrachtet wird.

Im Gegensatz zu Standardverfahren, die – wie in Satz B.33 – die Koordinaten zweier konjugierter Fluchtpunkte benötigen, reicht bei unserem Verfahren eine gute Schätzung von Fluchtpunktcoordinates sowie eine Raumrichtung aus. Bisher publizierte Verfahren zur Schätzung der Brennweite aus Fluchtpunkten versagen am Beispiel aus Abb. 4.5, da dort nur aus den Kantensegmenten in Y_w -Richtung eine gute Schätzung der Fluchtpunktcoordinates zu erwarten ist.

Nach Bestimmung der Rotationsparameter und der Brennweite kann man das Bild projektiv so transformieren, daß man eine Aufsicht auf die Fahrbahnebene erhält (siehe Abb. 4.6).

4.5.5 Iterative Verbesserung und Prädiktion von Rotationsparametern und Brennweite

Fluchtpunkte sollten nur zur Initialisierung von Rotationsparameter herangezogen werden, eine iterative Verbesserung der Rotationsparameter muß auf den Kantensegmenten, nicht aber auf den Fluchtpunkten basieren, da deren Koordinaten stärker fehlerbehaftet sind.

Nach Bestimmung der Rotationsparameter und der Brennweite der Kamera kann

man nach Lemma B.26 die 2D-Kantensegmente neu nach zugehörigen Fluchtpunkten gruppieren. Damit lassen sich iterativ neue Schätzwerte für Fluchtpunkte und Brennweite ausrechnen.

4.5.6 Translationsparameter

Liegt das Urbild $P(X_k, Y_k, Z_k)$ eines Bildpunktes mit den homogenen Koordinaten $\mathbf{p} = (x_b, y_b, f)^T$ in der Fahrbahnebene, die durch den Normalenvektor $\mathbf{E}_z = (E_{zx}, E_{zy}, E_{zz})^T$ und durch den Punkt mit den Kamerakoordinaten $\mathbf{T} = (0, 0, T_z)^T$ definiert wird, so gilt³:

$$\begin{pmatrix} X_k \\ Y_k \\ Z_k \end{pmatrix} = \frac{\mathbf{E}_z^T \mathbf{T}}{\mathbf{E}_z^T \mathbf{p}} \mathbf{p} = T_z \frac{E_{zz}}{E_{zx}x_b + E_{zy}y_b + E_{zz}f} \begin{pmatrix} x_b \\ y_b \\ f \end{pmatrix}. \quad (4.4)$$

Ohne Vorgabe der Raumkoordinaten ausgewählter Punkte, aber mit einer Annahme über die Maße eines Szenenkantensegments, etwa die Breite eines Haltestreifens, sowie den Bildkoordinaten $(x_b^{(1)}, y_b^{(1)})^T$ und $(x_b^{(2)}, y_b^{(2)})^T$ des Anfangs- bzw. Endpunktes der Projektion des Szenenkantensegments ins Bild, erhält man aus

$$\begin{aligned} d &= \left\| \begin{pmatrix} X_k^{(2)} \\ Y_k^{(2)} \\ Z_k^{(2)} \end{pmatrix} - \begin{pmatrix} X_k^{(1)} \\ Y_k^{(1)} \\ Z_k^{(1)} \end{pmatrix} \right\|_2 \\ &= |T_z| \left\| \frac{E_{zz}}{E_{zx}x_b^{(2)} + E_{zy}y_b^{(2)} + E_{zz}f} \begin{pmatrix} x_b^{(2)} \\ y_b^{(2)} \\ f \end{pmatrix} - \frac{E_{zz}}{E_{zx}x_b^{(1)} + E_{zy}y_b^{(1)} + E_{zz}f} \begin{pmatrix} x_b^{(1)} \\ y_b^{(1)} \\ f \end{pmatrix} \right\|_2 \end{aligned} \quad (4.5)$$

den Translationsparameter T_z . Er erhält ein positives Vorzeichen (siehe Lage von KKS und WKS in Abb. 4.1).

Die Berechnung des Translationsparameters basiert auf Verfahren, wie man die Raumkoordinaten eines Punktes in bekannter Raumebene aus seinen Bildkoordinaten [Haralick & Shapiro 93] bzw. eine Strecke vorgegebener Länge in einer bekannten Raumebene aus den Bildkoordinaten der Endpunkte ihrer Projektion berechnen kann. Bei bekannter Länge und bei bekannter Orientierung eines 3D-Segments, etwa bei gegebener Höhe eines Ampelmasts, kann man die Bestimmung des Translationsparameters nach [Caprile & Torre 90] bzw. nach Abschnitt 13.4.2 in [Haralick & Shapiro 93] durchführen. Dabei ist das Problem allerdings überbestimmt.

³Problem der Rückprojektion eines Punktes auf eine bekannte Ebene nach Gl. 13.13 in [Haralick & Shapiro 93].

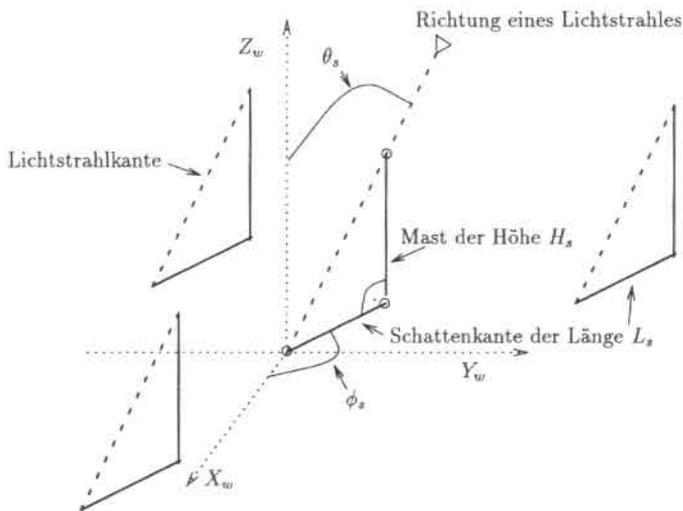


Abbildung 4.7: Schattenwurf im Szenenbereich: Unter der Annahme, die Lichtstrahlen fallen parallel ein, verlaufen alle virtuellen Lichtstrahlkanten und alle Schattenkanten senkrechter Masten jeweils parallel. Die Einfallrichtung des Sonnenlichts wird durch Polarwinkel θ_s und Azimutwinkel ϕ_s festgelegt. Den Azimutwinkel erhält man aus dem Schnittwinkel der Schattenkante eines senkrechten Mastes mit der positiven X_w -Achse, der Polarwinkel ergibt sich aus dem Schnittwinkel einer Lichtstrahlkante mit der positiven Z_w -Achse bzw. aus dem Tangens von Schattenkantenlänge L_s und Masthöhe H_s .

4.5.7 Berechnung der Einfallrichtung des Lichtes

Bei Sonnenschein werfen, wenn die Sonne nicht im Zenith steht, zur Fahrbahnebene senkrechte Masten einen Schatten auf die Fahrbahn. Die Lichteinfallrichtung läßt sich im WKS durch Angabe von Azimutwinkel ϕ_s und Polarwinkel θ_s darstellen. Wir schätzen die Lichteinfallrichtung im WKS, dagegen schätzen [Jiang & Ward 94] in einem 2D-Ansatz die Lichteinfallrichtung im Bildbereich: den Azimutwinkel der Lichteinfallrichtung im KKS.

Zunächst betrachten wir den Schattenwurf im Szenenbereich:

Definition 4.1 Den Schattenwurf eines senkrecht auf der Fahrbahn stehenden Mastes nennen wir (3D-)Schattenband, das von zwei Schattenkanten begrenzt wird, die Verbindung der Spitze des Mastes mit der Projektion der Spitze auf die Fahrbahnebene virtuelle (3D-)Lichtstrahlkante (siehe Abb. 4.7).

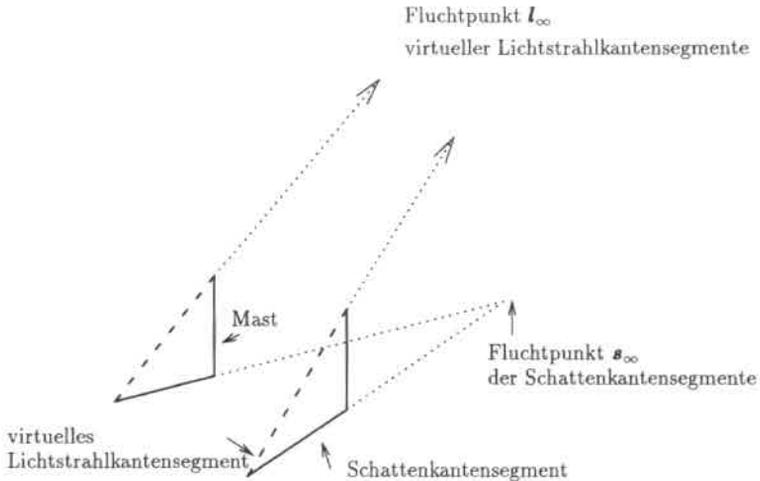


Abbildung 4.8: Schattenwurf im Bildbereich: Schattenkantensegmente sind die im Bild detektierten Abbilder der im Raum parallelen Schattenkanten. Zu Geraden verlängerte Schattenkantensegmente schneiden sich in einem Fluchtpunkt s_∞ . Virtuelle Lichtstrahlkantensegmente (gestrichelte Linien) sind Abbilder der im Raum parallelen virtuellen Lichtstrahlkanten. Zu Geraden verlängerte virtuelle Lichtstrahlkantensegmente schneiden sich in einem Fluchtpunkt l_∞ . Virtuelle Lichtstrahlkantensegmente erhält man mit Hilfe eines Eckendetektors als Verbindungssegment des Abbildes der Spitze eines Mastes und des Abbildes des Schattens der Spitze des Mastes.

Unter der Annahme, die Lichtstrahlen fallen parallel ein, sind im Szenenbereich

- alle virtuellen Lichtstrahlkanten parallel
- und alle Schattenbänder parallel.

Nun projizieren wir Schattenbänder und virtuelle Lichtstrahlkanten ins Bild:

Definition 4.2 *Im Bild detektierte Abbilder eines Schattenbandes heißen (2D-) Schattenkantensegmente. Ins Bild projizierte virtuelle Lichtstrahlkanten heißen virtuelle (2D-)Lichtstrahlkantensegmente.*

Ein virtuelles Lichtstrahlkantensegment verbindet das Abbild einer Mastspitze mit dem Abbild des Schattens der Mastspitze.

Zu Geraden verlängerte Schattenkantensegmente schneiden sich in einem Fluchtpunkt s_∞ . Zu Geraden verlängerte virtuelle Lichtstrahlkantensegmente schneiden sich in einem Fluchtpunkt l_∞ (siehe Abb. 4.8).

Satz 4.3 Bei bekannter Brennweite liefert der Fluchtpunkt \mathbf{l}_∞ der zu Geraden verlängerten virtuellen Lichtstrahlkantensegmente die Einfallsrichtung des Lichtes.

Beweis: Nach Lemma B.23 erhält man aus jedem Fluchtpunkt bei bekannter Brennweite bezüglich des KKSs den 3D-Richtungsvektor der parallelen Raumgeraden. \square

Bei bekannten Rotationsparametern kann man die nach Satz 4.3 bestimmte Einfallsrichtung des Lichtes vom KKS ins WKS transformieren. Geht man anschließend im WKS auf Kugelkoordinaten über, so erhält man Azimut- und Polarwinkel der Einfallsrichtung des Lichtes im WKS. Das Verfahren nach Satz 4.3 ist zum einen abhängig von der Brennweite, zum anderen abhängig von der Robustheit des Ecken-detektors, mit dessen Hilfe die virtuellen Lichtstrahlkantensegmente detektiert werden.

Die Einfallsrichtung des Lichtes kann auch ohne Kenntnis der Brennweite geschätzt werden. Den Azimutwinkel der Einfallsrichtung des Lichtes können wir wesentlich genauer mit Hilfe der Schattenkantensegmente berechnen; für die Schätzung des Polarwinkels der Einfallsrichtung des Lichtes werden allerdings die unter Umständen stark fehlerbehafteten Endpunkte der virtuellen Lichtstrahlkantensegmente benötigt.

Satz 4.4 (Schätzung des Azimutwinkels der Lichteinfallsrichtung im WKS) Sei die Bild-Ursprungsgerade $ax + by = 0$ mit dem homogenen Koordinatenvektor $\mathbf{g} = (a, b, 0)^T$ durch den Fluchtpunkt \mathbf{s}_∞ der Schattenkantensegmente geschätzt. Dann erhält man

1. die 3D-Orientierung der Schattenkanten bei Kenntnis des Normalenvektors \mathbf{E}_z der Fahrbahnebene aus

$$\mathbf{E}_z \times \mathbf{g} \quad (4.6)$$

2. und den Azimutwinkel ϕ_s der Einfallsrichtung des Lichtes bei Kenntnis des Normalenvektors der Fahrbahnebene und der 3D-Orientierung \mathbf{E}_z der X_w -Achse aus

$$\cos \phi_s = \frac{\mathbf{E}_z^T (\mathbf{E}_z \times \mathbf{g})}{\|\mathbf{E}_z \times \mathbf{g}\|_2} \quad (4.7)$$

ohne Kenntnis der Brennweite, des Skalierungsfaktors und der Höhe H_s der Masten.

Beweis: Behauptung nach Gl. 4.6 folgt aus Korollar B.47 mit Gl. B.52, Behauptung nach Gl. 4.7 folgt aus Lemma B.48 und der Definition eines Azimutwinkels. \square

Satz 4.5 (Schätzung des Polarwinkels der Lichteinfallsrichtung im WKS) Sei die Bild-Ursprungsgerade $ax + by = 0$ mit dem homogenen Koordinatenvektor $\mathbf{g} = (a, b, 0)^T$ durch den Fluchtpunkt \mathbf{l}_∞ der Lichtstrahlkantensegmente geschätzt. Das Urbild dieser Bild-Ursprungsgerade liegt in einer Raumebene mit dem Normalenvektor

$$\mathbf{N} = \mathbf{E}_z \times (\cos \phi_s \mathbf{E}_x + \sin \phi_s \underbrace{\mathbf{E}_z \times \mathbf{E}_z}_{\mathbf{E}_y}). \quad (4.8)$$

Kennt man den Azimutwinkel ϕ_s der Einfallsrichtung des Lichtes, die Normale \mathbf{E}_z der Fahrbahnebene und die 3D-Orientierung \mathbf{E}_z der X_w -Achse, dann erhält man

1. die 3D-Orientierung der Lichtstrahlkanten aus

$$\mathbf{N} \times \mathbf{g} \quad (4.9)$$

2. und den Polarwinkel θ_s der Einfallrichtung des Lichtes aus

$$\cos \theta_s = \frac{\mathbf{E}_z^T(\mathbf{N} \times \mathbf{g})}{\|\mathbf{N} \times \mathbf{g}\|_2} \quad (4.10)$$

ohne Kenntnis der Brennweite, des Skalierungsfaktors und der Höhe H_s der Masten.

Beweis: Die Urbilder der Lichtstrahlkantensegmente, die Lichtstrahlkanten, verlaufen parallel zu den Ebenen, die durch die Richtungsvektoren \mathbf{E}_z und $\cos \phi_s \mathbf{E}_x + \sin \phi_s \mathbf{E}_y$ aufgespannt werden. Hieraus folgt Gl. 4.8. Die Behauptung bzgl. Gl. 4.9 folgt aus Korollar B.47 mit Gl. B.52, Behauptung nach Gl. 4.10 folgt aus Lemma B.48 und der Definition eines Polarwinkels. \square

Kapitel 5

Modellbasierte Bildfolgenauswertung

5.1 Systematische Erweiterung des verwendeten Modellwissens

Dieses Kapitel zeigt, wie durch systematische Erweiterung des in den Auswertungsprozeß eingehenden Wissens signifikante Fortschritte bei der automatischen Auswertung von Bildfolgen zu erreichen sind. Im einzelnen werden folgende Schritte behandelt:

- **Kalibrierung der Kamera**

[Koller 92] kalibrierte seine Aufnahmen mit Hilfe eines Katasterplanes nach einem Verfahren von [Tsai 87]. Dazu wurden interaktiv Punkte im Plan zu ihren korrespondierenden Punkten in der Videoaufnahme der Kreuzung zugeordnet. Darüber hinaus war es notwendig, die Koordinaten ausgewählter Bildpunkte interaktiv mehrmals leicht zu verändern, um überhaupt eine brauchbare Kalibrierung zu erhalten. Sowohl die interaktive Zuordnung von Merkmalen aus dem Plan zu Merkmalen im Bild als auch die interaktive Iteration entfällt mit dem in dieser Arbeit neu vorgestellten Verfahren. Desweiteren gab [Koller 92] die Lichteinfallrichtung interaktiv vor. Wir schätzen Kalibrierdaten und Lichteinfallrichtung mit Hilfe von Annahmen über den Straßenverlauf aus den Bilddaten (vgl. Kapitel 4) und verringern damit die Zahl interaktiv einzustellender Parameter.

- **Entwurf und Anpassung von Modellen innerstädtischer Straßenkreuzungen an Bildfolgen**

Eine wesentliche Verbesserung im Vergleich zum System von [Koller 92] gelang durch die Modellierung und Anpassung für die in einer Bildfolge erfaßte Fahrbahn. Das im letzten Kapitel vorgestellte Kalibrierverfahren liefert eine Initialisierung für die Anpassung eines Fahrbahnmodells an die Bilddaten. Die

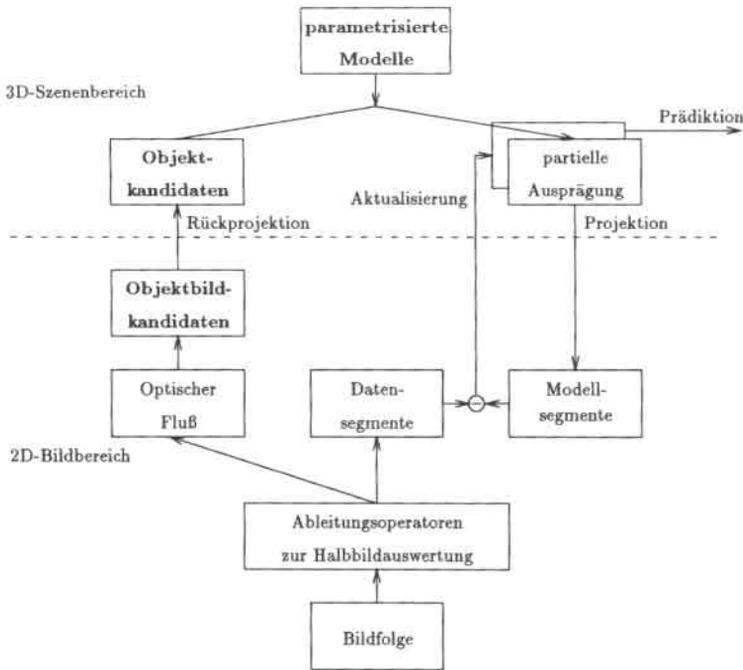


Abbildung 5.1: Detektion und Verfolgung von Fahrzeugen: Im Vergleich mit dem System aus Abb. 3.2 wird die Segmentierung des optischen Flußfeldes unter Einbezug von Wissen über die Fahrspuren durchgeführt. Die Fahrzeugmodelle wurden ergänzt.

Anpassung des Fahrbahnmodells ermöglicht eine Verbesserung der Kalibrierung. Die Fahrspuren des Fahrbahnmodells helfen, das optische Flußfeld zu segmentieren und Objektbildkandidaten zu detektieren. Bei der Berechnung von Bewegungsverben für Fahrzeugbewegungen wird ebenfalls auf das Fahrbahnmodell zurückgegriffen.

- **Erweiterung und Verfeinerung von Fahrzeugmodellen**

Das Modellwissen wurde auch bei den Fahrzeugen vergrößert. Zum einen wurden die Frontleuchten mit in das Fahrzeugmodell aufgenommen, zum anderen wurden für weitere Fahrzeugtypen, wie zum Beispiel für einen Bus, neue Modelle entworfen.

Die Abbildungen 5.1 und 5.2 stellen diese Schritte im Systemrahmen dar.

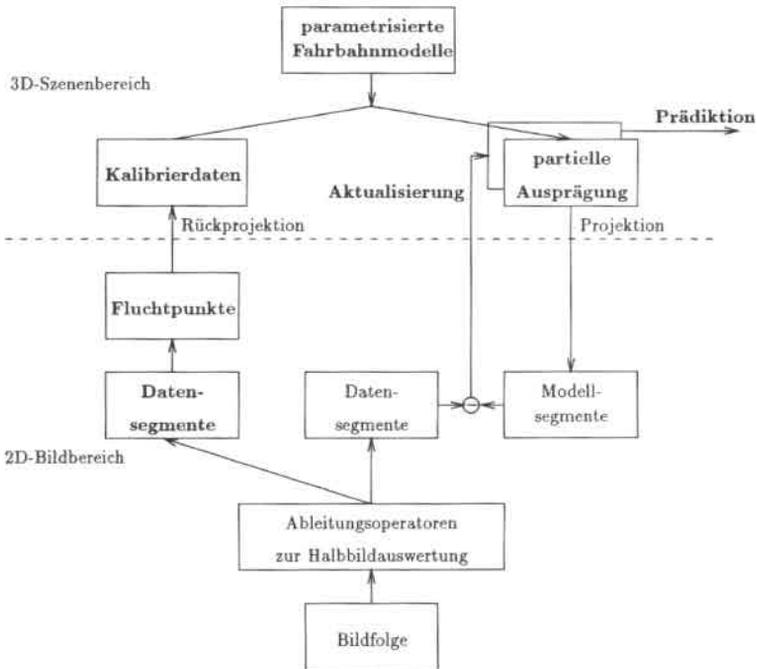


Abbildung 5.2: Initialisierung und Anpassung eines Fahrbahnmodells: Während im Teilsystem in Abb. 5.1 aus optischem Fluß aggregierte Objektbildkandidaten zur Initialisierung einer Modellausprägung herangezogen werden, wird in diesem Teilsystem eine Initialisierung aus Fluchtpunkten ermittelt, die ihrerseits aus Datensegmenten berechnet werden. Im Abschnitt 5.2 wird dargelegt, daß und wie die Systemkomponenten zur Datenassoziation sowie zur Aktualisierung bzw. Prädiktion einer Modellausprägung in Abb. 5.1 weiterentwickelt und modifiziert wurden, um statt eines Fahrzeugmodells ein Fahrbahnmodell an Datensegmente anpassen zu können: im Aktualisierungsschritt werden z.B. alle sechs – statt nur drei wie bei der Systemkomponente in Abb. 5.1 – Lagefreiheitsgrade geschätzt. Die Prädiktion nutzt die Voraussetzung einer festen Kamera.

Ein weiterer wesentlicher Fortschritt bei der Verfolgung von Fahrzeugen wurde durch eine Vergrößerung des in den Anpassungsprozeß eingehenden Modellwissens erreicht: in Kapitel 6 stellen wir ein neues Verfahren vor, bei dem nicht mehr Kantensegmente, sondern die Flächenstücke eines Polyedermodells an die Bilddaten angepaßt werden.

5.2 Entwurf und Anpassung von Modellen innerstädtischer Straßenkreuzungen an Bildfolgen

Als zentral für die weitere Vorgehensweise erweist sich die Auswertung von statischen Aspekten der in der Bildfolge aufgezeichneten Szene. Unsere Modelle innerstädtischer Straßenkreuzungen sind wesentlich umfangreicher und detaillierter als die Modelle von z.B. [Liou & Jain 87; Kanatani & Watanabe 90; Dickmanns & Mysliwetz 92; Kasprzak *et al.* 94]. Diese Autoren werten Bildfolgen aus, die aus einem bewegten Fahrzeug heraus aufgezeichnet wurden. Deshalb besteht bei ihnen die Fahrbahnstruktur nur aus bis zu drei parallelen Streifen: den beiden seitlichen Fahrbahngrenzungen und einem Mittelstreifen.

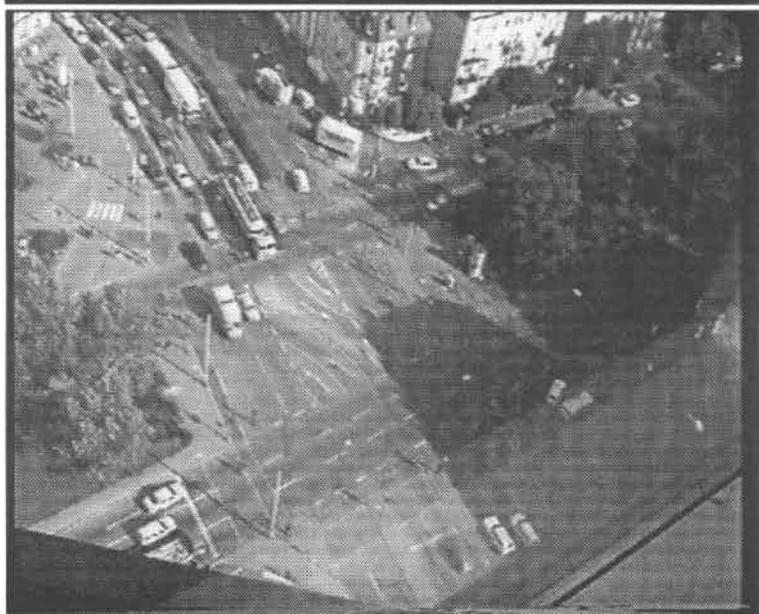
5.2.1 Probleme bei rein datengetriebener Segmentierung

Abb. 5.3 zeigt das komplexe Verkehrsgeschehen an einer Frankfurter Straßenkreuzung. Bei der Detektion bewegter Fahrzeugabbilder in diesem neuen Bildmaterial treten Probleme auf, wenn zwei Fahrzeuge dicht nebeneinander mit annähernd gleicher Geschwindigkeit fahren (vgl. Abb. 5.4): die Fahrzeugabbilder werden mit dem rein datengetriebenen Verfahren aus Abschnitt 3.3 zu einem einzigen Objektbildkandidaten zusammengefaßt. Durch Verringern der Schwellwerte, die in die Ballungsanalyse eingehen, kann man ein Aufbrechen der Objektbildkandidaten in mehrere Objektbildkandidaten erreichen. Jedoch zerfallen dann besonders bei langen Fahrzeugen, wie LKW oder Bus, die detektierten Objektbildkandidaten in viele kleine Objektbildkandidaten. Bei langen Fahrzeugen erhält man nämlich oft nur ein lückenhaftes Flußfeld, da die bei Abbildern langer Fahrzeuge teilweise vorhandenen homogenen Grauwertbereiche keine starke Grauwertvariation aufweisen, die zur Schätzung von Flußvektoren benötigt wird.

Anstatt weitere Untersuchungen zur Parameterwahl für den datengetriebenen Segmentierungsschritt durchzuführen, um einen geeigneten Parametersatz zur Detektion aller Fahrzeuge zu finden, der eventuell gar nicht existiert, oder zu versuchen, höhere Momente [Prokop & Reeves 92] aus den Flußvektoren zu berechnen und für eine Ballungsanalyse auszunutzen, vergrößern wir das an dieser Stelle eingehende Modellwissen: wir werden Wissen über die einzelnen Fahrspuren in der Ballungsanalyse einsetzen. Die Segmentierung unter Verwendung von Wissen über den 3D-Szenenbereich führt auf ein modellbasiertes Verfahren (Abschnitt 5.2.9), das sich dadurch von den in Kapitel 2 zitierten Arbeiten unterscheidet, die rein im 2D-Bildbereich arbeiten.



(a)



(b)

Abbildung 5.3: Jeweils die erste Aufnahme einer Folge des Verkehrsgeschehens an einer Frankfurter Straßenkreuzung.

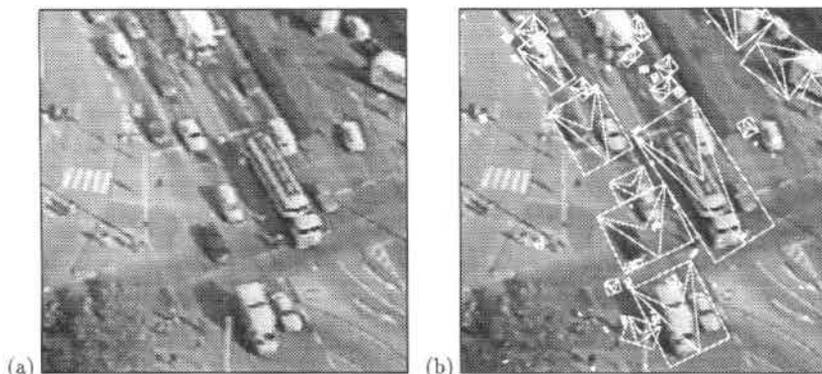


Abbildung 5.4: (a) zeigt einen Ausschnitt des oberen linken Quadranten der dritten Aufnahme der Frankfurter Straßenkreuzung aus Abb. 5.3(b).

(b) In enger Nachbarschaft in dieselbe Richtung fahrende Objekte werden bei dem rein datengetriebenen Segmentierungsverfahren aus Abschnitt 3.3 zu einem einzigen Objektbildkandidaten zusammengefaßt.

5.2.2 Einfaches geometrisches Fahrbahnmodell

Mit Hilfe von [Heinze 93] wurde in einem ersten Ansatz zusammen mit [Osswald 94; Leuck 94] ein einfaches geometrisches Fahrbahnmodell für die Frankfurter Straßenkreuzung aus Abb. 5.3 erstellt, siehe Abb. 5.5. Dabei wurde zunächst angenommen, die Straßenzüge schneiden sich senkrecht. Teile der Fahrbahnmarkierung und der Fahrbahnbegrenzungen wurden als unterbrochene Linien ins Modell aufgenommen. Auf Details wie z.B. Bürgersteige, Zebrastreifen oder auch Radwege wurde zunächst verzichtet.

5.2.3 Erweiterung zu schätzender Lageparameter

Allgemeine Hinweise zur Schätztheorie finden sich in den Lehrbüchern [Gelb 74; Maybeck 79; Krebs 80; Bar-Shalom & Fortmann 88], speziell für die Bildfolgenauswertung auch in [Faugeras 93; Daniilidis 92; Koller 92].

Zunächst werden Kantensegmente aus dem dritten Bild der Folge in Abb. 5.3(a) berechnet (siehe Abb. 5.6). Abb. 5.7 zeigt eine initiale, Abb. 5.8 eine nach Abschnitt 4.5.5 iterativ korrigierte Gruppierung der Kantensegmente aus Abb. 5.6 nach Hauptfluchtpunkten.

In unseren Untersuchungen stellte sich heraus, daß die aus Fluchtpunkten geschätzte Orientierung der Normalen der Fahrbahnebene relativ zur Kamera nicht so genau ist (siehe Abb. 5.9), daß man die Lageschätzung des Fahrbahnmodells auf die Schätzung dreier Freiheitsgrade – eines rotatorischen und zweier translatorischer Freiheitsgra-

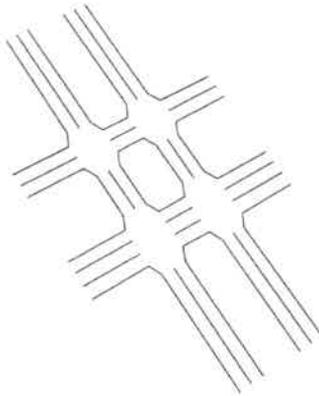


Abbildung 5.5: Einfaches geometrisches Fahrbahnmodell für die Frankfurter Straßenkreuzung aus Abb. 5.3, wobei zunächst angenommen wurde, daß sich die Straßenzüge im Kreuzungsbereich senkrecht schneiden. Teile der Fahrbahnmarkierung sowie der Fahrbahnbegrenzungen wurden als unterbrochene Linien ins Modell aufgenommen. Auf Details wie z.B. Bürgersteige, Zebrastrifen oder auch Radwege wurde in diesem ersten groben Modell verzichtet.

de, die die Lage des Modells in der Straßenebene angeben – beschränken könnte. Damit müssen bei einer Anpassung eines Fahrbahnmodells an Bilddaten alle sechs Lageparameter relativ zur Kamera geschätzt werden, d.h. drei translatorische und drei rotatorische. Diese Parameter werden in einem sechsdimensionalen Zustandsvektor zusammengefaßt, die Lageschätzung dann mit einem rekursiven Maximum-a-posteriori-Schätzer ausgeführt, in den als Messung aus der aktuellen Aufnahme extrahierte Kantensegmente eingehen.

Wir verzichten hier auf eine ausführliche Darstellung dieses Verfahrens, da es in seiner ersten Version nur eine leichte Modifikation des Verfahrens von [Koller 92] zur Lageschätzung von Fahrzeugen war, stellen aber ein verbessertes Verfahren in Anhang D.3 dar, das nach den Diskussionen der folgenden Abschnitte im Laufe zahlreicher Experimente entwickelt wurde. Es wurde im Laufe unserer Experimente klar, daß und vor allem auch wie – d.h. unter Ausnutzung von welchem Modellwissen – das Verfahren von [Koller 92] modifiziert werden mußte, um die Robustheit des Verfahrens deutlich zu steigern. Die hierzu notwendigen Modell- und Schätzprozeßverfeinerungsschritte werden in den folgenden Abschnitten näher erläutert.

5.2.4 Änderung der Abstandsfunktion zur Anpassung eines Fahrbahnmodells an Datenkantensegmente

Für die Anpassung eines Fahrbahnmodells an Bildkantensegmente ergaben sich Schwierigkeiten bei Anwendung der von [Koller 92] verwendeten Mittelpunkt-Re-

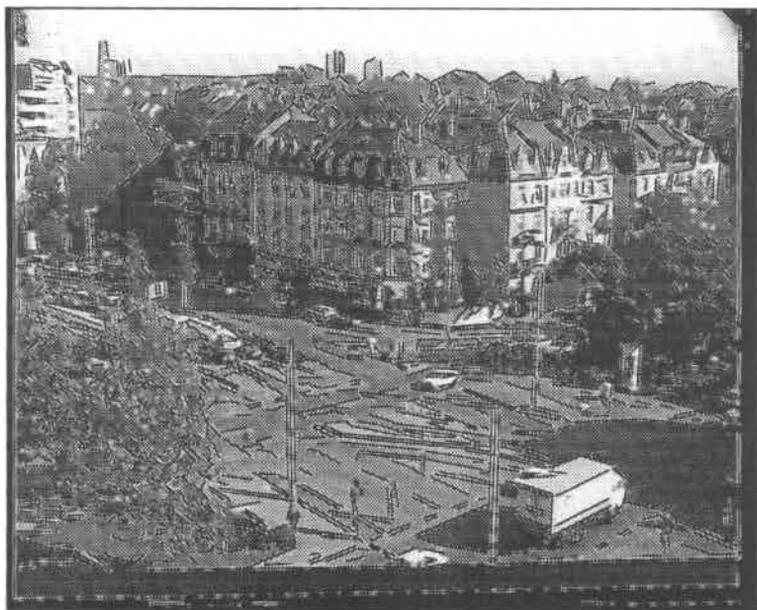


Abbildung 5.6: Berechnete Kantensegmente aus dem dritten Bild der Bildfolge aus Abb. 5.3(a).

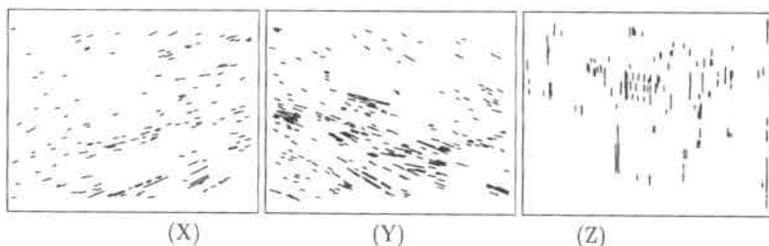


Abbildung 5.7: Die initiale Gruppierung der Kantensegmente aus Abb. 5.6 nach Hauptfluchtlinien der X -, Y - bzw. Z -Achse eines kartesischen Weltkoordinatensystems. Die Fahrbahnebene liegt in der XY -Ebene des Weltkoordinatensystems.

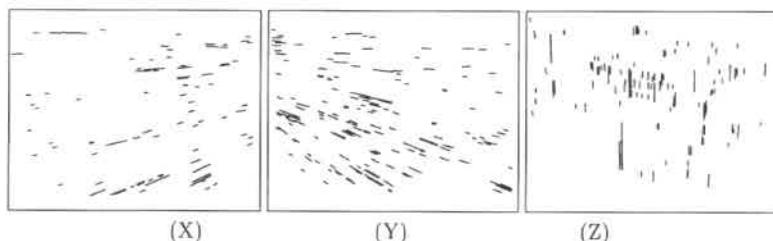


Abbildung 5.8: Nach einer Schätzung der Orientierung (relativ zur Fahrbahnebene) und Brennweite der Kamera werden die Kantensegmente aus Abb. 5.6 nochmals nach Hauptfluchtachsen der X -, Y - bzw. Z -Achse eines kartesischen Weltkoordinatensystems gruppiert.

präsentation für Kantensegmente [Deriche & Faugeras 90]. Die in [Koller 92] beschriebene Situation unterscheidet sich von der unseren dadurch, daß die Fahrzeuge bei [Koller 92] recht klein waren, wohingegen das Abbild einer Kreuzung das gesamte aufgenommene Bild umfaßt, und deshalb die Fahrbahnbegrenzungslinien sowie die Fahrspurtrennlinien viel länger sind als die Kantensegmente eines Fahrzeugabbildes. Darüber hinaus ist eine Fahrbahnbegrenzungslinie oder Fahrspurtrennlinie mit großer Wahrscheinlichkeit teilweise verdeckt, sei es durch Fahrzeuge, Verkehrsschilder oder Ampelmasten. Damit ist die Wahrscheinlichkeit viel geringer, ein Datenssegment aus der Bildfolge zu berechnen, das die volle Länge wie sein korrespondierendes Modellsegment hat. Die Mittelpunktposition und Länge der aus den Bilddaten berechneten Datenkantensegmente korreliert daher oft nur schwach mit der projizierten Mittelpunktposition und Länge des korrespondierenden Modellsegmentes.

Ein weiteres Problem führt zu ähnlichen Schwierigkeiten bei einer Zuordnung von Modell- zu Datenkantensegmenten. Im Falle unterbrochener Mittelstreifen, die zwei Fahrspuren trennen, können die Endpunkte und die Länge korrespondierender Datenkantensegmente meist gut aus den Bilddaten extrahiert werden. Aber in diesem Fall sind die genauen Positionen der Unterbrechungen im korrespondierenden Modellmittelstreifen nicht bekannt.

Damit erscheint es gerechtfertigt, die Position des Mittelpunktes und die Länge der Datenkantensegmente nicht mehr bei der Anpassung zu berücksichtigen. Die Mahalanobisdistanz zwischen Modell- und Datenkantensegment wird reduziert auf zwei Beiträge, nämlich:

- die Differenz der Orientierungen von Daten- und projiziertem Modellkantensegment in der Bildebene und
- den senkrechten Abstand zwischen dem Mittelpunkt eines Datenkantensegmentes und dem Modellkantensegment.

Da der senkrechte Abstand nichtlinear in den Parametern des Modellkantensegmentes ist, gehen wir im Rahmen der Datenassoziation zu einer verallgemeinerten

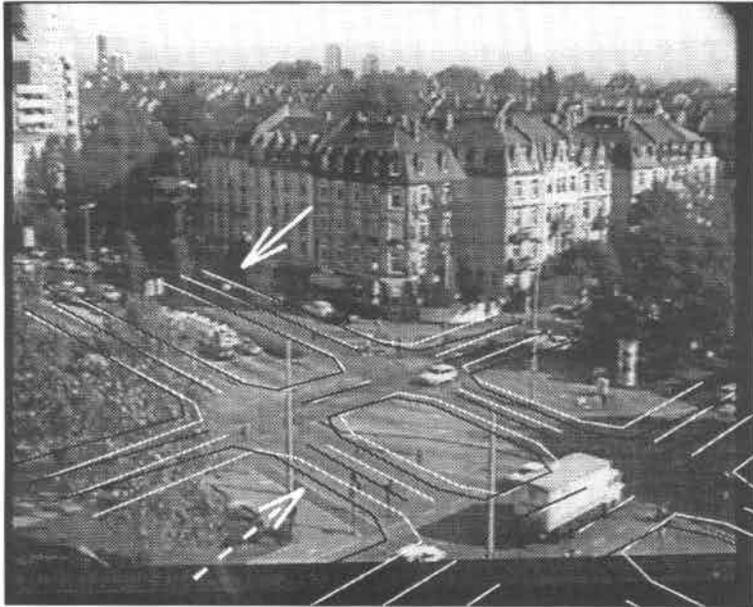


Abbildung 5.9: In dieser Modellanpassung wurden nur drei Freiheitsgrade geschätzt: zwei translatorische für die Position sowie ein rotatorischer für die Orientierung des Fahrbahnmodells in der Straßenebene. Die Projektion der initialen Lage des Kreuzungsmodells in die Bildebene wird mit weißen Linien dargestellt, die korrigierte Modellage nach der Anpassung mit dunklen Linien. Es ist notwendig, alle sechs Lagefreiheitsgrade im Anpassungsprozeß zu schätzen, da die aktualisierte Modellage nicht sehr genau ist (siehe durchgezogener heller Pfeil). Darüber hinaus werden wir das grobe Modell verfeinern, da z.B. die Markierungen des (abgetrennt von der Fahrbahn verlaufenden) Radweges im Bild fälschlicherweise den Fahrbahnbegrenzungen des Modells zugeordnet werden (siehe gestrichelter heller Pfeil).

Mahalanobisdistanz (siehe Anhang C.3.3 und Anhang D.2) über, wodurch die Meßfunktion für den Aktualisierungsschritt des Kalmanfilters nicht mehr in expliziter Form wie bei [Deriche & Faugeras 90; Crowley *et al.* 92; Koller 92], sondern in *impliziter* Form wie bei [Zhang & Faugeras 92a] vorliegt. Mathematische Einzelheiten hierzu finden sich in Anhang D.3.

5.2.5 Berücksichtigung der endlichen Breite von Fahrbahnmarkierungen

Ein anderer Effekt, der im Laufe unserer Versuche zu Problemen führte, trat bei der Extraktion paralleler Kantensegmente auf, die zu beiden Grauwertübergängen eines Abbildes einer Fahrbahnmarkierung gehören: ein Kantensegment gehört zum Übergang von dunklen zu hellen Grauwerten und das andere zum Übergang von hellen nach dunklen Grauwerten auf der anderen Seite der Markierung (siehe Abb. 5.10). Zunächst wurde im Anpassungsprozeß nur eines dieser beiden Kantensegmente benutzt, wobei eine Programmkomponente eingesetzt wurde, die ursprünglich von [Koller 92] zur Fahrzeugverfolgung entwickelt wurde. Die Einschränkung auf 1:1-Zuordnungen zwischen Modell- und Datensegmenten ist für Fahrzeuge gerechtfertigt, zumindest bei den Bildfolgen, die [Koller 92] auswertete. Doch diese Einschränkung ist für solche Fahrbahnmarkierungen nicht mehr gerechtfertigt, deren Abbilder so breit sind, daß alle beide Kantensegmente detektiert werden können, die das Streifenabbild begrenzen. Nun könnte man n:1-Zuordnungen zwischen mehreren Daten- und einem Modellsegment zulassen. Doch das Problem dabei ist, daß der datengetriebene Kantensegment-Extraktionsprozeß nicht immer die Detektion beider Datenkantensegmente für jedes Modellsegment garantiert (siehe nochmals Abb. 5.10), die darüber hinaus zum Modellsegment symmetrisch verlaufen sollen. Eine n:1-Zuordnung führt damit nicht unbedingt zu einer Verbesserung der Genauigkeit des Anpassungsverfahrens.

5.2.6 Berücksichtigung der Richtung von Datengrauwert- übergängen bei der Zuordnung von Modell- zu Datenkantensegmenten

Anstatt weiterhin eine 1:1-Zuordnung von Modell- zu Datenkantensegmenten durchzuführen, erschien es besser, mehr Wissen über die Struktur der Fahrbahnmarkierung in der Szene auszunutzen: eine Fahrbahnmarkierung besteht aus einem Streifen einer bestimmten Breite, der heller als der Fahrbahnbelag ist. Das Modell einer Fahrspurtrennmarkierung in der Szene besteht damit nicht mehr aus einer einzigen geraden Linie, die die Mittellinie der Markierung darstellt, sondern aus einem Streifen einer Breite von (0.12 ± 0.02) m, wie sie in Richtlinien für die Fahrbahnmarkierungen festgelegt ist (siehe u.a. in [Mück 94b]). Durch dieses so verfeinerte Fahrbahnmodell wurde es möglich, im Zuordnungsprozeß Dunkel-Hell- sowie Hell-Dunkel-Übergänge detektierter Datensegmente projizierten Modellsegmenten kor-

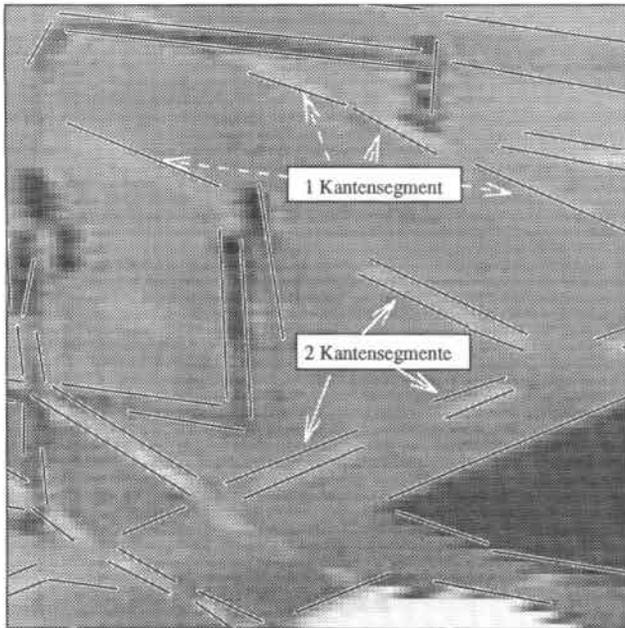


Abbildung 5.10: Vergrößerter Ausschnitt aus dem unterem mittlerem Bereich der Abb. 5.6: Bei einer kontrastreichen weißen Fahrbahnmarkierung endlicher Breite können zwei parallele Kantensegmente aus den Bilddaten berechnet werden (siehe durchgezogene Pfeile). Allerdings garantiert der Kantensegment-Extraktionsprozess nicht immer die Detektion beider Kantensegmente (siehe gestrichelte Pfeile).

rekt zuzuordnen. Es stellte sich heraus, daß mit dieser Modellverfeinerung der – mit teilweise beträchtlichem Suchaufwand verbundene – Einsatz eines Interpretationsbaumes [Grimson 90] vermieden werden konnte. Die Modellverfeinerung ermöglichte auch eine $n:1$ -Zuordnung mehrerer Datensegmente zu einem Modellsegment. Damit konnten deutlich mehr Datensegmente im Anpassungsprozeß verwendet werden, wodurch die Lageschätzung verbessert wurde.

5.2.7 Anpassung eines Fahrbahnmodells an Bildfolgen

Bisher wurde das Fahrbahnmodell nur an ein einzelnes Bild einer Folge angepaßt. Eine genauere Analyse zeigte, daß die Lageschätzung von Zuordnungen von Fahrbahnmodellsegmenten an Datensegmente von solchen Segmenten beeinflusst wird, die aus Bildbereichen berechnet werden, die das Abbild eines Fahrzeuges umfassen, siehe Abb. 5.11.

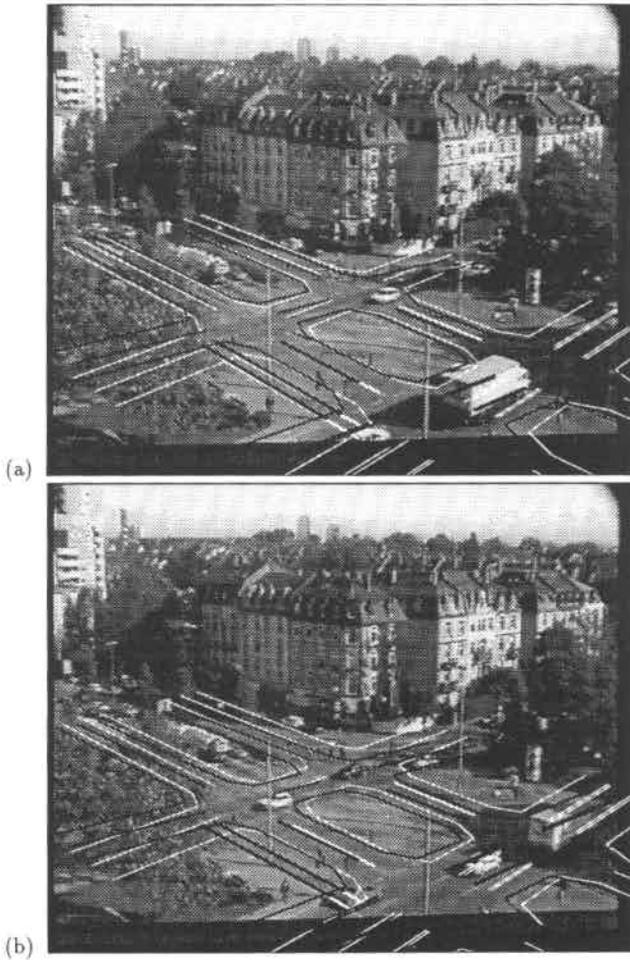


Abbildung 5.11: (a) Eine Anpassung des Kreuzungsmodells an das dritte Bild der Folge aus Abb. 5.3(a). Die Fehlzuordnungen zwischen den Modellsegmenten und den Datensegmenten des Abbildes des LKW im unteren rechten Teil der Aufnahme beeinflussen die Schätzung der Kreuzungsmodellage. Die initiale, ins Bild projizierte Modellage ist mit weißen Linien, die mit der Anpassung aktualisierte Lage mit dunklen Linien dargestellt. (b) Eine entsprechende Anpassung an ein Bild derselben Folge ungefähr 45 Aufnahmen später. Der Einfluß der Datensegmente des Abbildes des LKW beeinflusst die Anpassung nicht mehr so stark, so daß die aktualisierte, ins Bild projizierte Lageschätzung – dunkle Linien – mehr mit dem übereinstimmt, was ein menschlicher Beobachter als passend empfindet.

Prinzipiell kann man nicht ausschließen, daß unser Zuordnungsprozeß versucht, Fahrbahnmodellsegmente solchen Datensegmenten zuzuordnen, die weder zu Abbildern von Fahrbahnmarkierungen noch von Fahrbahnbegrenzungen gehören. Es gibt aber eine recht allgemeine und nichtsdestotrotz leistungsfähige Heuristik, mit deren Hilfe wir Probleme dieser Art lösen werden.

Die Grundüberlegung beruht auf der Feststellung, daß die Bildpositionen von *bewegten* Fahrzeugen sich mit jeder Aufnahme verschieben, wohingegen im Fall einer festen Kamera das Bild von Kantensegmenten, die zur Fahrbahnmarkierung oder -begrenzung gehören, an derselben Bildposition bleibt. Wir wenden einfach einen Verfolgungsansatz, der ursprünglich zur Verfolgung von Fahrzeugen entwickelt wurde [Koller 92], auf die Schätzung von Lageparametern eines Fahrbahnmodells an, indem wir als Bewegungsmodell des Fahrbahnmodells relativ zur Kamera eine konstante Geschwindigkeit mit Betrag 0 m/s annehmen. Wenn die erste Anpassung nicht zu sehr durch Fehlzusordnungen zwischen Modell- und Datensegmenten gestört wird, kann man erwarten, daß nach den Iterationen innerhalb der ersten Aufnahme als Ergebnis zumindest eine brauchbare Lageschätzung erzielt wird. Nutzt man von da an die in einer Aufnahme aktualisierte Schätzung zur Prädiktion der Lageschätzung des Fahrbahnmodells zum Zeitpunkt der nächsten Aufnahme mit einem Kalmanfilter, so wird auch die Kovarianzmatrix für den Fehler der Lageschätzung nach jeder Auswertung eines Bildes aktualisiert. Kantensegmente, die zu bewegten Fahrzeugabbildern gehören, haben nach ein paar Aufnahmen einen vergleichsweise geringen Einfluß auf die Lageschätzung, da es eine andauernde Bestätigung in Richtung ungestörter Werte durch all jene Zuordnungen gibt, bei denen Modell- den Datensegmenten zugeordnet werden, die zur Fahrbahnmarkierung und -begrenzung gehören.

Folglich werden die Einträge in der Kovarianzmatrix für die Lageschätzung nach und nach kleiner und damit auch der Einfluß der zeitlich variierenden Fehlzusordnungen zwischen Modell- und Datensegmenten von Fahrzeugabbildern. Damit mußte kein neues Verfahren eingeführt werden; eine Variante eines Verfahrens, das ursprünglich zur Fahrzeugverfolgung entwickelt wurde, war ausreichend (siehe Abb. 5.12).

Das Vertrauen, das wir für unsere Fähigkeit gewannen, systematische Abweichungen zwischen dem ins Bild projizierten Fahrbahnmodell und seiner Anpassung an Bilddaten zu erkennen und zu interpretieren, veranlaßte uns schließlich zu der Vermutung, daß sich die beiden Straßenzüge nicht rechtwinklig schneiden. Als wir schließlich einen Markierungsplan der Kreuzung erhielten, konnten wir uns davon sogar überzeugen.



Abbildung 5.12: Ergebnis nach Prädiktion und Aktualisierung der Fahrbahnlage-schätzung nach Auswertung von 89 Halbbildern, wobei eine relative Geschwindigkeit zwischen der Lage von Kamera und Fahrbahnmodell in der Szene von 0 m/s angenommen wurde. Die initiale projizierte Modellage ist mit weißen Linien, die nach der Anpassung aktualisierte Lage mit dunklen durchgezogenen Linien dargestellt. Weiterhin wurde mit dunklen gestrichelten Linien die Projektion der Modellage visualisiert, die sich aus der Auswertung von nur einem Bild ergab (vgl. Abb. 5.11b), ohne die mit einem Kalmanfilter akkumulierten Schätzungen vorangehender Aufnahmen zu nutzen. Die Fehler, die auf die Fehlzuordnungen von Modellsegmenten an Datensegmente bewegter Fahrzeugabbilder zurückzuführen waren, haben deutlich abgenommen.

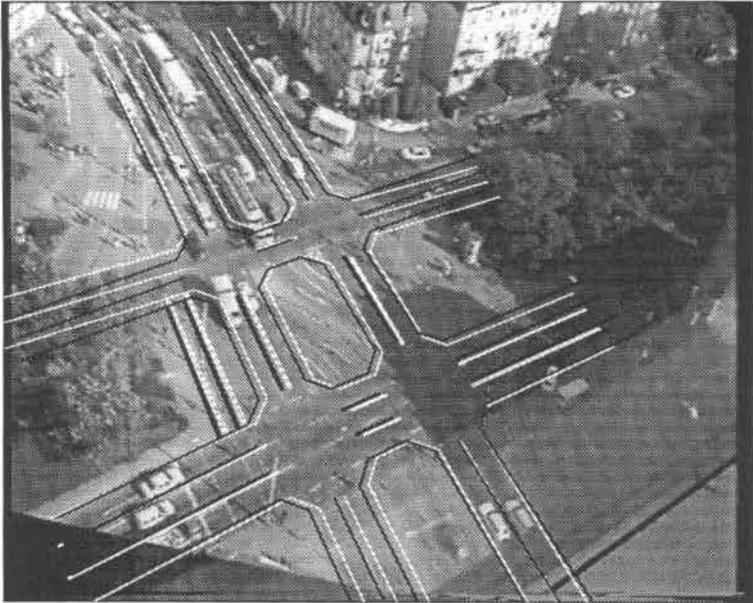


Abbildung 5.13: Anpassungen eines Kreuzungsmodells an das dritte Bild der Folge aus Abb. 5.3(b). Die initiale, ins Bild projizierte Modellage ist mit weißen Linien, die mit der Anpassung aktualisierte Lage mit dunklen Linien dargestellt.

5.2.8 Weitere Modellverfeinerungen

In der Aufnahme der Frankfurter Straßenkreuzung in Abb. 5.3 (b) ist im Gegensatz zur Aufnahme in Abb. 5.3(a) jene Straße deutlicher erfaßt, deren Abbild in der Aufnahme von links oben nach rechts unten verläuft. Das Fahrbahnmodell nach Abb. 5.5 repräsentiert diese Straße unzureichend: von den vier Spuren der Fahrbahn sind im Modell nur zwei dargestellt. Deshalb gab es auch bei der Anpassung dieses groben Modells an Bildkantensegmente Schwierigkeiten (siehe auch Abb. 5.13). [Leuck 95] stellt ein verfeinertes Modell vor, für dessen Entwicklung er jedoch auf den Markierungsplan zurückgriff. Doch die Zeit, die er in die Entwicklung eines feineren Modells steckte, sparte er dann beim Suchen nach geeigneten Parameterkombinationen zur Steuerung des Anpassungsprozesses. Eine Auswertung der Frankfurter Kreuzung mit Hilfe des verfeinerten Modells zeigt Abb. 5.14, eine Auswertung der Ettlinger-Tor-Szene aus Abb. 1.1 zeigt Abb. 5.15.

Zur Entwicklung solcher detaillierter Modelle wird allerdings auf einen Plan der Kreuzung zurückgegriffen. Im Gegensatz zu dem Verfahren von [Koller 92] geschieht die Zuordnung von Modellkomponenten an Bildmerkmale sowie die iterative Korrektur der Kalibrierungsparameter automatisch durch eine Systemkomponente.

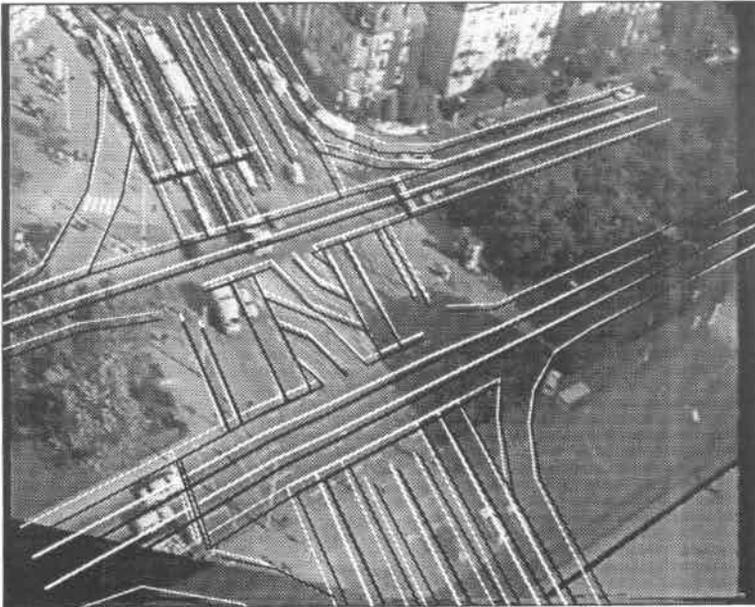


Abbildung 5.14: Anpassungen des verfeinerten Kreuzungsmodells an das dritte Bild der Folge aus Abb. 5.3(b). Die initiale, ins Bild projizierte Modellage ist mit weißen Linien, die mit der Anpassung aktualisierte Lage mit dunklen Linien dargestellt.

Ein weiterer Ansatz, der verspricht, daß man sich in Zukunft von einem Markierungsplan ganz lösen kann, wurde in Zusammenarbeit mit [Mück 94b] erarbeitet. Ausgehend von Untersuchungen über die in der Verkehrsfachliteratur üblichen Begriffe und ihre Beziehungen wurde zunächst ein begriffliches Kreuzungsmodell ermittelt. Hieraus läßt sich ein geometrisches Modell entwickeln, für dessen freie Formparameter zunächst Standardwerte aus den Richtlinien für den Straßenbau eingesetzt werden. In einem weiteren Schritt müssen die Formparameter durch Bildfolgenauswertung iterativ korrigiert werden. Im Gegensatz zu anderen Verfahren liegt damit ein geometrisches Modell vor, das zum einen an Bildfolgen angepaßt werden kann und bei dem zum anderen die Verbindung zu begrifflichen Beschreibungen noch besteht.

5.2.9 Modellgestützte Segmentierung

Nach Anpassung eines Fahrbahnmodells an die Bilddaten steht Wissen über den Verlauf der Fahrspuren zur Verfügung. Dieses Wissen läßt sich beim Segmentieren des optischen Flussfeldes sofort nutzen: Flußvektoren eines Ballungskeimes müssen nicht mehr nur die drei Prädikate aus Anhang A erfüllen, sondern die Rückprojektionen

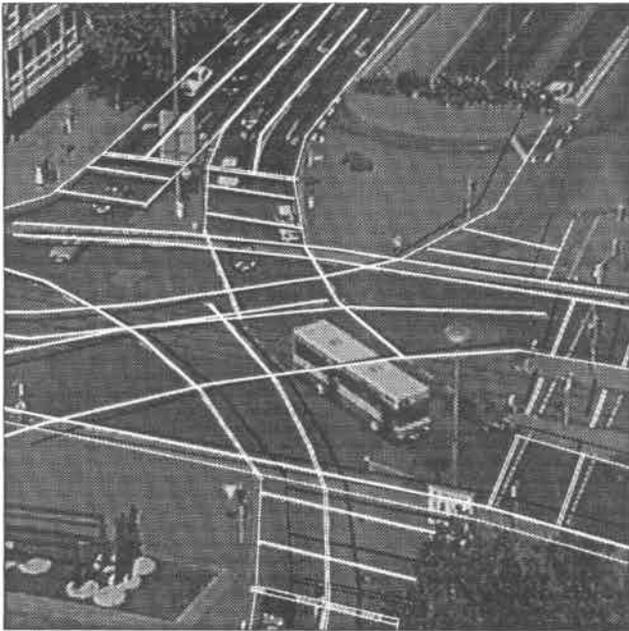


Abbildung 5.15: Anpassungen eines verfeinerten Kreuzungsmodells an das dritte Bild der Folge aus Abb. 1.1. Die initiale, ins Bild projizierte Modellage ist mit weißen Linien, die mit der Anpassung aktualisierte Lage mit dunklen Linien dargestellt. Die Diskrepanzen zwischen dem angepaßten Modell und dem Abbild des Ettlinger-Tor-Platzes im rechten Teil dieser Abbildung zeigen, wo weitere Verbesserungen erforderlich sind: der aus den Fluchtpunktkoordinaten berechnete initiale Schätzwert für die Brennweite sollte auch aktualisiert werden.

der Fußpunkte der Flußvektoren einer Ballung müssen außerdem auf derselben Fahrspur liegen. Ergebnisse dieses Verfahrens zur Detektion bewegter Fahrzeugabbilder finden sich in Abb. 5.16. Darüber hinaus konnte die Detektion bewegter Fahrzeugabbilder dadurch verbessert werden, daß wir nur solche Flußvektoren in der Ballungsanalyse berücksichtigen, deren Orientierung mit der Verkehrsrichtung der jeweiligen Fahrspur kompatibel ist. Dabei wurden von [Leuck 94] im Kreuzungsbereich mehrere sich überschneidende Fahrspuren zugelassen.

Die Struktur der Fahrbahnen bei der Detektion bewegter Fahrzeugabbilder einzusetzen, ist jedoch immer noch eine Heuristik und keine perfekte Regel: ein solches Vorgehen verhindert beispielsweise die Berechnung eines geeigneten Objektbildkandidaten, wenn ein Fahrzeug gerade die Fahrspur wechselt. Sofern aber dieses Fahrzeug nicht über einen längeren Zeitraum auf der Fahrspurmarkierung zwischen zwei

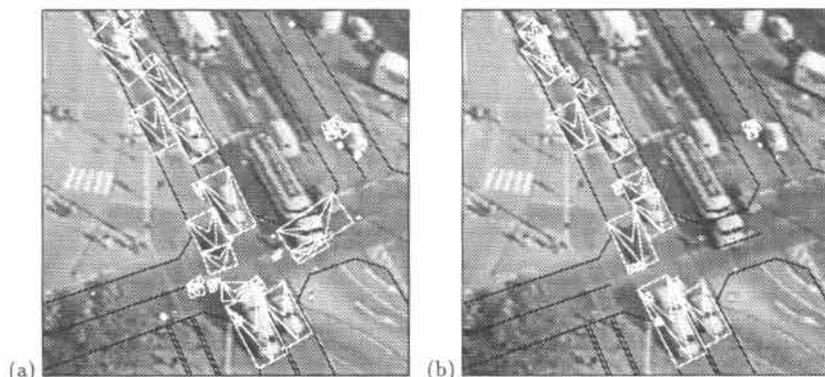


Abbildung 5.16: Ergebnis der modellgestützten Segmentierung eines Flußfeldes, das aus einem Ausschnitt – siehe Abb. 5.4(a) – der linken oberen Ecke einer Aufnahme der Frankfurter Straßenkreuzung in Abb. 5.3(b) aus dieser Bildfolge ermittelt wurde. (a) unterscheidet sich von Abb. 5.4(b) dadurch, daß in der Ballungsanalyse optischer Flußvektoren nicht nur die im Anhang A erklärten drei Prädikate erfüllt sein müssen, sondern die Rückprojektionen der Fußpunkte der Flußvektoren in die Fahrbahnebene auf derselben Fahrspur liegen müssen.

In (b) werden im Gegensatz zu (a) nur solche Flußvektoren in der Ballungsanalyse berücksichtigt, deren Orientierung mit der Verkehrsrichtung der jeweiligen Fahrspur kompatibel ist.

Spuren fährt, kann das Fahrzeug – vorausgesetzt, die Bildfolge ist lang genug – entweder detektiert werden, bevor es die Spur wechselt oder nachdem der Spurwechsel abgeschlossen ist.

5.3 Verfeinerung der Fahrzeugmodelle

5.3.1 Entwurf eines Busmodells

Im Zusammenhang mit Abb. 3.7 wurde diskutiert, daß das Abbild eines Busses mit dem einfachen Fahrzeugmodell, das durch Spezialisierung aus dem generischen Fahrzeugmodell von [Koller 92] entstand (vgl. Abb. 5.17), nur unzureichend beschrieben werden kann. Um die damit verbundenen Schwierigkeiten bei einer Modellanpassung zu beseitigen, erweitern wir das an dieser Stelle in den Auswertungsprozess eingehende Modellwissen. Abb. 5.18 zeigt ein Busmodell, das im Hinblick auf die Verfolgung eines Busses in der Ettlinger-Tor-Szene aus Datenblättern für Standardlinienbusse¹

¹An dieser Stelle sei Herrn Dehnert von den Verkehrsbetrieben Karlsruhe (VBK) gedankt, der auf diese Standardisierung hinwies und entsprechende Daten zur Verfügung stellte.

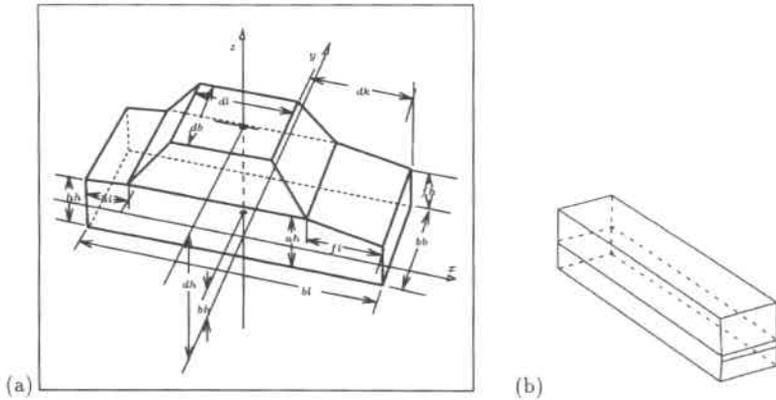


Abbildung 5.17: (a) Generisches polyedrisches 3D-Fahrzeugmodell aus 12 Flächen, 26 Kanten und 16 Ecken von [Koller 92], woraus durch Spezialisierung ein einfaches Busmodell (b) durch Nullsetzen der Formparameter für Frontlänge fl und Hecklänge hl und einem Gleichsetzen von Dachlänge dl und Bodenlänge bl entsteht.

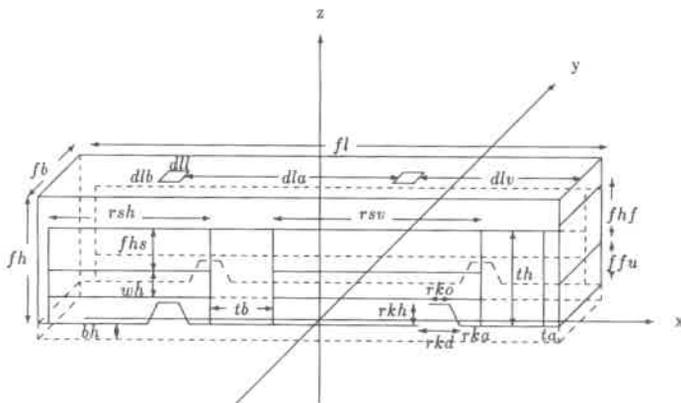


Abbildung 5.18: Abbildung eines Busmodells, das [Mück 94a] nach den Empfehlungen des Verbandes öffentlicher Verkehrsbetriebe aus einer Linienbusstandardisierung entwickelt. Die Formparameter werden in Tab. 5.1 erklärt.

	Beschreibung	min [m]	max [m]
<i>fl</i>	Fahrzeughöhe	11,00	11,48
<i>fh</i>	Fahrzeughöhe		2,68
<i>fb</i>	Fahrzeughöhe		2,50
<i>bh</i>	Bodenhöhe		0,32
<i>tb</i>	Türbreite		1,35
<i>th</i>	Türhöhe	2,05	2,11
<i>ta</i>	Türabstand	0,20	0,30
<i>fhf</i>	Fensterhöhe Frontseite		1,20
<i>fhs</i>	Fensterhöhe Seite	0,90	0,92
<i>wh</i>	Höhe der Werbefläche		0,81
<i>ffu</i>	Frontfenster Unterkante		0,75
<i>rsv</i>	Länge rechte Fensterseite vorne	4,20	4,50
<i>rsh</i>	Länge rechte Fensterseite hinten	3,40	3,50
<i>dll</i>	Dachlукenlänge		0,60
<i>dlb</i>	Dachlукenbreite		0,65
<i>dlv</i>	Dachluke vorne	2,88	3,09
<i>dla</i>	Abstand der Dachlукen	3,86	4,12
<i>rka</i>	Radkastenabstand		0,37
<i>rkh</i>	Radkastenhöhe		0,59
<i>rkd</i>	Radkastendurchmesser		0,89
<i>rko</i>	Radkasten obere Kante		0,59

Tabelle 5.1: Die 21 Formparameter des Busmodells aus Abb. 5.18.

von [Mück 94a] entwickelt wurde. Eine Standardisierung von Linienbussen führte zur Entwicklung des Standardlinienbusses SL I im Juli 1968. Die im Juli 1983 herausgegebene Empfehlung für den Standardlinienbus SL II folgte der erfolgreichen Vorgängerkonzeption SL I, so daß sich bis auf wenige Parameter lediglich das Aussehen zugunsten einer moderneren Gestaltung verändert hat. Tab. 5.1 listet die 21 Formparameter auf, mit denen das in Abb. 5.18 dargestellte Busmodell ausgeprägt werden kann. Aus den Formparametern lassen sich die Eckpunkte des Modells bestimmen. Dazu wird der Ursprung des Modellkoordinatensystems *bh* Meter unterhalb des Mittelpunktes der Busunterplatte festgelegt. Die Achsen eines Modellkoordinatensystems sind der Zeichnung in Abb. 5.18 entsprechend definiert. Das Busmodell besteht aus 64 Modelleckpunkten und 22 Modellflächenstücken (vgl. Abb. 5.19 sowie [Mück 94a]).

5.3.2 Modellierung der Frontleuchten eines Fahrzeugs

In Videoaufnahmen von Fahrzeugen sind die Frontleuchten meist deutlich sichtbar (vgl. Abb. 5.20). Eine Hinzunahme der Frontleuchten in die Fahrzeugmodellbeschreibung führt vor allem bei kontrastarmen Fahrzeugaufnahmen zu deutlich besseren Ergebnissen bei der Fahrzeugverfolgung. In Abb. 5.21 findet man das um Frontleuchten erweiterte Fahrzeugmodell für einen PKW.

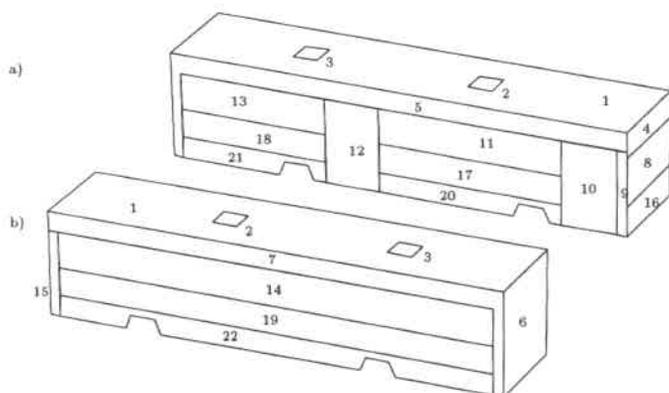


Abbildung 5.19: Numerierung der 22 Modellflächen des Busmodells aus Abb. 5.18 nach [Mück 94a]. Auf der (in Abb. 1.1 dem Betrachter) abgewandten Seite des Busses (b) wurden weniger Facetten als auf der zugewandten Seite (a) modelliert, da die Türen (Nr. 10 und Nr. 12) entfallen und die Fensterfront (Nr. 14), die Werbefläche (Nr. 19) und die Fläche Nr. 22 durchgängig statt unterbrochen wie die entsprechenden Facetten Nr. 11 und Nr. 13, Nr. 17 und Nr. 18 bzw. Nr. 20 und Nr. 21 repräsentiert sind.

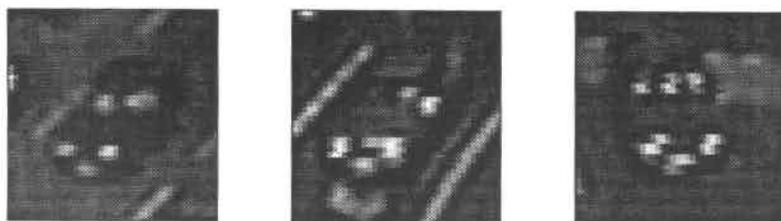


Abbildung 5.20: In kontrastarmen Fahrzeugabbildern sind die Frontleuchten deutlich sichtbar. Eine Hinzunahme der Frontleuchten in das Fahrzeugmodell stabilisiert die Fahrzeugverfolgung.

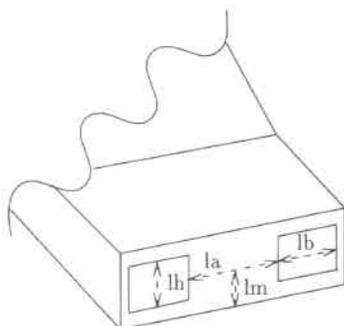


Abbildung 5.21: Das generische PKW-Modell aus Abbildung 5.17(a) wird um zwei Modellfacetten zur Beschreibung der Frontleuchten erweitert, wie an der Fahrzeugfront hier dargestellt. Dabei werden vier neue Formparameter eingeführt: Lampenhöhe lh , Lampenbreite lb , Abstand la der Lampen und die Höhe lm der Lampenmittelpunkte über der Fahrzeugunterseite.

Kapitel 6

Direkte Anpassung von Flächenmodellen an Grauwertgradienten

6.1 Problemstellung

In der Bildauswertung nutzen viele Verfahren zur Lageschätzung und Verfolgung von Objekten aus Bilddaten extrahierte Kantensegmente, zu deren Bestimmung meist eine ganze Reihe von Heuristiken notwendig ist, wie z.B. Kantenelemente zu Kantensegmenten aggregiert werden sollen [Henricsson & Heitger 94]. In der Arbeit von [Otte & Nagel 92] wurde versucht, dies etwas zu vermeiden (siehe Abschnitt 3.4). Solche Heuristiken sind auf das Zusammenwirken geeignet gewählter Schwellwertentscheidungen angewiesen, wobei die Schwellwerte vom Systembenutzer meist experimentell bestimmt werden müssen. Es ist eine keineswegs triviale, zumindest aber eine zeitraubende Aufgabe, alle notwendigen Parameter richtig einzustellen, insbesondere wenn die Parametereinstellungen nicht nur zur Auswertung eines einzelnen Testbildes, sondern zur Auswertung längerer Bildfolgen realer Szenen geeignet sein sollen. Viele Arbeiten (z.B. [Du *et al.* 93]) beschäftigen sich mit der Aufgabe, Bilder und Bildfolgen nach einer Extraktion von Kantensegmenten zu interpretieren, die mit Störungen behaftet sind, deren Endpunkte z.B. oft nur fehlerbehaftet zu bestimmen sind, etwa bedingt durch partielle Verdeckungen der zu verfolgenden Fahrzeuge. Aber auch bei vielen gut sichtbaren Kanten kann es Probleme bei einer Modellanpassung geben (vgl. Abb. 3.7). Die Zuordnung von Modell- zu Datenkantensegmenten erweist sich als eine schwierige Aufgabe. Auch ein menschlicher Beobachter kann oftmals nicht – ohne auf das Grauwertbild zu schauen – die richtige Zuordnung von einem Modell- zu einem Datenkantensegment angeben (vgl. Abb. 6.1). Weiterhin erfordern Verfahren, die eine Modellanpassung auf der Basis von Kantensegment-Zuordnungen durchführen, eine bestimmte Strategie, um die kombinatorische Explosion zu vermeiden [Grimson 90; Zhang 93b; Zhang & Faugeras 92a]. Die zur Bewertung für eine Auswahl von Zuordnungen *heuristisch* gewählte Abstandsfunktion (siehe Anhang C.3) liefert manchmal auch dann kleine Werte, wenn

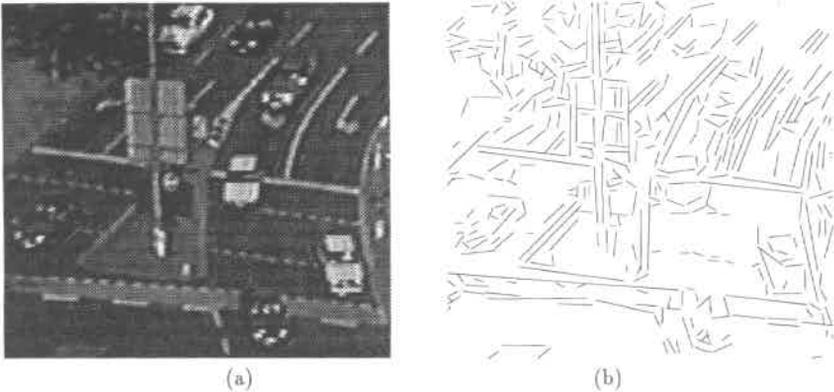


Abbildung 6.1: Ausschnitt aus Abb. 1.1 oben links: während im Grauwertbild (a) die Fahrzeuge für einen menschlichen Beobachter leicht zu erkennen sind, ist die Detektionsaufgabe im auf Kantensegmente reduzierten Datenmaterial (b) wesentlich schwieriger.

das ins Bild projizierte Modell nicht gut zum Datenmaterial paßt. Dies untersuchen wir an einem Beispiel in der Ettlinger-Tor-Szene.

6.2 Untersuchung der Mahalanobisdistanz an einem Beispiel in der Ettlinger-Tor-Szene

Erst die von [Koller 92] entwickelten Darstellungs- und Inspektionswerkzeuge sowie darauf aufbauende Erweiterungen ermöglichten ein rasches Erkennen von fragwürdigen Zuordnungen (vgl. Abb. 6.2): Zwischen langen Modellkantensegmenten und dazu fast senkrecht stehenden kurzen Datenkantensegmenten ergab sich eine kleine Mahalanobisdistanz, obwohl sowohl Winkel- als auch Längendifferenz beträchtlich waren. Es zeigte sich, daß die Anwendung der Mahalanobisdistanz zur Berechnung von Zuordnungen zwischen Modell- und Datenkantensegmenten oftmals auch Zuordnungen lieferte, die für den Beobachter des Systems nicht akzeptabel waren.

Daten- und Modellkantensegmente werden bei [Koller 92] in einer Mittelpunkt-Repräsentation mit Mittelpunkt $(m_\ell, m_n)^T$, Orientierung θ und Länge l dargestellt. Der Abstand zwischen einem Daten- und einem Modellkantensegment wird mit einer Mahalanobisdistanz (siehe Anhang C.3.2 und Anhang C.3.5) durchgeführt, in die eine Fehlerkovarianzmatrix Σ_D für das Datenkantensegment nach Gl. C.35 sowie eine Fehlerkovarianzmatrix Σ_M für das Modellkantensegment nach Gl. C.36 eingehen. Wir reduzieren unsere Untersuchungen der Mahalanobisdistanz zwischen zwei Kantensegmenten auf den Einfluß ihrer Winkel- und Längendifferenz auf ihren Abstand.

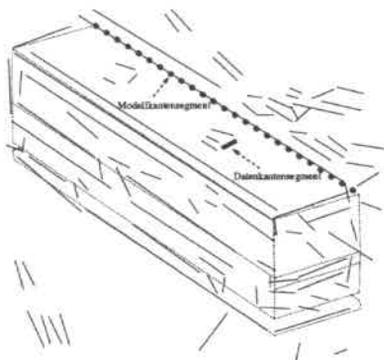


Abbildung 6.2: Berechnete Daten- und Modellkantensegmente aus einem Ausschnitt der 7. Aufnahme der Ettlinger-Tor-Szene (siehe Abb. 1.1), wobei das lange (Modell-) Dachkantensegment (Länge 104,7 Bildelemente) dem kurzen Datenkantensegment (Länge 5,3 Bildelemente) des Lüftungsfensters auf dem Dach des Busses zugeordnet wurde.

Die inverse Kovarianzmatrix $(\Sigma_D + \Sigma_M)^{-1}$ (siehe Gl. C.17) am Beispiel des langen (Modell-) Dachkantensegments und des kurzen Datenkantensegments des Lüftungsfensters auf dem Dach des Busses in der Ettlinger-Tor-Szene (vgl. Abb. 6.2) mit einer für den Betrachter nicht akzeptablen Zuordnung weist folgende Komponenten auf:

$$(\Sigma_D + \Sigma_M)^{-1} = \begin{pmatrix} 47,72 & 0,52 \\ 0,52 & 0,0069 \end{pmatrix}. \quad (6.1)$$

In Abb. 6.3 ist der Akzeptionsbereich (siehe Anhang C.3.2) mit Schwelle 100 dargestellt. Mit einer Winkeldifferenz von $\Delta\theta = +1,08 \text{ rad} \approx 62^\circ$ und einer Längendifferenz von $\Delta l = 99,39$ Bildelemente für die im Beispiel vorliegenden Modell- bzw. Datenkantensegmente erhält man folgenden Abstand:

$$\begin{aligned} & 47,72 * (\Delta\theta)^2 + 2 * 0,52 * \Delta\theta\Delta l + 0,0069 * (\Delta l)^2 \\ = & 55,66 + 111,6 + 68,11 \quad (6.2) \\ = & 235,37; \end{aligned}$$

bei negativer Winkeldifferenz von $\Delta\theta = -1,08 \text{ rad}$ erhält man:

$$\begin{aligned} & 47,72 * (\Delta\theta)^2 + 2 * 0,52 * \Delta\theta\Delta l + 0,0069 * (\Delta l)^2 \\ = & 55,66 - 111,6 + 68,11 \quad (6.3) \\ = & 12,17. \end{aligned}$$

Um den Einfluß der Winkeldifferenz und ihres Vorzeichens zu untersuchen, halten wir $\Delta l = 99,39$ fest. In Abb. 6.4 sind die einzelnen in die Mahalanobisdistanz

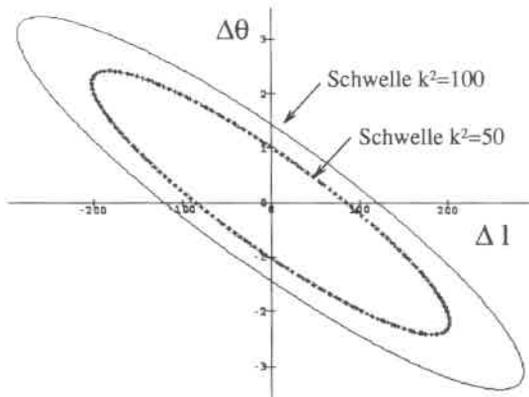


Abbildung 6.3: Akzeptationsbereich am Beispiel des langen (Modell-) Dachkantensegments und des Datenkantensegments des Lüftungsfensters auf dem Dach des Busses in der Ettlinger-Tor-Szene (vgl. Abb. 6.2) mit Schwelle von $k^2 = 50$ bzw. $k^2 = 100$; die Ellipse wird durch die Gleichung $47,72 * (\Delta\theta)^2 + 2 * 0,52 * \Delta\theta\Delta l + 0,0069 * (\Delta l)^2 = k^2$ beschrieben.

eingehenden Summanden in Abhängigkeit von $\Delta\theta$ aufgetragen, Abb. 6.5 zeigt dann die Werte der Mahalanobisdistanz in Abhängigkeit von $\Delta\theta$ bei Vernachlässigung und bei Berücksichtigung der gemischten Kovarianzen.

Berücksichtigt man die gemischten Kovarianzen, so werden positive Winkelabweichungen des Datenkantensegmentes vom Modellkantensegment auf andere Weise gewertet als negative. Das verbirgt sich aber genau hinter der Idee der gemischten Kovarianzen.

- Aus Sicht des Modellkantensegments werden negative Winkeldifferenzen den positiven vorgezogen. Dieses unsymmetrische Verhalten steckt in der Modellierung.
- Aus dem unstetigen Sprung der Mahalanobisdistanz bei einer Korrekturaddition kann man nicht schließen, der Winkeldifferenz kein Vorzeichen zuweisen zu können.
- Die Längendifferenz im Beispiel war groß, doch da es auch ihre Varianz war, reicht die Längendifferenz nicht zur Ablehnung der Zuordnung von Modell- zu Datenkantensegment aus. Gerade bei großen Längendifferenzen kann das negative Vorzeichen der gemischten Kovarianzen den durch die Längendifferenz eingeführten Abstand wieder ausgleichen.

Was bleibt übrig? Es ist die Modellbildung, deren Ergebnisse nicht mehr den in sie gesteckten Erwartungen entsprechen. Man hat hier ein statistisches Abstandsmaß

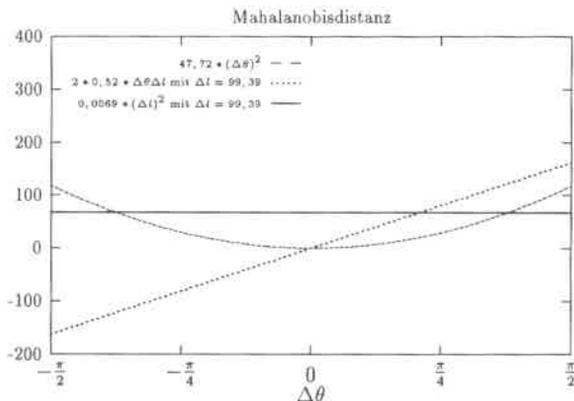


Abbildung 6.4: Untersuchung der einzelnen in die Mahalanobisdistanz zwischen einem Modell- und Datenkantensegment des Busses in der Ettliger-Tor-Szene (vgl. Abb. 6.2) eingehenden Summanden und die Abhängigkeit der Mahalanobisdistanz von $47,72 * (\Delta\theta)^2$, vom gemischten Term $2 * 0,52 * \Delta\theta \Delta l$ und von $0,0069 * (\Delta l)^2$, wobei $\Delta l = 99,39$ gesetzt wurde.

eingeführt, es korrekt durchgerechnet, aber sein Ergebnis entspricht nicht immer dem Wert, den der Betrachter wünscht. Man muß unterscheiden zwischen dem mit einer Kovarianzmatrix berechneten statistischen Abstand der Kantensegmente und dem Abstand, den der menschliche Beobachter der Kantensegmente sieht. Einem menschliche Beobachter der Kantensegmente springt eine Zuordnung von Kantensegmenten mit einer Winkeldifferenz von $\pm\pi/2$ als total falsch ins Auge.

Der Zahlenwert $\Delta\theta/\sigma_{\Delta\theta}$ der durch ihre Standardabweichung dividierten Winkeldifferenz erzielt keinen ausreichend großen Wert. Der Wert von $\Delta\theta/\sigma_{\Delta\theta}$ drückt die Mahalanobisdistanz bei einer Winkeldifferenz von $\pm\pi/2$ nicht über jede Schwelle — in Abbildung 6.5 bleibt die Mahalanobisdistanz bei Winkeldifferenzen von $-\pi/2$ bis $-\pi/4$ nahezu konstant, die negative Winkeldifferenz kann gar nicht so klein werden, daß die Mahalanobisdistanz in den Ablehnungsbereich gezwungen wird.

Die Varianz der Winkeldifferenz, die im wesentlichen die Breite der Akzeptionsellipse bestimmt, ist unabhängig von der Winkeldifferenz. Nach den Überlegungen in Abschnitt C.3.3 lassen sich die Kovarianzen von Daten- und Modellkantensegment getrennt transformieren. Der Beitrag eines Datenkantensegments zur Varianz der Winkeldifferenz ist sogar unabhängig vom Winkel selbst, er ist $2\sigma_1^2/l^2$ (vgl. Gl. C.35).

Man möchte eigentlich bei großem $\Delta\theta$ einen schmalen Akzeptionsbereich, was man, wenn man den Differenzwinkel selbst in die Abstandsberechnung eingehen läßt, wegen der Unabhängigkeit der Standardabweichung $\sigma_{\Delta\theta}$ von $\Delta\theta$ nicht erreicht.

In Anhang C.3.7 werden weitere Abstandsmaße untersucht. Keine der vorgeschla-

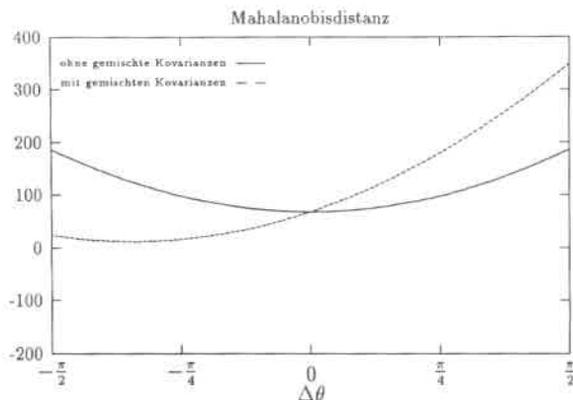


Abbildung 6.5: Mahalanobisdistanz zwischen einem Modell- und Datenkantensegment des Busses in der Ettlinger-Tor-Szene (vgl. Abb. 6.2) in Abhängigkeit von $\Delta\theta$ bei Vernachlässigung und bei Berücksichtigung der gemischten Kovarianzen.

genen Abstandsfunktionen brachte ein akzeptables Ergebnis. Das eigentliche Problem bei diesem Verfahren, das auf Zuordnungen von Kantensegmenten aufbaut, besteht darin, daß man zu wenig Modellwissen nutzt und sich daher permanent gezwungen fühlt, in den Auswertungsprozeß ad hoc Annahmen einzuführen, deren Abklärung aber neue Probleme aufzeigt, was sich beispielsweise in den Überlegungen in Anhang C.3 herauskristallisierte. Folglich wird für die hier zur Diskussion stehende Aufgabe der ganze Ansatz einer Zuordnung von Modellkantensegmenten zu Datenkantensegmenten verworfen und ein grundsätzlich anderer Weg besprochen.

Im folgenden wird gezeigt, daß der Anpassungsprozeß von ins Bild projizierten Modellausprägungen an das Bildmaterial deutlich besser gelingt, wenn man nicht nur Kantensegmente, sondern auch Flächenstücke des Polyedermodells in die Anpassung mit einbezieht. Ohne Annahmen über den Grauwert der Modellflächen selbst werden die Modellflächen — genauer: die in die Bildebene projizierten sichtbaren Modellflächenstücke — mit einer Vorschrift für den Verlauf des Grauwertgradienten als Funktion der Ortskoordinaten auf den Flächenstücken attribuiert und mit einem aus den Bilddaten berechneten (Daten-)Grauwertgradienten verglichen (vgl. Abb. 6.6). Die Modellanpassung wird auch gute Ergebnisse liefern, wenn die Fahrzeuge teilweise verdeckt sind, z.B. durch Laub. In einer solchen Situation versagen klassische Verfahren der Kantensegmentextraktion, denn es ist in diesem Fall schwierig, die zu den Fahrzeug gehörenden Kantensegmente zuverlässig zu ermitteln.

Durch den nun möglichen Verzicht auf die Berechnung von Kantensegmenten bei der Verfolgung von Fahrzeugen wurde der Satz einzustellender Parameter weiter reduziert.

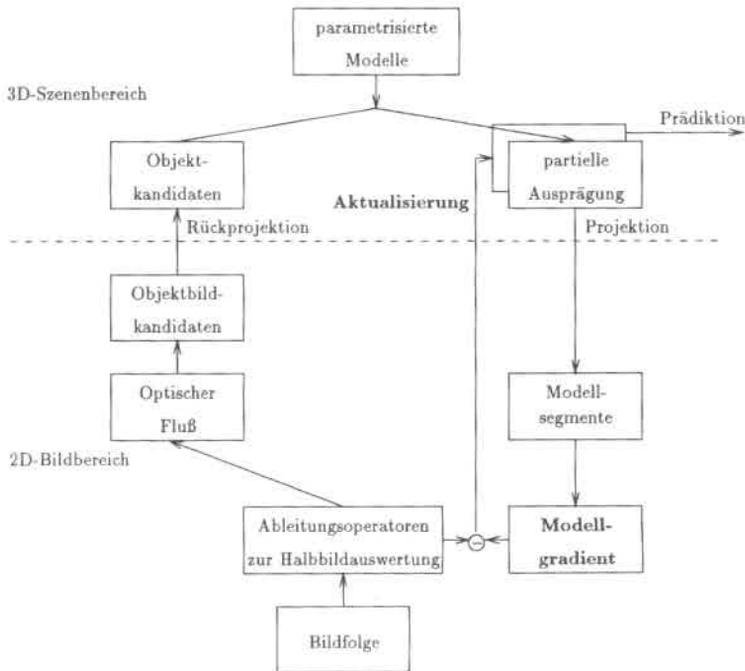


Abbildung 6.6: Detektion und Verfolgung von Fahrzeugen: Im Vergleich mit dem System aus Abb. 5.1 kann auf eine Datensegmentberechnung verzichtet werden: eine Modellausprägung wird mit Hilfe eines Vergleichs eines *Modellgradienten* mit aus den Bilddaten berechneten *Datengradienten* aktualisiert.

6.3 Synthetische Gradientennorm

Betrachtet man ein ins Bild projiziertes Modellkantensegment der Länge l auf der ξ -Achse mit dem Mittelpunkt im Ursprung als δ -Funktion und faltet diese mit einer bivariaten Gaußfunktion

$$\varphi(\xi, \eta) = Ae^{-\frac{1}{2} \left(\frac{\xi^2 + \eta^2}{\sigma^2} \right)}, \quad (6.4)$$

so erhält man als Fläche über der $\xi\eta$ -Ebene (vgl. Abb 6.7):

$$f(\xi, \eta, l) = \begin{cases} f_1(\xi, \eta, l) = \varphi(\xi + \frac{l}{2}, \eta), & \text{falls } \xi < -\frac{l}{2} \\ f_2(\xi, \eta, l) = \varphi(0, \eta), & \text{falls } -\frac{l}{2} \leq \xi < \frac{l}{2} \\ f_3(\xi, \eta, l) = \varphi(\xi - \frac{l}{2}, \eta), & \text{falls } \frac{l}{2} \leq \xi \end{cases} \quad (6.5)$$

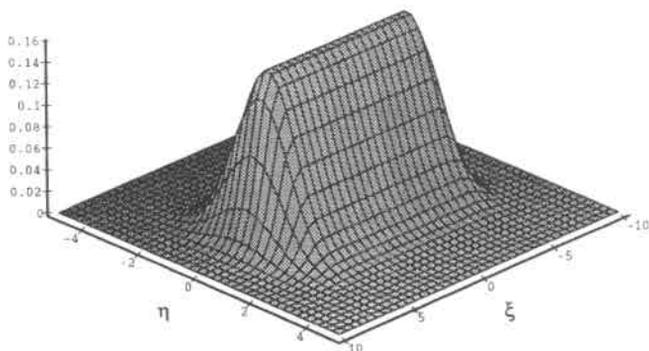


Abbildung 6.7: Ergebnis der Faltung eines auf der ξ -Achse liegenden Kantensegments mit Mittelpunkt $(0, 0)$ und Länge $l = 10$ mit einer 2D-Gaußfunktion, wobei $\sigma = 1$ gewählt wurde.

Diese Funktion $f(\xi, \eta, l)$ läßt sich als eine *synthetische* Gradientennorm auffassen. Wir bezeichnen die Funktion $f(\xi, \eta, l)$ im folgenden auch als *Modellgradientennorm*. Der synthetische Gradient wird mit dem Grauwertgradienten in der Nähe des projizierten Modellkantensegments verglichen. Betrachtet man die einzelnen in Gl. 6.5 eingehenden Funktionen f_1, f_2, f_3 jeweils als Fläche über der $\xi\eta$ -Ebene, so ist die durch f_1 bzw. f_3 definierte Fläche isotrop in einer Halbebene; die Funktion f_2 definiert eine Fläche mit glockenförmigem Durchschnitt senkrecht zur Richtung der ξ -Achse. Die Funktionen f_1 und f_3 dienen der Modellierung der Modellgradientennorm an Endpunkten des Kantensegmentes, die später u.a. als (ins Bild projizierte) Eckpunkte einer Polyedermodellausprägung auftreten. Diese zweidimensionale synthetische Gradientennorm modelliert – im Gegensatz zur Modellierung einer Gradientennorm als eindimensionale δ -Funktion von [Tan *et al.* 94c] – die Grauwertgradienten geeigneter, die durch Faltung der Bilddaten mit einer Gaußfunktion berechnet werden.

Lemma 6.1 Die Modellgradientennorm, d.h. die Funktion $f(\xi, \eta, l)$ nach Gl. 6.5, ist auf ganz \mathbb{R}^3 beliebig oft stetig differenzierbar.

Beweis: Die Gaußfunktion ist beliebig oft stetig differenzierbar. Die Behauptung des Lemmas folgt aus der Konstruktion der stückweise definierten Funktion $f(\xi, \eta, l)$ in Gl. 6.5. \square

Es sei ein in die Bildebene projiziertes Modellkantensegment in der Mittelpunkt-Repräsentation $(m_\xi, m_\eta, \theta, l)$ mit Mittelpunkt $\mathbf{m} = (m_\xi, m_\eta)^T$, Orientierung θ und Länge l (vgl. [Deriche & Faugeras 90]) gegeben. Nach Verschiebung des Kantenseg-

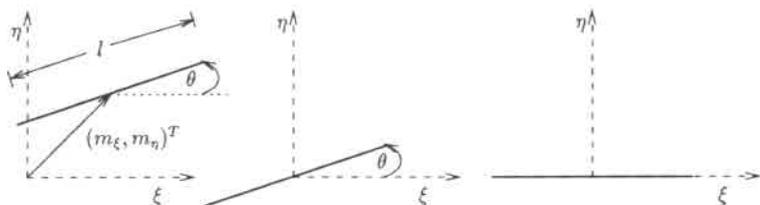


Abbildung 6.8: Gegeben sei ein Kantensegment mit Mittelpunkt $\mathbf{m} = (m_\xi, m_\eta)^T$, Orientierung θ und Länge l . Nach einer Verschiebung des Kantensegments um $-\mathbf{m}$ und anschließender Drehung um den Ursprung mit Drehwinkel $-\theta$ liegt es auf der ξ -Achse mit dem Mittelpunkt im Ursprung.

ments um $-\mathbf{m}$ und Drehung um $-\theta$ liegt es auf der ξ -Achse mit dem Mittelpunkt im Ursprung (vgl. Abb. 6.8). Dabei bezeichne

$$\mathcal{R}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (6.6)$$

die Abbildungsmatrix einer Drehung um den Ursprung um den Winkel θ mit $\mathcal{R}_{-\theta} = \mathcal{R}_\theta^T$. Aus diesen Vorbemerkungen ergibt sich sofort das folgende Lemma:

Lemma 6.2 *Betrachtet man ein Kantensegment mit Mittelpunkt $\mathbf{m} = (m_\xi, m_\eta)^T$, Orientierung θ und Länge l als δ -Funktion und faltet diese mit einer bivariaten Gaußfunktion, so erhält man als Faltungsergebnis*

$$f_{(m_\xi, m_\eta, \theta, l)}(\xi, \eta) = f(\mathcal{R}_\theta^T \begin{pmatrix} \xi - m_\xi \\ \eta - m_\eta \end{pmatrix}, l). \quad (6.7)$$

Die Funktion $f_{(m_\xi, m_\eta, \theta, l)}(\xi, \eta)$ ist für jeden festen Bildpunkt $(\xi, \eta)^T$ auf \mathbb{R}^4 beliebig oft nach ihren Komponenten $(m_\xi, m_\eta, \theta, l)^T$ stetig differenzierbar, da $f(x, y, l)$ und euklidische Transformationen diese Eigenschaften haben.

Wir definieren die *Modellgradientennorm* einer Menge von Kantensegmenten als das Maximum über alle Modellgradientennormen $f_{(m_\xi, m_\eta, \theta, l)}(\xi, \eta)$ der einzelnen Kantensegmente.

Lemma 6.3 *Die Modellgradientennorm einer Menge von Kantensegmenten ist für jeden Bildpunkt $(\xi, \eta)^T$ auf \mathbb{R}^4 stetig.*

Beweis: Jede Modellgradientennorm $f_{(m_\xi, m_\eta, \theta, l)}(\xi, \eta)$ der einzelnen Kantensegmente ist auf \mathbb{R}^4 stetig und bildet den \mathbb{R}^4 auf das Intervall $(0, A]$ ab (Surjektivität). Mit der Stetigkeit der einzelnen Funktionen ist auch ihr Maximum stetig (und surjektiv). \square

An den Punkten $(\xi, \eta) \in \mathbb{R}^2$, an denen sich die Modellgradientennormen verschiedener Kantensegmente gegenseitig durchdringen, ist die Modellgradientennorm der Menge der Kantensegmente *nicht* differenzierbar. Alle kritischen Stellen bilden zusammen jedoch nur einen *eindimensionalen* Teilraum des \mathbb{R}^2 . Die Modellgradientennorm einer Menge von Kantensegmenten ist bis auf diese wenigen Stellen beliebig oft stetig differenzierbar.

Wenn man die gesamte Szenenskizze mit einer Gaußfunktion faltet, dann ist das Ergebnis als Summe (und nicht als Maximum) einzelner Gaußfunktionen zwar überall differenzierbar, liefert aber bei Modellkantensegmenten, die in enger Nachbarschaft liegen, eine deutlich höhere Gradientennorm als bei weiter entfernt voneinander liegenden Modellkantensegmenten. Deshalb haben wir uns nicht für eine Modellierung durch die Summen der Faltungen einer Gaußfunktion mit den Modellsegmenten entschieden.

6.4 MAP-Schätzung

Durch das folgende Anpaßverfahren schätzen wir zu jedem Halbbildzeitpunkt t_k einen fünfdimensionalen Zustandsvektor \mathbf{z}_k , der aus der Position $(p_\xi, p_\eta)^T$ und Orientierung des Fahrzeugmodells in der Fahrbahnebene sowie den Beträgen seiner translatorischen Geschwindigkeit v und seiner Winkelgeschwindigkeit ω besteht.

Es bezeichne $(\hat{\mathbf{z}}_k^-, P_k^-)$ bzw. $(\hat{\mathbf{z}}_k^+, P_k^+)$ die geschätzten Zustandsvektoren und ihre Kovarianzmatrizen vor bzw. nach der Aktualisierung unter Auswertung der Messung zum Halbbildzeitpunkt t_k .

Sei ein Wert $\hat{\mathbf{z}}_k^-$ für den Zustandsvektor \mathbf{z}_k zum Halbbildzeitpunkt t_k bereits ermittelt, entweder als initiale Schätzung $\hat{\mathbf{z}}_0^-$ durch Segmentieren des Flußfeldes (siehe Kapitel 3) oder als präzidiertes Schätzwert $\hat{\mathbf{z}}_k^-$, der mit Hilfe einer Zustandsübergangsfunktion die Fahrzeugbewegung als eine stationäre Kreisfahrt beschreibt [Koller 92] und aus dem aktualisierten Schätzwert $\hat{\mathbf{z}}_{k-1}^+$ zum vorangehenden Zeitpunkt t_{k-1} gewonnen wird.

Man berechne aus dem Startzustand eine Modellausprägung und projiziere sie unter Berücksichtigung verdeckender Flächen in die Bildebene. Diese Szenenskizze besteht aus einer Menge von Modellkantensegmenten. Wir bezeichnen in Abhängigkeit vom Zustandsvektor \mathbf{z}_k mit $h_k(\xi, \eta, \mathbf{z}_k)$ an jeder Bildposition $(\xi, \eta)^T$ des interpolierten Halbbildes zum Zeitpunkt t_k die synthetische Gradientennorm der Menge von Modellkantensegmenten gemäß Abschnitt 6.3.

Wir nehmen an, die euklidische Norm $\|\nabla g_k(\xi, \eta)\|_2$ des Grauwertgradienten an jeder Bildposition $(\xi, \eta)^T$ des interpolierten Halbbildes zum Zeitpunkt t_k sei gleich dem Funktionswert der Meßfunktion $h_k(\xi, \eta, \mathbf{z}_k)$ plus einem weißen Gaußschen Rauschen v_k mit Varianz r :

$$z_k(\xi, \eta) = \|\nabla g_k(\xi, \eta)\|_2 = h_k(\xi, \eta, \mathbf{z}_k) + v_k. \quad (6.8)$$

Die Meßfunktion $h_k(\xi, \eta, \mathbf{z}_k)$ ist nichtlinear im Zustandsvektor \mathbf{z}_k , denn die perspektivische Projektion ist nichtlinear.

Unter der weiteren Annahme, der Zustand \mathbf{x}_k sei normalverteilt um den Schätzwert $\hat{\mathbf{x}}_k^-$ mit Kovarianz P_k^- , kann eine Maximum-a-posteriori-Schätzung (MAP-Schätzung) formuliert werden, bei der folgende Bewertungsfunktion zu minimieren ist:

$$\frac{1}{2nr} \sum_{\xi, \eta} (\|\nabla g_k(\xi, \eta)\|_2 - h_k(\xi, \eta, \mathbf{x}_k))^2 + \frac{1}{2} (\mathbf{x}_k - \hat{\mathbf{x}}_k^-)^T P_k^{-1} (\mathbf{x}_k - \hat{\mathbf{x}}_k^-) \rightarrow \min_{\mathbf{x}_k},$$

was zu einem Aktualisierungsschritt eines *iterativen erweiterten Kalmanfilters* (IEKF) [Bar-Shalom & Fortmann 88; Gelb 74] führt. n bezeichnet die Größe des Summationsbereiches.

Die partielle Ableitung der letzten Gleichung nach den Komponenten des Zustandsvektors \mathbf{x}_k liefert:

$$-\frac{1}{nr} \sum_{\xi, \eta} (\|\nabla g_k(\xi, \eta)\|_2 - h_k(\xi, \eta, \mathbf{x}_k)) \left. \frac{\partial h_k(\xi, \eta, \mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_k} + P_k^{-1} (\mathbf{x}_k - \hat{\mathbf{x}}_k^-). \quad (6.9)$$

Die Jakobimatrix

$$H_k(\xi, \eta, \mathbf{x}) = \frac{\partial h_k(\xi, \eta, \mathbf{x})}{\partial \mathbf{x}} \quad (6.10)$$

der Meßfunktion wird in Anhang D.4 hergeleitet.

In unserer Implementierung wird die Minimierung mit Hilfe des Gauß-Newton-Verfahrens durchgeführt, das mit der Levenberg-Marquardt-Iteration modifiziert wurde [Scales 85; Daniilidis 92; Koller 92], um Konvergenz im Falle von Singularitäten der nichtlinearen Meßfunktion zu sichern.

Am Ende der Gauß-Newton-Iteration aktualisieren wir die Kovarianzmatrix P_k^- wie folgt:

$$P_k^+ = \left(\frac{1}{nr} \sum_{\xi, \eta} \left(\{H_k(\xi, \eta, \hat{\mathbf{x}}_k^+)\}^T H_k(\xi, \eta, \hat{\mathbf{x}}_k^+) \right) + P_k^{-1} \right)^{-1}, \quad (6.11)$$

wobei die Jakobimatrix $H_k(\xi, \eta, \mathbf{x})$ der Meßfunktion an der Stelle $\hat{\mathbf{x}}_k^+$ der aktualisierten Zustandsschätzung ausgewertet wird. Anschließend werden der Zustandsvektor $\hat{\mathbf{x}}_k^+$ und seine Kovarianzmatrix P_k^+ unter der Annahme einer stationären Kreisfahrt als Bewegungsmodell in das nächste Halbbild präzidiert.

6.5 Diskussion der Abstimmungsparameter des Verfahrens

6.5.1 Modellgradientennorm

Die Modellgradientennorm wird mit einem Gaußschen Faltungskern mit $\sigma = 1.5$ [Bildpunkten] (siehe Gl. 6.4) konstruiert. Der Summationsbereich und seine Größe

werden wie folgt bestimmt: wir summieren über alle Bildpunkte $(\xi, \eta)^T$, an deren Stelle der Wert der Modellgradientennorm ein Prozent seines Maximalwerts überschreitet.

Der Parameter A bezeichnet den Kontrast unserer Modellgradientennorm (siehe Gl. 6.4). Er wird aus den Grauwerten des aktuellen Bildes durch $\frac{1}{n} \sqrt{\sum_{\xi, \eta} \|\nabla g_k(\xi, \eta)\|_2^2}$ geschätzt. In einer ersten Implementierung stellte sich heraus, daß dieser Parameter an den Bildausschnitt angepaßt werden muß, in dem sich das Abbild des Fahrzeuges befindet, dessen Lageparameter geschätzt werden soll. Ein falscher Wert für A liefert eine verzerrte Differenz für $(\|\nabla g_k(\xi, \eta)\|_2 - h_k(\xi, \eta, \mathbf{z}_k))^2$. In unseren Bilddaten variiert der Parameter A zwischen 10 und 20.

6.5.2 Grauwertgradient

Die Bildfolge wird mit trivariaten Gaußfunktionen mit einer räumlichen Skala von $\sigma_x = \sigma_y = 1.25$ [Bildpunkte] und einer zeitlichen Skala von $\sigma_t = 1$ [1/50s] gefaltet, wobei die Zeilenverschränkung und der zeitliche Versatz einzelner Halbbilder (vgl. Abschnitt 3.2) berücksichtigt werden.

6.5.3 Kalmanfilter

Da die Zustandsaktualisierung invariant gegenüber einer Reskalierung von r und P_k^- mit demselben Faktor ist, können wir die Varianz r des Meßrauschens auf einen festen Wert setzen, z.B. auf 10. Unter der Annahme, daß die Elemente außerhalb der Hauptdiagonalen von der Startkovarianzmatrix verschwinden, benutzen wir als Startwerte $\sigma_{p_{x_0}} = \sigma_{p_{y_0}} = 2$ m, $\sigma_{\phi_0} = 0.1$ rad, $\sigma_{v_0} = 0.1$ m/Halbbild, $\sigma_{\omega_0} = 0.01$ rad/Halbbild. Wir benutzen ein Prozeßrauschen von $\sigma_v = \sqrt{2 \cdot 10^{-3}}$ m/Halbbild und $\sigma_\omega = \sqrt{2 \cdot 10^{-7}}$ rad/Halbbild.

6.6 Ergebnisse experimenteller Untersuchungen

6.6.1 Ettlinger Tor

Zunächst testen wir das neu entwickelte Verfahren mit der Ettlinger-Tor-Bildfolge aus Abb. 1.1, die aus 50 Aufnahmen besteht. Die Anpassung des Busmodells aus Abb. 5.18 an die Grauwertgradienten des dritten Bildes¹ dieser Folge wird in Abb. 6.9 erläutert.

¹Die zur Ermittlung der Grauwertableitungen erforderlichen diskreten Faltungskerne erstrecken sich in der Zeitachse über insgesamt 7 Halbbilder, deren Abstand jeweils 20 msec beträgt. In die Auswertung eines Halbbildes werden damit drei Halbbilder zusätzlich zu dem untersuchten Halbbild aus vorangehenden Aufnahmen benötigt. Deshalb beginnen die folgenden experimentellen Untersuchungen nicht mit dem ersten Bild der jeweiligen Folge.

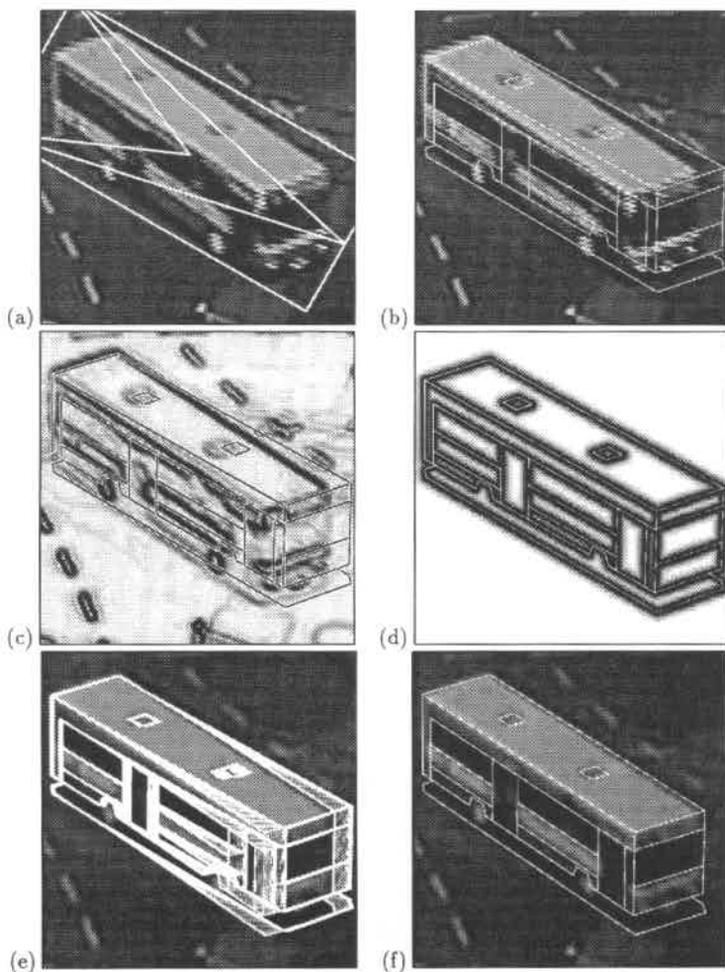


Abbildung 6.9: Beispiel einer Anpassung des Busmodells aus Abb. 5.18 an die Grauwertgradienten, die für den Zeitpunkt des zweiten Halbbildes der dritten Aufnahme der Ettliger-Tor-Folge berechnet wurden: (a) zeigt einen für das Busabbild detektierten Objektkandidaten; (b) die daraus berechnete initiale Modellausprägung; (c) den aus den Bilddaten berechneten Datengradienten, wobei große Werte der Grauwertgradientennorm mit dunklen Punkten visualisiert werden; (d) die aus (b) berechnete synthetische Gradientennorm (vgl. Abschnitt 6.3, insbesondere Abb. 6.7); (e) eine iterative Anpassung aufgrund der MAP-Schätzung nach Gl. 6.9; (f) die endgültige Aktualisierung der Modellausprägung, wobei ein tiefpaßgefiltertes Bild unterlegt ist, bei dem die fehlenden Zeilen des Halbbildes räumlich-zeitlich interpoliert wurden (siehe auch Abb. 6.10).

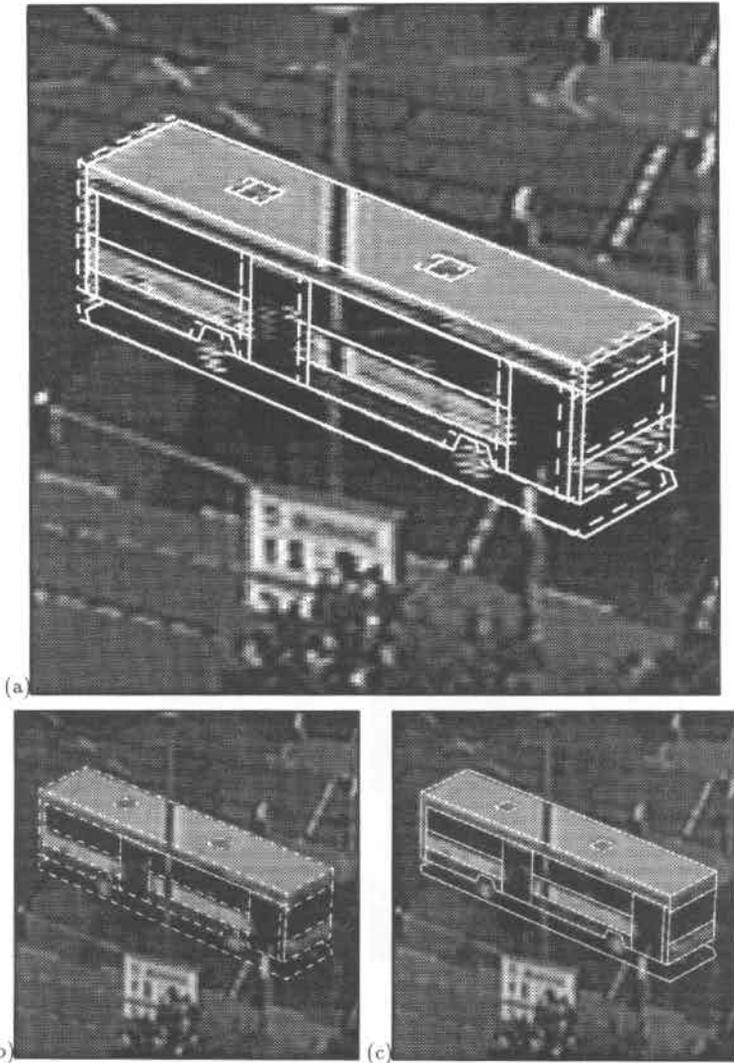


Abbildung 6.10: (a) Vollbild, das zwei zeitlich versetzte Halbbilder umfaßt. Überlagert sind die Lageschätzungen für das Busmodell zum Zeitpunkt des ersten (gestrichelte Linien) sowie des zweiten Halbbildes (durchgezogene Linien); (b) räumlich-zeitlich interpoliertes erstes Halbbild mit überlagerten Lageschätzung zum Zeitpunkt des ersten Halbbildes; (c) räumlich-zeitlich interpoliertes zweites Halbbild mit überlagerten Lageschätzung zum Zeitpunkt des zweiten Halbbildes.

2. Halbbild der 3. Aufnahme



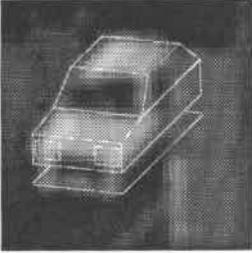
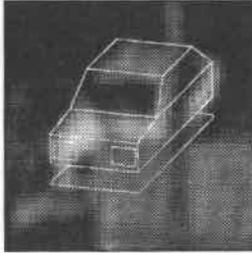
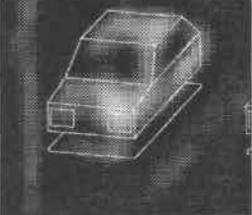
1. Halbbild der 25. Aufnahme



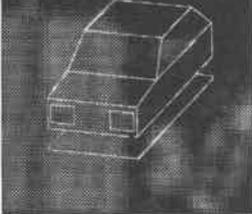
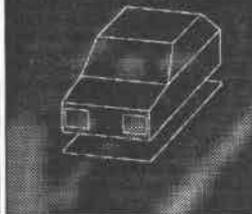
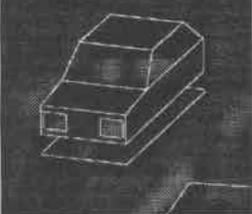
1. Halbbild der 48. Aufnahme



#1



#2



#3

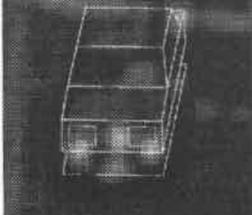
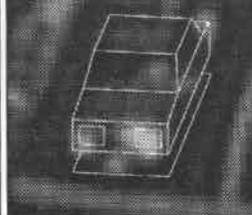
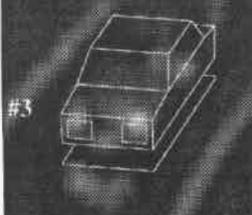


Abbildung 6.11: Die Auswertung der Ettliger-Tor-Bildfolge, visualisiert am Beispiel der 3., 25. und 48. Aufnahme: Die Anpassung der drei kleinen Fahrzeuge, die sich im oberen linken Bereich der Folge bewegen, wird in vergrößerten Ausschnitten dargestellt: Probleme bei der Auswertung des Abbildes von Fahrzeug Nr. 11, das sich im unteren rechten Teil der aufgezeichneten Bildfolge vor dem Bus bewegt, werden genauer in den Abbildungen 6.12-6.13 diskutiert.

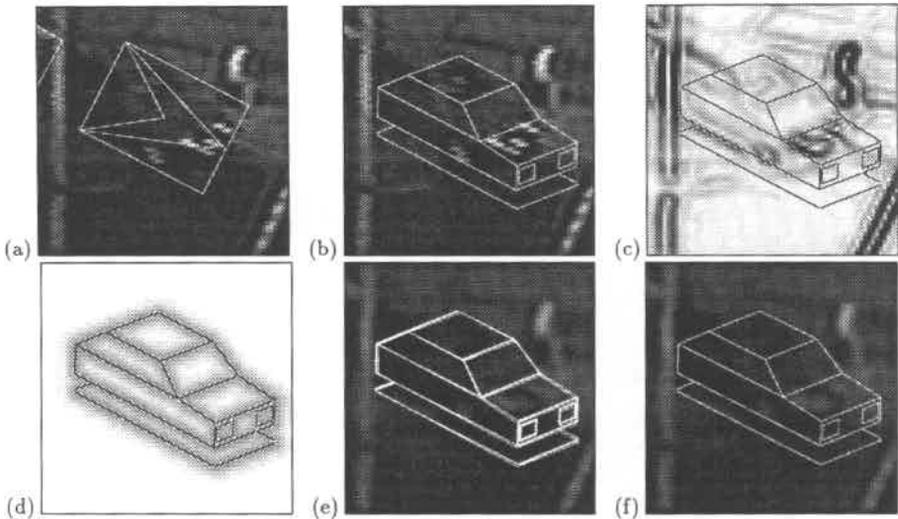


Abbildung 6.12: Der initial detektierte bewegte Bildbereich (a) überdeckt nur teilweise das Fahrzeugabbild, denn das Fahrzeug ist teilweise vom Laternenpfosten verdeckt (vgl. Abb. 6.11). Damit ist der Startwert für die Zustandsschätzung (b) nicht sehr genau. Aus den Grauwertgradienten (c) wird der durchschnittliche Bildkontrast zu $A = 9.5$ (siehe Gl. 6.4 und Abschnitt 6.5.1) geschätzt und zur Erzeugung der synthetischen Gradientennorm (d) herangezogen. Ausgehend von der schlechten initialen Lageschätzung konvergiert die MAP-Iteration (e) nicht zum korrekten Zustandsvektor; (f) zeigt die Aktualisierung der Modellausprägung.

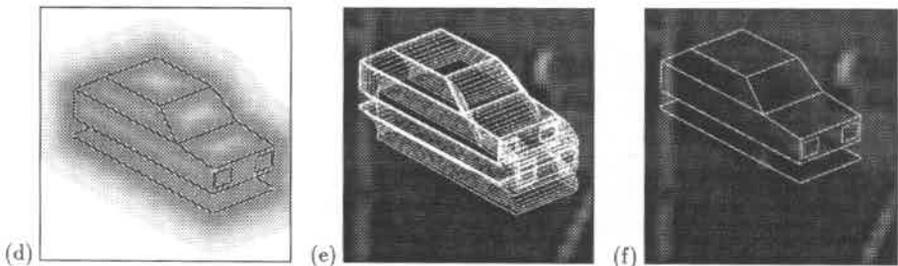


Abbildung 6.13: Erhöht man die Skala σ (siehe Abschnitt 6.5.1 sowie Gl. 6.4) in Vergleich mit der Auswertung in Abb. 6.12, konvergiert die Iteration in der Nähe des korrekten Zustandsvektors.

Die in Abb. 6.9(e) dargestellten Verarbeitungsschritte werden im folgenden näher erläutert: In unserem System werden zeilenverschränkte Halbbilder interpoliert, indem die Bildfolge räumlich-zeitlich mit Gaußmasken gefaltet wird. Die Auswertung zeilenverschränkter Halbbilder zeigt auch Abb. 6.10.

Weitere Anpaßergebnisse finden sich in Abb. 6.11.

Probleme gibt es – wie auch schon in Abschnitt 3.5 – bei der Auswertung des Abbildes von Fahrzeug Nr. 11, das sich im unteren rechten Teil der aufgezeichneten Bildfolge vor dem Bus bewegt, da der initial detektierte bewegte Bildbereich aufgrund des Laternenpostens das Fahrzeug Nr. 11 nicht ganz überdeckt. Verwendet man diese nicht sehr genaue Lageschätzung als Startwert, konvergiert die MAP-Schätzung nicht zum korrekten Zustandsvektor, siehe Abb. 6.12.

Erhöht man die Skala σ (siehe Abschnitt 6.5.1 sowie Gl. 6.4), so konvergiert die MAP-Iteration in der Nähe des korrekten Zustandsvektors. Jedoch können mit diesem vergrößerten σ dann die anderen Fahrzeuge nicht korrekt verfolgt werden. Daher initialisieren wir die Verfolgung von Fahrzeug Nr. 11 mit einem Startwert, der fünf Halbbilder später berechnet wird, wenn das Fahrzeug nicht mehr so stark verdeckt ist wie in den ersten Halbbildern. Dann konvergiert das Verfahren auch für diesen kritischen Fall mit denselben Verfahrensparametern wie für alle übrigen Fahrzeuge.

Die Verfolgung der Fahrzeuge liefert ähnliche Trajektorien wie in Abb. 3.5. Jedoch können die Trajektorien jetzt mit einem einheitlichen Parametersatz ermittelt werden, der auch für die anderen Bildfolgen gültig ist, die wir in den folgenden Abschnitten zeigen.

6.6.2 Frankfurter Kreuzung

Als zweites Beispiel zeigen wir die Ergebnisse einer Auswertung einer Bildfolge von 200 Aufnahmen, die an einer Frankfurter Kreuzung aufgenommen wurde. Die Abbilder eines Fahrzeugs überdecken einen Bildausschnitt von 50×30 Bildpunkten eines (720×576) Bildes. Das dritte Bild dieser Folge wird in Abb. 6.14 dargestellt, die Fahrzeuge dort auch numeriert. Die berechneten Anpassungsergebnisse in Abb. 6.15

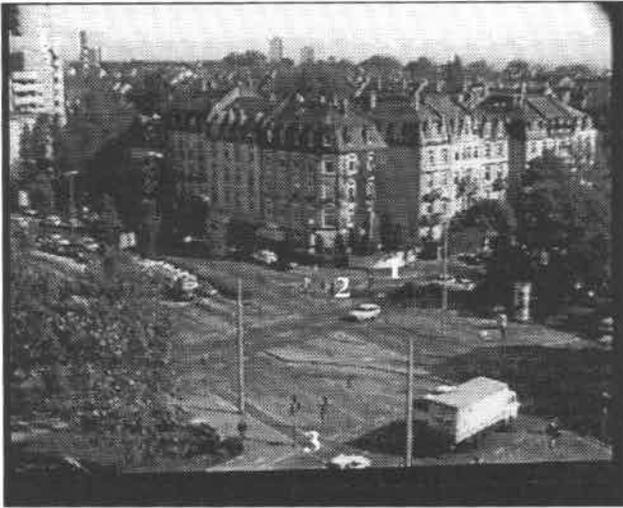


Abbildung 6.14: Dritte Aufnahme der Bildfolge einer Frankfurter Kreuzung.

zeigen die Auswertung der 50., 90. und 96. Aufnahme der Folge, insbesondere die Fahrzeuge mit Nr. 1 bis Nr. 3. Trajektorien sind in Abb. 6.16 dargestellt. Fahrzeug Nr. 4 tritt erst später in das Gesichtsfeld der Kamera ein und wird dann verfolgt. Alle vier Fahrzeuge konnten verfolgt werden, obwohl sie teilweise durch einen Fahnenmast oder durch Blätter verdeckt waren. Darüber hinaus wurde Fahrzeug Nr. 2 korrekt verfolgt, wenn es das Gesichtsfeld der Kamera verläßt. Der LKW in Abb. 6.14 wurde nicht verfolgt, weil für ihn noch kein geeignetes Modell entworfen worden ist.

6.6.3 Durlacher Tor

Die Durlacher-Tor-Bildfolge wurde von [Koller 92; Koller *et al.* 93a] aufgezeichnet und ausgewertet. [Koller *et al.* 93a] haben die Fahrzeuge Nr. 6, Nr. 7 und Nr. 9 verfolgt. Mit ihrem Ansatz konnten diese Autoren allerdings die Fahrzeuge Nr. 1, Nr. 2 und Nr. 10 nicht verfolgen. Die anderen Fahrzeuge konnten verfolgt werden, wenn man für jedes Fahrzeug einen eigenen Parametersatz bestimmt hat. Unser

1. Halbbild der 50. Aufnahme 1. Halbbild der 90. Aufnahme 1. Halbbild der 96. Aufnahme

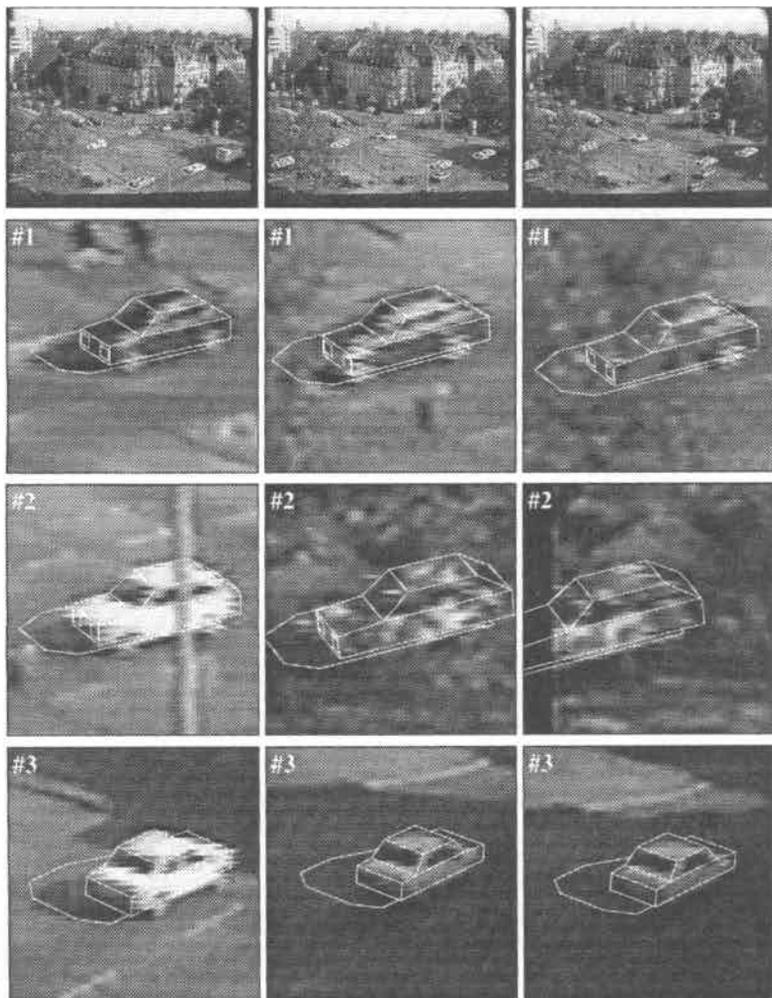


Abbildung 6.15: Auswertungsergebnisse einer Bildfolge, die an einer Frankfurter Kreuzung aufgezeichnet wurde: Fahrzeug Nr. 2 ist teilweise durch einen Fahnenmast verdeckt (erste Spalte). In der 90. Aufnahme fährt Fahrzeug Nr. 1 in den Teil der Straße, der teilweise durch Blätter verdeckt ist (zweite Spalte). In der 96. Aufnahme verläßt Fahrzeug Nr. 2 das Gesichtsfeld der Kamera (dritte Spalte).

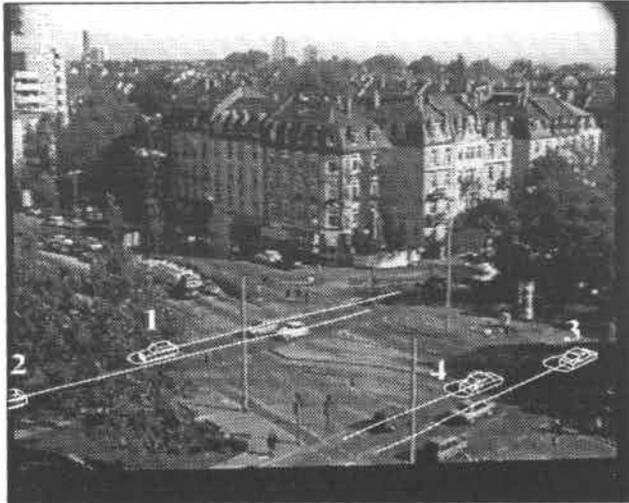


Abbildung 6.16: 100. Aufnahme der Frankfurter Kreuzung mit berechneten Trajektorien nach Auswertung von 200 Halbbildern.

neuer Ansatz erlaubt es, alle Fahrzeuge mit demselben Parametersatz zu verfolgen. Ergebnisse unserer Auswertung zeigt Abb. 6.17.

Wir gehen kurz auf die schwieriger zu verfolgenden Fahrzeuge ein: Um Fahrzeug Nr. 9 zu verfolgen, waren [Koller *et al.* 93a] gezwungen, die Initialisierung leicht von Hand zu korrigieren. Unsere Initialisierung ist besser als jene, die [Koller *et al.* 93a] zur Verfügung stand, da wir ein dichtes optisches Flussfeld segmentieren, vgl. Abschnitt 3.3. Die initiale Lageschätzung für Fahrzeug Nr. 2 ist immer noch nicht korrekt, da es in der dritten Aufnahme der Folge gerade erst in das Gesichtsfeld der Kamera eintritt. Das Bild von Fahrzeug Nr. 10 ist teilweise durch die eingeblendete Zeitangabe verdeckt, siehe erste Spalte in Abb. 6.17. Nach der Auswertung von 15 Halbbildern stabilisierte sich die Verfolgung von Fahrzeug Nr. 10, nach Auswertung von 25 Halbbildern sogar auch für Fahrzeug Nr. 2 (siehe zweite Spalte in Abb. 6.17).

Abb. 6.18 zeigt nochmals die Halbbildauswertung, indem die Lageschätzungen von Fahrzeug Nr. 10, das sich von links nach rechts bewegt, für beide Halbbilder dieser Aufnahme eingeblendet sind. Die Positionsschätzungen sind damit so genau, daß sogar die Bewegung von Halbbild zu Halbbild zuverlässig erfaßt und beschrieben werden kann.

Abb. 6.19 zeigt überlagerte Trajektorien nach Auswertung der gesamten Bildfolge.

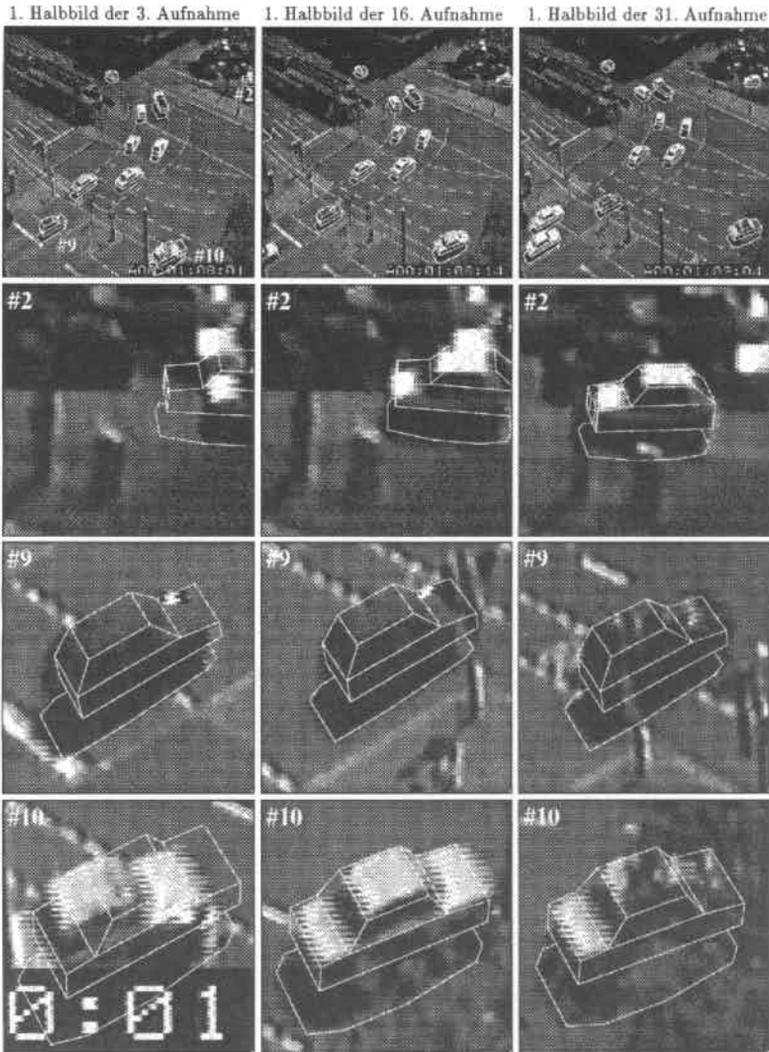


Abbildung 6.17: Die Auswertung einer Bildfolge, die [Koller 92] am Durlacher-Tor-Platz in Karlsruhe aufgenommen hat. Die initiale Lageschätzung für Fahrzeug Nr. 2 ist nicht korrekt, da es gerade in das Gesichtsfeld der Kamera eintritt. Das Bild von Fahrzeug Nr. 10 ist teilweise durch die eingblendete Zeitangabe verdeckt, siehe erste Spalte. Nach der Auswertung von 15 Halbbildern stabilisierte sich die Verfolgung von Fahrzeug Nr. 10, nach Auswertung von 25 Halbbildern sogar auch für Fahrzeug Nr. 2 (siehe zweite Spalte). Im unteren rechten Quadranten der 31. Aufnahme fährt Fahrzeug Nr. 10 gerade in den Bereich der Straße, der von den Blättern verdeckt wird.

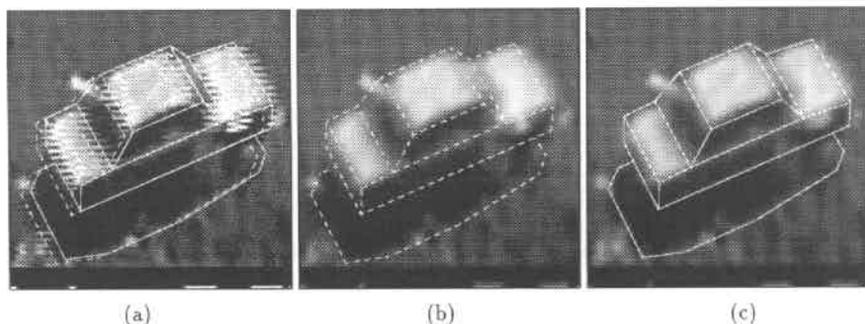


Abbildung 6.18: Auswertung zeilenverschränkter Halbbilder: (a) zeigt das 16. Bild der Durlacher-Tor-Folge, wobei für beide Halbbilder dieser Aufnahme die Lageschätzungen dargestellt sind; (b) Auswertung des ersten Halbbildes; (c) Auswertung des zweiten Halbbildes.

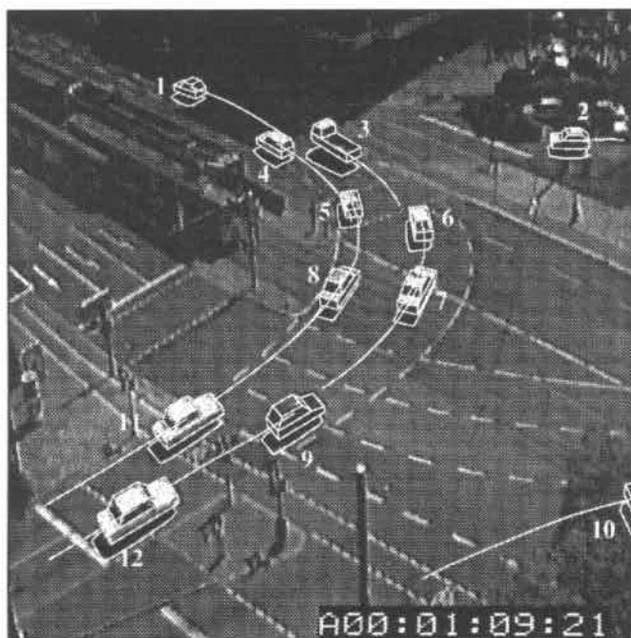


Abbildung 6.19: 48. Bild der Durlacher-Tor-Folge: Mit unserem neuen Ansatz können wir alle 12 Fahrzeuge mit denselben Parametern verfolgen.

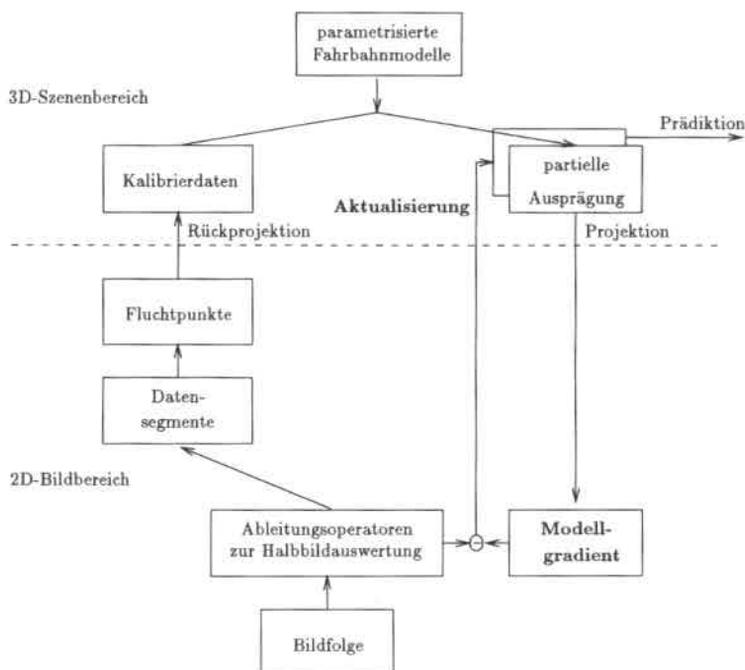


Abbildung 6.20: Initialisierung und Anpassung eines Fahrbahnmodells: Wie in Abb. 5.2 wird in diesem Teilsystem eine Initialisierung aus Fluchtpunkten ermittelt, die ihrerseits aus Datensegmenten berechnet werden. Im Gegensatz zu Abb. 5.2 werden Datensegmente jedoch nur für die Initialisierung des modellgestützten Anpassungsverfahrens benötigt, eine Modellausprägung wird – wie in Abb. 6.6 für Fahrzeuge – mit Hilfe eines Vergleichs eines *Modellgradienten* mit aus den Bilddaten berechneten *Datengradienten* aktualisiert.

6.7 Fahrbahn

Eine Flächenanpassung wird auch bei der Anpassung des Fahrbahnmodells durchgeführt (vgl. Abb. 6.20). Abb. 6.21 zeigt eine iterative Anpassung und Abb. 6.22 die aktualisierte Lageschätzung. Allerdings ist der Fangbereich dieses Verfahrens z. Zt. noch kleiner als bei dem auf einer Extraktion und Zuordnung von Kantensegmenten beruhenden Verfahren, so daß zu ungenaue Initialisierungen nicht kompensiert werden können.

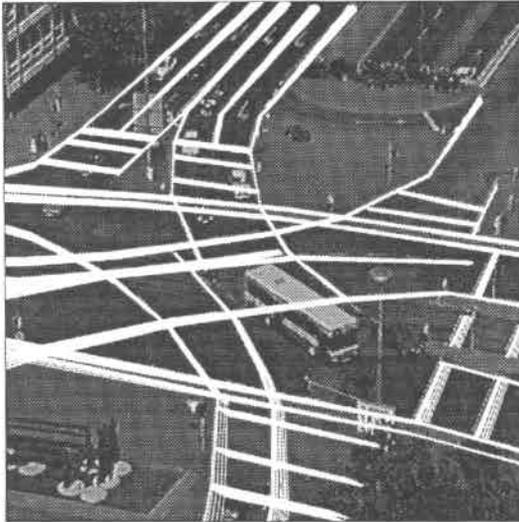


Abbildung 6.21: Eine iterative Anpassung mit der MAP-Schätzung nach Gl. 6.9

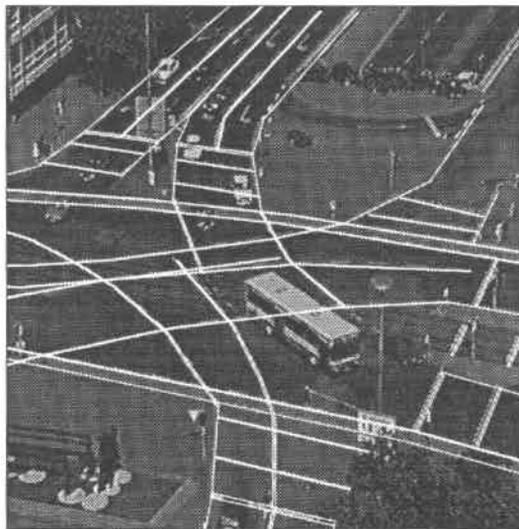


Abbildung 6.22: Die endgültige Aktualisierung der Fahrbahnmodellausprägung.

Kapitel 7

Begriffliche Beschreibungen

Im Vergleich zum System von [Koller 92] berechnet das weiterentwickelte System auch begriffliche Beschreibungen von Verkehrsgeschehen aus einer Bildfolge. Je nach Komplexität der Vorgänge in der Szene kann man auf der begrifflichen Ebene verschiedene Abstraktionsstufen unterscheiden. Ein *Ablauf* als generischer Oberbegriff für kombinierte zeitliche und räumliche Veränderungen in der Szene ist eine Generalisierung von *Geschehen*, von *einfachem Ablauf*, von *Episode* und von *Geschichte* (vgl. auch [Nagel 88; Nagel 91; Walter 89; Kollnig & Nagel 93]). Ein Geschehen hat die kleinste und eine Geschichte die größte Komplexität unter den Abläufen. Eine Geschichte ist aus Episoden zusammengesetzt (Aggregation); diese bestehen aus einfachen Abläufen, welche ihrerseits aus Geschehen aufgebaut sind. Das in dieser Arbeit vorgestellte System gibt inkrementell im Halbbild-Zeitraaster der Bildfolge Geschehensbeschreibungen in Form von deutschen [Kollnig & Nagel 93; Kollnig 92] und englischen [Kollnig *et al.* 94; Damm 93] Bewegungsverben¹ und Verbalphrasen aus. Zur Zeit sind 67 Verben bzw. Verbalphrasen für Geschehen an einer innerstädtischen Kreuzung [Kollnig *et al.* 94] sowie weitere 28 Verben bzw. Verbalphrasen für Geschehen an einer Tankstelle [Damm 94] im System repräsentiert. Abb. 7.1 zeigt berechnete begriffliche Beschreibungen für die Bustrajektorie aus Abb. 3.5. Einheiten zur Berechnung von Geschehensbeschreibungen wurden bereits in [Kollnig & Nagel 93] publiziert, so daß wir die Ergebnisse hier nur knapp darstellen, weitere Auswertungsergebnisse finden sich in [Kollnig *et al.* 94; Nagel 94b; Nagel 95b; Storkebaum 92; Maier 93; Damm 93; Damm 94]. Zur Abstraktion von Details quantitativer Art, die nach automatischer Auswertung von Bildfolgen bei geometrischen Beschreibungen anfallen, werden bei [Kollnig & Nagel 93] unscharfe Menge (fuzzy sets) eingesetzt. Um die Systemausgabe – diese wird auch schon in einfachen Beispielen unübersichtlich, siehe Abb. 7.1 – überprüfen zu können [Nagel 77], wird die Fahrzeugtrajektorie des Agens mit den berechneten Geschehen eingefärbt, siehe Abb. 7.2 [Kollnig & Nagel 93; Kollnig *et al.* 94; Maier 93].

¹Untersuchungen wurden auch mit griechischen (zusammen mit Yannis Stamatou), chinesischen (zusammen mit Dr. Jinyou Zhang), spanischen (zusammen mit Ricardo Orcero), bulgarischen (zusammen mit Maria Lilova) und persischen (zusammen mit Nazanin Baumgärtner) Verben durchgeführt. Die Ergebnisse wurden jeweils in einem internen Bericht zusammengestellt.

Agensgeschehen		
7 :	93	normal_fahren(obj_10), λ 1
11 :	93	nach_links_fahren(obj_10), λ 1
10 :	93	nach_links_fahren(obj_10), λ 0,9
9 :	93	nach_links_fahren(obj_10), λ 0,53
8 :	93	nach_links_fahren(obj_10), λ 0,25
7 :	93	vorwaertsfahren(obj_10), λ 1

Objektgeschehen		
26 :	37	devonfahren(obj_10, obj_9), λ 0,93
14 :	37	dakonfahren(obj_10, obj_9), λ 0,79
13 :	37	devonfahren(obj_10, obj_9), λ 0,64
12 :	37	devonfahren(obj_10, obj_9), λ 0,41
12 :	38	devonfahren(obj_10, obj_9), λ 0,2
11 :	38	devonfahren(obj_10, obj_9), λ 0,19
7 :	38	devonfahren(obj_10, obj_9), λ 0,15
7 :	62	vorausfahren(obj_10, obj_9), λ 1

Fahrbahngeschehen		
62 :	65	abbiegen(obj_10), λ 1
61 :	66	abbiegen(obj_10), λ 0,75
73 :	74	abbiegen(obj_10), λ 1
72 :	75	abbiegen(obj_10), λ 0,75
7 :	93	befahren_fahr(obj_10), λ 1

Ortsgeschehen		
36 :	42	ort_erreichen(obj_10, ort_0), λ 0,69
36 :	43	ort_erreichen(obj_10, ort_0), λ 0,79
36 :	44	ort_erreichen(obj_10, ort_0), λ 0,88
36 :	45	ort_erreichen(obj_10, ort_0), λ 0,91
7 :	45	hin_fahren_zu_ort(obj_10, ort_0), λ 1
34 :	46	ort_erreichen(obj_10, ort_0), λ 1
45 :	51	ort_ueberfahren(obj_10, ort_0), λ 0,92
44 :	51	ort_ueberfahren(obj_10, ort_0), λ 0,88
44 :	52	ort_ueberfahren(obj_10, ort_0), λ 0,82
43 :	52	ort_ueberfahren(obj_10, ort_0), λ 0,79
43 :	53	ort_ueberfahren(obj_10, ort_0), λ 0,72
42 :	53	ort_ueberfahren(obj_10, ort_0), λ 0,69
42 :	54	ort_ueberfahren(obj_10, ort_0), λ 0,62
41 :	54	ort_ueberfahren(obj_10, ort_0), λ 0,5
41 :	55	ort_ueberfahren(obj_10, ort_0), λ 0,52
40 :	55	ort_ueberfahren(obj_10, ort_0), λ 0,5
40 :	56	ort_ueberfahren(obj_10, ort_0), λ 0,41
39 :	56	ort_ueberfahren(obj_10, ort_0), λ 0,4
51 :	57	ort_verlassen(obj_10, ort_0), λ 0,69
39 :	57	ort_ueberfahren(obj_10, ort_0), λ 0,31
38 :	57	ort_ueberfahren(obj_10, ort_0), λ 0,3
51 :	58	ort_verlassen(obj_10, ort_0), λ 0,8
38 :	58	ort_ueberfahren(obj_10, ort_0), λ 0,2
37 :	58	ort_ueberfahren(obj_10, ort_0), λ 0,2
51 :	59	ort_verlassen(obj_10, ort_0), λ 0,91
51 :	60	ort_verlassen(obj_10, ort_0), λ 0,92
49 :	61	ort_verlassen(obj_10, ort_0), λ 1
51 :	93	wegfahren_von_ort(obj_10, ort_0), λ 1

Abbildung 7.1: Berechnete begriffliche Beschreibungen für die Bustrajektorie aus Abb. 3.5: Agensgeschehen beziehen sich nur auf das Agens selbst, Ortsgeschehen zusätzlich auf einen Ort, Fahrbahngeschehen auch auf die Fahrbahn, und Objektgeschehen beziehen neben dem Agens noch ein weiteres Fahrzeug als Patiens ein. Der Bus als Objekt Nr. 10 wurde als Agens, Objekt Nr. 9 als Patiens gewählt, ein punktförmiger Ort befindet sich in der Nähe des Mastes der Straßenbeleuchtung. Die in Form temporallogischer Fakten [Schäfer 93] ausgegebenen Geschehensbeschreibungen enthalten einen Gültigkeitszeitraum (vor dem Ausrufungszeichen), Bezugsobjekte und Orte (in runden Klammern als Argumente) sowie einen Zusicherungsgrad λ mit $0 < \lambda \leq 1$, der die vom Erkennungssystem quantifizierte Akzeptanz der Geschehensbeschreibung im Gültigkeitszeitraum ausdrückt. Zur Abstraktion von Details quantitativer Art, die nach automatischer Auswertung von Bildfolgen bei geometrischen Beschreibungen anfallen, werden bei [Kollnig & Nagel 93] unscharfe Menge zur Modellierung der geometrischen Trajektorien eingesetzt, woraus ein Zusicherungsgrad für jede Geschehensbeschreibung berechnet wird.

Geschehensbeschreibungen können einem temporallogischen Beweiser [Brzoska 93] als Fakten eingegeben werden, der von [Schäfer 93] implementiert wurde. Aufbauend auf einem Ansatz von [Nagel 91], benutzt [Schäfer 93] Übergangsdiagramme, um Geschehen zu einfachen Abläufen zu verketten. Dabei geht allerdings die inkrementelle Auswertungsstrategie verloren.

Eine weiterführende inkrementelle Auswertung wurde dagegen von [Haag 94; Damm 94; Haag 95] untersucht, die auf Arbeiten und Werkzeugen von [Nagel 91; Krüger 91] zurückgreifen. Am Beispiel einer Tankstellenszene wertet [Damm 94] eine Bildfolge bis hin zu Geschehensbeschreibungen aus (siehe Abb. 7.3). Eine Erweiterung dieser (reinen) Zustandsbeschreibungen der Objekte und ihrer Beziehungen um Hand-

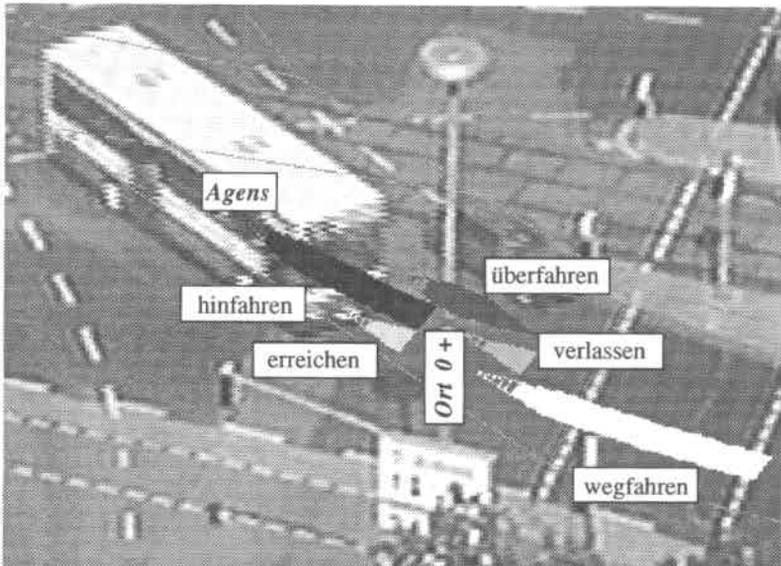


Abbildung 7.2: Zur Überprüfung der Systemausgabe wird die Trajektorie des Agens mit berechneten Geschehen eingefärbt. Die Dicke der Trajektorie visualisiert den Zuverlässigkeitsgrad. Treten mehrere Beschreibungen gleichzeitig auf, so werden die mit Geschehen eingefärbten Trajektorien nebeneinander ins Bild projiziert. Diese Abbildung visualisiert die Ortsgeschehen des Busses aus Abb. 7.1.

lungsmöglichkeiten und -absichten der Akteure führt zu einem Situationsbegriff. Analog zur Ausprägung von Fahrzeug- und Fahrbahnmodellen auf geometrischer Ebene prägt [Haag 95] Situationsschemata aus, die zusätzliches Wissen enthalten, z.B. über übliches Verhalten von Akteuren im Straßenverkehr. Damit lassen sich zusätzlich zur Prädiktion geometrischer Beschreibungen mit einem Kalmanfilter auch auf begrifflicher Ebene Prädiktionen durchführen [Haag 95]. Berechnete Folgesituationen können dann die modellgestützte Bildauswertung unterstützen. Dieses Vorgehen ermöglicht z.B. eine selektive Bildauswertung, bei der man die Auswertung auf in der Situation wichtige Objekte fokussiert.

Eine Zusammenfassung der relevanten Literatur zur Extraktion begrifflicher Beschreibungen aus Bildfolgen findet sich in [Nagel 88], weitere Quellen wurden in [Heinze *et al.* 91; Kollnig & Nagel 93; Kollnig *et al.* 94] aufgeführt. Im Gegensatz zu den Ergebnissen der vorliegenden Arbeit greifen [Neumann & Novak 86; Mohnhaupt 90; Mohnhaupt & Neumann 90; Mohnhaupt 92; Walter 89] auf synthetische Daten zurück, ohne ein einziges Bildsignal auszuwerten. [Schirra *et al.* 87; Herzog *et al.* 89; Retz-Schmidt 91; Retz-Schmidt 92; Herzog 92; Howarth & Buxton 92; Schirra 94] ergänzen fehlende automatische Auswertungsergebnisse interaktiv.

Im Gegensatz zu den unscharfen Mengen [Kollnig & Nagel 93], der temporalen Logik [Brzoska 93; Schäfer 93] sowie einem ersten Ansatz für eine unscharfe temporale Logik [Baumgärtner 94] verwenden [Huang *et al.* 94] oder auch [Howarth & Buxton 92; Howarth & Buxton 93] Bayessche Netze zur rechnerinternen Darstellung begrifflicher Beschreibungen. Ein Vergleich der Ergebnisse unserer Forschungsgruppe mit den Ergebnissen von [Huang *et al.* 94] oder auch von [Howarth & Buxton 92; Howarth & Buxton 93] ist schwierig, da sich in den Arbeiten dieser Autoren nur wenige Ergebnisse mit realen Daten finden. [Huang *et al.* 94] interpretieren Fahrzeugtrajektorien, die [Koller *et al.* 94] aus Bildfolgen berechnen, die eine feste Kamera von einer Autobahnbrücke aufgezeichnet hat. [Howarth & Buxton 92] berechnet aus interaktiv ergänzten Trajektoriendaten Fahrbahngeschehen, wobei das Fahrbahnmodell nicht an Bilddaten angepaßt wird. Auf der Basis räumlicher Relationen zwischen zwei Fahrzeugen werden bei [Howarth & Buxton 93] Geschehen wie „vorausfahren“, „überholt werden“ und „hinterherfahren“ ausgegeben. Diese Forschungsgruppe verwenden teilweise geometrische Eingangsdaten, die Baker, Sullivan und Mitarbeiter aus Bildfolgen berechnen (vgl. Abschnitt 2.2), verfügt aber nicht über ein Gesamtsystem, in dem alle Auswertungsstufen vom Bildsignal bis zu begrifflichen Beschreibungen implementiert sind. Daher sind [Gong & Buxton 92] gezwungen, auf synthetische Daten zurückzugreifen, um auf geometrischer Auswertungsstufe Prädiktionen über die Bewegung eines Fahrzeugs zu machen.

Weiterhin gibt es Arbeiten, die eine umgekehrte Kette vom Begriff zur Bildfolge untersuchen: Das System ANTLIMA (ANTicipation of Listeners' IMagery) von [Schirra 94] erzeugt synthetische Bildfolgen zur Visualisierung begrifflicher Beschreibungen. Dabei werden Bildvorstellungen, die sich Hörer von mitgeteilten Geschehen machen, für bewegte Objekte und einen statischen Hintergrund visualisiert. [Messer 92] stellt einen Ansatz zur modellbasierten Synthese eines Bildauswertungssystems zur Erkennung von 2D-Objekten vor. Am Beispiel einer Szene aus dem Straßenverkehr zeigt [Messer 92], wie man die Erkennung von Randpfosten begrifflich spezifiziert, woraus sein System automatisch durch Kombination einfacher signalnaher Operatoren einen Segmentierungsalgorithmus generiert. Die Algorithmen werden in der PROLOG-ähnlichen Sprache TRIAS ausgegeben, die von [Eckstein 92] als Werkzeug für eine Entwicklung von Bildauswertungsanwendungen entwickelt wurde. Die in dieser Sprache auftretenden Operatoren wurden in dem dazugehörenden Bildauswertungssystem HORUS [Eckstein *et al.* 93] implementiert.

Im Rahmen eines Fahrerunterstützungssystems stellt [Nagel 94a] eine Kette von begrifflicher Auftragspezifikation bis hin zur Auftragsausführung unter Einbeziehung von Sensordaten vor.

Begriffliche Beschreibungen bilden auch bei [Stopp *et al.* 94] eine geeignete Mensch-Maschine-Schnittstelle: beispielsweise läßt sich eine Auftragspezifikation für einen mobilen Roboter natürlichsprachlich formulieren. Im von [Stopp *et al.* 94] vorgestellten Ansatz schlagen die Autoren die Verwendung statischer räumlicher Relationen als Verbindungsglied zwischen der natürlichen Sprache auf der einen und der signalnahen Auswertung auf der anderen Seite vor. Die Verbindung zwischen beiden Ebenen ist jedoch noch nicht geschlossen.

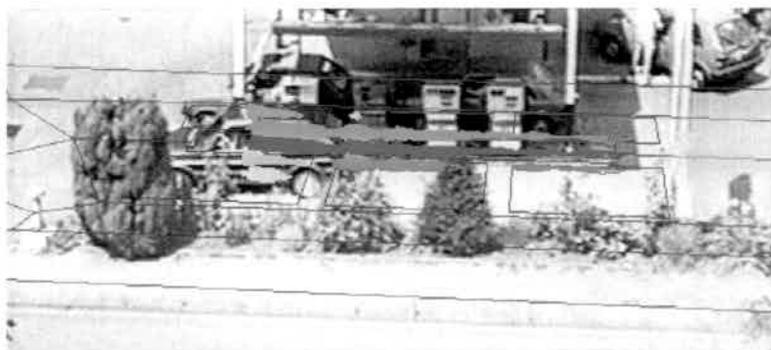


Abbildung 7.3: Die Agentrajektorie ist mit Tankstellengeschehen eingefärbt. Dabei entsprechen die Farben der Trajektorienabschnitte den Farben der Geschehen aus Abb. 7.4. Mit grünen Linien ist ein Modell der Tankspuren und Tankplätze überlagert. Hinterlegt ist das letzte Bild der Folge: das Fahrzeug hat auf dem letzten Tankplatz der vorderen Tankspur angehalten [Damm 94].

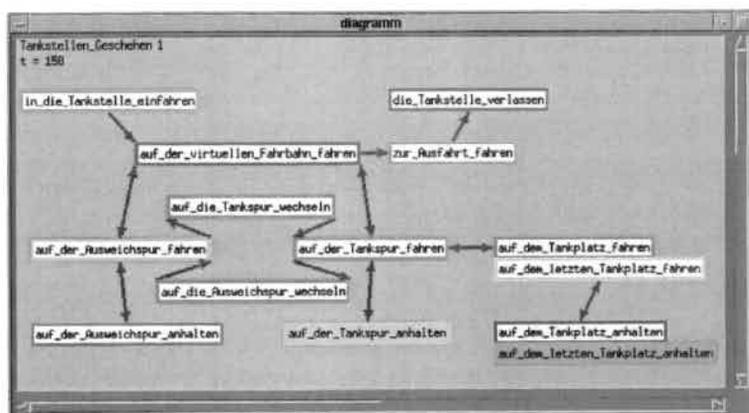


Abbildung 7.4: Darstellung (eines Teils) der Tankstellengeschehen nach [Damm 94]. Zum Zeitpunkt der letzten Aufnahme dieser Sequenz wurden die Geschehen „auf der Tankspur anhalten“ und „auf dem letzten Tankplatz anhalten“ erkannt.

Kapitel 8

Diskussion und Ausblick

8.1 Zusammenfassung und Diskussion

In der vorliegenden Arbeit wurden Bildfolgen des innerstädtischen Straßenverkehrs automatisch ausgewertet. Dabei beschränkte sich das vorgestellte System nicht auf die Lösung einer Teilaufgabe, sondern realisierte *erstmalig* eine ganze Kette der Auswertungsstufen von *sensorisch* erfaßten Bilddaten über *geometrische* bis hin zu *begrifflichen* Beschreibungen von Verkehrsgeschehen. Das System wertet sowohl *dynamische* als auch *statische* Aspekte der in der Bildfolge aufgezeichneten Szene aus. Ein Vergleich mit konkurrierenden Ansätzen wurde ausführlich in den Abschnitten 2.5 und 4.4.3 sowie in Kapitel 7 durchgeführt. In der vorliegenden Arbeit wurden folgende Fortschritte erreicht:

- Der Einsatz leistungsfähigerer signalnaher Komponenten zur optischen Flußschätzung erwies sich als eine notwendige Voraussetzung zur Detektion bewegter Objekte. Die somit verbesserten Initialisierungen ermöglichen erst eine robuste Verfolgung von Fahrzeugen, denn der Fangbereich der modellgestützten Verfolgung braucht nun nicht mehr in gleichem Maße wie früher die Unzulänglichkeiten einer zu groben Initialisierung zu kompensieren.
- Mit der Annahme, die in einer Bildfolge aufgenommene Kreuzung bestehe aus zwei Scharen jeweils zu den Fahrbahnbegrenzungen paralleler Geraden und einer dritten Geradenschar parallel zu den Ampelmasten, schätzen wir Fluchtpunkte aus den Kantensegmenten einer Aufnahme. Es wurde erstmals umfassend diskutiert, wie Kalibrierungsparameter und Fluchtpunktkoordinaten zusammenhängen (Tab. B.1). Die Fluchtpunktkoordinaten werden dann benutzt, um in einem neuen Verfahren eine initiale Lage der Kamera relativ zur Fahrbahnebene zu schätzen. In einem modellbasierten Verfahren wird diese Lageschätzung anschließend iterativ korrigiert.
- In Zusammenhang mit der Fahrbahnanpassung wurden eine ganze Reihe von Modellverfeinerungsschritten dargestellt. Die Untersuchungen der Unterschiede zwischen unseren Erwartungen und dem Verhalten des Systems nach einer

Veränderung zeigten, daß wir mehr Annahmen über die Szene einbringen mußten als ursprünglich erwartet. Die Zeit, die man in die Entwicklung geeigneter Modelle steckt, spart man dann aber bei der Suche nach geeigneten Abstimmparametern und Schwellwerten zur Steuerung des Anpassungsprozesses. Bei unzureichenden Modellen wird es oft unmöglich, passende Abstimmparameter zu finden, die nicht nur zur Auswertung eines einzelnen Bildes, sondern einer gesamten Bildfolge einsetzbar sind. Deshalb wurden mehrere Fahrspuren, Fahrbahnmarkierungen, wie z.B. Mittelstreifen, und auch Radwege in das Kreuzungsmodell aufgenommen. Außerdem wurden Markierungen nicht mehr als einfache Linien, sondern als Streifen einer bestimmten Breite dargestellt. Damit war es erstmals möglich, ein Fahrbahnmodell an Aufnahmen komplexer innerstädtischer Straßenkreuzungen anzupassen.

- Ein signifikanter Fortschritt bei der Verfolgung von Fahrzeugen gelang durch die direkte Anpassung von Polyedermodellen an Grauwertgradienten. Dadurch konnte auf die Extraktion von Kantensegmenten, die damit verbundenen Heuristiken und Probleme verzichtet werden. Es wurde erstmals möglich, Fahrzeuge trotz beachtlicher Variation von Typ, Größe, Position und Orientierung im Bild sowie von Fahrtrichtung mit Beispielen aus sehr unterschiedlichen Realweltkreuzungen durch das robustere Verfahren mit demselben Parametersatz zu bearbeiten. Die Positionsschätzungen sind damit so genau, daß sogar Bewegungen von Halbbild zu Halbbild zuverlässig erfaßt und beschrieben werden können. Aufgrund der guten Ergebnisse dieses neuen Ansatzes zur Lageschätzung von Objekten fanden im Rahmen dieser Arbeit hergeleitete neue Invarianten (vgl. Gl. C.12 und Gl. D.8) für die Anpassung von Kantensegmenten nur im Anhang Platz.
- Mit dem realisierten Anschluß der Auswertungsergebnisse an die begriffliche Beschreibungsebene lassen sich erstmals Experimente auf mehreren, unterschiedlichen Auswertungsstufen machen. Begriffliche Beschreibungen ermöglichen eine Kopplung von aus Bildfolgen ermittelten Auswertungsergebnissen an Datenbanksysteme oder terminologische Systeme.
- Ein wichtiger Punkt war das Zusammenspiel verschiedener Systemkomponenten: Ziel der Arbeit war nicht, einzelne Verfahren isoliert zu optimieren, sondern sie sollten im Systemrahmen zuverlässigere Auswertungsergebnisse liefern. Auch mußte bei vom Benutzer entdeckten Schwachpunkten das gesamte System daraufhin untersucht werden, an welcher Stelle aufgrund welcher unzureichenden Komponente das System keine zufriedenstellenden Ergebnisse lieferte. Meistens lag die Ursache von Problemen bei der Auswertung darin, daß zu wenig Modellwissen in der Bildauswertung ausgenutzt wurde.

Das System hat mittlerweile einen Reifegrad erreicht, der es gestattet, Systemverbesserungen durch einen Gradientenanstieg im Verständnisraum und nicht nur auf der Ebene des Programmcodes durchzuführen.

Das Gesamtsystem wird in Abb. 8.1 dargestellt.

8.2 Ausblick

Der vorgestellte Systemrahmen bildet einen Ausgangspunkt für zahlreiche interessante, aus Zeitgründen in der vorliegenden Arbeit jedoch nicht mehr aufgegriffene Folgearbeiten, die auf verschiedenen Auswertungsstufen möglich sind. Mit dem zur Verfügung stehenden System können rasch Experimente auch zunächst konkurrierender Ansätze auf verschiedenen Auswertungsstufen durchgeführt werden. Wir diskutieren kurz einige weiterführende Möglichkeiten.

- **Klassifikation:** Mit der Zielsetzung, zunächst eine zuverlässige Verfolgung von Fahrzeugen zu erreichen, wurde eine automatische Klassifikation in dieser Arbeit zurückgestellt. Eine Klassifikation von Fahrzeugen ist auf verschiedenen Auswertungsstufen möglich: zunächst kann auf der Basis von Objektkandidaten eine grobe Klassifikation in LKW und PKW erfolgen; eine Unterscheidung zwischen PKW mit Stufenheck bzw. PKW mit Fließheck hat [Koller 92] durch Freigabe von Formparametern im modellgestützten Verfolgungsprozeß mit drei Fahrzeugen durchgeführt, eine Weiterentwicklung dieses Ansatzes für einen Routineeinsatz steht noch aus. Das in der vorliegenden Arbeit vorgestellte Gesamtsystem liefert zudem die Voraussetzungen, auch das Kreuzungsmodell in der Klassifikation von Fahrzeugen zu verwenden, da zum Beispiel Straßenbahnen oder teilweise auch Busse auf abgetrennten Fahrspuren fahren. Mit einer Klassifikation von Fahrzeugen können als Ergänzung zu den im Rahmen dieser Arbeit vorgestellten Geschehensbeschreibungen begriffliche Beschreibungen für die in einer Bildfolge aufgezeichneten Objekte berechnet werden.
- **Makroskopische Verkehrsanalyse:** Ein wichtiger Punkt im Rahmen einer Verkehrsanalyse ist der Übergang von der Auswertung einzelner Fahrzeuge zur Interpretation der Bewegung einer Fahrzeuggruppe, um Kenngrößen wie mittlere Geschwindigkeiten von Fahrzeugen zu ermitteln sowie eine Analyse typischen Anfahr- und Abbiegeverhaltens an der Kreuzung durchzuführen. Hierfür bietet es sich beispielsweise an, auf einer datengetriebenen Auswertungsstufe Flußvektoren zu verfolgen oder auch mehrere Objektbildkandidaten zu einer Ballung zusammenzufassen. Weiterhin ist auch im Szenenbereich auf einer höheren Auswertungsstufe eine Ballung von Trajektorien einzelner Fahrzeuge möglich. Von Interesse sind auch Untersuchungen, wie aus mikroskopischen begrifflichen Beschreibungen einzelner Fahrzeugbewegungen makroskopische Beschreibungen für Fahrzeugpuls aggregiert werden können [Früauf 95].
- **Modellverfeinerungen:** Bei Auswertung anderer Bildfolgen, in denen die Fahrzeugabbilder größer als bei den im Rahmen dieser Arbeit untersuchten Aufnahmen sind, kann es notwendig sein, Fenster, Türen und Räder in dem Modell darzustellen, um eine robuste Verfolgung sicherzustellen. Diese Details können dann auch im Rahmen der Erzeugung begrifflicher Beschreibungen

aufgegriffen werden. Notwendig erscheint eine Modellierung einer Straßenbahn oder auch von LKWs. In diesem Zusammenhang tritt auch in einem stärkeren Maße die Problemstellung auf, wie man Fahrzeugmodelle automatisch akquirieren kann. Modellanpassungen gegliederter Objekte, z.B. von Fußgängern [Kellner 93; Schürmann 94], sind denkbar. Das vorgestellte Kreuzungsmodell repräsentiert zur Zeit nur die Fahrbahnen an einer Kreuzung. Eine Erweiterung des Kreuzungsmodells auf Beschreibungen von Ampeln, Verkehrsschildern, Häusern und Bäumen ist wünschenswert. In diesem Zusammenhang erscheint es notwendig, statt eines 2D-Kreuzungsmodells ein 3D-Kreuzungsmodell zu entwickeln und auch andere Verfahren zur Berechnung von Initialisierungen zu testen (s. [Barnard & Fischler 82]). Eine weitere Stoßrichtung wird sein, sich von einem polygonalen Kreuzungsmodell zu trennen, um konische Segmente oder gar eine Klothoide [Dickmanns & Mysliwetz 92] an Bilddaten anzupassen. Neue interessante Fragestellungen treten auch bei Auswertung von anderem Bildmaterial auf, z.B. bei Aufnahmen von Tankstellengeschehen [Damm 94]. Eine weitere Herausforderung bleibt die Auswertung von Bildfolgen, die eine bewegte Kamera aufzeichnet [Mourad 94].

- **Systemuntersuchungen:** Das Bildauswertungssystem kann man auch als eine Meßeinrichtung auffassen, deren Stabilität mit Methoden der Regelungstechnik nachzuweisen ist. Bei Verfahren, die auf Kantensegmentzuordnungen basieren, ist dies schwierig, da in die Ermittlung der Kantensegmente viele Schwellwertentscheidungen eingehen, so daß der Fehler der Auswertungsschritte sicherlich nicht normalverteilt ist. Es bleibt zu untersuchen, inwieweit bei Anpassung der Fahrzeugmodelle an Grauwertgradienten analytische Stabilitätsuntersuchungen möglich werden.
- **Erweiterung begrifflicher Beschreibungen:** Ein wichtiger Gesichtspunkt bei Berechnung begrifflicher Beschreibungen ist eine Rückkopplung der begrifflichen Auswertungsergebnisse auf Bildfolgen. Zum einen müssen vom System erzeugte begriffliche Beschreibungen visualisiert werden, um sie überprüfen zu können. Dies wurde im Rahmen der vorliegenden Arbeit bei Geschehensbeschreibungen durchgeführt. Möglich ist auch eine Visualisierung der Ausgaben des temporallogischen Beweisers [Schäfer 93] oder eine Visualisierung der Ausprägungen von Situationsschemata [Haag 95; Krüger 91] in der Bildfolge. Zum zweiten können begriffliche Beschreibungen genutzt werden, um die geometrischen Auswertungsstufen zu verbessern. Damit lassen sich auch globalere Optimierungskriterien für Fahrzeugtrajektorien aufstellen, was den Auswertungsprozeß gegenüber nie auszuschließenden stochastischen Störeinflüssen noch stärker stabilisiert.

Ein weiteres Ziel bleibt die Berechnung reichhaltiger begrifflicher Beschreibungen. [Nagel 94a] schlägt in diesem Zusammenhang vor, mit der Diskurs-Repräsentationstheorie [Kamp & Reyle 93] eine Verbindung zwischen Bildfolgenauswertung, Logik und natürlicher Sprache herzustellen, um Algorithmen, die ursprünglich zur rechnerinternen Darstellung und maschinellen Interpre-

tation geschriebener natürlicher Sprache entwickelt wurden, auf begriffliche Beschreibungen anzuwenden, die ein Automat aus Bildfolgen aufbaut [Gerber 94].

Anhang A

Ballung optischer Flußvektoren

[Koller 92] benutzte zur Detektion bewegter Fahrzeugabbilder die folgende Ballungsanalyse, die auf Arbeiten von [Sung 88] zurückgeht: In einem ersten Schritt wird jeder Flußvektor als ein Ballungskeim betrachtet. Einem solchen Keim werden alle Vektoren zugeordnet, für die die Konjunktion folgender drei Prädikate erfüllt ist:

- Zwei Vektoren erfüllen das Prädikat **benachbart**, wenn der euklidische Abstand ihrer Fußpunkte einen Schwellwert t_n nicht überschreitet.
- Zwei Vektoren erfüllen das Prädikat **parallel**, wenn die absolute Differenz ihrer Orientierungen einen Schwellwert t_p nicht überschreitet.
- Zwei Vektoren erfüllen das Prädikat **gleich lang**, wenn ihre relative - bezüglich des ersten Vektors - Längendifferenz einen Schwellwert t_l nicht überschreitet.

In einem zweiten Schritt werden in einem rekursiven Verschmelzungsprozeß maximal disjunkte Mengen aus den Ballungen gebildet.

In einem dritten Schritt werden die Fußpunkte aller Vektoren einer solchen Menge mit einem umschreibenden Rechteck (siehe z.B. Abb. 3.3) eingeschlossen.

In einem vierten Schritt wird überprüft, ob ein Rechteck einen Vektor aus einem anderen Rechteck enthält. In diesem Fall werden die Ballungen vereinigt und ein umschreibendes Rechteck für die Vereinigung gebildet.

Im Beispiel der Ettlinger-Tor-Szene (vgl. Abb. 1.1 und Abb. 3.3) wählten wir als Parameter $t_n = 1$, $t_p = 15^\circ$, $t_l = 15\%$. Das Abbild des Busses wurde zusammenhängend detektiert, da im vierten Schritt auch Vektoren zusammengefaßt werden, die nicht benachbart sind.

Anhang B

Kalibrierung

B.1 Projektive Geometrie

Im Vergleich zur euklidischen stehen in der projektiven Geometrie keine Metrik und damit auch keine Längen und Winkel, sondern Kollinearität von Punkten und Zusammenlaufen (engl. concurrency) von Geraden im Vordergrund [Kanatani 91]. Eine Einführung in die projektive Geometrie für die Bildauswertung findet sich in [Duda & Hart 73], eine ausführlichere Darstellung geben [Mundy & Zisserman 92; Faugeras 93]. Einen guten Einstieg vermittelt auch [Kanatani 91; Kanatani 93a], woran wir uns zunächst orientieren.

B.1.1 Homogene Koordinaten von Punkten

Wir fassen den Bildbereich als zweidimensionalen projektiven Raum auf und definieren *homogene Koordinaten von Punkten*:

Definition B.1 Jeder Punkt des Bildbereiches wird durch ein homogenes Koordinatentripel $(m_1, m_2, m_3)^T$ beschrieben, wobei mindestens eine der drei Koordinaten von null verschieden sein muß.

Mit einer Konstanten f , die mit der Brennweite identifiziert wird, definieren wir den Zusammenhang zwischen projektiven, homogenen und euklidischen, inhomogenen Koordinaten:

Definition B.2 Ein Punkt mit den homogenen Koordinaten $(m_1, m_2, m_3)^T$ wird im Fall

- $m_3 \neq 0$ mit dem eigentlichen Punkt $(fm_1/m_3, fm_2/m_3)^T$ der Bildebene identifiziert, wobei $x = fm_1/m_3, y = fm_2/m_3$ inhomogene Koordinaten genannt werden,
- $m_3 = 0$ mit einem im Unendlichen liegenden uneigentlichen Punkt oder Fernpunkt identifiziert.

Definition B.3 Die (Halb-)Gerade durch optisches Zentrum und Bildpunkt heißt Sehstrahl des Bildpunktes.

Die folgenden Lemmata verdeutlichen den Zusammenhang von homogenen Koordinaten und der Bildaufnahme zugrundeliegenden Projektion.

Lemma B.4 Auf dem durch den homogenen Koordinatenvektor eines Bildpunktes gegebenen Sehstrahl findet sich das Urbild des Bildpunktes in der Szene [Kanatani 91].

Beweis: Eigenschaft der Zentralprojektion. □

Lemma B.5 Der zu einem Bildpunkt mit den inhomogenen Koordinaten $(x, y)^T$ gehörende homogene Koordinatenvektor $(x, y, f)^T$ und der Richtungsvektor des Sehstrahls des Bildpunktes sind linear abhängig [Kanatani 91].

Beweis: Folgt sofort aus Definition B.2. □

B.1.2 Homogene Koordinaten von Geraden

Definition B.6 Eine Gerade wird durch ein vom Nullvektor verschiedenes homogenes Koordinatentripel $(n_1, n_2, n_3)^T$ definiert.

Auch hier unterscheiden wir zwischen eigentlichen und uneigentlichen Geraden:

Definition B.7 Eine Gerade mit den homogenen Koordinaten $(n_1, n_2, n_3)^T$ stellt im Fall

1. $n_1 \neq 0$ oder $n_2 \neq 0$ die eigentliche Gerade $n_1x + n_2y + n_3f = 0$ dar,
2. $n_1 = n_2 = 0$ eine uneigentliche Gerade im Unendlichen, die Ferngerade [Kanatani 91] dar.

Als die Menge aller Fernpunkte ergänzt die Ferngerade die euklidische Ebene zur projektiven Ebene.

Definition B.8 Die Ebene durch optisches Zentrum und eine Bildgerade heißt Projektionsebene der Geraden.

Lemma B.9 Die Projektionsebene einer Bildgeraden enthält das Urbild der Gerade im Raum.

Beweis: Eigenschaft der Zentralprojektion. □

Lemma B.10 *Der zu einer Bildgeraden $ax + by + c = 0$ gehörende homogene Koordinatenvektor $(a, b, c/f)^T$ ist zugleich ein Normalenvektor der Projektionsebene der Geraden.*

Beweis: Folgt aus Definition B.8. □

Kennzeichnend für die projektive Geometrie ist auch die Dualität von Punkten und Geraden: die Definitionen und Lemmata B.6-B.10 entsprechen jeweils den Definitionen und Lemmata B.1-B.5.

Lemma B.11 *Der 3D-Richtungsvektor einer Raumgeraden ist orthogonal zum homogenen Koordinatenvektor der zugehörigen Bildgeraden.*

Beweis: Nach Lemma B.9 liegt die 3D-Raumgerade, die Urbild der Bildgeraden ist, in der Projektionsebene der Bildgeraden. Die Behauptung folgt aus Lemma B.10. □

Korollar B.12 *Der Normalenvektor der Projektionsebene einer Bildgeraden ist orthogonal zur Raumgeraden [Echigo 90].*

Definition B.13 *Auf Eins normierte homogene Koordinaten heißen N -Vektoren [Kanatani 91].*

Damit erhält man eine Bildsphäre [Kanatani 91]. Dabei sind die mit Fernpunkten vervollständigte Bildebene und die Bildsphäre mathematisch äquivalent, denn der homogene Koordinatenvektor eines Punktes der Bildebene und sein normierter N -Vektor enthalten dieselbe Information: die Richtung des Sehstrahls.

B.2 Fluchtpunkte

Der folgende Satz zeigt die Bedeutung von Fluchtpunkten im Zusammenhang mit Fragen zur Kamerakalibrierung:

Satz B.14 *Der Richtungsvektor einer Geraden im Raum ist gleichzeitig der homogene Koordinatenvektor des Fluchtpunktes seiner Projektion [Kanatani 91; Lavest et al. 93].*

Beweis: Zu der Raumgeraden mit der Gleichung $(X_k, Y_k, Z_k)^T = (P, Q, R)^T + \lambda(U, V, W)^T$ mit $W \neq 0$ gehört der Fluchtpunkt $(x_\infty, y_\infty)^T$, wobei sich seine inhomogenen Bildkoordinaten mit den Abbildungsgleichungen 4.1 wie in [Haralick 80; Haralick & Shapiro 93; Caprile & Torre 90; Lavest et al. 93; Schuster et al. 93] aus

$$\begin{pmatrix} x_\infty \\ y_\infty \end{pmatrix} = \lim_{\lambda \rightarrow \infty} \begin{pmatrix} f \frac{P + \lambda U}{R + \lambda W} \\ f \frac{Q + \lambda V}{R + \lambda W} \end{pmatrix} = \begin{pmatrix} f \frac{U}{W} \\ f \frac{V}{W} \end{pmatrix} \quad (\text{B.1})$$

ergeben. Seine homogenen Bildkoordinaten lauten $(f \frac{U}{W}, f \frac{V}{W}, f)^T = \frac{1}{W}(U, V, W)^T$. Zu der im Fall $W = 0$ parallel zur Bildebene verlaufenden Raumgeraden gehört der uneigentliche Fluchtpunkt mit den homogenen Koordinaten $(U, V, 0)^T$ [Lavest *et al.* 93]. \square

Die Aussage von Satz B.14 ist äquivalent mit folgendem Korollar:

Korollar B.15 *Der Verbindungsvektor vom optischen Zentrum zum Fluchtpunkt ist kollinear zum Richtungsvektor aller Raumgeraden, deren Bilder sich im Fluchtpunkt schneiden* [Collins & Weiss 90; Caprile & Torre 90].

Lemma B.16 *Die Lage der Fluchtpunkte im Bild ist unabhängig von der Position der Kamera im Raum, aber abhängig von der Orientierung und Brennweite der Kamera* [Schuster *et al.* 93].

Beweis: Beim Grenzübergang in Gl. B.1 spielt der Aufpunkt $(P, Q, R)^T$ der Raumgeraden keine Rolle. \square

B.3 Schätzung von Fluchtpunktkoordinaten

Sei eine Schar von Bildgeraden mit den Gleichungen

$$a_i x + b_i y + c_i = 0, \quad i = 1, \dots, n \quad (\text{B.2})$$

gegeben, deren Urbilder parallel im Raum verlaufen.

B.3.1 Fluchtpunktschätzung mit inhomogenen Bildkoordinaten

Der Fluchtpunkt (x_∞, y_∞) soll auf allen Geraden der Schar liegen, dies führt auf das auch von [Straforini *et al.* 93] formulierte Minimierungsproblem:

Problem B.17 (Fluchtpunktschätzung mit inhomogenen Koordinaten)

$$\text{Minimiere: } \sum_{i=1}^n (a_i x + b_i y + c_i)^2. \quad (\text{B.3})$$

Satz B.18 *Das Problem B.17 wird durch den Fluchtpunkt*

$$\begin{pmatrix} x_\infty \\ y_\infty \end{pmatrix} = \begin{pmatrix} \sum a_i^2 & \sum a_i b_i \\ \sum a_i b_i & \sum b_i^2 \end{pmatrix}^{-1} \begin{pmatrix} -\sum a_i c_i \\ -\sum b_i c_i \end{pmatrix} \quad (\text{B.4})$$

gelöst. Darüber hinaus gilt:

$$\tan \bar{\phi} = \frac{y_\infty}{x_\infty} = \frac{\sum a_i^2 \sum b_i c_i - \sum a_i b_i \sum a_i c_i}{\sum b_i^2 \sum a_i c_i - \sum a_i b_i \sum b_i c_i}. \quad (\text{B.5})$$

Beweis: Differentiation der Funktion in Problem B.17 liefert das Gleichungssystem:

$$\begin{aligned} 2 \sum_{i=1}^n a_i(a_i x + b_i y + c_i) &= 0 \\ 2 \sum_{i=1}^n b_i(a_i x + b_i y + c_i) &= 0. \end{aligned} \quad (\text{B.6})$$

Dieses System besitzt die Lösung aus Gl. B.4. \square

Im Fall paralleler Bildgeraden hat nicht nur die oben gestellte Minimierungsaufgabe keinen Sinn mehr, sondern es steht auch in Gl. B.5 auf der rechten Seite ein unbestimmter Ausdruck. Bei gestörten Eingangsdaten ist die Matrix in Gl. B.4 in diesem Fall schlecht konditioniert, so daß die Fluchtpunktschätzung sehr ungenau wird, obwohl der 3D-Richtungsvektor der Raumgeraden eigentlich genau zu ermitteln wäre. Die Güte der Fluchtpunktschätzung hängt stark von der Position der Bildkoordinaten des Fluchtpunktes ab.

Ein weiterer Ansatz besteht darin, einen Fluchtpunkt als Schwerpunkt der Schnittpunkte der zu Geraden verlängerten Kantensegmente zu bestimmen [Caprile & Torre 90]; dies führt zu ähnlichen Problemen.

B.3.2 Fluchtpunktschätzung mit homogenen Koordinaten

Nach Lemma B.25 ist der homogene Koordinatenvektor des Fluchtpunktes orthogonal zu den homogenen Koordinatenvektoren der Bildgeraden der Schar. Dies führt auf das Minimierungsproblem:

Problem B.19

$$\text{Minimiere } \sum_{i=1}^n (a_i x + b_i y + \frac{c_i}{f} z)^2 \text{ unter der Nebenbedingung } x^2 + y^2 + z^2 = 1. \quad (\text{B.7})$$

Mit den Bezeichnungen

$$\mathbf{A} = \begin{pmatrix} \vdots \\ a_i & b_i & \frac{c_i}{f} \\ \vdots \end{pmatrix}, \quad \mathbf{E} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (\text{B.8})$$

kann man dieses Minimierungsproblem folgendermaßen formulieren:

$$\text{minimiere } \|\mathbf{A}\mathbf{E}\|_2 \text{ unter der Nebenbedingung } \|\mathbf{E}\|_2 = 1. \quad (\text{B.9})$$

Dazu äquivalent ist die Minimierung von

$$\|\mathbf{A}\mathbf{E}\|_2^2 - \lambda(\|\mathbf{E}\|_2^2 - 1)$$

mit dem Lagrangeschen Multiplikator $-\lambda$. Damit wäre ein Einheitsvektor \mathbf{E} gesucht, der

$$(\mathcal{A}^T \mathcal{A} - \lambda \mathcal{I}) \mathbf{E} = 0$$

erfüllt, der folglich Eigenvektor von

$$\mathcal{A}^T \mathcal{A} = \begin{pmatrix} \sum_{i=1}^n a_i^2 & \sum_{i=1}^n a_i b_i & \sum_{i=1}^n a_i \frac{c_i}{f} \\ \sum_{i=1}^n a_i b_i & \sum_{i=1}^n b_i^2 & \sum_{i=1}^n b_i \frac{c_i}{f} \\ \sum_{i=1}^n a_i \frac{c_i}{f} & \sum_{i=1}^n b_i \frac{c_i}{f} & \sum_{i=1}^n \frac{c_i^2}{f^2} \end{pmatrix} \quad (\text{B.10})$$

ist.

Satz B.20 Das Minimierungsproblem B.19 wird gelöst durch den Einheitsvektor von $\mathcal{A}^T \mathcal{A}$ zum kleinsten Eigenwert [Chen & Jiang 91; Kanatani 91; Kanatani 92; Kanatani 93b].

B.3.3 Schätzung des Azimutwinkels der 3D-Orientierung einer Fluchtpunkttrichtung

Denkbar ist auch die *schärfere* Nebenbedingung $x^2 + y^2 = 1$, wenn der Fluchtpunkt nicht im Hauptpunkt liegt:

Problem B.21

$$\text{Minimiere } \sum_{i=1}^n (a_i x + b_i y + \frac{c_i}{f} z)^2 \text{ unter der Nebenbedingung } x^2 + y^2 = 1. \quad (\text{B.11})$$

Satz B.22 Das Problem B.21 führt mit der Substitution $x = \cos \bar{\phi}$, $y = \sin \bar{\phi}$ auf die Lösung

$$\tan 2\bar{\phi} = 2 \frac{\sum a_i b_i \sum c_i^2 - \sum a_i c_i \sum b_i c_i}{\sum a_i^2 \sum c_i^2 - (\sum a_i c_i)^2 + (\sum b_i c_i)^2 - \sum b_i^2 \sum c_i^2} \quad (\text{B.12})$$

für das Verhältnis von Bildkoordinaten eines Fluchtpunktes bzw. für den Azimutwinkel $\bar{\phi}$ im KKS einer zugehörigen Raumgeraden.

Beweis: Mit der Substitution $x = \cos \bar{\phi}$, $y = \sin \bar{\phi}$ erhalten wir aus Problem B.21:

$$\text{Minimiere } \sum_{i=1}^n (a_i \cos \bar{\phi} + b_i \sin \bar{\phi} + \frac{c_i}{f} z)^2. \quad (\text{B.13})$$

Differentiation liefert das System:

$$\frac{\partial}{\partial \bar{\phi}} : \quad 2 \sum (a_i \cos \bar{\phi} + b_i \sin \bar{\phi} + \frac{c_i}{f} z) (-a_i \sin \bar{\phi} + b_i \cos \bar{\phi}) = 0, \quad (\text{B.14a})$$

$$\frac{\partial}{\partial z} : \quad 2 \sum (a_i \cos \bar{\phi} + b_i \sin \bar{\phi} + \frac{c_i}{f} z) \frac{c_i}{f} = 0. \quad (\text{B.14b})$$

Hierzu ist äquivalent:

$$\sum(-a_i^2 \sin \bar{\phi} \cos \bar{\phi} + a_i b_i (\cos^2 \bar{\phi} - \sin^2 \bar{\phi}) + b_i^2 \sin \bar{\phi} \cos \bar{\phi}) \quad (\text{B.15a})$$

$$+ z \sum(-a_i \frac{c_i}{f} \sin \bar{\phi} + b_i \frac{c_i}{f} \cos \bar{\phi}) = 0$$

$$z = - \frac{\sum(a_i \frac{c_i}{f} \cos \bar{\phi} + b_i \frac{c_i}{f} \sin \bar{\phi})}{\sum \frac{c_i^2}{f^2}}. \quad (\text{B.15b})$$

Setzt man Gl. B.15b in Gl. B.15a ein, ergibt sich:

$$\begin{aligned} & \sum[(b_i^2 - a_i^2) \sin \bar{\phi} \cos \bar{\phi} + a_i b_i (\cos^2 \bar{\phi} - \sin^2 \bar{\phi})] \\ & + \frac{(\sum a_i \frac{c_i}{f})^2 \sin \bar{\phi} \cos \bar{\phi} - \sum a_i \frac{c_i}{f} \sum b_i \frac{c_i}{f} (\cos^2 \bar{\phi} - \sin^2 \bar{\phi}) - (\sum b_i \frac{c_i}{f})^2 \sin \bar{\phi} \cos \bar{\phi}}{\sum \frac{c_i^2}{f^2}} = 0. \end{aligned} \quad (\text{B.16})$$

Eine Anwendung der trigonometrischen Additionstheoreme liefert:

$$\begin{aligned} & - \sum a_i^2 \sin 2\bar{\phi} + 2 \sum a_i b_i \cos 2\bar{\phi} + \sum b_i^2 \sin 2\bar{\phi} \\ & + \frac{(\sum a_i \frac{c_i}{f})^2}{\sum \frac{c_i^2}{f^2}} \sin 2\bar{\phi} - 2 \frac{\sum a_i \frac{c_i}{f} \sum b_i \frac{c_i}{f}}{\sum \frac{c_i^2}{f^2}} \cos 2\bar{\phi} - \frac{(\sum b_i \frac{c_i}{f})^2}{\sum \frac{c_i^2}{f^2}} \sin 2\bar{\phi} = 0 \end{aligned} \quad (\text{B.17})$$

und somit:

$$\begin{aligned} & \left(- \sum a_i^2 + \sum b_i^2 + \frac{(\sum a_i \frac{c_i}{f})^2 - (\sum b_i \frac{c_i}{f})^2}{\sum \frac{c_i^2}{f^2}} \right) \sin 2\bar{\phi} \\ & + 2 \left(\sum a_i b_i - \frac{\sum a_i \frac{c_i}{f} \sum b_i \frac{c_i}{f}}{\sum \frac{c_i^2}{f^2}} \right) \cos 2\bar{\phi} = 0, \end{aligned} \quad (\text{B.18})$$

woraus die Behauptung folgt. \square

Nach Gleichung B.12 ist die Lösung des Problems B.21 in Satz B.22 unabhängig von f . Dieser Ansatz liefert auch ein brauchbares Ergebnis bei in der Bildebene parallelen oder fast parallelen Geraden: Mit $a_i = a$ und $b_i = b$ ergibt sich aus Gl. B.12:

$$\begin{aligned} \tan 2\bar{\phi} &= 2 \frac{nab \sum c_i^2 - ab (\sum c_i)^2}{na^2 \sum c_i^2 - a^2 (\sum c_i)^2 + b^2 (\sum c_i)^2 - nb^2 \sum c_i^2} \\ &= 2 \frac{ab (n \sum c_i^2 - (\sum c_i)^2)}{(a^2 - b^2) (n \sum c_i^2 - (\sum c_i)^2)} \\ &= 2 \frac{ab}{a^2 - b^2} \end{aligned} \quad (\text{B.19})$$

B.4 Fluchtpunkte und Kalibrierung

In diesem Abschnitt wird untersucht, wie die Lage von Fluchtpunkten zur Kalibrierung der Kamera verwendet werden kann. Im einzelnen bestehen folgende Möglichkeiten:

B.4.1 Fluchtpunkt und Brennweite

Lemma B.23 *Bei bekannter Brennweite lässt sich die Orientierung der im Raum parallelen Geradenschar aus dem Fluchtpunkt ihrer Projektion ins Bild bestimmen [Barnard 83].*

Beweis: Folgt aus Satz B.14. □

Definition B.24 *Ist die Geradenschar im Raum parallel zu einer der drei Koordinatenachsen des Weltkoordinatensystems, so spricht man von einem Hauptfluchtpunkt [Magee & Aggarwal 84].*

Zur Berechnung eines Fluchtpunktes werden oft folgende Lemmata ausgenutzt:

Lemma B.25 *Der homogene Koordinatenvektor eines Fluchtpunktes ist orthogonal zum homogenen Koordinatenvektor jeder Bildgeraden, die durch diesen Fluchtpunkt verläuft [Collins & Weiss 90].*

Beweis: Nach Lemma B.11 sind der homogene Koordinatenvektor (siehe Def. B.7) einer Bildgeraden und der Richtungsvektor der zugehörigen Raumgeraden orthogonal. Der Richtungsvektor der Raumgeraden und der homogene Koordinatenvektor des zur Geradenrichtung gehörenden Fluchtpunktes sind nach Satz B.14 linear abhängig. □

Kennt man die Orientierung $(U, V, W)^T$ einer im Raum parallelen Geradenschar und die Brennweite f , so kann man für jedes Kantensegment im Bild entscheiden, ob es auf dem Bild einer Geraden dieser Schar liegt [Lebègue & Aggarwal 93]:

Lemma B.26 *Ein Kantensegment auf der Bildgeraden $(ax + by + c = 0)$ hat als Urbild ein 3D-Kantensegment mit der 3D-Orientierung $(U, V, W)^T$, wenn*

$$(aU + bV + \frac{c}{f}W)^2 < \text{Schwellwert} \quad (\text{B.20})$$

gilt.

Beweis: Im Fall ungestörter Bildgeraden gilt nach der Orthogonalitätsaussage von Lemma B.25 sogar: $aU + bV + \frac{c}{f}W = 0$. □

B.4.2 Horizont und Brennweite

Lemma B.27 Die Menge aller Fluchtpunkte einer Menge A von Raumgeraden liegt genau dann auf einer Geraden g im Bildbereich, wenn alle Raumgeraden aus A parallel zu einer Ebene E im Raum verlaufen.

Beweis: [Caprile & Torre 90] □

Definition B.28 Die Gerade g aus Lemma B.27 heißt Fluchtlinie der Ebene E .

Satz B.29 Der Normalenvektor einer Ebene im Raum und der homogene Koordinatenvektor der Fluchtlinie dieser Ebene sind linear abhängig [Kanatani 91].

Beweis: Satz B.14 und Lemma B.27. □

Die Ebene durch das optische Zentrum und die Fluchtlinie haben dieselbe Ebenenormale wie eine die Fluchtlinie erzeugende Ebene in der 3D-Welt.

Korollar B.30 Die Fluchtlinie einer Ebene bestimmt bei bekannter Brennweite den Normalenvektor der sie erzeugenden Ebene in der 3D-Welt.

Zu in der Szene parallelen Ebenen gehört dieselbe Fluchtlinie im Bildbereich.

Definition B.31 Die Fluchtlinie einer zur Fahrbahnebene parallelen Raumebene heißt Horizont.

Den Horizont erhält man damit als Schnitt der Bildebene mit der Ebene, die parallel zur Fahrbahnebene durch das Zentrum der Linse verläuft [Duda & Hart 73].

B.4.3 Zwei konjugierte Fluchtpunkte

Definition B.32 Zwei zu im Raum orthogonalen Geraden gehörende Fluchtpunkte heißen konjugierte Fluchtpunkte [Duda & Hart 73].

Satz B.33 Aus den Bildkoordinaten $(x_\infty^{(1)}, y_\infty^{(1)})^T, (x_\infty^{(2)}, y_\infty^{(2)})^T$ zweier konjugierter Fluchtpunkte erhält man die Brennweite f aus:

$$f = \sqrt{-x_\infty^{(1)}x_\infty^{(2)} - y_\infty^{(1)}y_\infty^{(2)}}. \quad (\text{B.21})$$

Beweis: Die homogenen Koordinatenvektoren $(x_\infty^{(1)}, y_\infty^{(1)}, f)^T, (x_\infty^{(2)}, y_\infty^{(2)}, f)^T$ (vgl. Def. B.2) der Fluchtpunkte sind nach Satz B.14 jeweils kollinear zu den Richtungsvektoren der Raumgeraden. Nach Voraussetzung sind die beiden Raumgeraden orthogonal:

$$\begin{pmatrix} x_\infty^{(1)} \\ y_\infty^{(1)} \\ f \end{pmatrix}^T \cdot \begin{pmatrix} x_\infty^{(2)} \\ y_\infty^{(2)} \\ f \end{pmatrix} = 0, \quad (\text{B.22})$$

woraus man die Brennweite f erhält (vgl. [Duda & Hart 73; Caprile & Torre 90; Chen & Jiang 91]). \square

Im Gegensatz zu Verfahren, die inhomogene Koordinaten von Fluchtpunkten im Bildbereich berechnen, sind Verfahren zu sehen, die sofort homogene Koordinaten von Fluchtpunkten auf der Bildsphäre liefern. Hierfür muß aber die Brennweite als bekannt vorausgesetzt werden. Ist die Brennweite dagegen unbekannt, so berechnet man zunächst die homogenen Koordinaten der Fluchtpunkte mit einer vorläufigen, fest gewählten Brennweite f^* :

Lemma B.34 Sei $(m_1^*, m_2^*, m_3^*)^T$ der homogene Koordinatenvektor eines Bildpunktes bezüglich der Brennweite f^* und $(m_1, m_2, m_3)^T$ der homogene Koordinatenvektor desselben Bildpunktes bezüglich einer anderen Brennweite f , so gilt [Kanatani & Onodera 91; Kanatani 92]:

$$\begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \begin{pmatrix} m_1^* \\ m_2^* \\ \frac{f}{f^*} m_3^* \end{pmatrix}. \quad (\text{B.23})$$

Beweis: Die Bildkoordinaten von $(m_1^*, m_2^*, m_3^*)^T$ lauten $x = f^* m_1^* / m_3^*$, $y = f^* m_2^* / m_3^*$. Der homogene Koordinatenvektor dieses Bildpunktes bezüglich einer anderen Brennweite f lautet $(x, y, f)^T = (f^* m_1^* / m_3^*, f^* m_2^* / m_3^*, f)^T$, woraus mit der Homogenität des Koordinatenvektors die Behauptung folgt. \square

Bei homogenen Koordinaten konjugierter Fluchtpunkte kann die richtige Brennweite anschließend ermittelt werden:

Satz B.35 Aus — eventuell mit einer noch unbekanntem, jedoch fest gewählten Brennweite f^* berechneten — homogenen Bildkoordinaten $(u_1, v_1, w_1)^T, (u_2, v_2, w_2)^T$ zweier konjugierter Fluchtpunkte erhält man die richtige Brennweite f aus:

$$f = f^* \sqrt{-\frac{u_1 u_2 + v_1 v_2}{w_1 w_2}}, \quad (\text{B.24})$$

wobei allerdings keiner der homogenen Vektoren parallel zur Bildebene verlaufen darf: $w_i \neq 0$.

Beweis: Mit Hilfe von Lemma B.34 korrigiert man die homogenen Bildkoordinaten, für die nach Voraussetzung gelten soll (vgl. [Kanatani & Onodera 91; Kanatani 92]):

$$\begin{pmatrix} u_1 \\ v_1 \\ \frac{f}{f^*} w_1 \end{pmatrix}^T \cdot \begin{pmatrix} u_2 \\ v_2 \\ \frac{f}{f^*} w_2 \end{pmatrix} = 0 \quad (\text{B.25})$$

\square

Lemma B.36 Aus zwei konjugierten Fluchtpunkten erhält man den Normalenvektor von zugehörigen komplanaren Raumgeraden [Chen & Jiang 91; Kanatani & Onodera 91; Li-Qun et al. 92; Caprile & Torre 90; Collins & Beveridge 93] und damit die Rotationsparameter.

Beweis: Zunächst bestimmt man die Brennweite f . Damit erhält man auch die homogenen Koordinaten der Fluchtpunkte, d.h. die Richtungsvektoren zugehöriger Sehstrahlen. Das Kreuzprodukt dieser Richtungsvektoren liefert den gesuchten Normalenvektor. Äquivalent hierzu ist die Berechnung einer Fluchtlinie als Verbindungsgerade zweier Fluchtpunkte im Bildbereich. Mit bereits berechneter Brennweite f erhält man aus den homogenen Koordinaten der Fluchtlinie sofort den Normalenvektor der Raumebene. \square

B.4.4 Zwei Fluchtpunkte und Brennweite bzw. Schnittwinkel

Die bisher eingehende Orthogonalität von Geradenscharen ist nicht immer gegeben, obgleich die in der Literatur dargestellten Verfahren zur Kalibrierung mittels Kantensegmenten Orthogonalität meist voraussetzen. Falls sich zwei Straßen an einer Kreuzung nicht in einem rechten Winkel schneiden, so liefern die Geraden der Fahrbahnbegrenzungen keine zwei konjugierten Fluchtpunkte. Wir zeigen im folgenden, daß fehlende Winkelinformation durch Kenntnis über die Brennweite berechnet werden kann und umgekehrt.

Satz B.37 Aus — eventuell mit einer noch unbekanntem, jedoch fest gewählten Brennweite f^* berechneten — homogenen Bildkoordinaten $(u_1, v_1, w_1)^T, (u_2, v_2, w_2)^T$ zweier nicht notwendigerweise konjugierter Fluchtpunkte erhält man entweder

1. (nichtlinear) die richtige Brennweite f bei bekanntem Schnittwinkel $\Delta\gamma$ der zu den Fluchtpunkten gehörenden Raumrichtungen oder
2. (auf einfache Weise) den Schnittwinkel $\Delta\gamma$ bei bekannter Brennweite f .

Außerdem ergibt sich aus dem Kreuzprodukt

$$\begin{pmatrix} u_1 \\ v_1 \\ \frac{f}{f^*} w_1 \end{pmatrix} \times \begin{pmatrix} u_2 \\ v_2 \\ \frac{f}{f^*} w_2 \end{pmatrix} \quad (\text{B.26})$$

der homogene Koordinatenvektor der durch beide Fluchtpunkte festgelegten Fluchtlinie (beispielsweise der Normalenvektor der Fahrbahnebene).

Beweis: Mit Hilfe von Lemma B.34 korrigiert man die homogenen Bildkoordinaten, für die nach Voraussetzung gelten soll:

$$\frac{\begin{pmatrix} u_1 \\ v_1 \\ \frac{f}{f^*}w_1 \end{pmatrix}^T \cdot \begin{pmatrix} u_2 \\ v_2 \\ \frac{f}{f^*}w_2 \end{pmatrix}}{\left\| \begin{pmatrix} u_1 \\ v_1 \\ \frac{f}{f^*}w_1 \end{pmatrix} \right\|_2 \left\| \begin{pmatrix} u_2 \\ v_2 \\ \frac{f}{f^*}w_2 \end{pmatrix} \right\|_2} = \cos \Delta\gamma \quad (\text{B.27})$$

Quadrieren liefert eine biquadratische Gleichung in f , die algebraisch nach f auflösbar ist. \square

Brennweite f der Kamera und Schnittwinkel $\Delta\gamma$ horizontaler Raumgeraden sind damit unmittelbar gekoppelt.

B.4.5 Drei Fluchtpunkte

Liegen die Bildkoordinaten $(x_\infty^{(1)}, y_\infty^{(1)})^T, (x_\infty^{(2)}, y_\infty^{(2)})^T, (x_\infty^{(3)}, y_\infty^{(3)})^T$ dreier gegenseitig konjugierter Fluchtpunkte vor, so erhält man mit Satz B.33 drei Gleichungen für die Brennweite. Bei vertauschten Eingangsdaten ist das Problem jetzt überbestimmt: Die nach Satz B.33 aus den Bildkoordinaten zweier konjugierter Fluchtpunkte berechnete Brennweite führt zwar zu einem Orthogonalsystem für die homogenen Koordinatenvektoren dieser beiden Fluchtpunkte, nicht aber zu einem Orthogonalsystem für die homogenen Koordinatenvektoren aller drei Fluchtpunkte.

Lemma B.38 *Das Verhältnis der Bildkoordinaten eines Punktes ist unabhängig von der Brennweite [Kumar & Hanson 90].*

Beweis: Gleichung 4.1. \square

Satz B.39 *Die Lage einer Bild-Ursprungsgeraden durch einen Fluchtpunkt hängt nur von den Rotationsparametern, nicht aber von der Brennweite ab [Echigo 90]. Die Brennweite gibt nur die Lage des Fluchtpunktes auf der Ursprungsgeraden an (vgl. Abb. B.1).*

Beweis: Nach Lemma B.16 wird die Lage eines Fluchtpunktes nur durch die Orientierung und die Brennweite der Kamera bestimmt. Mit Lemma B.38 folgt die Behauptung. \square

Mit P bezeichnen wir die orthographische Projektion, die einen Punkt mit den Kamerakoordinaten $(X_k, Y_k, Z_k)^T$ auf den Punkt mit den Bildkoordinaten $(x_b, y_b)^T$

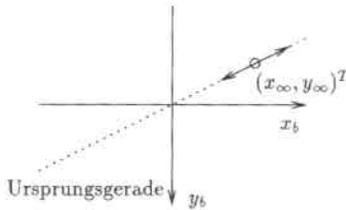


Abbildung B.1: Die Ursprungsgerade durch einen Fluchtpunkt hängt nach [Echigo 90] nur von der Orientierung der Kamera, nicht aber von der Brennweite ab. Die Brennweite gibt nur die Lage des Fluchtpunktes auf der Ursprungsgeraden an, hier angedeutet durch die Pfeile. Liegt der Fluchtpunkt im Hauptpunkt, so ist die Ursprungsgerade undefiniert.

wie folgt abbildet:

$$\begin{pmatrix} x_b \\ y_b \end{pmatrix} = P\left(\begin{pmatrix} X_k \\ Y_k \\ Z_k \end{pmatrix}\right) = \begin{pmatrix} X_k \\ Y_k \end{pmatrix}. \quad (\text{B.28})$$

Korollar B.40 Während ein 3D-Richtungsvektor \mathbf{E} durch die Angabe zweier Freiheitsgrade bestimmt wird, werden

- seine orthographische Projektion $P(\mathbf{E})$,
- die Ursprungsgerade durch seinen Fluchtpunkt im Bild und
- die Bildpunktverhältnisse seines Fluchtpunktes

nur durch einen Freiheitsgrad bestimmt: durch den Azimutwinkel des 3D-Richtungsvektors im KKS.

Lemma B.41 Sei eine Ursprungsgerade $ax + by = 0$ durch einen zu der 3D-Richtung \mathbf{E} gehörenden Fluchtpunkt gegeben. Mit dem homogenen Koordinatenvektor $\mathbf{n} = (a, b, 0)^T$ der Ursprungsgeraden gilt dann

- $\mathbf{n}^T \cdot \mathbf{E} = 0$,
- $P(\mathbf{n})^T \cdot P(\mathbf{E}) = 0$ und
- $P(\mathbf{E}) = (b, -a)^T$ oder — falls $\mathbf{E} = (0, 0, 1)^T$ — $P(\mathbf{E}) = (0, 0)^T$.

Beweis: Die erste Behauptung folgt aus Lemma B.25. □

Wir untersuchen im folgenden, wie man aus den Projektionen $P(\mathbf{E}_x)$, $P(\mathbf{E}_y)$ und $P(\mathbf{E}_z)$ dreier Raumrichtungen \mathbf{E}_x , \mathbf{E}_y und \mathbf{E}_z diese Raumrichtungen berechnen kann. Mehrdeutigkeiten treten dann auf, wenn zwei der drei Projektionen linear abhängig sind.

Lemma B.42 *Unter den Voraussetzungen:*

1. \mathbf{E}_x und \mathbf{E}_y sind orthogonal: $\mathbf{E}_x^T \cdot \mathbf{E}_y = 0$,
2. \mathbf{E}_z ist der Normalenvektor der von \mathbf{E}_x und \mathbf{E}_y aufgespannten Ebene:
 $\mathbf{E}_z = \mathbf{E}_x \times \mathbf{E}_y$,
3. $P(\mathbf{E}_x)^T \cdot P(\mathbf{E}_y) = 0$,
4. $P(\mathbf{E}_x)$ und $P(\mathbf{E}_y)$ sind beide vom Nullvektor verschieden,

sind $P(\mathbf{E}_x)$, $P(\mathbf{E}_y)$ oder $P(\mathbf{E}_z)$, $P(\mathbf{E}_z)$ linear abhängig.

Beweis: Nach Voraussetzung 4 sind $P(\mathbf{E}_x)$, $P(\mathbf{E}_y)$ vom Nullvektor verschieden. Nach der Orthogonalität in Voraussetzung 3 und Lemma B.41 kann man

$$\begin{aligned}\mathbf{E}_x &= (\cos \beta, \sin \beta, Z_x)^T, \\ \mathbf{E}_y &= (-\sin \beta, \cos \beta, Z_y)^T\end{aligned}\quad (\text{B.29})$$

mit $Z_x Z_y = 0$ ansetzen, denn nach Voraussetzung 1 ist

$$0 = \mathbf{E}_x^T \cdot \mathbf{E}_y = \underbrace{-\cos \beta \sin \beta + \sin \beta \cos \beta}_0 + Z_x Z_y = Z_x Z_y. \quad (\text{B.30})$$

Mit Voraussetzung 2

$$\mathbf{E}_z = \mathbf{E}_x \times \mathbf{E}_y = (-Z_x \cos \beta + Z_y \sin \beta, -Z_x \sin \beta - Z_y \cos \beta, 1)^T \quad (\text{B.31})$$

verschwindet das Produkt der Determinanten

$$\begin{aligned}\det(P(\mathbf{E}_x), P(\mathbf{E}_z)) \cdot \det(P(\mathbf{E}_y), P(\mathbf{E}_z)) \\ &= \underbrace{\begin{vmatrix} \cos \beta & -Z_x \cos \beta + Z_y \sin \beta \\ \sin \beta & -Z_x \sin \beta - Z_y \cos \beta \end{vmatrix}}_{=-Z_y(\cos^2 \beta + \sin^2 \beta)} \cdot \underbrace{\begin{vmatrix} -\sin \beta & -Z_x \cos \beta + Z_y \sin \beta \\ \cos \beta & -Z_x \sin \beta - Z_y \cos \beta \end{vmatrix}}_{=Z_x(\cos^2 \beta + \sin^2 \beta)} \\ &= 0.\end{aligned}\quad (\text{B.32})$$

□

Satz B.43 (Schätzung des Normalenvektors einer Ebene) *Unter den Voraussetzungen:*

1. durch homogene Koordinatenvektoren \mathbf{n}_x , \mathbf{n}_y und \mathbf{n}_z sind drei Bild-Ursprungsgeraden durch drei zu den 3D-Richtungsvektoren \mathbf{E}_x , \mathbf{E}_y und \mathbf{E}_z gehörende Fluchtpunkte gegeben:

$$P(\mathbf{n}_x)^T \cdot P(\mathbf{E}_x) = 0, \quad (\text{B.33})$$

$$P(\mathbf{n}_y)^T \cdot P(\mathbf{E}_y) = 0, \quad (\text{B.34})$$

$$P(\mathbf{n}_z)^T \cdot P(\mathbf{E}_z) = 0, \quad (\text{B.35})$$

2. \mathbf{E}_z ist der Normalenvektor der von \mathbf{E}_x und \mathbf{E}_y aufgespannten Ebene:

$$\mathbf{E}_z = \mathbf{E}_x \times \mathbf{E}_y,$$

3. die Vektoren $P(\mathbf{E}_x)$ und $P(\mathbf{E}_y)$ sind linear unabhängig¹,

4. der Vektor $P(\mathbf{E}_z)$ ist vom Nullvektor verschieden²,

5. der Kosinus des Schnittwinkels

$$\cos \Delta\gamma = \frac{\mathbf{E}_x^T \cdot \mathbf{E}_y}{\|\mathbf{E}_x\| \cdot \|\mathbf{E}_y\|} \quad (\text{B.36})$$

ist bekannt und

6. im Fall $\cos \Delta\gamma = 0$ sind $P(\mathbf{E}_x)$ und $P(\mathbf{E}_z)$ sowie $P(\mathbf{E}_y)$ und $P(\mathbf{E}_z)$ linear unabhängig³,

kann man \mathbf{E}_z bestimmen.

Beweis: Nach Voraussetzung 1, Voraussetzung 2 und Lemma B.46 gilt

$$\begin{aligned} \mathbf{E}_x &= \mathbf{E}_z \times \mathbf{n}_x, \\ \mathbf{E}_y &= \mathbf{E}_z \times \mathbf{n}_y. \end{aligned} \quad (\text{B.37})$$

Mit $P(\mathbf{E}_x)$ und $P(\mathbf{E}_y)$ (Voraussetzung 3) sind auch \mathbf{n}_x und \mathbf{n}_y linear unabhängig, so daß wir ähnlich wie im Anhang A.2.2 von [Horn 86] ansetzen können:

$$\mathbf{E}_z = \lambda \mathbf{n}_x + \mu \mathbf{n}_y + \nu (\mathbf{n}_x \times \mathbf{n}_y). \quad (\text{B.38})$$

Da P eine lineare Funktion ist und $P(\mathbf{n}_x \times \mathbf{n}_y) = 0$ gilt, folgt:

$$P(\mathbf{E}_z) = \lambda P(\mathbf{n}_x) + \mu P(\mathbf{n}_y), \quad (\text{B.39})$$

woraus man, da nach Lemma B.41 und Voraussetzung 4 mit \mathbf{n}_x auch $P(\mathbf{E}_z)$ bekannt ist, folgendes Gleichungssystem erhält:

$$\begin{pmatrix} P(\mathbf{n}_x) & P(\mathbf{n}_y) \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = P(\mathbf{E}_z). \quad (\text{B.40})$$

Nach Voraussetzung 3 existiert die inverse Matrix $(P(\mathbf{n}_x), P(\mathbf{n}_y))^{-1}$, so daß man λ und μ bestimmen kann. Aus Gl. B.38 folgt weiterhin:

$$\begin{aligned} \mathbf{E}_z \times \mathbf{n}_x &= -\mu (\mathbf{n}_x \times \mathbf{n}_y) + \nu (\mathbf{n}_x \times \mathbf{n}_y) \times \mathbf{n}_x, \\ \mathbf{E}_z \times \mathbf{n}_y &= \lambda (\mathbf{n}_x \times \mathbf{n}_y) + \nu (\mathbf{n}_x \times \mathbf{n}_y) \times \mathbf{n}_y \end{aligned} \quad (\text{B.41})$$

¹Insbesondere sind dann $P(\mathbf{E}_x)$ sowie $P(\mathbf{E}_y)$ vom Nullvektor verschieden.

²Diese Forderung folgt nicht aus Voraussetzung 3: zeigen $\mathbf{E}_x, \mathbf{E}_y, \mathbf{E}_z$ beispielsweise jeweils in die Richtungen der entsprechenden Achsen des Kamerakoordinatensystems, so sind $P(\mathbf{E}_x)$ und $P(\mathbf{E}_y)$ orthogonal und linear unabhängig, $P(\mathbf{E}_z)$ jedoch der Nullvektor.

³In diesem Fall bilden $\mathbf{E}_x, \mathbf{E}_y, \mathbf{E}_z$ ein Orthogonalsystem. Sind darüber hinaus die Projektionen $P(\mathbf{E}_x), P(\mathbf{E}_y)$ orthogonal, dann sind nach Lemma B.42 die Projektionen $P(\mathbf{E}_x), P(\mathbf{E}_z)$ oder $P(\mathbf{E}_y), P(\mathbf{E}_z)$ linear abhängig.

und mit der Beziehung $(\mathbf{a} \times \mathbf{b})^T(\mathbf{c} \times \mathbf{d}) = (\mathbf{a}^T \mathbf{c})(\mathbf{b}^T \mathbf{d}) - (\mathbf{a}^T \mathbf{d})(\mathbf{b}^T \mathbf{c})$ (vgl. S. 99 in [Kanatani 93a]) demnach:

$$\begin{aligned} & (\mathbf{E}_z \times \mathbf{n}_x)^T \cdot (\mathbf{E}_z \times \mathbf{n}_y) \\ &= -\lambda\mu \|(\mathbf{n}_x \times \mathbf{n}_y)\|^2 + \nu^2 ((\mathbf{n}_x \times \mathbf{n}_y) \times \mathbf{n}_z)^T \cdot ((\mathbf{n}_x \times \mathbf{n}_y) \times \mathbf{n}_y) \\ &= -\lambda\mu \|(\mathbf{n}_x \times \mathbf{n}_y)\|^2 + \nu^2 \|(\mathbf{n}_x \times \mathbf{n}_y)\|^2 \mathbf{n}_z^T \cdot \mathbf{n}_y. \end{aligned} \quad (\text{B.42})$$

Desweiteren ist

$$\|\mathbf{E}_z \times \mathbf{n}_x\|^2 = \mu^2 \|(\mathbf{n}_x \times \mathbf{n}_y)\|^2 + \nu^2 \|(\mathbf{n}_x \times \mathbf{n}_y)\|^2 \|\mathbf{n}_z\|^2 \quad (\text{B.43})$$

$$\|\mathbf{E}_z \times \mathbf{n}_y\|^2 = \lambda^2 \|(\mathbf{n}_x \times \mathbf{n}_y)\|^2 + \nu^2 \|(\mathbf{n}_x \times \mathbf{n}_y)\|^2 \|\mathbf{n}_y\|^2. \quad (\text{B.44})$$

Nach Voraussetzung 5 kennt man $\cos \Delta\gamma$. So erhält man aus

$$\frac{(\mathbf{E}_z \times \mathbf{n}_x)^T \cdot (\mathbf{E}_z \times \mathbf{n}_y)}{\|\mathbf{E}_z \times \mathbf{n}_x\| \|\mathbf{E}_z \times \mathbf{n}_y\|} = \cos \Delta\gamma \quad (\text{B.45})$$

die Gleichung

$$\frac{-\lambda\mu + \nu^2 \mathbf{n}_z^T \cdot \mathbf{n}_y}{\sqrt{\mu^2 + \nu^2 \|\mathbf{n}_z\|^2} \sqrt{\lambda^2 + \nu^2 \|\mathbf{n}_y\|^2}} = \cos \Delta\gamma. \quad (\text{B.46})$$

Nun gilt:

- aus Voraussetzung 3 folgt $\|\mathbf{n}_x\|^2 \neq 0$ und $\|\mathbf{n}_y\|^2 \neq 0$,
- im Fall $\cos \Delta\gamma \neq 0$ erhält man aus Gl. B.46 ν und folglich \mathbf{E}_z ,
- im Fall $\cos \Delta\gamma = 0$ erhält man nur mit der Zusatzbedingung $\mathbf{n}_z^T \cdot \mathbf{n}_y \neq 0$ aus Gl. B.46 eine Lösung für ν . Die Zusatzbedingung $\mathbf{n}_z^T \cdot \mathbf{n}_y \neq 0$ folgt aus Voraussetzung 6 mit Lemma B.42.

□

Korollar B.44 Sind durch homogene Koordinatenvektoren $\mathbf{n}_x, \mathbf{n}_y$ und \mathbf{n}_z drei Bild-Ursprungsgeraden durch drei zu den 3D-Richtungsvektoren $\mathbf{E}_x, \mathbf{E}_y$ und \mathbf{E}_z gehörende Fluchtpunkte gegeben:

$$P(\mathbf{n}_x)^T \cdot P(\mathbf{E}_z) = 0, \quad (\text{B.47})$$

$$P(\mathbf{n}_y)^T \cdot P(\mathbf{E}_y) = 0, \quad (\text{B.48})$$

$$P(\mathbf{n}_z)^T \cdot P(\mathbf{E}_z) = 0, \quad (\text{B.49})$$

gilt

$$\mathbf{E}_z = \mathbf{E}_x \times \mathbf{E}_y, \quad (\text{B.50})$$

$$\mathbf{E}_x^T \cdot \mathbf{E}_y = 0, \quad (\text{B.51})$$

und sind je zwei der Vektoren $P(\mathbf{E}_x)$, $P(\mathbf{E}_y)$ und $P(\mathbf{E}_z)$ linear unabhängig — insbesondere zeigt dann keiner der 3D-Richtungsvektoren in Richtung der optischen Achse —, dann kann man \mathbf{E}_z bestimmen.

Korollar B.45 *Mit den Voraussetzungen von Korollar B.44 lassen sich aus den drei Ursprungsgeraden durch die Hauptfluchtpunkte die Rotationsparameter — unabhängig von der Brennweite — vollständig berechnen.*

Tab. B.1 faßt verschiedene Möglichkeiten zur Berechnung von Kalibrierungsparametern aus Fluchtpunkten zusammen.

B.5 Rückprojektion von Geradenorientierungen von der Bild- in eine Raumebene

Zur Berechnung einer Rückprojektion von Orientierungen von Bildgeraden aus dem Bildbereich in die Fahrbahnebene in der Szene reicht es, wenn man den Normalenvektor der Fahrbahnebene kennt:

Lemma B.46 (Gerade in bekannter Ebene) *Wenn das Urbild einer Bildgeraden $ax + by + c = 0$ in einer bekannten Ebene $AX + BY + CZ + D = 0$ liegen soll, dann erhält man den Richtungsvektor der Raumgeraden unabhängig von der Tiefe D aus [Haralick & Shapiro 93]:*

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} \times \begin{pmatrix} a \\ b \\ c/f \end{pmatrix}. \quad (\text{B.52})$$

Beweis: Der Richtungsvektor des Urbilds der Geraden muß zum einen orthogonal zum Normalenvektor der Ebene, zum anderen nach Lemma B.11 orthogonal zum homogenen Koordinatenvektor der Geraden sein. \square

Korollar B.47 *Die Berechnung der Richtung des Urbilds einer Bild-Ursprungsgeraden $ax + by = 0$, die in einer Ebene $AX + BY + CZ + D = 0$ liegt, ist bei Kenntnis der Ebenennormale $(A, B, C)^T$ unabhängig von der Brennweite f und der Tiefe D möglich.*

Lemma B.48 *Die Berechnung des Schnittwinkels der in einer Ebene $AX + BY + CZ + D = 0$ liegenden Urbilder zweier Bildgeraden $a_1x + b_1y + c_1 = 0$ und $a_2x + b_2y + c_2 = 0$ ist bei Kenntnis der Brennweite f und der Ebenennormale $(A, B, C)^T$ sogar ohne Kenntnis der Tiefe D möglich. Gilt $c_1 = c_2 = 0$, so kann der Schnittwinkel unabhängig von der Brennweite f berechnet werden.*

Beweis: Der Kosinus des Schnittwinkels der in der Ebene $AX + BY + CZ + D = 0$ liegenden Urbilder ergibt sich aus dem Skalarprodukt der nach Lemma B.46 mit Gl. B.52 berechneten 3D-Richtungsvektoren der Urbilder. \square

	Normalenvektor der Fahrbahn und damit die Rotationsmatrix	Brenn- weite	Schnittwinkel der Straßen	Orientierung der Masten
Fluchtpunkt vertikaler Geraden und Brennweite	ja, bei senk- rechten Masten Lemma B.23			ja Lemma B.23
Horizont und Brennweite	ja Korollar B.30			
2 Fluchtpunkte von zwei aufeinander senkrechten Fahrbahnbegrenzungen	ja Lemma B.36	ja Satz B.33	90 Grad	
2 Fluchtpunkte der Fahrbahnbegrenzung und Schnittwinkel	ja Satz B.37	ja Satz B.37		
2 Fluchtpunkte der Fahrbahnbegrenzung und Brennweite	ja Satz B.37		ja Satz B.37	
3 Hauptfluchtpunkte: sich senkrecht kreuzende Straßen, senkrechte Masten	ja Lemma B.36 (über- bestimmt)	ja Satz B.33 (über- bestimmt)	90 Grad	90 Grad
3 Fluchtpunkte sowie unbekannter Schnittwinkel sich kreuzender Straßen, senkrechte Masten	ja Satz B.37	ja Satz B.33	ja Satz B.37	90 Grad
3 Fluchtpunkte sowie keine Information über den Schnittwinkel sich kreuzender Straßen bzw. Orientierung der Masten, aber bekannte Brennweite	ja Satz B.37		ja Satz B.37	ja Lemma B.23
Ursprungsgeraden durch 3 Hauptfluchtpunkte: senkrechte Straßen, senkrechte Masten	ja Korollar B.45		90 Grad	90 Grad
Ursprungsgeraden durch 3 Fluchtpunkte, Schnitt- winkel der Straßen, senkrechte Masten	ja Satz B.43			90 Grad

Tabelle B.1: Diskutierte Möglichkeiten zur Berechnung der Rotationsmatrix, der Brennweite, des Schnittwinkels der Straßenzüge und der Orientierung von Masten. In den letzten beiden Zeilen der Tabelle wird nur die Kenntnis der Bildpunktverhältnisse der Fluchtpunkte, nicht aber der Fluchtpunkte selbst vorausgesetzt.

Anhang C

Abstand zwischen Modell- und Datenkantensegmenten

In diesem Anhang behandeln wir Überlegungen, die im Zusammenhang mit experimentellen Untersuchungen zur Zuordnung von Modell- an Datenkantensegmenten durchgeführt wurden. Zunächst wird dargestellt, wie aus dem Anpaßprozeß von Kantenelementen an ein Geradensegment Unsicherheiten für dieses Geradensegment zu schätzen sind.

Anschließend werden Probleme bei heuristisch gewählten Abstandsfunktionen zwischen Modell- und Datenkantensegmenten aufgezeigt, deren Abklärung aber zu neuen Problemen führte. Folglich wird der ganze Ansatz einer Zuordnung von Modellkantensegmenten zu Datenkantensegmenten verworfen und in Kapitel 6 ein grundsätzlich anderer Weg beschritten, um die Lage eines Objektes in der Szene aus Bilddaten zu schätzen.

C.1 Schätzung von σ_{\perp}^2 aus den Bilddaten

Gegeben sei eine Menge von n Punkten $\{P(\xi_1, \eta_1), \dots, P(\xi_n, \eta_n)\}$, an die wir eine Gerade anpassen möchten. Wir suchen die Gerade, die die Summe der Quadrate der lotrechten Distanzen zwischen den Punkten und der Geraden minimiert. Im Gegensatz zur von [Koller 92] verwendeten Methode der kleinsten Fehlerquadrate, die als Fehler die Differenz entlang der Ordinatenachse zwischen einem Punkt und der Geraden betrachtet, ist das folgende Verfahren unabhängig von der Lage des Koordinatensystems.

Nach [Duda & Hart 73; Pavlidis & Horowitz 74] zeigt die gesuchte Gerade in Richtung des Eigenvektors zum größten Eigenwert λ_{\max} der Matrix:

$$\Sigma_S = \begin{pmatrix} V_{\xi\xi} & V_{\xi\eta} \\ V_{\xi\eta} & V_{\eta\eta} \end{pmatrix} \quad (\text{C.1})$$

mit

$$\begin{aligned}
 E_{\xi} &= \frac{1}{n} \sum_{i=1}^n \xi_i, & V_{\xi\xi} &= \frac{1}{n} \sum_{i=1}^n (\xi_i - E_{\xi})^2, \\
 E_{\eta} &= \frac{1}{n} \sum_{i=1}^n \eta_i, & V_{\eta\eta} &= \frac{1}{n} \sum_{i=1}^n (\eta_i - E_{\eta})^2, \\
 & & V_{\xi\eta} &= \frac{1}{n} \sum_{i=1}^n (\xi_i - E_{\xi})(\eta_i - E_{\eta}).
 \end{aligned} \tag{C.2}$$

Die Eigenwerte der Matrix Σ_S sind

$$\begin{aligned}
 \lambda_{\max} &= \frac{V_{\xi\xi} + V_{\eta\eta} + \sqrt{(V_{\xi\xi} - V_{\eta\eta})^2 + 4V_{\xi\eta}^2}}{2} \quad (\text{größerer Eigenwert}), \\
 \lambda_{\min} &= \frac{V_{\xi\xi} + V_{\eta\eta} - \sqrt{(V_{\xi\xi} - V_{\eta\eta})^2 + 4V_{\xi\eta}^2}}{2} \quad (\text{kleinerer Eigenwert}).
 \end{aligned} \tag{C.3}$$

Die angepaßte Gerade hat nach [Pavlidis & Horowitz 74] den Winkel

$$\phi = \frac{1}{2} \arctan \frac{2V_{\xi\eta}}{V_{\xi\xi} - V_{\eta\eta}} \tag{C.4}$$

mit der positiven Richtung der ξ -Achse und dem Abstand

$$d = E_{\xi} \sin \phi - E_{\eta} \cos \phi \tag{C.5}$$

vom Ursprung des Koordinatensystems.

Wir setzen

$$\sigma_{\perp}^2 = \lambda_{\min} \quad (\text{kleinerer Eigenwert}), \tag{C.6}$$

womit der durch σ_{\perp}^2 modellierte Fehler senkrecht zum Kantensegment mit dem in [Pavlidis & Horowitz 74] angegebenen Fehlermaß übereinstimmt.

Zunächst wurde in unseren Experimenten durch die Wahl von $\sigma_{\parallel}^2 = \lambda_{\max}$ die Ausdehnung entlang der Länge eines Kantensegments mit einer Gaußverteilung modelliert; systematische Fehler in der Länge wurden jedoch nicht erfaßt. Da ein langes Datenkantensegment, etwa aufgrund von Verdeckungen, eventuell nur halb so lang wie das zugehörige Modellkantensegment ist, erschien dieses große σ_{\parallel}^2 schon vertretbar. [Deriche & Faugeras 90] geben an, daß, mit der Ausnahme von kurzen Kantensegmenten, σ_{\parallel}^2 viel größer als σ_{\perp}^2 sein soll.

[Zhang & Faugeras 92b] beschreiben den Mittelpunkt \mathbf{m} eines Kantensegments mit Endpunkten M_1 und M_2 als

$$\mathbf{m} = \frac{M_1 + M_2}{2} + n\mathbf{u}, \tag{C.7}$$

wobei \mathbf{u} ein Einheitsvektor in der Richtung des Kantensegments ist und n eine (skalare) normalverteilte Zufallsvariable mit Mittelwert 0 und Standardabweichung $\sigma = \kappa l$ mit der Länge l des Kantensegments sowie einem experimentell bestimmten, vom Kantensegment unabhängigen $\kappa = 0,2$ ist.

In unseren Experimenten ergab die Wahl von $\sigma_{\parallel}^2 = \lambda_{\max}$ keine brauchbaren Ergebnisse: zwar wurden durch Verdeckung aufgebrochene Datenkantensegmente noch richtig zugeordnet, doch auch zahlreiche Datenkantensegmente der Fahrbahnmarkierung wurden oft als Zuordnungskandidaten für Modellkantensegmente von Fahrzeugen ausgewählt, denn σ_{\parallel}^2 nahm gerade bei großen Kantensegmenten sehr große Werte an. Daher wurde σ_{\parallel}^2 nicht mehr auf diese Weise aus den Bilddaten geschätzt, sondern wie bei [Koller 92] fest auf $\sigma_{\parallel}^2 = 2,4$ eingestellt.

C.2 Unsicherheit in der Orientierung eines Kantensegments

Gegeben seien n Punkte $\{P(\xi_1, \eta_1), \dots, P(\xi_n, \eta_n)\}$, die an ein Kantensegment angepaßt werden sollen.

Wir berechnen nun, wie sich Fehler der Punkte $\{P(\xi_1, \eta_1), \dots, P(\xi_n, \eta_n)\}$ zu einem Fehler von ϕ fortpflanzen. Ein analoges Ergebnis wurde inzwischen von [Daucher *et al.* 93] publiziert, allerdings ohne eine Herleitung. Untersuchungen zu dieser Fragestellung finden sich auch bei [Kanatani 93a], der eine Fehlerrechnung aber nicht in der Bildebene, sondern auf einer Bildsphäre durchführt.

Wir nehmen an, die Fehler in den Positionen der Punkte $P(\xi_i, \eta_i)$ seien unkorreliert und haben die gleiche Kovarianzmatrix

$$\Sigma_{P(\xi_i, \eta_i)} = \begin{pmatrix} \sigma_{\xi\xi}^2 & \sigma_{\xi\eta}^2 \\ \sigma_{\xi\eta}^2 & \sigma_{\eta\eta}^2 \end{pmatrix}. \quad (\text{C.8})$$

Mit den Ableitungen

$$\begin{aligned} \frac{\partial V_{\xi\xi}}{\partial \xi_k} &= \frac{\partial}{\partial \xi_k} \frac{1}{n} \sum_{i=1}^n (\xi_i - E_{\xi})^2 & \frac{\partial V_{\xi\xi}}{\partial \eta_k} &= 0 \\ &= \frac{2}{n} \sum_{i=1}^n \left((\xi_i - E_{\xi})(\delta_{ik} - \frac{\partial}{\partial \xi_k} E_{\xi}) \right) \\ &= \frac{2}{n} \sum_{i=1}^n \left((\xi_i - E_{\xi})(\delta_{ik} - \frac{1}{n}) \right) \\ &= \frac{2}{n} (\xi_k - E_{\xi}) \end{aligned} \quad (\text{C.9})$$

$$\begin{aligned} \frac{\partial V_{\eta\eta}}{\partial \xi_k} &= 0 & \frac{\partial V_{\eta\eta}}{\partial \eta_k} &= \frac{2}{n} (\eta_k - E_{\eta}) \\ \frac{\partial V_{\xi\eta}}{\partial \xi_k} &= \frac{1}{n} (\eta_k - E_{\eta}) & \frac{\partial V_{\xi\eta}}{\partial \eta_k} &= \frac{1}{n} (\xi_k - E_{\xi}) \end{aligned}$$

erhält man aus Gl. C.4

$$\begin{aligned}
 \frac{\partial \phi}{\partial \xi_k} &= \frac{\partial}{\partial \xi_k} \frac{1}{2} \arctan \frac{2V_{\xi\eta}}{V_{\xi\xi} - V_{\eta\eta}} \\
 &= \frac{1}{2} \frac{1}{1 + \left(\frac{2V_{\xi\eta}}{V_{\xi\xi} - V_{\eta\eta}}\right)^2} 2 \frac{(V_{\xi\xi} - V_{\eta\eta}) \frac{\partial V_{\xi\eta}}{\partial \xi_k} - V_{\xi\eta} \frac{\partial (V_{\xi\xi} - V_{\eta\eta})}{\partial \xi_k}}{(V_{\xi\xi} - V_{\eta\eta})^2} \\
 &= \frac{1}{1 + \tan^2 2\phi} \frac{(V_{\xi\xi} - V_{\eta\eta}) \frac{1}{n} (\eta_k - E_\eta) - V_{\xi\eta} \frac{2}{n} (\xi_k - E_\xi)}{(V_{\xi\xi} - V_{\eta\eta})^2} \\
 &= \frac{1}{n} \frac{\cos^2 2\phi}{V_{\xi\xi} - V_{\eta\eta}} ((\eta_k - E_\eta) - \tan 2\phi (\xi_k - E_\xi)) \quad (C.10)
 \end{aligned}$$

und

$$\frac{\partial \phi}{\partial \eta_k} = \frac{1}{n} \frac{\cos^2 2\phi}{V_{\xi\xi} - V_{\eta\eta}} ((\xi_k - E_\xi) + \tan 2\phi (\eta_k - E_\eta)). \quad (C.11)$$

Mit Hilfe der Fehlerfortpflanzungsregel erhält man für den Fehler σ_ϕ^2 des Winkels ϕ

$$\begin{aligned}
 \sigma_\phi^2 &= \sum_{k=1}^n \left(\begin{array}{cc} \frac{\partial \phi}{\partial \xi_k} & \frac{\partial \phi}{\partial \eta_k} \end{array} \right) \left(\begin{array}{cc} \sigma_{\xi\xi}^2 & \sigma_{\xi\eta}^2 \\ \sigma_{\xi\eta}^2 & \sigma_{\eta\eta}^2 \end{array} \right) \left(\begin{array}{c} \frac{\partial \phi}{\partial \xi_k} \\ \frac{\partial \phi}{\partial \eta_k} \end{array} \right) \\
 &= \sum_{k=1}^n \left(\left(\frac{\partial \phi}{\partial \xi_k} \right)^2 \sigma_{\xi\xi}^2 + 2 \frac{\partial \phi}{\partial \xi_k} \frac{\partial \phi}{\partial \eta_k} \sigma_{\xi\eta}^2 + \left(\frac{\partial \phi}{\partial \eta_k} \right)^2 \sigma_{\eta\eta}^2 \right) \\
 &= \frac{1}{n} \frac{\cos^4 2\phi}{(V_{\xi\xi} - V_{\eta\eta})^2} \left\{ \begin{array}{l} [V_{\eta\eta} - 2V_{\xi\eta} \tan 2\phi + V_{\xi\xi} \tan^2 2\phi] \sigma_{\xi\xi}^2 \\ + 2[(-\tan 2\phi)V_{\xi\xi} + (1 - \tan^2 2\phi)V_{\xi\eta} + (\tan 2\phi)V_{\eta\eta}] \sigma_{\xi\eta}^2 \\ + [V_{\xi\xi} + 2V_{\xi\eta} \tan 2\phi + V_{\eta\eta} \tan^2 2\phi] \sigma_{\eta\eta}^2 \end{array} \right\} \\
 &= \frac{1}{n} \frac{1}{(V_{\xi\xi} - V_{\eta\eta})^2 \left(1 + \frac{4V_{\xi\eta}^2}{(V_{\xi\xi} - V_{\eta\eta})^2}\right)^2} \left\{ \begin{array}{l} \left[V_{\eta\eta} - \frac{4V_{\xi\eta}^2}{V_{\xi\xi} - V_{\eta\eta}} + \frac{4V_{\xi\xi} V_{\xi\eta}^2}{(V_{\xi\xi} - V_{\eta\eta})^2} \right] \sigma_{\xi\xi}^2 \\ + 2 \left[-\frac{2V_{\xi\eta} V_{\xi\xi}}{V_{\xi\xi} - V_{\eta\eta}} + \left(1 - \frac{4V_{\xi\eta}^2}{(V_{\xi\xi} - V_{\eta\eta})^2}\right) V_{\xi\eta} + \frac{2V_{\xi\eta} V_{\eta\eta}}{V_{\xi\xi} - V_{\eta\eta}} \right] \sigma_{\xi\eta}^2 \\ + \left[V_{\xi\xi} + \frac{4V_{\xi\eta}^2}{V_{\xi\xi} - V_{\eta\eta}} + \frac{4V_{\eta\eta} V_{\xi\eta}^2}{(V_{\xi\xi} - V_{\eta\eta})^2} \right] \sigma_{\eta\eta}^2 \end{array} \right\} \\
 &= \frac{1}{n} \frac{(V_{\xi\xi} - V_{\eta\eta})^2}{(V_{\xi\xi}^2 - 2V_{\xi\xi} V_{\eta\eta} + V_{\eta\eta}^2 + 4V_{\xi\eta}^2)^2} \left\{ \begin{array}{l} \left[\frac{V_{\eta\eta} (V_{\xi\xi}^2 - 2V_{\xi\xi} V_{\eta\eta} + V_{\eta\eta}^2 + 4V_{\xi\eta}^2)}{(V_{\xi\xi} - V_{\eta\eta})^2} \right] \sigma_{\xi\xi}^2 \\ + 2 \left[\frac{-V_{\xi\eta} (V_{\xi\xi}^2 - 2V_{\xi\xi} V_{\eta\eta} + V_{\eta\eta}^2 + 4V_{\xi\eta}^2)}{(V_{\xi\xi} - V_{\eta\eta})^2} \right] \sigma_{\xi\eta}^2 \\ + \left[\frac{V_{\xi\xi} (V_{\xi\xi}^2 - 2V_{\xi\xi} V_{\eta\eta} + V_{\eta\eta}^2 + 4V_{\xi\eta}^2)}{(V_{\xi\xi} - V_{\eta\eta})^2} \right] \sigma_{\eta\eta}^2 \end{array} \right\}
 \end{aligned}$$

$$= \frac{1}{n} \frac{V_{\eta\eta} \sigma_{\xi\xi}^2 - 2V_{\xi\eta} \sigma_{\xi\eta}^2 + V_{\xi\xi} \sigma_{\eta\eta}^2}{(V_{\xi\xi} - V_{\eta\eta})^2 + 4V_{\xi\eta}^2} \quad (C.12)$$

Mit den Annahmen $\sigma_{\xi\eta}^2 = 0$ und $\sigma_{\xi\xi}^2 = \sigma_{\eta\eta}^2 = \sigma^2$ erhält man aus der letzten Zeile unter Berücksichtigung von Gl. C.3

$$\begin{aligned} \sigma_{\phi}^2 &= \frac{1}{n} \frac{V_{\xi\xi} + V_{\eta\eta}}{(V_{\xi\xi} - V_{\eta\eta})^2 + 4V_{\xi\eta}^2} \sigma^2 \\ &= \frac{1}{n} \frac{\lambda_{\max} + \lambda_{\min}}{(\lambda_{\max} - \lambda_{\min})^2} \sigma^2. \end{aligned} \quad (C.13)$$

Damit ist σ_{ϕ}^2 unabhängig vom Winkel ϕ und invariant gegenüber einer gemeinsamen Rotation der Kantenelemente. Da die Matrix Σ_S positiv (semi)definit ist, sind ihre beiden Eigenwerte λ_{\max} und λ_{\min} nicht negativ. Für große $\Delta\lambda = \lambda_{\max} - \lambda_{\min}$ ist $\sigma_{\phi}^2 \approx \frac{1}{n} \frac{1}{\lambda_{\max}} \sigma^2$, für $\Delta\lambda = \lambda_{\max} - \lambda_{\min} \rightarrow +\infty$ konvergiert $\sigma_{\phi}^2 \rightarrow 0$. Für $\Delta\lambda = \lambda_{\max} - \lambda_{\min} \rightarrow 0$ konvergiert $\sigma_{\phi}^2 \rightarrow +\infty$; in diesem Fall sollten die eingehenden Kantenelemente statt an ein Geradensegment besser an einen Kreis angepaßt werden.

Außerdem hängt σ_{ϕ}^2 antiproportional von der Anzahl n der in die Anpassung eingehenden Kantenelemente ab. Ein Nachteil dieser Schätzung von σ_{ϕ}^2 besteht in der Einführung eines neuen Parameters σ^2 , der die Unsicherheit der einzelnen Pixelpositionen beschreibt.

C.3 Diskussion der Mahalanobisdistanz zwischen Modell- und Datensegmenten

C.3.1 Definition der Mahalanobisdistanz

Für eine n -dimensionale, normalverteilte Zufallsvariable \mathbf{x} mit Erwartungswert $\boldsymbol{\mu}$ und Kovarianzmatrix Σ heißt die Größe

$$d^2(\mathbf{x}, \boldsymbol{\mu}) = (\mathbf{x} - \boldsymbol{\mu})^T (\Sigma)^{-1} (\mathbf{x} - \boldsymbol{\mu}) \quad (C.14)$$

die (quadratische) Mahalanobisdistanz von \mathbf{x} zu $\boldsymbol{\mu}$ [Duda & Hart 73; Tou & Gonzales 74]. Sie genügt einer χ^2 -Verteilung mit n Freiheitsgraden [Bar-Shalom & Fortmann 88]. Mit einer Konstanten k definiert die Gleichung

$$(\mathbf{x} - \boldsymbol{\mu})^T (\Sigma)^{-1} (\mathbf{x} - \boldsymbol{\mu}) = k^2 \quad (C.15)$$

ein Hyperellipsoid mit „Radius“ k [Niemann 83]. Mit Hilfe einer Vertafelung der χ^2 -Verteilung läßt sich die Wahrscheinlichkeit, daß \mathbf{x} im Hyperellipsoid liegt, ermitteln.

Da die Addition einer Konstanten \mathbf{c} zu einer Zufallsvariablen die Kovarianzmatrix unverändert läßt, jedoch den Erwartungswert verschiebt, gilt für die Zufallsvariable $\mathbf{y} = \mathbf{x} + \mathbf{c}$:

$$d^2(\mathbf{y}, \boldsymbol{\mu} + \mathbf{c}) = d^2(\mathbf{x}, \boldsymbol{\mu}) \quad (C.16)$$

die Addition einer Konstanten hat folglich keinen Einfluß auf die Mahalanobisdistanz einer Zufallsvariablen \mathbf{x} von ihrem Erwartungswert $\boldsymbol{\mu}$.

C.3.2 Mahalanobisdistanz zweier Zufallsvariablen

Mit der Mahalanobisdistanz kann man auch die Differenz zweier Zufallsvariablen — beispielsweise eines Datenkantensegments \mathbf{x}_D und eines Modellkantensegments \mathbf{x}_M — beschreiben, indem man ihre Differenz mit Hilfe der Kovarianzmatrix ihrer Differenz normalisiert: hat \mathbf{x}_D die Kovarianzmatrix Σ_D sowie \mathbf{x}_M die Kovarianzmatrix Σ_M , so hat ihre Differenz $\mathbf{x}_D - \mathbf{x}_M$ bei stochastischer Unabhängigkeit die Kovarianzmatrix $\Sigma_D + \Sigma_M$. Die (quadratische) Mahalanobisdistanz zwischen \mathbf{x}_D und \mathbf{x}_M ist definiert als:

$$d^2(\mathbf{x}_D, \mathbf{x}_M) = (\mathbf{x}_D - \mathbf{x}_M)^T (\Sigma_D + \Sigma_M)^{-1} (\mathbf{x}_D - \mathbf{x}_M) \quad . \quad (\text{C.17})$$

Diese hat jedoch nur dann eine χ^2 -Verteilung, wenn der Erwartungswert von $\mathbf{x}_D - \mathbf{x}_M$ verschwindet, wenn demnach \mathbf{x}_D und \mathbf{x}_M denselben Erwartungswert haben.

Mit einem festen k bzw. einer vorgegebenen Wahrscheinlichkeit — [Bar-Shalom & Fortmann 88] schlagen 95% vor — definiert die Mahalanobisdistanz ein Hyperellipsoid wie in Gl. C.15. Das Innere des Hyperellipsoids nennen wir Akzeptionsbereich (für eine mögliche Zuordnung von \mathbf{x}_D zu \mathbf{x}_M), das Äußere Ablehnungsbereich.

Die Mahalanobisdistanz zweier Zufallsvariablen ist unabhängig von der Wahl des Koordinatensystems. Geht durch Drehung mit Drehmatrix R und Translation \mathbf{t} eine Zufallsvariable \mathbf{x} mit Kovarianzmatrix Σ in die Zufallsvariable $R\mathbf{x} + \mathbf{t}$ mit Kovarianzmatrix $R\Sigma R^T$ über, so erhält man für die Mahalanobisdistanz:

$$\begin{aligned} d^2(R\mathbf{x}_D + \mathbf{t}, R\mathbf{x}_M + \mathbf{t}) &= (R\mathbf{x}_D - R\mathbf{x}_M)^T (R\Sigma_D R^T + R\Sigma_M R^T)^{-1} (R\mathbf{x}_D - R\mathbf{x}_M) \\ &= (R(\mathbf{x}_D - \mathbf{x}_M))^T (R^T)^{-1} (\Sigma_D + \Sigma_M)^{-1} R^{-1} R(\mathbf{x}_D - \mathbf{x}_M) \\ &= (\mathbf{x}_D - \mathbf{x}_M)^T R^T R (\Sigma_D + \Sigma_M)^{-1} R^T R (\mathbf{x}_D - \mathbf{x}_M) \\ &= (\mathbf{x}_D - \mathbf{x}_M)^T (\Sigma_D + \Sigma_M)^{-1} (\mathbf{x}_D - \mathbf{x}_M) \\ &= d^2(\mathbf{x}_D, \mathbf{x}_M) \quad . \end{aligned} \quad (\text{C.18})$$

Diese Aussage läßt sich verallgemeinern auf lineare Transformationen [Niemann 83]: Geht durch lineare Transformationen mit einer regulären Matrix S und einer Translation um den Vektor \mathbf{t} eine Zufallsvariable \mathbf{x} mit Kovarianzmatrix Σ in die Zufallsvariable $S\mathbf{x} + \mathbf{t}$ mit Kovarianzmatrix $S\Sigma S^T$ über, so erhält man für die Mahalanobisdistanz:

$$\begin{aligned} d^2(S\mathbf{x}_D + \mathbf{t}, S\mathbf{x}_M + \mathbf{t}) &= (S\mathbf{x}_D - S\mathbf{x}_M)^T (S\Sigma_D S^T + S\Sigma_M S^T)^{-1} (S\mathbf{x}_D - S\mathbf{x}_M) \\ &= (S(\mathbf{x}_D - \mathbf{x}_M))^T (S^T)^{-1} (\Sigma_D + \Sigma_M)^{-1} S^{-1} S(\mathbf{x}_D - \mathbf{x}_M) \\ &= (\mathbf{x}_D - \mathbf{x}_M)^T S^T (S^T)^{-1} (\Sigma_D + \Sigma_M)^{-1} (\mathbf{x}_D - \mathbf{x}_M) \\ &= (\mathbf{x}_D - \mathbf{x}_M)^T (\Sigma_D + \Sigma_M)^{-1} (\mathbf{x}_D - \mathbf{x}_M) \\ &= d^2(\mathbf{x}_D, \mathbf{x}_M) \quad . \end{aligned} \quad (\text{C.19})$$

Die Mahalanobisdistanz zweier Zufallsvariablen ist demnach invariant gegenüber Kongruenz- und Ähnlichkeitsabbildungen sowie — auch unterschiedlichen — Skalierungen der Koordinatenachsen. Die Unabhängigkeit der Mahalanobisdistanz gegenüber Skalierung ist nur ein Spezialfall der durch die Mahalanobisdistanz ja gerade beabsichtigten Unabhängigkeit des Abstandsmaßes von Schwankungen, die durch eine „Division“ durch die Kovarianzmatrix der Differenz erzwungen wird.

C.3.3 Verallgemeinerte Mahalanobisdistanz zweier Zufallsvariablen

Sei $f(\mathbf{x}_D, \mathbf{x}_M)$ eine (vektorwertige) Funktion zweier unabhängiger Zufallsvariablen \mathbf{x}_D und \mathbf{x}_M mit Kovarianzmatrix Σ_D bzw. Σ_M . Dann wird durch

$$d_f^2(\mathbf{x}_D, \mathbf{x}_M) = \mathbf{f}^T(\mathbf{x}_D, \mathbf{x}_M) \Sigma_f^{-1} \mathbf{f}(\mathbf{x}_D, \mathbf{x}_M) \quad (\text{C.20})$$

eine verallgemeinerte (quadratische) Mahalanobisdistanz zwischen \mathbf{x}_D und \mathbf{x}_M definiert, wobei sich die Kovarianzmatrix Σ_f aus

$$\Sigma_f = \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial(\mathbf{x}_D, \mathbf{x}_M)} \end{pmatrix} \begin{pmatrix} \Sigma_D & | & 0 \\ \hline & & \\ 0 & | & \Sigma_M \end{pmatrix} \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial(\mathbf{x}_D, \mathbf{x}_M)} \end{pmatrix}^T \quad (\text{C.21})$$

$$= \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial \mathbf{x}_D} & | & \frac{\partial \mathbf{f}}{\partial \mathbf{x}_M} \end{pmatrix} \begin{pmatrix} \Sigma_D & | & 0 \\ \hline & & \\ 0 & | & \Sigma_M \end{pmatrix} \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial \mathbf{x}_D} \\ \hline \frac{\partial \mathbf{f}}{\partial \mathbf{x}_M} \end{pmatrix} \quad (\text{C.22})$$

$$= \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial \mathbf{x}_D} \end{pmatrix} \Sigma_D \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial \mathbf{x}_D} \end{pmatrix}^T + \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial \mathbf{x}_M} \end{pmatrix} \Sigma_M \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial \mathbf{x}_M} \end{pmatrix}^T \quad (\text{C.23})$$

berechnen läßt. Wir halten fest, daß sich — ohne weitere Forderungen an die die Distanz definierende Funktion f — die Kovarianzmatrizen von \mathbf{x}_D bzw. \mathbf{x}_M getrennt transformieren lassen, was zu einer Halbierung der Zeilen- bzw. Spaltenzahl der auftretenden Matrizen führt.

Die Funktion d_f^2 ist als quadratische Form in \mathbf{f} nur positiv semidefinit. Ist die Funktion f in ihren beiden Argumenten symmetrisch, d.h. gilt $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$, so ist auch die Funktion d_f^2 symmetrisch. Gilt für die Funktion f die Gleichung $f(\mathbf{x}, \mathbf{x}) = 0$, dann gilt auch $d_f^2(\mathbf{x}, \mathbf{x}) = 0$. Außerdem gilt die Dreiecksungleichung $d_f^2(\mathbf{x}, \mathbf{y}) \leq d_f^2(\mathbf{x}, \mathbf{z}) + d_f^2(\mathbf{z}, \mathbf{y})$. Demnach ist d_f^2 eine Halbmetrik¹, wenn f symmetrisch ist und die Gleichung $f(\mathbf{x}, \mathbf{x}) = 0$ erfüllt.

¹ Eine Halbmetrik ρ ordnet jedem Paar \mathbf{x}, \mathbf{y} eine reelle Zahl zu, so daß folgende Eigenschaften erfüllt sind:

- $\rho(\mathbf{x}, \mathbf{y}) \geq 0$;

C.3.4 Transformation der verallgemeinerten Mahalanobisdistanz zweier Zufallsvariablen bei Verkettung ihrer definierenden Funktionen

Verkettet man die eine verallgemeinerte Mahalanobisdistanz zweier unabhängiger Zufallsvariablen \mathbf{x}_D und \mathbf{x}_M mit Kovarianzmatrix Σ_D bzw. Σ_M definierende Funktion $\mathbf{f}(\mathbf{x}_D, \mathbf{x}_M)$ mit einer Funktion \mathbf{g} :

$$d_{g \circ \mathbf{f}}^2(\mathbf{x}_D, \mathbf{x}_M) = (\mathbf{g} \circ \mathbf{f})^T(\mathbf{x}_D, \mathbf{x}_M) \Sigma_{g \circ \mathbf{f}}^{-1}(\mathbf{g} \circ \mathbf{f})(\mathbf{x}_D, \mathbf{x}_M), \quad (\text{C.24})$$

so erhält man die zur transformierten verallgemeinerten Mahalanobisdistanz gehörende Kovarianzmatrix $\Sigma_{g \circ \mathbf{f}}$ zum einen aus:

$$\Sigma_{g \circ \mathbf{f}} = \left(\frac{\partial(\mathbf{g} \circ \mathbf{f})}{\partial \mathbf{x}_D} \right) \Sigma_D \left(\frac{\partial(\mathbf{g} \circ \mathbf{f})}{\partial \mathbf{x}_D} \right)^T + \left(\frac{\partial(\mathbf{g} \circ \mathbf{f})}{\partial \mathbf{x}_M} \right) \Sigma_M \left(\frac{\partial(\mathbf{g} \circ \mathbf{f})}{\partial \mathbf{x}_M} \right)^T, \quad (\text{C.25})$$

zum anderen aber mit Hilfe der Kettenregel

$$\left(\frac{\partial(\mathbf{g} \circ \mathbf{f})}{\partial(\mathbf{x}_D, \mathbf{x}_M)} \right) = \left(\frac{\partial \mathbf{g}}{\partial \mathbf{f}} \right) \Big|_{\mathbf{f}=\mathbf{f}(\mathbf{x}_D, \mathbf{x}_M)} \cdot \left(\frac{\partial \mathbf{f}}{\partial(\mathbf{x}_D, \mathbf{x}_M)} \right) \quad (\text{C.26})$$

auch aus:

$$\Sigma_{g \circ \mathbf{f}} = \left(\frac{\partial \mathbf{g}}{\partial \mathbf{f}} \right) \Sigma_{\mathbf{f}} \left(\frac{\partial \mathbf{g}}{\partial \mathbf{f}} \right)^T \Big|_{\mathbf{f}=\mathbf{f}(\mathbf{x}_D, \mathbf{x}_M)}; \quad (\text{C.27})$$

d.h., hat man bereits $\Sigma_{\mathbf{f}}$ berechnet und möchte man das Abstandsmaß mit Hilfe von \mathbf{g} verändern, so braucht man nicht mehr auf die ursprünglichen Kovarianzmatrizen Σ_D bzw. Σ_M zurückzugreifen, sondern kann $\Sigma_{\mathbf{f}}$ und Gl. C.27 benutzen. Außerdem folgt für die Funktion $\mathbf{h} = \mathbf{g} \circ \mathbf{f}$ die Gleichung

$$\Sigma_{\mathbf{h}} = \Sigma_{g \circ \mathbf{f}}. \quad (\text{C.28})$$

Die letzte Aussage betrachten wir nochmals an einem Beispiel: Sei eine Ursprungsgerade durch einen Punkt $P = (\xi, \eta)$ mit Varianzen σ_{ξ}^2 sowie σ_{η}^2 gegeben, wobei wir $\xi > 0$ und $\eta > 0$ voraussetzen. Der Winkel ϕ , den die Gerade mit der positiven Richtung der x -Achse einschließt, läßt sich u.a. durch eine der folgenden Gleichungen bestimmen:

$$\phi_{\arctan}(\xi, \eta) = \arctan \frac{\eta}{\xi}, \quad (\text{C.29})$$

$$\phi_{\arcsin}(\xi, \eta) = \arcsin \frac{\eta}{\sqrt{\xi^2 + \eta^2}}, \quad (\text{C.30})$$

$$\phi_{\arccos}(\xi, \eta) = \arccos \frac{\xi}{\sqrt{\xi^2 + \eta^2}}. \quad (\text{C.31})$$

-
- $\rho(\mathbf{x}, \mathbf{y}) = 0$, falls $\mathbf{x} = \mathbf{y}$;
 - $\rho(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{y}, \mathbf{x})$;
 - $\rho(\mathbf{x}, \mathbf{y}) \leq \rho(\mathbf{x}, \mathbf{z}) + \rho(\mathbf{z}, \mathbf{y})$ für alle $\mathbf{x}, \mathbf{y}, \mathbf{z}$.

Bei einer Metrik gilt zusätzlich: aus $\rho(\mathbf{x}, \mathbf{y}) = 0$ folgt $\mathbf{x} = \mathbf{y}$.

Da die Ableitungen

$$\frac{\partial \phi_{\arctan}}{\partial(\xi, \eta)} = \left(\frac{\eta}{\xi^2 \left(1 + \frac{\eta^2}{\xi^2}\right)}, \frac{1}{\xi \left(1 + \frac{\eta^2}{\xi^2}\right)} \right) = \left(-\frac{\eta}{\xi^2 + \eta^2}, \frac{\xi}{\xi^2 + \eta^2} \right) \quad (\text{C.32})$$

$$\frac{\partial \phi_{\arcsin}}{\partial(\xi, \eta)} = \left(-\frac{\eta\xi}{\sqrt{1 - \frac{\eta^2}{\xi^2 + \eta^2}} (\xi^2 + \eta^2)^{3/2}}, \frac{\frac{1}{\sqrt{\xi^2 + \eta^2}} - \frac{\eta^2}{(\xi^2 + \eta^2)^{3/2}}}{\sqrt{1 - \frac{\eta^2}{\xi^2 + \eta^2}}} \right) = \left(-\frac{\eta}{\xi^2 + \eta^2}, \frac{\xi}{\xi^2 + \eta^2} \right)$$

$$\frac{\partial \phi_{\arccos}}{\partial(\xi, \eta)} = \left(\frac{\frac{1}{\sqrt{\xi^2 + \eta^2}} - \frac{\xi^2}{(\xi^2 + \eta^2)^{3/2}}}{\sqrt{1 - \frac{\xi^2}{\xi^2 + \eta^2}}}, \frac{\xi\eta}{\sqrt{1 - \frac{\xi^2}{\xi^2 + \eta^2}} (\xi^2 + \eta^2)^{3/2}} \right) = \left(-\frac{\eta}{\xi^2 + \eta^2}, \frac{\xi}{\xi^2 + \eta^2} \right)$$

alle übereinstimmen, erhält man — unabhängig von der Wahl einer der drei Berechnungsvorschriften aus Gln. C.29 - C.31 — den Fehler des Winkels aus

$$\sigma_\phi^2 = \left(-\frac{\eta}{\xi^2 + \eta^2} \right)^2 \sigma_\xi^2 + \left(\frac{\xi}{\xi^2 + \eta^2} \right)^2 \sigma_\eta^2. \quad (\text{C.33})$$

Da aber mit

$$f_{\sin}(t) = \frac{t}{\sqrt{1+t^2}} \qquad f_{\cos}(t) = \frac{1}{\sqrt{1+t^2}} \quad (\text{C.34})$$

zum einen $\arctan = \arcsin \circ f_{\sin}$, zum anderen $\arctan = \arccos \circ f_{\cos}$ gilt, folgt die Unabhängigkeit des Fehlers des Winkels von seiner Berechnungsvorschrift direkt aus Gl. C.28. Das heißt aber auch, daß man die Kovarianzmatrix des Kosinus eines Schnittwinkels sowohl direkt mit Hilfe des Skalarprodukts als auch indirekt mit Hilfe der Berechnung des Schnittwinkels als Differenz zweier Arcustangens erhält.

C.3.5 Untersuchung der Mahalanobisdistanz von [Deriche & Faugeras 90] bzw. [Koller 92]

Im folgenden verwenden wir die *Mittelpunkt-Repräsentation* (m_x, m_y, θ, l) nach [Deriche & Faugeras 90] zur Darstellung von Modell- und Datenkantensegmenten und untersuchen ihre Mahalanobisdistanz nach Abschnitt C.3.2 bzw. nach Abschnitt C.3.3, wenn für die die Distanz definierende Funktion f die komponentenweise Differenz der Mittelpunkt-Repräsentationen von Modell- und Datenkantensegment

gewählt wird. Die Fehlerkovarianz von Datenkantensegmenten ist durch

$$\Sigma_D = \Sigma_D(\theta) = \begin{pmatrix} \frac{\sigma_{\parallel}^2 \cdot \cos^2 \theta + \sigma_{\perp}^2 \cdot \sin^2 \theta}{2} & \frac{(\sigma_{\parallel}^2 - \sigma_{\perp}^2) \cdot \sin \theta \cos \theta}{2} & 0 & 0 \\ \frac{(\sigma_{\parallel}^2 - \sigma_{\perp}^2) \cdot \sin \theta \cos \theta}{2} & \frac{\sigma_{\perp}^2 \cdot \cos^2 \theta + \sigma_{\parallel}^2 \cdot \sin^2 \theta}{2} & 0 & 0 \\ 0 & 0 & \frac{2\sigma_{\perp}^2}{l^2} & 0 \\ 0 & 0 & 0 & 2\sigma_{\parallel}^2 \end{pmatrix} \quad (\text{C.35})$$

gegeben, wobei wiederum σ_{\parallel} den Fehler der Kantenendpunkte in bzw. σ_{\perp} den Fehler senkrecht zur Kantenrichtung erfäßt.

Die Fehlerkovarianzmatrix Σ_M eines Modellkantensegmentes wird aus der Fehlerkovarianz P des geschätzten Modellzustands \mathbf{x} , bestehend aus Position, Orientierung, Geschwindigkeit und Winkelgeschwindigkeit des Fahrzeugmodells, und mit Hilfe der Jakobimatrix der Meßfunktion \mathbf{h} berechnet:

$$\Sigma_M = \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right) P \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right)^T. \quad (\text{C.36})$$

Geometrische Überlegungen, wie die Forderung nach Unabhängigkeit des Schnittwinkels zweier Kantensegmente von der Wahl des Koordinatensystems — hier speziell: der aus den Orientierungen zweier Kantensegmente berechnete Schnittwinkel soll sich nicht ändern, wenn man die Kantensegmente gleichzeitig in der Bildebene dreht —, erfordern die im folgenden beschriebene Korrektur des aus ihren Winkeln θ_D bzw. θ_M berechneten Differenzwinkels $\Delta\theta$. Ist der Wertebereich des Winkel eines Kantensegmentes das Intervall $(-\pi, \pi]$ — wir sprechen dann von orientierten Segmenten —, so lautet die Korrektur für den Schnittwinkel (vgl. [Zhang & Faugeras 92b; Zhang & Faugeras 92a]):

$$\Delta\theta = \theta_D - \theta_M = \begin{cases} \theta_D - \theta_M, & \text{falls } \theta_D - \theta_M \in (-\pi, \pi] \\ \theta_D - \theta_M - 2\pi, & \text{falls } \theta_D - \theta_M > \pi \\ \theta_D - \theta_M + 2\pi, & \text{falls } \theta_D - \theta_M \leq -\pi \end{cases} \quad (\text{C.37})$$

$$= \theta_D - \theta_M \bmod 2\pi. \quad (\text{C.38})$$

Der geometrische Gedanke dabei ist, als Schnittwinkel von Halbgeraden bzw. gerichteten Strecken den kleineren der beiden Schnittwinkel zu wählen.

Ist der Wertebereich eines Winkels θ das Intervall $(-\pi/2, \pi/2]$, sind die Segmente also nicht orientiert, so lautet die Korrektur:

$$\Delta\theta = \theta_D - \theta_M = \begin{cases} \theta_D - \theta_M, & \text{falls } \theta_D - \theta_M \in (-\pi/2, \pi/2] \\ \theta_D - \theta_M - \pi, & \text{falls } \theta_D - \theta_M > \pi/2 \\ \theta_D - \theta_M + \pi, & \text{falls } \theta_D - \theta_M \leq -\pi/2 \end{cases} \quad (\text{C.39})$$

$$= \theta_D - \theta_M \bmod \pi. \quad (\text{C.40})$$

Der geometrische Gedanke dabei ist, als Schnittwinkel von Geraden den kleineren der beiden Schnittwinkel zu wählen. Das Vorzeichen des so berechneten Schnittwinkels induziert keine Ordnungsrelation zwischen den einzelnen Winkeln: die durch

$$\theta_D \geq \theta_M : \iff \theta_D - \theta_M \geq 0 \quad (\text{C.41})$$

definierte Relation \geq ist weder transitiv noch antisymmetrisch².

Eine Drehung eines Kantensegmentes in der Bildebene um einen Winkel ϕ läßt sich durch folgende Transformation im (m_x, m_y, θ, l) -Raum darstellen

$$S\mathbf{x} + \mathbf{t} = \begin{pmatrix} \cos \phi & -\sin \phi & 0 & 0 \\ \sin \phi & \cos \phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} m_x \\ m_y \\ \theta \\ l \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \phi \\ 0 \end{pmatrix}. \quad (\text{C.42})$$

Da für jedes um den Winkel ϕ gedrehte Datenkantensegment \mathbf{x}_D die transformierte Kovarianzmatrix $S\Sigma_D(\theta)S^T$ gleich der aus Gl. C.35 für ein Datenkantensegment mit Winkel $\theta + \phi$ direkt berechneten Kovarianzmatrix $\Sigma_D(\theta + \phi)$ ist, folgt mit Gl. C.19 für die Mahalanobisdistanz zweier gedrehter Kantensegmente bei Berücksichtigung der Korrekturaddition bei Bildung der Winkeldifferenz:

$$d^2(S\mathbf{x}_D + \mathbf{t}, S\mathbf{x}_M + \mathbf{t}) = d^2(\mathbf{x}_D, \mathbf{x}_M); \quad (\text{C.43})$$

die mit der Korrekturaddition erhaltene Unabhängigkeit des Wertes des Schnittwinkels zweier Kantensegmente von ihrer Lage hat demnach die Rotationsunabhängigkeit der Mahalanobisdistanz zweier Kantensegmente in der Bildebene zur Folge.

Unter der Annahme, der mathematisch positive Drehsinn entspreche einer Linksdrehung, ist das Vorzeichen des Schnittwinkels von Modell- und Datenkantensegment korrekt gewählt: in Abb. C.1 (a) sowie (a') findet sich ein Beispiel von orientierten Modell- und Datenkantensegmenten, in Abb. C.1 (b) und (b') von nicht orientierten Modell- und Datenkantensegmenten. Man hat zwar — aufgrund der Periodizität des Schnittwinkels — für den Schnittwinkel von Modell- und Datenkantensegment die Wahl zwischen den Winkeln α und β bzw. zwischen α' und β' , nicht jedoch zwischen α und α' , denn diese drücken — mit ihrem verschiedenen Vorzeichen — jeweils eine andere Lage von Daten- zu Modellkantensegment aus. Die Entscheidung für α statt für β bzw. für α' statt für β' ist äquivalent mit der oben aufgeführten Berechnung des kleineren Schnittwinkels.

In Ringen oder auch in endlichen Körpern lassen sich die Elemente ohne Probleme mit einem Vorzeichen versehen, in Körpern der Charakteristik 2 spielt das Vorzeichen sogar keine Rolle. Die Schnittwinkel sind nach obiger Definition Elemente des Rings $\mathbb{R} \bmod 2\pi$ bzw. $\mathbb{R} \bmod \pi$. Doch in der Distanzberechnung geht diese Information nicht ein. Beim dort auftretenden Term:

$$2\sigma \cdot \underbrace{\Delta\theta}_{\substack{\in \mathbb{R} \bmod 2\pi \\ \text{bzw. } \in \mathbb{R} \bmod \pi}} \cdot \underbrace{\Delta l}_{\in \mathbb{R}} \quad (\text{C.44})$$

²Bei endlichen Gruppen gibt es auch keine Ordnung, die mit ihrer Verknüpfung verträglich ist.

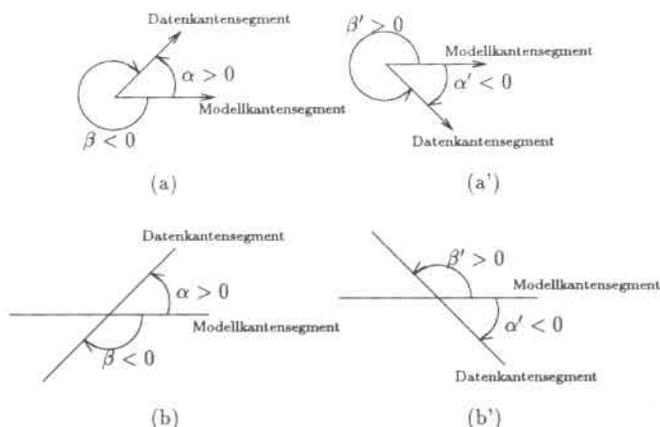


Abbildung C.1: Positive (a) und negative (a') Schnittwinkel von orientierten Modell- und Datenkantensegmenten; positive (b) und negative (b') Schnittwinkel von nicht orientierten Modell- und Datenkantensegmenten

geht man bei der Bildung des Produkts $\Delta\theta \cdot \Delta l$ davon aus, daß jeder Faktor eine reelle Zahl ist. Man macht dabei den Fehler, daß man eine eigentlich erst noch zu definierende multiplikative Verknüpfung

$$\cdot : \mathbb{R} \bmod 2\pi \times \mathbb{R} \rightarrow \mathbb{R} \text{ bzw. } \cdot : \mathbb{R} \bmod \pi \times \mathbb{R} \rightarrow \mathbb{R} \quad (\text{C.45})$$

durch die gewöhnliche multiplikative Verknüpfung

$$\cdot : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \quad (\text{C.46})$$

reeller Zahlen ersetzt³. Deshalb schlagen wir als Heuristik vor, die Mahalanobisdistanz mit beiden Vorzeichen für den Differenzwinkel — d.h. eine Wahl zwischen α und α' — durchzurechnen und sich anschließend für das Maximum der beiden Werte zu entscheiden — mit der Begründung, bei kleinen Winkeldifferenzen mache dies nicht viel aus, bei großen, die im Zuordnungsprozeß sowieso verworfen werden sollen, führe dies dazu, daß zumindest eines der Vorzeichen den Wert der Distanz in den Ablehnungsbereich bringt.

Da der Wertebereich eines Winkels in der Mittelpunkt-Repräsentation der bei uns vorliegenden Modellkantensegmenten nur das Intervall $(-\pi/2, \pi/2]$ umfaßt, beschränken wir im folgenden die Orientierung der Datenkantensegmenten auch auf dieses Intervall.

³Es geht dabei verloren, daß das Vorzeichen im geordneten Körper \mathbb{R} auf einfache Weise mit der Ordnung zusammenhängt, das Vorzeichen im Ring $\mathbb{R} \bmod \pi$ jedoch keine mit der Addition im Ring verträgliche Ordnung induziert. Die Vorzeichen bei Δl und $\Delta\theta$ haben verschiedene Bedeutungen und sind nicht auf die gleiche Weise zu behandeln.

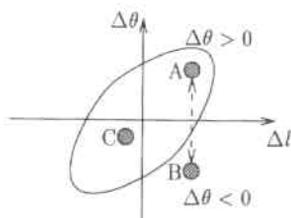


Abbildung C.2: Das Innere der Ellipse stellt den Akzeptionsbereich, das Äußere den Ablehnungsbereich dar. Die Abbildung soll verdeutlichen, wie gerade bei großen Längen- und Winkeldifferenzen das Vorzeichen des Winkels eine wichtige Rolle im Hinblick auf Akzeption (Punkt A) oder Ablehnung (Punkt B) spielt. Bei kleinen Längen- und Winkeldifferenzen befindet man sich in der Nähe des Ursprungs (Punkt C), wo die Wahl des Vorzeichens von $\Delta\theta$ weniger bedeutungsvoll ist.

C.3.6 Auswirkungen der Korrekturaddition auf die Berechnung des Abstands zweier Kantensegmente

Wir reduzieren unsere Untersuchungen der Mahalanobisdistanz zwischen zwei Kantensegmenten auf die letzten beiden Komponenten von $\mathbf{x}_D - \mathbf{x}_M$: auf den $\Delta\theta$ - Δl -Raum. Wenn die gemischten Kovarianzen nicht verschwinden, ist der Akzeptionsbereich eine nicht achsenparallele Ellipse (s. Abb. C.2). Da die Längendifferenz mit einem Vorzeichen versehen ist, muß auch die Winkeldifferenz mit einem Vorzeichen versehen werden. Gerade bei großen $\Delta\theta$ und Δl spielt die Wahl des Vorzeichens von $\Delta\theta$ die entscheidende Rolle im Hinblick auf Akzeption oder Ablehnung. Liegen keine gemischten Kovarianzen vor, so spielt das Vorzeichen von $\Delta\theta$ und Δl keine Rolle, denn die Ellipse ist symmetrisch zu den Koordinatenachsen.

Durch Vernachlässigung gemischter Kovarianzen (vgl. [Crowley *et al.* 92]) oder durch geschickte Wahl des Koordinatensystems, in dem die Kovarianzmatrix Diagonalform annimmt [Deriche & Faugeras 90], lassen sich Fragestellungen im Zusammenhang mit dem Vorzeichen der Winkeldifferenz vermeiden. Bei uns treten jedoch bei den Modellkantensegmenten gemischte Kovarianzen auf.

Die Korrekturaddition hat — wie auch [Zhang & Faugeras 92a] erwähnen — zwar keine Auswirkung auf die Kovarianzmatrix der Differenz, wohl aber auf den Erwartungswert der Differenz. Bei kleinen Winkeldifferenzen ist die Annahme, der Erwartungswert von $\Delta\theta$ sei null, noch gerechtfertigt. Bei großen Winkeldifferenzen erscheint das fragwürdig.

Die Korrekturaddition bewirkt einen Vorzeichenwechsel bei der Winkeldifferenz. Unter der Annahme, vor der Korrekturaddition seien Winkel- und Längendifferenz positiv sowie Winkel und Länge negativ korreliert, womit aufgrund der Invertierung der Kovarianzmatrix die gemischten Kovarianzen zu einer Verstärkung des Abstands führen, hat der durch die Korrekturaddition bewirkte Vorzeichenwechsel bei der

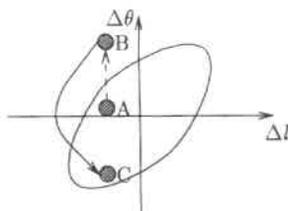


Abbildung C.3: Wächst die Winkeldifferenz $\Delta\theta$, beginnend bei null (Punkt A), so kommt man relativ rasch in den Ablehnungsbereich (Punkt B). Ein weiteres Anwachsen von $\Delta\theta$ bringt aber — aufgrund der Korrekturaddition — wieder eine Akzeption (Punkt C). Dieses Verhalten widerspricht der Normalverteilungsannahme bei $\Delta\theta$, die nicht erfaßt, daß $\Delta\theta$ einen beschränkten — genauer: zyklischen — Wertebereich hat.

Winkeldifferenz eine starke Verminderung des Abstandes zur Folge (vgl. Abb. C.3).

C.3.7 Untersuchung anderer Abstandsmaße

In Abschnitt 6.2 wird die Mahalanobisdistanz zweier Kantensegmente an einem Beispiel in der Ettlinger-Tor-Szene untersucht. Wir setzen diese Untersuchung an dieser Stelle fort, um prüfen, ob statt der komponentenweisen Differenz der Mittelpunkt-Repräsentationen von Modell- und Datenkantensegmenten — oben mit der Funktion f bezeichnet — beispielsweise mit dem Kosinus des Schnittwinkels Zahlenwerte zu erzielen sind, die den Abstand zweier Kantensegmente mit großen Winkeldifferenzen stärker zum Ausdruck bringen als die Winkeldifferenz selbst.

Mit den Überlegungen in Abschnitt 6.2 suchen wir eine Abstandsfunktion $g(\Delta\theta)$ mit der Eigenschaft:

$$\left(\frac{g(\Delta\theta)}{g'(\Delta\theta)}\right)^2 \text{ ist monoton wachsend in } \Delta\theta$$

oder sogar

$$\frac{g(\Delta\theta)}{g'(\Delta\theta)} \text{ ist monoton wachsend in } \Delta\theta$$

Auf einem anderen Weg gelangen auch [Taubin 93] bzw. [Leonardis 93] mit einem von [Taubin 91] vorgeschlagenen, nicht statistischen approximierten quadratischen Abstand von Punkten zu Kurven bzw. Flächen auf Terme der Form $\left(\frac{g(\Delta\theta)}{g'(\Delta\theta)}\right)^2$. [Taubin 93] verfeinert das Abstandsmaß unter Verwendung höherer Ableitungen.

Das Ersetzen der Winkeldifferenz durch den Kosinus des Schnittwinkels, genauer:

durch $1 - \cos \Delta\theta$, entspricht einer Verkettung $\mathbf{g}_1 \circ \mathbf{f}$, wobei

$$\mathbf{g}_1(\Delta\theta, \Delta l) = \begin{pmatrix} 1 - \cos \Delta\theta \\ \Delta l \end{pmatrix}. \quad (\text{C.47})$$

Die zugehörige Kovarianzmatrix ergibt sich aus

$$\Sigma_{\mathbf{g}_1 \circ \mathbf{f}} = \begin{pmatrix} \sin \Delta\theta & 0 \\ 0 & 1 \end{pmatrix} (\Sigma_D + \Sigma_M) \begin{pmatrix} \sin \Delta\theta & 0 \\ 0 & 1 \end{pmatrix}^T. \quad (\text{C.48})$$

Man erhält — nach den Überlegungen in Abschnitt C.3.4 — dieselbe Kovarianzmatrix $\Sigma_{\mathbf{g}_1 \circ \mathbf{f}}$, wenn man den Kosinus des Schnittwinkels direkt über die bekannte Formel mit Skalarprodukt durch Normprodukt berechnet.

Im Beispiel der Ettlinger-Tor-Szene ergeben sich folgende Zahlenwerte:

$$\begin{aligned} d_{\mathbf{g}_1 \circ \mathbf{f}}^2(\Delta\theta, \Delta l) &= 47,72 * \frac{(1 - \cos \Delta\theta)^2}{\sin^2 \Delta\theta} + 2 * 0,52 * \frac{(1 - \cos \Delta\theta)}{\sin \Delta\theta} \Delta l + 0,0069 * (\Delta l)^2 \\ &= \begin{cases} 23,6 & \text{für } \Delta\theta = -1,08, \quad \Delta l = 99,39 \\ 146,3 & \text{für } \Delta\theta = +1,08, \quad \Delta l = 99,39 \end{cases} \end{aligned} \quad (\text{C.49})$$

Damit erhält man die transformierte Mahalanobisdistanz wie in Abbildung C.4 wiedergegeben. Wir diskutieren den auf den ersten Blick nicht sehr vorteilhaften Verlauf dieses Distanzmaßes in Abhängigkeit vom Schnittwinkel $\Delta\theta$ weiter unten.

Mit der Begründung, der Sinus — im Gegensatz zum Kosinus — ist für kleine Argumente linear mit Steigung 1, wird im folgenden der Sinus des Schnittwinkels untersucht. Die Kovarianzmatrix nach der Transformation mit $\mathbf{g}_2(\Delta\theta, \Delta l) = (\sin(\Delta\theta), \Delta l)^T$ lautet

$$\Sigma_{\mathbf{g}_2 \circ \mathbf{f}} = \begin{pmatrix} \cos \Delta\theta & 0 \\ 0 & 1 \end{pmatrix} (\Sigma_D + \Sigma_M) \begin{pmatrix} \cos \Delta\theta & 0 \\ 0 & 1 \end{pmatrix}^T. \quad (\text{C.50})$$

Im Beispiel der Ettlinger-Tor-Szene ergeben sich folgende Zahlenwerte:

$$\begin{aligned} d_{\mathbf{g}_2 \circ \mathbf{f}}^2(\Delta\theta, \Delta l) &= 47,72 * \frac{\sin^2 \Delta\theta}{\cos^2 \Delta\theta} + 2 * 0,52 * \frac{\sin \Delta\theta}{\cos \Delta\theta} \Delta l + 0,0069 * (\Delta l)^2 \\ &= \begin{cases} 43,3 & \text{für } \Delta\theta = -1,08, \quad \Delta l = 99,39 \\ 426,52 & \text{für } \Delta\theta = +1,08, \quad \Delta l = 99,39. \end{cases} \end{aligned} \quad (\text{C.51})$$

Damit erhält man den in Abbildung C.5 dargestellten Verlauf der transformierten Mahalanobisdistanz.

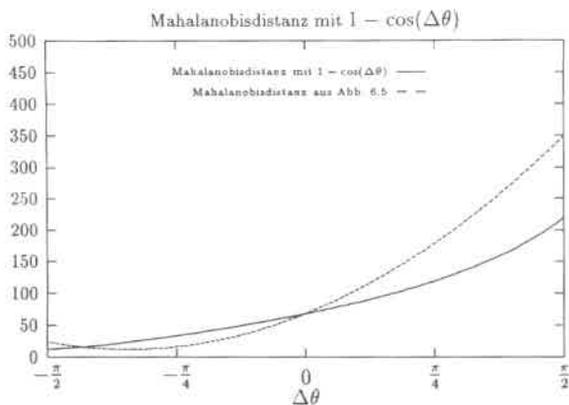


Abbildung C.4: Vergleich zwischen der bisherigen Mahalanobisdistanz, die auf der Winkeldifferenz aufbaut, und der den Kosinus des Schnittwinkels einziehenden verallgemeinerten Mahalanobisdistanz am Beispiel der Ettlinger-Tor-Szene (vgl. Abb. 6.2). Die Längendifferenz wurde auf 99,39 gesetzt.

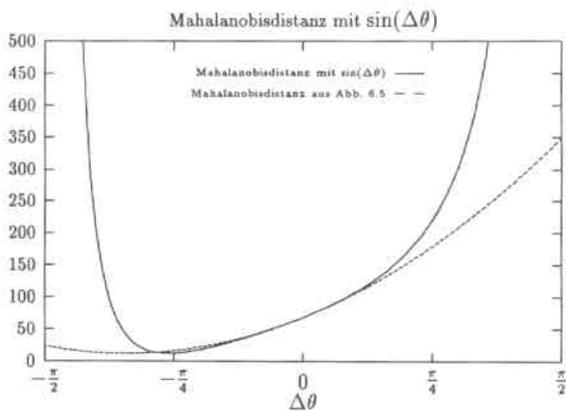


Abbildung C.5: Vergleich zwischen der bisherigen Mahalanobisdistanz, die auf der Winkeldifferenz aufbaut, und der verallgemeinerten Mahalanobisdistanz mit dem Sinus des Schnittwinkels am Beispiel der Ettlinger-Tor-Szene (vgl. Abb. 6.2). Die Längendifferenz wurde auf 99,39 gesetzt.

Die Abbildungen C.4 und C.5 bzw. die Abstandsfunktionen in Gleichungen C.49 und C.51 hängen auf einfache Weise zusammen, wenn man die trigonometrische Formel $\tan(x/2) = (1 - \cos x) / \sin x$ verwendet:

$$d_{g \circ f}^2(\Delta\theta, \Delta l) = d_{g_2 \circ f}^2(\Delta\theta/2, \Delta l); \quad (\text{C.52})$$

damit entpuppt sich der Sinus des Schnittwinkels $\Delta\theta$ als dasselbe Abstandsmaß wie $1 - \cos 2\Delta\theta$ mit dem Kosinus des doppelten Schnittwinkels.

Der Sinus in der Abstandsberechnung ist demnach auch nur ein versteckter Schwellwertoperator, wenn man den Funktionsverlauf in der Nähe von $\pm\pi/2$ betrachtet.

Im allgemeinen erhält man nach einer Verkettung $g \circ f$ mit $g(\Delta\theta, \Delta l) = (g(\Delta\theta), \Delta l)^T$ folgendes Abstandsmaß:

$$d_{g \circ f}^2(\Delta\theta, \Delta l) = 47,72 * \frac{g^2(\Delta\theta)}{g'^2(\Delta\theta)} + 2 * 0,52 * \frac{g(\Delta\theta)}{g'(\Delta\theta)} \Delta l + 0,0069 * (\Delta l)^2. \quad (\text{C.53})$$

Wir untersuchen im folgenden den Ausdruck

$$\frac{g(\Delta\theta)}{g'(\Delta\theta)}. \quad (\text{C.54})$$

Für $g(\Delta\theta)$ läßt sich mit der Absicht, nicht zwischen positiven und negativen Winkeldifferenzen $\Delta\theta$ unterscheiden zu wollen, eine gerade Funktion wählen. Mit der Begründung, das Vorzeichen des Schnittwinkels entsprechend zu berücksichtigen, kann man auch eine ungerade Funktion für $g(\Delta\theta)$ wählen. Da die Ableitung einer geraden Funktion ungerade, die einer ungeraden Funktion gerade ist, erweist sich der Quotient $\frac{g(\Delta\theta)}{g'(\Delta\theta)}$ in beiden Fällen als ungerade, womit das Vorzeichen von $\Delta\theta$ im gemischten Term

$$\frac{g(\Delta\theta)}{g'(\Delta\theta)} \Delta l \quad (\text{C.55})$$

immer eine Rolle spielt.

Wir drehen die Betrachtung jetzt um. Wir setzen $\frac{g(\Delta\theta)}{g'(\Delta\theta)}$ gleich einer geraden Funktion $\gamma(\Delta\theta)$, um uns bei der Berechnung der Distanz in Gl. C.53 vom Vorzeichen der Winkeldifferenz zu lösen. Die Differentialgleichung

$$\frac{g(\Delta\theta)}{g'(\Delta\theta)} = \gamma(\Delta\theta) \quad (\text{C.56})$$

hat als Lösung

$$g(\Delta\theta) = k e^{\int \frac{\gamma(\Delta\theta)}{\Delta\theta} d\Delta\theta} \quad (\text{C.57})$$

mit einer Konstanten k . Dieses Abstandsmaß läßt sich jedoch geometrisch *nicht* begründen: beispielsweise mit $\gamma(\Delta\theta) = (\Delta\theta)^2$ erhält man

$$g(\Delta\theta) = k e^{-\frac{1}{\Delta\theta}}, \quad (\text{C.58})$$

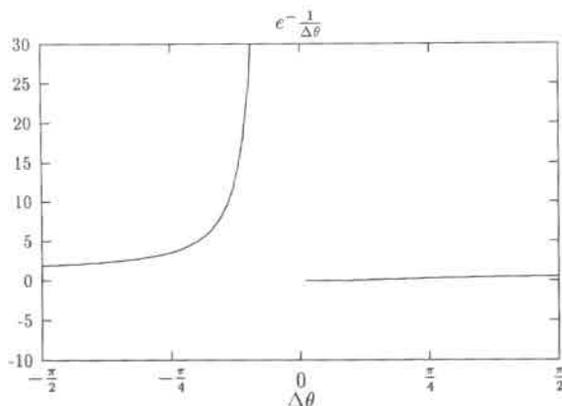


Abbildung C.6: Graph der Funktion $g(\Delta\theta) = ke^{-\frac{1}{\Delta\theta}}$ für $k = 1$.

der Graph dieser Funktion findet sich in Abb. C.6; weder für die Polstelle noch für den nicht symmetrischen Verlauf läßt sich eine Begründung geben, die den Einsatz dieser Funktion für unsere Zwecke rechtfertigt. Einen zur Achse $\Delta\theta = 0$ symmetrischen Verlauf der Mahalanobisdistanz kann man nicht erzwingen. Folglich scheidet $\gamma(\Delta\theta) = \tan^2(\Delta\theta)$ aus.

Jedoch mit $\gamma(\Delta\theta) = \tan^3(\Delta\theta)$ und somit

$$\begin{aligned} g(\Delta\theta) &= e^{\int \frac{d\Delta\theta}{\gamma(\Delta\theta)}} \\ &= \frac{e^{-\frac{\cos(\Delta\theta)}{2}}}{\sin(\Delta\theta)} \end{aligned} \quad (\text{C.59})$$

erhält man die transformierte Mahalanobisdistanz wie in Abb. C.7; mit $\gamma(\Delta\theta) = (\tan(\Delta\theta))^{\frac{1}{2}}$ und mit Hilfe des vom Computer-Algebra-System **Maple** berechneten

$$\begin{aligned} g(\Delta\theta) &= e^{\int \frac{d\Delta\theta}{\gamma(\Delta\theta)}} \\ &= \frac{e^{\frac{\ln(1+\cos(\Delta\theta)^{2/3})}{2} - \frac{\ln(\cos(\Delta\theta)^{4/3} - \cos(\Delta\theta)^{2/3} + 1)}{4}} \cdot \sqrt{5} \arctan\left(\frac{(2\cos(\Delta\theta)^{2/3} - 1)\sqrt{5}}{2}\right)}{\sqrt{\sin(\Delta\theta)^{2/3} + \cos(\Delta\theta)^{2/3}} e^{\frac{\sqrt{5} \arctan\left(\frac{(\sin(\Delta\theta)^{2/3} - 2\cos(\Delta\theta)^{2/3})\sqrt{5}}{2\sin(\Delta\theta)^{2/3}}\right)}{2}}} \\ &= \frac{\sqrt{\sin(\Delta\theta)^{2/3} + \cos(\Delta\theta)^{2/3}}}{\sqrt{\cos(\Delta\theta)^{4/3} - \cos(\Delta\theta)^{2/3} \sin(\Delta\theta)^{2/3} + \sin(\Delta\theta)^{4/3}}} \quad (\text{C.60}) \end{aligned}$$

erhält man die transformierte Mahalanobisdistanz wie in Abb. C.8.

Störend in allen untersuchten Abstandsmaßen ist die Lage des Minimums. Das Minimum des Abstandes verläuft bei $\Delta\theta \neq 0$, nicht bei $\Delta\theta = 0$. Gerade bei großen Δl verläuft das Minimum des Abstandes sogar *weit* weg von $\Delta\theta = 0$. Ursache sind die

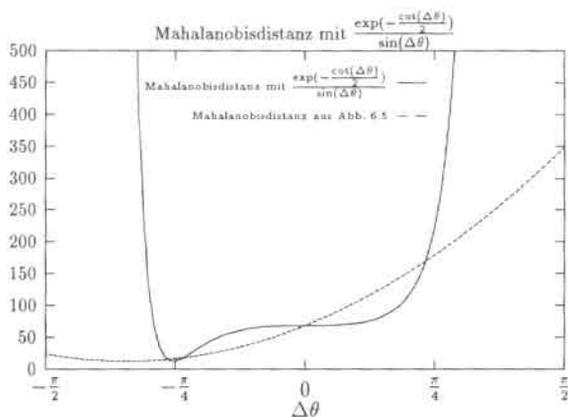


Abbildung C.7: Vergleich zwischen der bisherigen Mahalanobisdistanz, die auf der Winkeldifferenz aufbaut, und der verallgemeinerten Mahalanobisdistanz nach Gl. C.59 am Beispiel der Ettlinger-Tor-Szene (vgl. Abb. 6.2). Die Längendifferenz wurde auf 99,39 gesetzt. Die noch ausgeprägtere Lage des Minimums ist störend.

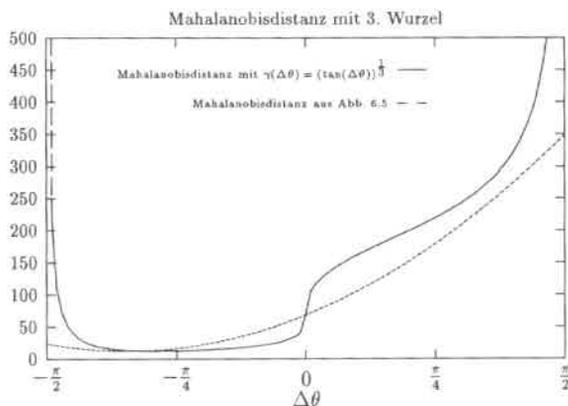


Abbildung C.8: Vergleich zwischen der bisherigen Mahalanobisdistanz, die auf der Winkeldifferenz aufbaut, und der verallgemeinerten Mahalanobisdistanz nach Gl. C.60 am Beispiel der Ettlinger-Tor-Szene (vgl. Abb. 6.2). Die Längendifferenz wurde auf 99,39 gesetzt.

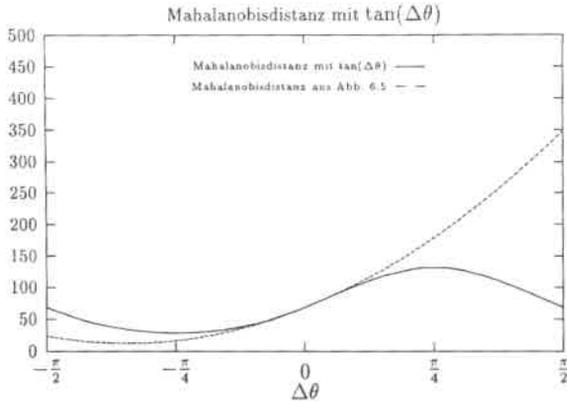


Abbildung C.9: Vergleich zwischen der bisherigen Mahalanobisdistanz, die auf der Winkeldifferenz aufbaut, und der verallgemeinerten Mahalanobisdistanz mit dem Tangens des Schnittwinkels am Beispiel der Ettlinger-Tor-Szene (vgl. Abb. 6.2). Die Längendifferenz wurde auf 99,39 gesetzt.

gemischten Kovarianzen. Man kann für die Winkeldifferenz einen eigenen Schwellwert einführen und damit größere Winkeldifferenzen gar nicht mehr in der Mahalanobisdistanz mit anderen Parametern zusammenbringen. Damit berechnet man die Mahalanobisdistanz nicht mehr für den ganzen Parametervektor, sondern man berechnet mehrere — mit jeweils eigenen Schwellen — Mahalanobisdistanzen für Teile des Parametervektors, unter Vernachlässigung eventuell vorhandener gemischter Kovarianzen. Mit der Begründung, die gleiche Orientierung zweier Kantensegmente sei das wichtigste Kriterium für eine potentielle Zuordnung zwischen ihnen, kann man für die Winkeldifferenz eine eigene Schwelle einführen, wie beispielsweise bei [Crowley *et al.* 92; Zhang & Faugeras 92a]. Ist diese Schwelle überschritten, so werden die Differenzen des anderen Parameters gar nicht mehr verglichen.

Mit dem Tangens des Schnittwinkels, $g(\Delta\theta) = \tan(\Delta\theta)$, erhält man die transformierte Mahalanobisdistanz wie in Abb. C.9. Da jedoch

$$\gamma(\Delta\theta) = \frac{g(\Delta\theta)}{g'(\Delta\theta)} = \frac{\tan(\Delta\theta)}{1 + \tan^2(\Delta\theta)} = \sin(\Delta\theta) \cos(\Delta\theta) = \frac{\sin(2\Delta\theta)}{2} \quad (\text{C.61})$$

nicht das am Ende des letzten Abschnitts aufgestellte Monotoniekriterium erfüllt, scheidet dieser Ansatz sofort aus.

Keine der vorgeschlagenen Abstandsfunktionen brachte ein überzeugendes Ergebnis. Wir brechen an dieser Stelle die theoretischen Untersuchungen zur Winkeldifferenz ab und stellen in Kapitel 6 einen grundsätzlich anderen Weg vor, um die Lage eines Objektes in der Szene aus Bilddaten zu schätzen.

Anhang D

Berechnung von Jakobimatrizen

D.1 Vorbemerkungen

Im allgemeinen sind die Operatoren Differentiation und Substitution nicht vertauschbar:

$$\left(\frac{\partial}{\partial x} f(x)\right)\Big|_{x=c} \neq \frac{\partial}{\partial x}(f(x)|_{x=c}) = \frac{\partial}{\partial x} f(c) = 0 \quad (\text{D.1})$$

Lemma D.1 Ist die Funktion $f(x, y)$ separabel, d.h.

$$f(x, y) = g(x)h(y), \quad (\text{D.2})$$

so ist

$$\left(\frac{\partial}{\partial y} f(x, y)\right)\Big|_{y=c} = \frac{\partial}{\partial y}(f(x, y)|_{y=c}) \quad (\text{D.3})$$

Die Differentiation einer m -dimensionalen vektorwertigen Funktion nach einem n -dimensionalen Vektor \mathbf{x} ist eine $m \times n$ -Jakobimatrix:

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \frac{\partial f_m(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{pmatrix}, \quad (\text{D.4})$$

deren Zeilen die partiellen Ableitungen der Komponenten f_i von \mathbf{f} nach den Komponenten x_j von \mathbf{x} enthalten (siehe Anhang A.3.2 [Horn 86]). Mit dieser Notation ist die Differentiation einer skalarwertigen Funktion nach einem n -dimensionalen Vektor eine $1 \times n$ -Matrix, d.h. ein Zeilenvektor.

D.2 Datenassoziation im Beispiel der Fahrbahnmodellanpassung an Datenkantensegmente

Für die Anpassung eines Fahrbahnmodells an Datenkantensegmente umfaßt ein sechsdimensionaler Zustandsvektor \mathbf{x} die Position und Orientierung des Fahrbahnmodells. Den Abstand eines Datenkantensegments in der Mittelpunkt-Repräsentation \mathbf{X}_i^D mit Mittelpunkt $(\xi_i^D, \eta_i^D)^T$, Orientierung θ_i^D und Länge l_i^D von einem in die Bildebene projizierten Modellkantensegment in der Mittelpunkt-Repräsentation $\mathbf{X}_i^M = \mathbf{X}_i^M(\mathbf{x})$ mit Mittelpunkt $(\xi_i^M, \eta_i^M)^T$, Orientierung θ_i^M und Länge l_i^M berechnen wir mit einer verallgemeinerten Mahalanobisdistanz

$$\mathbf{f}_i(\mathbf{X}_i^D, \mathbf{X}_i^M) = \begin{pmatrix} \theta_i^D - \theta_i^M \\ -\sin(\theta_i^M)(\xi_i^D - \xi_i^M) + \cos(\theta_i^M)(\eta_i^D - \eta_i^M) \end{pmatrix}. \quad (\text{D.5})$$

Die erste Komponente der vektorwertigen Funktion $\mathbf{f}_i(\mathbf{X}_i^D, \mathbf{X}_i^M)$ ist die Differenz der Orientierungen von Daten- und Modellkantensegment, die zweite stellt den senkrechten Abstand des Mittelpunktes des Datenkantensegmentes vom Modellkantensegment dar (vgl. Abschnitt 5.2.4).

Nach Abschnitt C.3.3 müssen bei Einführung einer verallgemeinerten Mahalanobisdistanz die beiden 2×4 -Jakobimatrizen

$$\left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^M} \right) = \begin{pmatrix} 0 & 0 & -1 & 0 \\ \sin(\theta_i^M) & -\cos(\theta_i^M) & -\cos(\theta_i^M)(\xi_i^D - \xi_i^M) - \sin(\theta_i^M)(\eta_i^D - \eta_i^M) & 0 \end{pmatrix} \quad (\text{D.6})$$

und

$$\left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^D} \right) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -\sin(\theta_i^M) & \cos(\theta_i^M) & 0 & 0 \end{pmatrix} \quad (\text{D.7})$$

bestimmt werden. Mit Σ_D aus Gl. C.35 transformiert man die Kovarianzmatrix eines Datenkantensegmentes in eine 2×2 -Matrix:

$$\left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^D} \right) \Sigma_D \left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^D} \right)^T = \begin{pmatrix} 2\frac{\sigma_l^2}{l^2} & 0 \\ 0 & \begin{pmatrix} \frac{1}{2} \sin^2(\theta_i^M) \sigma_{\parallel}^2 \cos^2(\theta_i^D) + \frac{1}{2} \sin^2(\theta_i^M) \sigma_{\perp}^2 \\ -\frac{1}{2} \sin^2(\theta_i^M) \sigma_{\perp}^2 \cos^2(\theta_i^D) \\ -\sin(\theta_i^M) \cos(\theta_i^M) \cos(\theta_i^D) \sin(\theta_i^D) \sigma_{\parallel}^2 \\ + \sin(\theta_i^M) \cos(\theta_i^M) \cos(\theta_i^D) \sin(\theta_i^D) \sigma_{\perp}^2 \\ + \frac{1}{2} \cos^2(\theta_i^M) \sigma_{\perp}^2 \cos^2(\theta_i^D) + \frac{1}{2} \cos^2(\theta_i^M) \sigma_{\parallel}^2 \\ - \frac{1}{2} \cos^2(\theta_i^M) \sigma_{\parallel}^2 \cos^2(\theta_i^D) \end{pmatrix} \end{pmatrix}$$

$$\begin{aligned}
&= \begin{pmatrix} \frac{2\sigma_{\perp}^2}{\rho^2} & 0 \\ 0 & \frac{1}{4}(\sigma_{\perp}^2 - \sigma_{\parallel}^2) \cos 2(\theta_i^M - \theta_i^D) + \frac{1}{4}(\sigma_{\perp}^2 + \sigma_{\parallel}^2) \end{pmatrix} \\
&= \begin{pmatrix} \frac{2\sigma_{\perp}^2}{\rho^2} & 0 \\ 0 & \frac{1}{2}\sigma_{\perp}^2 \cos^2(\theta_i^M - \theta_i^D) + \frac{1}{2}\sigma_{\parallel}^2 [1 - \cos^2(\theta_i^M - \theta_i^D)] \end{pmatrix} \quad (D.8)
\end{aligned}$$

Dieser Ausdruck hängt nur vom Schnittwinkel $\theta_i^M - \theta_i^D$ ab, nicht von der Lage im Koordinatensystem (rotationsinvariant). Der Ausdruck im unteren rechten Element der Matrix in Gl. D.8 stellt eine Konvexkombination aus $\sigma_{\perp}^2/2$ und $\sigma_{\parallel}^2/2$ dar. Im Fall $\theta_i^M - \theta_i^D = 0$ ist er gleich $\sigma_{\perp}^2/2$, im Fall $\theta_i^M - \theta_i^D = \pi/2$ gleich $\sigma_{\parallel}^2/2$.

Mit Σ_M nach Gl. C.36 transformiert man die Kovarianzmatrix eines Modellkantensegmentes in eine (i.a. vollbesetzte) 2×2 -Matrix

$$\left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^M} \right) \Sigma_M \left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^M} \right)^T \quad (D.9)$$

Mit

$$\Sigma_{f_i} = \left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^D} \right) \Sigma_D \left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^D} \right)^T + \left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^M} \right) \Sigma_M \left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^M} \right)^T \quad (D.10)$$

werden Zuordnungen zwischen (u.U. mehreren) Datenkantensegmenten und einem Modellkantensegment zugelassen, wenn

$$d_{f_i}^2(\mathbf{X}_i^D, \mathbf{X}_i^M) = \mathbf{f}_i^T(\mathbf{X}_i^D, \mathbf{X}_i^M) \Sigma_{f_i}^{-1} \mathbf{f}_i(\mathbf{X}_i^D, \mathbf{X}_i^M) \quad (D.11)$$

eine fest eingestellte Schwelle nicht überschreitet.

D.3 Anpassung eines Fahrbahnmodells an Kantensegmente

Wenn unter Ausnutzung der verallgemeinerten Mahalanobisdistanz aus dem letzten Abschnitt in der aktuellen Aufnahme zum Zeitpunkt t_k zwischen Daten- und Modellkantensegmenten m Zuordnungen gefunden wurden, kann man die Datenkantensegmente zu einem $4m$ -dimensionalen *Meßvektor*

$$\mathbf{z}_k = (\mathbf{X}_1^D, \dots, \mathbf{X}_m^D)^T \quad (D.12)$$

sowie die Modellkantensegmente zu einer $4m$ -dimensionalen vektorwertigen *Meßfunktion*

$$\mathbf{h}_k(\mathbf{z}_k) = (\mathbf{X}_1^M(\mathbf{z}_k), \dots, \mathbf{X}_m^M(\mathbf{z}_k))^T \quad (D.13)$$

zusammenfassen. Mehrfachzuordnungen verschiedener Datensegmente zu einem Modellkantensegment führen zu mehrfachen Einträgen desselben Modellkantensegmentes, aber in verschiedenen Komponenten der Meßfunktion $\mathbf{h}_k(\mathbf{z}_k)$.

Im Rahmen eines Kalmanfilteraktualisierungsschrittes nutzen wir nicht wie [Koller 92; Deriche & Faugeras 90] die Differenz $\mathbf{z}_k - \mathbf{h}_k(\mathbf{z}_k)$ selbst, sondern – passend zur verallgemeinerten Mahalanobisdistanz bei der Datenassoziation des letzten Abschnittes – eine neue, implizit definierte Meßfunktion

$$\bar{\mathbf{h}}_k(\mathbf{z}_k, \mathbf{z}_k) = \mathbf{f}(\mathbf{z}_k, \mathbf{h}_k(\mathbf{z}_k)) = \left(\mathbf{f}_1(\mathbf{X}_1^D, \mathbf{X}_1^M(\mathbf{z}_k)), \dots, \mathbf{f}_m(\mathbf{X}_m^D, \mathbf{X}_m^M(\mathbf{z}_k)) \right)^T \quad (\text{D.14})$$

mit $\bar{\mathbf{h}}_k(\mathbf{z}_k, \mathbf{z}_k) = 0$ für den wahren, aber unbekanntenen Zustandsvektor \mathbf{z}_k .

Die blockdiagonale Meßkovarianzmatrix R_k , die aus m Blöcken von 4×4 -Kovarianzmatrizen der Form wie in Gl. C.35 besteht, wird in eine blockdiagonale Meßkovarianzmatrix \bar{R}_k überführt, die aus m Blöcken von 2×2 -Kovarianzmatrizen wie in Gl. D.8 besteht. Die Meßkovarianzmatrix R_k hängt von der Aufnahme zum Zeitpunkt t_k ab, da das in die Berechnung der Matrix R_k eingehende σ_1^2 für jedes Datenkantensegment nach Anhang C.1 aus der jeweiligen Aufnahme geschätzt wird.

Unter Ausnutzung einer initialen Zustandsvektorschätzung $\hat{\mathbf{z}}_k^-$ mit Kovarianzmatrix P_k^- minimieren wir im Rahmen einer MAP-Schätzung zur kalmanfilterbasierten Zustandsaktualisierung im Gegensatz zu [Koller 92] nicht mehr

$$\frac{1}{2} (\mathbf{z}_k - \mathbf{h}_k(\mathbf{z}_k))^T R_k^{-1} (\mathbf{z}_k - \mathbf{h}_k(\mathbf{z}_k)) + \frac{1}{2} (\mathbf{z}_k - \hat{\mathbf{z}}_k^-)^T P_k^{-1} (\mathbf{z}_k - \hat{\mathbf{z}}_k^-) \rightarrow \min_{\mathbf{z}_k} \quad , \quad (\text{D.15})$$

sondern die Zielfunktion:

$$\frac{1}{2} \bar{\mathbf{h}}(\mathbf{z}_k, \mathbf{z}_k)^T \bar{R}_k^{-1} \bar{\mathbf{h}}(\mathbf{z}_k, \mathbf{z}_k) + \frac{1}{2} (\mathbf{z}_k - \hat{\mathbf{z}}_k^-)^T P_k^{-1} (\mathbf{z}_k - \hat{\mathbf{z}}_k^-) \rightarrow \min_{\mathbf{z}_k} \quad . \quad (\text{D.16})$$

Die Ableitung der letzten Zeile nach den Komponenten des Zustandsvektors \mathbf{z} liefert

$$\bar{\mathbf{H}}^T(\mathbf{z}_k) \bar{R}_k^{-1} \bar{\mathbf{h}}(\mathbf{z}_k, \mathbf{z}_k) + P_k^{-1} (\mathbf{z}_k - \hat{\mathbf{z}}_k^-) \quad (\text{D.17})$$

mit der Jakobimatrix

$$\bar{\mathbf{H}}(\mathbf{z}) = \frac{\partial \bar{\mathbf{h}}(\mathbf{z}_k, \mathbf{z}_k)}{\partial \mathbf{z}} \quad (\text{D.18})$$

der neuen Meßfunktion $\bar{\mathbf{h}}(\mathbf{z}_k, \mathbf{z}_k)$. Hieraus und aus Gl. D.14 erhält man mit Hilfe der Kettenregel

$$\frac{\partial \bar{\mathbf{h}}(\mathbf{z}_k, \mathbf{z}_k)}{\partial \mathbf{z}} = \begin{pmatrix} \vdots \\ \frac{\partial \mathbf{f}_i(\mathbf{X}_i^D, \mathbf{X}_i^M)}{\partial \mathbf{X}_i^M} \Big|_{\mathbf{X}_i^M = \mathbf{X}_i^M(\mathbf{z})} \cdot \frac{\partial \mathbf{X}_i^M(\mathbf{z})}{\partial \mathbf{z}} \\ \vdots \end{pmatrix} \quad . \quad (\text{D.19})$$

Die Jakobimatrix

$$\left(\frac{\partial \mathbf{f}_i}{\partial \mathbf{X}_i^M} \right) \quad (\text{D.20})$$

findet sich in Gl. D.6, die Herleitung der Jakobimatrix

$$\frac{\partial \mathbf{X}_i^M(\mathbf{x})}{\partial \mathbf{x}} \quad (\text{D.21})$$

findet man im Anhang von [Koller 92], wobei eine einfache Modifikation der dort aufgeführten Rechnung die Schätzung aller sechs (und nicht nur dreier) Lagefreiheitsgrade erlaubt [Osswald 94].

Die Zustandsübergangsfunktion ist im Rahmen der Fahrbahnanpassung die identische Funktion auf \mathbb{R}^6 , die zugehörige Jakobimatrix damit eine Einheitsmatrix.

D.4 Jakobimatrix der Meßfunktion bei direkter Anpassung an die Grauwertgradienten

Der Kalmanfilteraktualisierungsschritt benötigt die Jakobimatrix der Meßfunktion. Die Jakobimatrix der Zustandsübergangsfunktion findet sich in [Koller 92].

Es sei ein in die Bildebene projiziertes Modellkantensegment mit Mittelpunkt $\mathbf{m} = (m_\xi, m_\eta)^T$, Orientierung θ und Länge l gegeben (vgl. Abschnitt 6.3). Mit $\mathbf{X}_M = (m_\xi, m_\eta, \theta, l)^T$ bezeichnen wir seine Mittelpunkt-Repräsentation nach [Deriche & Faugeras 90].

Weiterhin seien $\boldsymbol{\xi} = (\xi, \eta)^T$ die Bildkoordinaten eines Punktes. Dann berechnen wir die Ableitung

$$\frac{\partial f_{(m_\xi, m_\eta, \theta, l)}(\boldsymbol{\xi})}{\partial \mathbf{x}} \quad (\text{D.22})$$

der synthetischen Gradientennorm nach dem Zustandsvektor \mathbf{x} , d.h. der Position, Orientierung, translatorischen Geschwindigkeit und Winkelgeschwindigkeit der geschätzten Fahrzeuglage in der Straßenebene. Für jeden Bildpunkt $\boldsymbol{\xi}$ muß diese 1×5 -Matrix berechnet werden.

Die Anwendung der Kettenregel auf die letzte Gleichung liefert:

$$\frac{\partial f_{(m_\xi, m_\eta, \theta, l)}(\boldsymbol{\xi})}{\partial \mathbf{x}} = \frac{\partial f_{(m_\xi, m_\eta, \theta, l)}(\boldsymbol{\xi})}{\partial \mathbf{X}_M} \frac{\partial \mathbf{X}_M}{\partial \mathbf{x}} \quad (\text{D.23})$$

Die 4×5 Jakobimatrix

$$\frac{\partial \mathbf{X}_M}{\partial \mathbf{x}} \quad (\text{D.24})$$

die ausdrückt, wie ein projiziertes Modellkantensegment vom Zustand \mathbf{x} abhängt, wurde bereits von [Koller 92] hergeleitet.

Betrachten wir Lemma 6.2, so können wir den ersten Ausdruck auf der rechten Seite von Gl. D.23 schreiben als:

$$\frac{\partial f_{(m_\xi, m_\eta, \theta, l)}(\boldsymbol{\xi})}{\partial \mathbf{X}_M} = \frac{\partial f(\mathcal{R}_\theta^T(\boldsymbol{\xi} - \mathbf{m}), l)}{\partial \mathbf{X}_M} \quad (\text{D.25})$$

Eine Anwendung der Kettenregel auf die letzte Gleichung liefert ein Produkt einer 1×3 -Matrix mit einer 3×4 -Matrix:

$$\frac{\partial f(\mathcal{R}_\theta^T(\boldsymbol{\xi} - \mathbf{m}), l)}{\partial \mathbf{X}_M} = \left(\frac{\partial f(\boldsymbol{\xi}, l)}{\partial (\boldsymbol{\xi}, l)} \right) \Big|_{\substack{\boldsymbol{\xi} \\ l}} = \left(\mathcal{R}_\theta^T(\boldsymbol{\xi} - \mathbf{m}) \right) \cdot \left(\frac{\partial (\mathcal{R}_\theta^T(\boldsymbol{\xi} - \mathbf{m}), l)}{\partial \mathbf{X}_M} \right). \quad (\text{D.26})$$

Die stückweise definierte Funktion $f(\boldsymbol{\xi}, l) = f(\xi, \eta, l)$ aus Gl. 6.5 ist nicht nur stetig, sondern beliebig oft stetig differenzierbar. Ihre ersten partiellen Ableitungen nach den drei Variablen ξ, η, l ergeben sich mit den Ableitungen der Gaußfunktion

$$\begin{aligned} \varphi_\xi &= -\frac{\xi}{\sigma^2} \varphi, \\ \varphi_\eta &= -\frac{\eta}{\sigma^2} \varphi \end{aligned} \quad (\text{D.27})$$

aus:

$$\frac{\partial f(\xi, \eta, l)}{\partial \xi} = \begin{cases} \varphi_\xi(\xi + \frac{1}{2}, \eta), & \text{falls } \xi < -\frac{1}{2} \\ 0, & \text{falls } -\frac{1}{2} \leq \xi < \frac{1}{2} \\ \varphi_\xi(\xi - \frac{1}{2}, \eta), & \text{falls } \frac{1}{2} \leq \xi \end{cases}, \quad (\text{D.28})$$

$$\frac{\partial f(\xi, \eta, l)}{\partial \eta} = \begin{cases} \varphi_\eta(\xi + \frac{1}{2}, \eta), & \text{falls } \xi < -\frac{1}{2} \\ \underbrace{\varphi_\eta(0, \eta)}_{\text{Lemma D.1}}, & \text{falls } -\frac{1}{2} \leq \xi < \frac{1}{2} \\ \varphi_\eta(\xi - \frac{1}{2}, \eta), & \text{falls } \frac{1}{2} \leq \xi \end{cases}, \quad (\text{D.29})$$

$$\frac{\partial f(\xi, \eta, l)}{\partial l} = \begin{cases} \frac{1}{2} \varphi_\xi(\xi + \frac{1}{2}, \eta), & \text{falls } \xi < -\frac{1}{2} \\ 0, & \text{falls } -\frac{1}{2} \leq \xi < \frac{1}{2} \\ -\frac{1}{2} \varphi_\xi(\xi - \frac{1}{2}, \eta), & \text{falls } \frac{1}{2} \leq \xi \end{cases}. \quad (\text{D.30})$$

Mit der 2×1 -Jakobimatrix

$$\frac{\partial \mathcal{R}_\theta^T(\boldsymbol{\xi} - \mathbf{m})}{\partial \mathbf{m}} = -\mathcal{R}_\theta^T, \quad (\text{D.31})$$

der 2×1 -Jakobimatrix

$$\frac{\partial \mathcal{R}_\theta^T(\boldsymbol{\xi} - \mathbf{m})}{\partial \theta} = \frac{\partial \mathcal{R}_\theta^T}{\partial \theta}(\boldsymbol{\xi} - \mathbf{m}) = \begin{pmatrix} -\sin \theta & \cos \theta \\ -\cos \theta & -\sin \theta \end{pmatrix} (\boldsymbol{\xi} - \mathbf{m}), \quad (\text{D.32})$$

und mit

$$\frac{\partial l}{\partial m_\xi} = \frac{\partial l}{\partial m_\eta} = \frac{\partial l}{\partial \theta} = 0, \quad \frac{\partial l}{\partial l} = 1, \quad (\text{D.33})$$

können wir das Produkt der 1×3 -Matrix mit der 3×4 -Matrix aus Gl. D.26 auch schreiben als:

$$\begin{aligned} & \frac{\partial f(\mathcal{R}_\theta^T(\boldsymbol{\xi} - \mathbf{m}), l)}{\partial \mathbf{X}_M} && \text{(D.34)} \\ & = \left(\frac{\partial f}{\partial \xi} \quad \frac{\partial f}{\partial \eta} \quad \frac{\partial f}{\partial l} \right) \Big|_{(\boldsymbol{\xi}) = (\mathcal{R}_\theta^T(\boldsymbol{\xi} - \mathbf{m}))} \cdot \begin{pmatrix} -\mathcal{R}_\theta^T & \frac{\partial \mathcal{R}_\theta^T}{\partial \theta}(\boldsymbol{\xi} - \mathbf{m}) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Literaturverzeichnis

- [Attwood *et al.* 93] C.I. Attwood, G.D. Sullivan, K.D. Baker, Model Construction from a Single Perspective View Using Shape from Symmetry, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 155-164.
- [Barnard 83] S.T. Barnard, Interpreting Perspective Images, *Artificial Intelligence* 21 (1983) 435-462.
- [Barnard & Fischler 82] S.T. Barnard, M.A. Fischler, Computational Stereo, *ACM Computing Surveys* 14:4 (1982) 553-572.
- [Barron *et al.* 92] J. L. Barron, D. J. Fleet, S. S. Bauchemin, *Performance of Optical Flow Techniques*, Technical Report (RPL-TR-9107), Robotics and Perception Laboratory, Queen's University Kingston, Ontario/Canada, July 1992. (Revised July 1993).
- [Barron *et al.* 94] J. L. Barron, D. J. Fleet, S. S. Bauchemin, Performance of Optical Flow Techniques, *International Journal of Computer Vision* 12:1 (1994) 43-77. Weitere Hinweise in [Barron *et al.* 92].
- [Bar-Shalom & Fortmann 88] Y. Bar-Shalom, T. E. Fortmann, *Tracking and Data Association*, Academic Press, Inc., Boston/MA, Orlando/FL, San Diego/CA, New York/NY, Austin/TX, London/UK, Sydney, Tokyo, Toronto/Canada, 1988.
- [Batista *et al.* 93] J. Batista, J. Dias, H. Araújo, T. de Almeida, Monoplanar Camera Calibration: Iterative Multi-Step Approach, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 479-488.
- [Baumberg & Hogg 94] A. Baumberg, D. Hogg, Learning Flexible Models from Image Sequences, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on Computer Vision (ECCV '94)*, Vol. I, Stockholm, Sweden, May 2-6, 1994, Lecture Notes in Computer Science 800, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 299-308.
- [Baumgärtner 94] N. Baumgärtner, *Entwicklung eines Inferenzsystems für das unscharfe temporallogische Schlußfolgern von Abläufen in Straßenverkehrsszenen*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Oktober 1994.

- [Beardsley & Murray 92] P. Beardsley, D. Murray, Camera Calibration Using Vanishing Points, in *Proc. British Machine Vision Conference*, Leeds/UK, Sept. 22-24, 1992, pp. 416-425.
- [Beardsley et al. 92] P. Beardsley, D. Murray, A. Zisserman, Camera Calibration Using Multiple Views, in G. Sandini (Ed.), *Proc. Second European Conference on Computer Vision (ECCV '92)*, S. Margherita Ligure, Italy, May 18-23, 1992, Lecture Notes in Computer Science **588**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1992, pp. 312-320.
- [Bernasch & Koutny 93] J. Bernasch, R. Koutny, Stabile Objektverfolgung und Detektion von nicht-vorhersagbarem Verhalten in komplexen Bildfolgen, in S.J. Pöppel, H. Handels (Hrsg.), *Mustererkennung 1993*, 15. DAGM-Symposium, Lübeck, 27.-29. September 1993, Springer-Verlag, Berlin, Heidelberg, New York/NY u.a., 1993, pp. 19-26.
- [Birnbaum et al. 93] L. Birnbaum, M. Brand, P. Cooper, Looking for Trouble: Using Causal Semantics to Direct Focus of Attention, in *Proc. Fourth International Conference on Computer Vision (ICCV '93)*, Berlin, Germany, May 11-14, 1993, pp. 49-56.
- [Blake et al. 93] A. Blake, R. Curwen, A. Zisserman, A Framework for Spatiotemporal Control in the Tracking of Visual Contours, *International Journal of Computer Vision* **11:2** (1993) 127-145.
- [Bouthemy & François 93] P. Bouthemy, E. François, Motion Segmentation and Qualitative Dynamic Scene Analysis from an Image Sequence, *International Journal of Computer Vision* **10** (1993) 157-182.
- [Bouthemy & Santillana Rivero 87] P. Bouthemy, J. Santillana Rivero, A Hierarchical Likelihood Approach for Region Segmentation According to Motion-Based Criteria, in *Proc. First International Conference on Computer Vision (ICCV '87)*, London, UK, June 8-11, 1987, pp. 463-467.
- [Brzoska 93] C. Brzoska, *Temporal-logisches Programmieren*, Dissertation, Fakultät für Informatik der Universität Karlsruhe, Dezember 1993.
- [Campani & Verri 92] M. Campani, A. Verri, Motion Analysis from First-Order Properties of Optical Flow, *Computer Vision, Graphics, and Image Processing* **56** (1992) 90-107.
- [Caprile & Torre 90] B. Caprile, V. Torre, Using Vanishing Points for Camera Calibration, *International Journal of Computer Vision* **4** (1990) 127-140.
- [Cédras & Shah 94] C. Cédras, M. Shah, A Survey of Motion Analysis from Moving Light Displays, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '94)*, Seattle/WA, June 21-23, 1994, pp. 214-221. Weitere Hinweise in [Cédras & Shah 95].

- [Cédras & Shah 95] C. Cédras, M. Shah, Motion-Based Recognition: A Survey, *Image and Vision Computing* **13:2** (1995) 129-155.
- [Chen & Jiang 91] W. Chen, B.C. Jiang, 3-D Camera Calibration Using Vanishing Point Concept, *Pattern Recognition* **24** (1991) 57-67.
- [Collins & Beveridge 93] R.T. Collins, J.R. Beveridge, Matching Perspective Views of Coplanar Structures Using Projective Unwarping and Similarity Matching, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '93)*, New York/NY, June 15-17, 1993, pp. 240-245.
- [Collins & Weiss 90] R.T. Collins, R.S. Weiss, Vanishing Point Calculation as a Statistical Inference on the Unit Sphere, in *Proc. Third International Conference on Computer Vision (ICCV '90)*, Osaka, Japan, Dec. 4-7, 1990, pp. 400-403.
- [Cooper et al. 94] P. Cooper, L. Birnbaum, D. Halabe, M. Brand, P.N. Prokopenko, Divided We Fall: Resolving Occlusions Using Causal Reasoning, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on Computer Vision (ECCV '94)*, Vol. I, Stockholm, Sweden, May 2-6, 1994, Lecture Notes in Computer Science **800**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 535-540.
- [Crowley et al. 92] J. Crowley, P. Stelmaszyk, T. Skordas, P. Puget, Measurement and Integration of 3-D Structures by Tracking Edge Lines, *International Journal of Computer Vision* **8** (1992) 29-52.
- [Damm 93] H. Damm, *Zur Übertragung eines Systems zur Extraktion von Bewegungsverbänden aus Bildfolgen in die englische Sprache*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, November 1993.
- [Damm 94] H. Damm, *Entwurf und Implementierung von Situations- und Handlungsschemata zur Beschreibung von Abläufen einer Tankstellenszene*, Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Dezember 1994.
- [Daniilidis 92] K. Daniilidis, *Zur Fehlerempfindlichkeit in der Ermittlung von Objektbeschreibungen und relativen Bewegungen aus monokularen Bildfolgen*, Dissertation, Fakultät für Informatik, Universität Karlsruhe (TH), Juli 1992.
- [Daucher et al. 93] N. Daucher, M. Dhome, J.T. Lapresté, G. Rives, Modelled Object Pose Estimation and Tracking by Monocular Vision, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 247-257.
- [Del Bimbo & Nesi 93] A. Del Bimbo, P. Nesi, *A Vision System for Estimating People Flow*, Technical Report DSI-RT 15/93, Department of Systems and Informatics, Faculty of Engineering, University of Florence/Italy, 1993.

- [Deriche & Faugeras 90] R. Deriche, O. Faugeras, Tracking Line Segments, *Image and Vision Computing* 8:4 (1990) 261-270.
- [Dickmanns & Mysliwetz 92] E.D. Dickmanns, B.D. Mysliwetz, Recursive 3-D Road and Relative Ego-State Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-14:2 (1992) 199-213.
- [Du et al. 93] L. Du, G.D. Sullivan, K.D. Baker, Quantitative Analysis of the View-point Consistency Constraint in Model-Based Vision, in *Proc. Fourth International Conference on Computer Vision (ICCV '93)*, Berlin, Germany, May 11-14, 1993, pp. 632-639.
- [Dubuisson & Jain 93] M.-P. Dubuisson, A.K. Jain, Object Contour Extraction Using Color and Motion, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '93)*, New York/NY, June 15-17, 1993, pp. 471-476.
- [Dubuisson & Jain 94] M.-P. Dubuisson, A.K. Jain, 2D Matching of 3D Moving Objects in Color Outdoor Scenes, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '94)*, Seattle/WA, June 21-23, 1994, pp. 887-891.
- [Duda & Hart 73] R. Duda, P. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York/NY, Chichester, Brisbane, Toronto, Singapore, 1973.
- [Echigo 90] T. Echigo, A Camera Calibration Technique Using Three Sets of Parallel Lines, *Machine Vision and Applications* 3 (1990) 159-167.
- [Eckstein 92] W. Eckstein, *Die Bildanalyseprache TRIAS*, Dissertation, Fakultät für Mathematik und Informatik, Technische Universität München, Dezember 1992. Erschienen in der Reihe Dissertationen zur Künstlichen Intelligenz (DIS-KI) 31, infix-Verlag, Sankt Augustin, 1993.
- [Eckstein et al. 93] W. Eckstein, G. Lohmann, U. Meyer-Gruhl, R. Riemer, L. Altamirano Robles, J. Wunderwald, Benutzerfreundliche Bildanalyse mit HORUS: Architektur und Konzepte, in S.J. Pöppel, H. Handels (Hrsg.), *Mustererkennung 1993*, 15. DAGM-Symposium, Lübeck, 27.-29. September 1993, Springer-Verlag, Berlin, Heidelberg, New York/NY u.a., 1993, pp. 332-340.
- [Enkelmann 87] W. Enkelmann, Untersuchungen zur Hindernisdetektion durch Auswertung von Verschiebungsvektorfeldern, in E. Paulus (Hrsg.), *9. DAGM-Symposium Mustererkennung 1987*, Braunschweig, 29.9.-1.10.1987, Informatik-Fachberichte 149, Springer-Verlag, Berlin, Heidelberg, New York/NY u.a., 1987, pp. 282-286.
- [Enkelmann 91] W. Enkelmann, Obstacle Detection by Evaluation of Optical Flow Fields from Image Sequences, *Image and Vision Computing* 9 (1991) 160-168.

- [Faugeras 93] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, The MIT Press, Cambridge/MA, London/UK, 1993.
- [Freeman 93] W.T. Freeman, Exploiting the Generic View Assumption to Estimate Scene Parameters, in *Proc. Fourth International Conference on Computer Vision (ICCV '93)*, Berlin, Germany, May 11-14, 1993, pp. 347-356.
- [Früauf 95] T. Früauf, *Berechnung begrifflicher Beschreibungen für Ansammlungen von Fahrzeugen durch Bildfolgenauswertung*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Januar 1995.
- [Gelb 74] A. Gelb (Ed.), *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.
- [Gerber 94] R. Gerber, *Diskurs-Repräsentationstheorie zur Darstellung zeitlicher begrifflicher Beschreibungen von Abläufen in Straßenverkehrsszenen*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Dezember 1994.
- [Gong & Buxton 92] S. Gong, H. Buxton, On the Visual Expectation of Moving Objects, in B. Neumann (Ed.), *Proc. 10th European Conference on Artificial Intelligence (ECAI '92)*, Vienna/Austria, August 3-7, 1992, pp. 781-784.
- [Gong & Buxton 93] S. Gong, H. Buxton, From Contextual Knowledge to Computational Constraints, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 229-238.
- [Grimson 90] W.E.L. Grimson, *Object Recognition by Computer: The Role of Geometric Constraints*, MIT Press, Cambridge, MA, 1990.
- [Haag 94] M. Haag, *Modellierung eines Situationsgraphen am Beispiel einer Tankstellenszene*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Mai 1994.
- [Haag 95] M. Haag, *Situationsgraphen zur Modellierung von Zuständen und Handlungen im Diskursbereich Straßenverkehr*, Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, März 1995.
- [Haralick 80] R.M. Haralick, Using Perspective Transformations in Scene Analysis, *Computer Graphics and Image Processing* **13** (1980) 191-221.
- [Haralick 89] R.M. Haralick, Determining Camera Parameters from the Perspective Projection of a Rectangle, *Pattern Recognition* **22:3** (1989) 223-230.
- [Haralick & Shapiro 93] R.M. Haralick, L.G. Shapiro, *Computer and Robot Vision*, Volume 2, Addison-Wesley Publishing Company, Reading/MA and others, 1993.

- [Heinze *et al.* 91] N. Heinze, W. Krüger, H.-H. Nagel, Berechnung von Bewegungs-
verben zur Beschreibung von aus Bildfolgen gewonnenen Fahrzeugtrajektorien
in Straßenverkehrsszenen, *Informatik - Forschung und Entwicklung* 6 (1991)
51-61.
- [Heinze 93] N. Heinze, *Beschreibung der Benutzerschnittstellen der Werkzeuge für
die Situationsmodellierung*, Interner Bericht, Fraunhofer-Institut für Informa-
tions- und Datenverarbeitung (ITB), Karlsruhe, Deutschland, August 1993.
- [Henricsson & Heitger 94] O. Henricsson, F. Heitger, The Role of Key-Points in Find-
ing Contours, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on
Computer Vision ECCV '94*, Vol. II, Stockholm, Sweden, May 2-6, 1994, Lec-
ture Notes in Computer Science 801, Springer-Verlag, Berlin, Heidelberg, New
York/NY and others, 1994, pp. 371-382.
- [Herslund *et al.* 93] M.-B. Herslund, N.O. Jørgensen, G. Petersen, Videobasierte
Verhaltensstudien, in H.H. Topp (Hrsg.), *Videobildverarbeitung in der Ver-
kehrsanalyse*, Grüne Reihe Nr. 28, Fachgebiet Verkehrswesen, Fakultät für
Bauingenieurwesen der Universität Kaiserslautern, 1993, pp. 57-69.
- [Herzog 92] G. Herzog, Utilizing Interval-Based Event Representations for Incre-
mental High-Level Scene Analysis, in *Proc. Fourth European Workshop on
Semantics of Time, Space and Movement and Spatio-Temporal Reasoning*,
Château de Bonas, France, September 4-8, 1992, pp. 425-435.
- [Herzog *et al.* 89] G. Herzog, C.-K. Sung, E. André, W. Enkelmann, H.-H. Nagel,
Th. Rist, W. Wahlster, G. Zimmermann, Incremental Natural Language De-
scription of Dynamic Imagery, in Ch. Freksa, W. Brauer (Hrsg.), *Wissensba-
sierte Systeme, 3. Internationaler GI-Kongress*, München, 16. - 17. Oktober
1989, Informatik-Fachberichte 227, Springer-Verlag, Berlin, Heidelberg, New
York/NY u.a., 1989, pp. 153-162.
- [Horn 86] B.K.P. Horn, *Robot Vision*, MIT Press, Cambridge, MA, 1986.
- [Howarth & Buxton 92] R. Howarth, H. Buxton, Analogical Representation of Spa-
ce and Time, *Image and Vision Computing* 10:7 (1992) 467-478.
- [Howarth & Buxton 93] R. Howarth, H. Buxton, Selective Attention in Dyna-
mic Vision, in *Proc. 13th Intern. Joint Conf. on Artificial Intelligence*,
Chambéry/France, Aug. 28 - Sep. 3, 1993, pp. 1579-1584.
- [Hsu *et al.* 84] Y.Z. Hsu, H.-H. Nagel, G. Rekers, New Likelihood Test Methods for
Change Detection in Image Sequences, *Computer Vision, Graphics, and Image
Processing* 26 (1984) 73-106.
- [Huang *et al.* 94] T. Huang, D. Koller, J. Malik, G. Ogasawara, B. Rao, S. Russell,
J. Weber, Automatic Symbolic Traffic Scene Analysis Using Belief Networks,
in *Proc. of the 12th National Conference on Artificial Intelligence*, Seattle/WA,
Jul. 31-Aug. 4, 1994, pp. 966-972.

- [Hupfer 93] Ch. Hupfer, Videoverkehrsanalyse (ViVA), in H.H. Topp (Hrsg.), *Videobildverarbeitung in der Verkehrsanalyse*, Grüne Reihe Nr. 28, Fachgebiet Verkehrswesen, Fakultät für Bauingenieurwesen der Universität Kaiserslautern, 1993, pp. 83–95.
- [Huttenlocher *et al.* 93] D.P. Huttenlocher, G.A. Klandermann, W.J. Rucklidge, Comparing Images Using the Hausdorff Distance, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-15:9** (1993) 850–863.
- [Irani *et al.* 94a] M. Irani, B. Rousso, S. Peleg, Computing Occluding and Transparent Motion, *International Journal of Computer Vision* **12:1** (1994) 5–16.
- [Irani *et al.* 94b] M. Irani, B. Rousso, S. Peleg, Recovery of Ego-Motion Using Image Stabilization, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '94)*, Seattle/WA, June 21–23, 1994, pp. 454–460.
- [Jiang & Ward 94] C. Jiang, M.O. Ward, Shadow Segmentation and Classification in a Constrained Environment, *CVGIP: Image Understanding* **59:2** (1994) 213–225.
- [Kamp & Reyle 93] H. Kamp, U. Reyle, *From Discourse to Logic*, Kluwer Academic Publishers, Dordrecht, Boston/MA, London, 1993.
- [Kanade 78] T. Kanade, Region Segmentation: Signal vs. Semantics, in *Proc. 4th Int. Joint Conf. on Pattern Recognition*, Kyoto, Japan, Nov. 7–10, 1978, pp. 95–105.
- [Kanade 80] T. Kanade, Region Segmentation: Signal vs. Semantics, *Computer Graphics and Image Processing* **13** (1980) 279–297.
- [Kanatani 91] K. Kanatani, Computational Projective Geometry, *CVGIP: Image Understanding* **54** (1991) 333–348. Vgl. Abschnitte 2.1–2.3, 3.1–3.3, 4.1, 4.3 in [Kanatani 93a].
- [Kanatani 92] K. Kanatani, Statistical Analysis of Focal-Length Calibration Using Vanishing Points, *IEEE Trans. Robotics and Automation* **8:6** (1992) 767–775. Vgl. Abschnitt 10.1 in [Kanatani 93a].
- [Kanatani 93a] K. Kanatani, *Geometric Computation for Machine Vision*, Clarendon Press, Oxford/UK, 1993.
- [Kanatani 93b] K. Kanatani, Renormalization for Unbiased Estimation, in *Proc. Fourth International Conference on Computer Vision (ICCV '93)*, Berlin, Germany, May 11–14, 1993, pp. 599–606. Vgl. Abschnitte 8.4, 9.3, 9.4, 10.3 in [Kanatani 93a].
- [Kanatani & Onodera 91] K. Kanatani, Y. Onodera, Anatomy of Camera Calibration Using Vanishing Points, *IEICE Transactions* **E-74:10** (1991) 3369–3378. Vgl. Abschnitte 3.2, 3.4 in [Kanatani 93a].

- [Kanatani & Watanabe 90] K. Kanatani, K. Watanabe, Reconstruction of 3-D Road Geometry from Images for Autonomous Land Vehicles, *IEEE Trans. Robotics and Automation* **6** (1990) 127-132.
- [Karmann & von Brandt 90] K.-P. Karmann, A. von Brandt, Moving Object Recognition Using Adaptive Background Memory, in V. Cappellini (Ed.), *Proc. Third Intern. Workshop on Time-Varying Image Processing and Moving Object Recognition*, Florence, Italy, May 29-31, 1989, Elsevier, Amsterdam, Oxford, New York/NY, Tokyo, 1990, pp. 725-732.
- [Kasprzak et al. 94] W. Kasprzak, H. Niemann, D. Wetzel, Adaptive Road Parameter Estimation in Monocular Image Sequences, in *Proc. British Machine Vision Conference*, York/UK, Sept. 13-16, 1994, pp. 691-700.
- [Kaufmann 93] B. Kaufmann, *Ballungsanalyse und 2D-Verfolgung bewegter Objekte zur Initialisierung einer Modellanpassung*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Mai 1993.
- [Kellner 93] A. Kellner, *Detektion sich bewegender Fußgänger in monokularen Bildfolgen*, Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), durchgeführt am Fraunhofer-Institut für Informations- und Datenverarbeitung (IITB), Karlsruhe, Juli 1993.
- [Kilger 93a] M. Kilger, Ein videobasiertes System zur Erfassung von Verkehrsdaten, in H.H. Topp (Hrsg.), *Videobildverarbeitung in der Verkehrsanalyse*, Grüne Reihe Nr. 28, Fachgebiet Verkehrswesen, Fakultät für Bauingenieurwesen der Universität Kaiserslautern, 1993, pp. 22-39.
- [Kilger 93b] M. Kilger, Heavy Traffic Monitoring in Real-Time, in *Proc. 8th Scandinavian Conf. on Image Analysis*, Tromsø, Norway, May 25-28, 1993, pp. 1077-1083.
- [Kilger 93c] M. Kilger, Interpretationsgetriebene Parameteradaption und Verfahrenswahl in der Szenenanalyse, in S.J. Pöpl, H. Handels (Hrsg.), *Mustererkennung 1993*, 15. DAGM-Symposium, Lübeck, 27.-29. September 1993, Springer-Verlag, Berlin, Heidelberg, New York/NY u.a., 1993, pp. 725-732.
- [Kinzel 94] W. Kinzel, *Aufgabenabhängige Bildbeschreibungen zur Personenerkennung - Kanten-, Regionenbilder und Erkennungsergebnisse*, Vortrag, Forschungskolloquium Bildverstehen, Studienstiftung des deutschen Volkes, 22.-24.07.1994, Fischbachau/Bayern, 1994. Siehe auch: W. Kinzel, *Präattentive und attentive Bildverarbeitungsschritte zur visuellen Erkennung von Fußgängern*, Dissertation, angenommen am 8. Juni 1994 durch die Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München, Neubiberg.

- [Koller *et al.* 91] D. Koller, N. Heinze, H.-H. Nagel, Algorithmic Characterization of Vehicle Trajectories from Image Sequences by Motion Verbs, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '91)*, Lahaina, Maui, Hawaii/HI, June 3-6, 1991, pp. 90-95.
- [Koller 92] D. Koller, *Detektion, Verfolgung und Klassifikation bewegter Objekte in monokularen Bildfolgen am Beispiel von Straßenverkehrsszenen*, Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Juni 1992. Erschienen in der Reihe Dissertationen zur Künstlichen Intelligenz (DISKI) 13, infix-Verlag, Sankt Augustin, 1992.
- [Koller *et al.* 93a] D. Koller, K. Daniilidis, H.-H. Nagel, Model-Based Object Tracking in Monocular Image Sequences of Road Traffic Scenes, *International Journal of Computer Vision* 10:3 (1993) 257-281.
- [Koller *et al.* 93b] D. Koller, J. Weber, J. Malik, *Robust Multiple Car Tracking with Occlusion Reasoning*, Technical Report UCB/CSD-93-780, Computer Science Department, University of California at Berkeley, October 1993.
- [Koller *et al.* 94] D. Koller, J. Weber, J. Malik, Robust Multiple Car Tracking with Occlusion Reasoning, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on Computer Vision (ECCV '94)*, Vol. I, Stockholm, Sweden, May 2-6, 1994, Lecture Notes in Computer Science 800, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 189-196. Weitere Hinweise in [Koller *et al.* 93b].
- [Kollnig 92] H. Kollnig, *Berechnung von Bewegungsverben und Ermittlung einfacher Abläufe*, Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Januar 1992.
- [Kollnig & Nagel 93] H. Kollnig, H.-H. Nagel, Ermittlung von begrifflichen Beschreibungen von Geschehen in Straßenverkehrsszenen mit Hilfe unscharfer Mengen, *Informatik - Forschung und Entwicklung* 8:4 (1993) 186-196.
- [Kollnig *et al.* 94] H. Kollnig, H.-H. Nagel, M. Otte, Association of Motion Verbs with Vehicle Movements Extracted from Dense Optical Flow Fields, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on Computer Vision ECCV '94*, Vol. II, Stockholm, Sweden, May 2-6, 1994, Lecture Notes in Computer Science 801, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 338-347.
- [Kories & Zimmermann 84] R. Kories, G. Zimmermann, Motion Detection in Image Sequences: an Evaluation of Feature Detectors, in *Proc. Int. Conf. on Pattern Recognition*, Montreal, Canada, July 30 - Aug. 2, 1984, pp. 778-780.
- [Kories & Zimmermann 86] R. Kories, G. Zimmermann, A Versatile Method for the Estimation of Displacement Vector Fields from Image Sequences, in *Proc. of IEEE Workshop on Motion: Representation and Analysis*, Kiawah Island Resort, Charleston/SC, May 7-9, 1986, pp. 101-106.

- [Krebs 80] V. Krebs, *Nichtlineare Filterung*, R. Oldenbourg Verlag, München und Wien, 1980.
- [Krein 90] B. Krein, *Entwurf und Implementierung einer interaktiven Kamerakalibrierung mit Hilfe eines digitalisierten Grundrißplanes der dargestellten Außenraumszene*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, April 1990.
- [Krüger 91] W. Krüger, *Begriffsgraphen zur Situationsmodellierung in der Bildfolgenauswertung*, Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), Januar 1991. erschienen in: Informatik-Fachberichte **311**. Berlin, Heidelberg, New York/NY, London, Paris, Tokyo: Springer-Verlag 1992.
- [Kumar & Hanson 90] R. Kumar, A.R. Hanson, Sensitivity of the Pose Refinement Problem to Accurate Estimation of Camera Parameters, in *Proc. Third International Conference on Computer Vision (ICCV '90)*, Osaka, Japan, Dec. 4-7, 1990, pp. 365-369.
- [Lavest *et al.* 93] J.-M. Lavest, G. Rives, M. Dhome, Three-Dimensional Reconstruction by Zooming, *IEEE Trans. Robotics and Automation* **9:2** (1993) 196-207.
- [Lebègue & Aggarwal 93] X. Lebègue, J.K. Aggarwal, Significant Line Segments for an Indoor Mobile Robot, *IEEE Trans. Robotics and Automation* **9:6** (1993) 801-815.
- [Lenz 87] R. Lenz, Linsenfehlerkorrigierte Eichung von Halbleiterkameras mit Standardobjektiven für hochgenaue 3D-Messungen in Echtzeit, in E. Paulus (Hrsg.), *9. DAGM-Symposium Mustererkennung 1987*, Braunschweig, 29.9.-1.10.1987, Informatik-Fachberichte **149**, Springer-Verlag, Berlin, Heidelberg, New York/NY u.a., 1987, pp. 212-215.
- [Lenz & Tsai 88] R. Lenz, R.Y. Tsai, Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Metrology, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-10** (1988) 713-720.
- [Leonardis 93] A. Leonardis, *Image Analysis Using Parametric Models*, Dissertation, Faculties of Electrical Engineering and Computer Science, University of Ljubljana, Slovenia, March 1993.
- [Leuck 94] H. Leuck, *Segmentierung von optischen Flussfeldern unter Verwendung eines Fahrbahnmodells zur Initialisierung einer Modellanpassung*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Juli 1994.
- [Leuck 95] H. Leuck, *Verfolgung und Klassifikation von Fahrzeugen in Bildfolgen*, Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, März 1995.

- [Liou & Jain 87] S.P. Liou, R.C. Jain, Road Following Using Vanishing Points, *Computer Vision, Graphics, and Image Processing* **39** (1987) 116-130.
- [Li-Qun *et al.* 92] X. Li-Qun, D. Young, D. Hogg, Building a Model of a Road Junction Using Moving Vehicle Information, in *Proc. British Machine Vision Conference*, Leeds/UK, Sept. 22-24, 1992, pp. 443-452.
- [Liu & Huang 93] Y. Liu, T.S. Huang, Vehicle-Type Motion Estimation from Multi-frame Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-15:8** (1993) 802-808.
- [Liu *et al.* 90] Y. Liu, T.S. Huang, O.D. Faugeras, Determination of Camera Location from 2-D to 3-D Line and Point Correspondences, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-12:1** (1990) 28-37.
- [Lowe 92a] D.G. Lowe, Robust Model-Based Motion Tracking Through the Integration of Search and Estimation, *International Journal of Computer Vision* **8:2** (1992) 113-122.
- [Lowe 92b] D. Lowe, Solving for 3-D Model Parameters from the Location of Image Features, in O. Khatib, J.J. Craig, T. Lozano-Pérez (Eds.), *The Robotics Review 2*, The MIT Press, Cambridge, MA, 1992, pp. 137-143.
- [Magee & Aggarwal 84] M.J. Magee, J.K. Aggarwal, Determining Vanishing Points From Perspective Images, *Computer Vision, Graphics, and Image Processing* **26** (1984) 256-267.
- [Maier 93] M. Maier, *Werkzeuge zur Bewertung von Beschreibungen einfacher Abläufe in Bildfolgen*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Juni 1993.
- [Marslin *et al.* 91] R.F. Marslin, G.D. Sullivan, K.D. Baker, Kalman Filters in Constrained Model-Based Tracking, in *Proc. British Machine Vision Conference*, Glasgow/UK, Sept. 24-26, 1991, pp. 371-374.
- [Mäusel 91] R. Mäusel, *Fernsehtechnik: von der Kamera bis zum Bildschirm*, Huethig, Heidelberg, 1991.
- [Maybank 92] S. Maybank, *Theory of Reconstruction from Image Motion*, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1992.
- [Maybeck 79] P.S. Maybeck, *Stochastic Models, Estimation and Control*, Volume 1 & 2, Academic Press, New York/NY and London/UK, 1979.
- [Messer 92] T. Messer, *Wissensbasierte Synthese von Bildanalyseprogrammen*, Dissertation, Fakultät für Mathematik und Informatik, Technische Universität München, August 1992. Erschienen in der Reihe Dissertationen zur Künstlichen Intelligenz (DISKI) **22**, infix-Verlag, Sankt Augustin, 1992.

- [Meyer & Bouthemy 94] F. Meyer, P. Bouthemy, Region-Based Tracking Using Affine Motion Models in Long Image Sequences, *CVGIP: Image Understanding* **60:2** (1994) 119-140.
- [Mohnhaupt 90] M. Mohnhaupt, Eine hybride Repräsentation von Objektbewegungen: Von analogen zu propositionalen Beschreibungen, in Ch. Freksa, Ch. Habel (Hrsg.), *Repräsentation und Verarbeitung räumlichen Wissens*, Informatik-Fachberichte **245**, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1990, pp. 143-155.
- [Mohnhaupt 92] M. Mohnhaupt, *Prinzipien piktorieller Repräsentationssysteme*, Informatik-Fachberichte **300**, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1992.
- [Mohnhaupt & Neumann 90] M. Mohnhaupt, B. Neumann, On the Use of Motion Concepts for Top-Down Control in Traffic Scenes, in O. Faugeras (Ed.), *Proc. Second European Conference on Computer Vision (ECCV '92)*, Antibes, France, April 23-26, 1990, Lecture Notes in Computer Science **427**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1990, pp. 598-600.
- [Mourad 94] S. Mourad, *Anpassung der Lage eines Kreuzungsmodells durch Auswertung von Bildfolgen bei bewegter Kamera*, Diplomarbeit, Fakultät für Elektrotechnik der Universität Karlsruhe (TH), durchgeführt am Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Dezember 1994.
- [Mück 94a] K. Mück, *Entwicklung und Implementierung eines Busmodells*, Interner Bericht, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, März 1994.
- [Mück 94b] K. Mück, *Entwurf und Anpassung begrifflicher und geometrischer Modelle an Bildfolgen von innerstädtischen Straßenkreuzungen*, Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, November 1994.
- [Mundy & Zisserman 92] J. Mundy, A. Zisserman (Eds.), *Geometric Invariance in Computer Vision*, The MIT Press, Cambridge/MA, London/UK, 1992.
- [Nagel 77] H.-H. Nagel, Analysing Sequences of TV-Frames, in *Proc. 5th Intern. Joint Conf. on Artificial Intelligence*, Cambridge/MA, 22-25 August 1977, p. 626.
- [Nagel 79] H.-H. Nagel, Über die Repräsentation von Wissen zur Auswertung von Bildern, in J.P. Foith (Hrsg.), *DAGM-Symposium Mustererkennung 1979*, Karlsruhe, Informatik-Fachberichte **20**, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1979, pp. 3-21.

- [Nagel 81] H.-H. Nagel, Image Sequence Analysis: What Can We Learn from Applications?, in T. Huang (Ed.), *Image Sequence Analysis*, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1981, pp. 19–228.
- [Nagel 85] H.-H. Nagel, Analyse und Interpretation von Bildfolgen, *Informatik-Spektrum* 8:4 (1985) 178–200 und *Informatik-Spektrum* 8:6 (1985) 312–327.
- [Nagel 87] H.-H. Nagel, On the Estimation of Optical Flow: Relations between Different Approaches and Some New Results, *Artificial Intelligence* 33 (1987) 299–324.
- [Nagel 88] H.-H. Nagel, From Image Sequences towards Conceptual Descriptions, *Image and Vision Computing* 6:2 (1988) 59–74.
- [Nagel 89] H.-H. Nagel, Zur Erkennung von Situationen durch Auswertung von Bildfolgen, *FhG-Berichte 1/89* (1989) 25–33.
- [Nagel 91] H.-H. Nagel, La représentation de situations et leur reconnaissance à partir de séquences d'images — The Representation of Situations and their Recognition from Image Sequences, in 8^e *Congrès Reconnaissance des Formes et Intelligence Artificielle*, Lyon-Villeurbanne, 25–29 Novembre 1991, AFCET, 1991, pp. 1221–1229.
- [Nagel 92] H.-H. Nagel, Direct Estimation of Optical Flow and Its Derivatives, in G.A. Orban, H.-H. Nagel (Eds.), *Artificial and Biological Vision Systems*, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1992, pp. 193–224.
- [Nagel 94a] H.-H. Nagel, AI Approaches towards Sensor-Based Driver Support in Road Vehicles, in B. Nebel, L. Dreschler-Fischer (Eds.), *KI-94: Advances in Artificial Intelligence, Proc. 18th German Annual Conference on Artificial Intelligence*, Saarbrücken, Germany, September 18–23, 1994, Lecture Notes in Artificial Intelligence 861, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 1–15.
- [Nagel 94b] H.-H. Nagel, A Vision of 'Vision and Language' Comprises Action: an Example from Road Traffic, *Artificial Intelligence Review, Special Volume on Integration of Natural Language and Vision Processing* 8:5/6 (1994) 189–214.
- [Nagel 95a] H.-H. Nagel, Optical Flow Estimation and the Interaction Between Measurement Errors at Adjacent Pixel Positions, *International Journal of Computer Vision* (1995). To appear.
- [Nagel 95b] H.-H. Nagel, Auswertung von Bildfolgen: Forschungsgegenstand und Werkzeug (1995). Eingereicht zur Veröffentlichung.
- [Nagel et al. 94] H.-H. Nagel, G. Socher, H. Kollnig, M. Otte, Motion Boundary Detection in Image Sequences by Local Stochastic Tests, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on Computer Vision ECCV '94*, Vol.

- II, Stockholm, Sweden, May 2-6, 1994, Lecture Notes in Computer Science **801**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 305-315.
- [Neumann & Novak 86] B. Neumann, H.-J. Novak, **NAOS**: Ein System zur natürlich-sprachlichen Beschreibung zeitveränderlicher Szenen, *Informatik - Forschung und Entwicklung* **1** (1986) 83-92.
- [Niemann 83] H. Niemann, *Klassifikation von Mustern*, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1983.
- [Niyogi & Adelson 94] S.A. Niyogi, E.H. Adelson, Analyzing and Recognizing Walking Figures in XYT, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '94)*, Seattle/WA, June 21-23, 1994, pp. 469-474.
- [Osswald 94] T.-H. Osswald, *Anpassung generischer Modelle innerstädtischer Straßenkreuzungen an Bildfolgen*, Diplomarbeit, Fakultät für Elektrotechnik der Universität Karlsruhe (TH), durchgeführt am Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, April 1994.
- [Otte 94] M. Otte, *Extraktion von linienförmigen Merkmalen und Ermittlung des optischen Flusses mit seinen Ableitungen aus Voll- und Halbbildfolgen*, Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Deutschland, Februar 1994. Dissertationen zur Künstlichen Intelligenz (DIS-KI), Band 63, Sankt Augustin: infix-Verlag 1994.
- [Otte & Nagel 92] M. Otte, H.-H. Nagel, Extraction of Line Drawings from Gray Value Images by Non-Local Analysis of Edge Element Structures, in G. Sandini (Ed.), *Proc. Second European Conference on Computer Vision (ECCV '92)*, S. Margherita Ligure, Italy, May 18-23, 1992, Lecture Notes in Computer Science **588**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1992, pp. 687-695.
- [Palmer & Tai 93] P. Palmer, A. Tai, An Optimized Vanishing Point Detector, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 529-538.
- [Pavlidis & Horowitz 74] T. Pavlidis, S.L. Horowitz, Segmentation of Plane Curves, *IEEE Transactions on Computer* **C-23:8** (1974) 860-870.
- [Penna 91] M.A. Penna, Camera Calibration: A Quick and Easy Way to Determine the Scale Factor, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-13:12** (1991) 1240-1245.
- [Poelman & Kanade 94] C.J. Poelman, T. Kanade, A paraperspective Factorization Method for Shape and Motion Recovery, in J.-O. Eklundh (Ed.), *Proc. Third*

- European Conference on Computer Vision ECCV '94*, Vol. II, Stockholm, Sweden, May 2-6, 1994, Lecture Notes in Computer Science **801**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 97-108.
- [Proesmans *et al.* 94] M. Proesmans, L. Van Gool, E. Pauwels, A. Oosterlinck, Determination of Optical Flow and Its Discontinuities Using Non-Linear Diffusion, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on Computer Vision ECCV '94*, Vol. II, Stockholm, Sweden, May 2-6, 1994, Lecture Notes in Computer Science **801**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 295-304.
- [Prokop & Reeves 92] R.J. Prokop, A.P. Reeves, A Survey of Moment-Based Techniques for Unoccluded Object Representation and Recognition, *CVGIP: Graphical Models and Image Processing* **54:5** (1992) 438-460.
- [Quan & Mohr 89] L. Quan, R. Mohr, Determining Perspective Structures Using Hierarchical Hough Transform, *Pattern Recognition Letters* **9:4** (1989) 279-286.
- [Regensburger 93] U. Regensburger, *Zur Erkennung von Hindernissen in der Bahn eines Straßenfahrzeugs durch maschinelles Echtzeitschichten*, Dissertation, Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München, Neubiberg, Dezember 1993.
- [Rehg & Kanade 94] J.M. Regh, T. Kanade, Visual Tracking of High DOF Articulated Structures: an Application to Human Hand Tracking, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on Computer Vision ECCV '94*, Vol. II, Stockholm, Sweden, May 2-6, 1994, Lecture Notes in Computer Science **801**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 35-46.
- [Retz-Schmidt 91] G. Retz-Schmidt, Recognizing Intentions, Interactions and Causes of Plan Failures, *User Modeling and User-Adapted Interaction* **1:2** (1991) 173-202.
- [Retz-Schmidt 92] G. Retz-Schmidt, *Die Interpretation des Verhaltens mehrerer Akteure in Szenenfolgen*, *Informatik-Fachberichte* **308**, Springer-Verlag, Berlin, Heidelberg, New York/NY u.a., 1992.
- [Roberts & Charnley 93] J. Roberts, D. Charnley, Attentive Visual Tracking, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 459-468.
- [Scales 85] L.E. Scales, *Introduction to Non-Linear Optimization*, Macmillan, London/UK, 1985.
- [Schäfer 93] K. Schäfer, *Entwicklung einer temporallogischen Sprache zur Beschreibung von Abläufen in Straßenverkehrsszenen*, Diplomarbeit, Institut für Algorithmen und Kognitive Systeme sowie Institut für Logik, Komplexität und Deduktionssysteme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Januar 1993.

- [Schick & Dickmanns 91] J. Schick, E. D. Dickmanns, Simultaneous Estimation of 3D Shape and Motion of Objects by Computer Vision, in *Proc. IEEE Workshop on Visual Motion*, Nassau Inn, Princeton/NJ, October 7-9, 1991, IEEE Computer Society Press, Los Alamitos/CA, 1991, pp. 256-261.
- [Schirra 94] J.R.J. Schirra, *Bildbeschreibung als Verbindung von visuellem und sprachlichem Raum. Eine interdisziplinäre Untersuchung von Bildvorstellungen in einem Hörermodell*, Dissertation, Universität des Saarlandes, Saarbrücken, 1994. Erschienen in der Reihe Dissertationen zur Künstlichen Intelligenz (DIS-KI) 71, infix-Verlag, Sankt Augustin, 1994.
- [Schirra et al. 87] J.R.J. Schirra, G. Bosch, C.K. Sung, G. Zimmermann, From Image Sequences to Natural Language: A First Step towards Automatic Perception and Description of Motion, *Applied Artificial Intelligence* 1 (1987) 287-307.
- [Schirra & Stopp 93] J.R.J. Schirra, E. Stopp, ANTLIMA - A Listener Model with Mental Images. in *Proc. 13th Intern. Joint Conf. on Artificial Intelligence*, Vol. I, Chambéry/France, Aug. 28 - Sep. 3, 1993, pp. 175-180.
- [Schürmann 94] J. Schürmann, *Detektion sich bewogender Fußgänger in monokularen Bildfolgen*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), durchgeführt am Fraunhofer-Institut für Informations- und Datenverarbeitung (IITB), Karlsruhe, Juli 1994.
- [Schuster et al. 93] R. Schuster, N. Ansari, A. Bani-Hashemi, Steering a Robot with Vanishing Points, *IEEE Trans. Robotics and Automation* 9:4 (1993) 491-498.
- [Shigang et al. 90] L. Shigang, S. Tsuji, M. Imai, Determining of Camera Rotation from Vanishing Points of Lines on Horizontal Planes, in *Proc. Third International Conference on Computer Vision (ICCV '90)*, Osaka, Japan, Dec. 4-7, 1990, pp. 499-502.
- [Socher 93] G. Socher, *Test zur Beurteilung der Güte optischer Flußvektoren und Entwicklung eines Verfahrens zur Segmentierung von optischen Flußfeldern*, Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, November 1993.
- [Solder 92] U. Solder, *Echtzeitfähige Entdeckung von Objekten in der weiten Vorausschau eines Straßenfahrzeugs*, Dissertation, Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München, Neubiberg, Dezember 1992.
- [Solder & Graefe 93] U. Solder, V. Graefe, Visual Detection of Distant Objects, in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '93)*, Yokohama/Japan, July 26-30, 1993, pp. 1042-1049.

- [Stopp *et al.* 94] E. Stopp, K.-P. Gapp, G. Herzog, Th. Längle, T.C. Lüth, Utilizing Spatial Relations for Natural Language Access to an Autonomous Mobile Robot, in B. Nebel, L. Dreschler-Fischer (Eds.), *KI-94: Advances in Artificial Intelligence, Proc. 18th German Annual Conference on Artificial Intelligence*, Saarbrücken, Germany, September 18-23, 1994, Lecture Notes in Artificial Intelligence **861**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 39-50.
- [Storkebaum 92] A. Storkebaum, *Untersuchung von Bewegungsabläufen in Bildfolgen*, Studienarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Juli 1992.
- [Straforini *et al.* 93] M. Straforini, C. Coelho, M. Campani, Extraction of Vanishing Points from Images of Indoor and Outdoor Scenes, *Image and Vision Computing* **11:2** (1993) 91-99.
- [Sullivan 92] G.D. Sullivan, Visual Interpretation of Known Objects in Constrained Scenes, *Philosophical Transactions Royal Society London (B)* **337** (1992) 361-370.
- [Sung 88] C.-K. Sung, Extraktion von typischen und komplexen Vorgängen aus einer Bildfolge einer Verkehrsszene, in H. Bunke, O. Kübler, P. Stucki (Hrsg.), *Mustererkennung 1988*, Zürich, Informatik-Fachberichte **180**, Springer-Verlag, Berlin, Heidelberg, New York/NY u.a., 1988, pp. 90-96.
- [Sung & Zimmermann 86] C.-K. Sung, G. Zimmermann, Detektion und Verfolgung mehrerer Objekte in Bildfolgen, in G. Hartmann (Ed.), *Mustererkennung 1986*, Paderborn, Informatik-Fachberichte **125**, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1986, pp. 181-184.
- [Szeliski & Kang 94] R. Szeliski, S.B. Kang, Recovering 3D Shape and Motion from Image Streams Using Nonlinear Least Squares, *Journal of Visual Communication and Image Representation* **5:1** (1994) 10-28.
- [Tai *et al.* 93] A. Tai, J. Kittler, M. Petrou, T. Winderatt, Vanishing Point Detection, *Image and Vision Computing* **11:4** (1993) 240-245.
- [Tan *et al.* 92a] T.N. Tan, G.D. Sullivan, K.D. Baker, Structure from Motion Using the Ground Plane Constraint, in G. Sandini (Ed.), *Proc. Second European Conference on Computer Vision (ECCV '92)*, S. Margherita Ligure, Italy, May 18-23, 1992, Lecture Notes in Computer Science **588**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1992, pp. 277-281.
- [Tan *et al.* 92b] T.N. Tan, G.D. Sullivan, K.D. Baker, Linear Algorithms for Object Pose Estimation, in *Proc. British Machine Vision Conference*, Leeds/UK, Sept. 22-24, 1992, pp. 600-609.

- [Tan *et al.* 92c] T.N. Tan, G.D. Sullivan, K.D. Baker, 3D Structure and Motion Estimation from 2D Image Sequences, in *Proc. British Machine Vision Conference*, Leeds/UK, Sept. 22-24, 1992, pp. 67-78.
- [Tan *et al.* 93] T.N. Tan, G.D. Sullivan, K.D. Baker, Recognising Objects on the Ground Plane, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 85-94.
- [Tan *et al.* 94a] T.N. Tan, G.D. Sullivan, K.D. Baker, Recognising Objects on the Ground Plane, *Image and Vision Computing* 12:3 (1994) 164-172.
- [Tan *et al.* 94b] T.N. Tan, G.D. Sullivan, K.D. Baker, Pose Determination and Recognition of Vehicles in Traffic Scenes, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on Computer Vision (ECCV '94)*, Vol. I, Stockholm, Sweden, May 2-6, 1994, Lecture Notes in Computer Science 800, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 501-506.
- [Tan *et al.* 94c] T.N. Tan, G.D. Sullivan, K.D. Baker, Fast Vehicle Localization and Recognition without Line Extraction and Matching, in *Proc. British Machine Vision Conference*, York/UK, Sept. 13-16, 1994, pp. 85-94.
- [Tan *et al.* 94d] T.N. Tan, G.D. Sullivan, K.D. Baker, Linear Algorithms for Multi-Frame Structure from Constrained Motion, in *Proc. British Machine Vision Conference*, York/UK, Sept. 13-16, 1994, pp. 589-598.
- [Taubin 91] G. Taubin, Estimation of Planar Curves, Surfaces, and Nonplanar Space Curves Defined by Implicit Equations with Applications to Edge and Range Image Segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-13:11 (1991) 1115-1138.
- [Taubin 93] G. Taubin, An Improved Algorithm for Algebraic Curve and Surface Fitting, in *Proc. Fourth International Conference on Computer Vision (ICCV '93)*, Berlin, Germany, May 11-14, 1993, pp. 658-665.
- [Toal & Buxton 92] A. F. Toal, H. Buxton, Spatio-temporal Reasoning within a Traffic Surveillance System, in G. Sandini (Ed.), *Proc. Second European Conference on Computer Vision (ECCV '92)*, S. Margherita Ligure, Italy, May 18-23, 1992, Lecture Notes in Computer Science 588, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1992, pp. 884-892.
- [Tomasi & Kanade 92] C. Tomasi, T. Kanade, Shape and Motion from Image Streams under Orthography: A Factorization Method, *International Journal of Computer Vision* 9 (1992) 137-154.
- [Tou & Gonzales 74] J.T. Tou, R.C. Gonzales, *Pattern Recognition Principles*, Addison-Wesley Publishing Company, Reading/MA, London/UK, 1974.

- [Tsai 87] R. Tsai, A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses, *IEEE Trans. Robotics and Automation* **3** (1987) 323-344.
- [von Holt 94] V. von Holt, Tracking and Classification of Overtaking Vehicles on Autobahnen, in M. de Saint Blancard, I. Masaki (Eds.), *Proc. Intelligent Vehicles '94 Symposium*, Paris/France, October 24-26, 1994, pp. 314-319.
- [Wachter 93] S. Wachter, *Zur automatischen Lokalisierung von Kraftfahrzeugen in Bildfolgen bei bewegter Kamera*, Diplomarbeit, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), durchgeführt am Fraunhofer-Institut für Informations- und Datenverarbeitung (IITB), Karlsruhe, Juli 1993.
- [Walter 89] I. Walter, *Datenbankgestützte Repräsentation und Extraktion von Episodenbeschreibungen aus Bildfolgen*, Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), Februar 1989. Auch erschienen als: Informatik-Fachberichte **213**, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1989.
- [Wang & Tsai 90] L.L. Wang, W.H. Tsai, Computing Camera Parameters Using Vanishing Line Information from a Rectangular Parallelepiped, *Machine Vision and Applications* **3** (1990) 129-141.
- [Wang & Tsai 91] L.L. Wang, W.H. Tsai, Camera Calibration by Vanishing Lines for 3D Computer Vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-13** (1991) 370-376.
- [Wegmann & Spring 93] M. Wegmann, M. Spring, TRAVIS - Verkehrsanalyse mit Videobildverarbeitung, in H.H. Topp (Hrsg.), *Videobildverarbeitung in der Verkehrsanalyse*, Grüne Reihe Nr. 28, Fachgebiet Verkehrswesen, Fakultät für Bauingenieurwesen der Universität Kaiserslautern, 1993, pp. 57-69.
- [Weiss et al. 90] R.S. Weiss, H. Nakatani, E.M. Riseman, An Error Analysis for Surface Orientation from Vanishing Points, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-12:12** (1990) 1179-1185.
- [Worrall et al. 91] A.D. Worrall, R.F. Marslin, G.D. Sullivan, K.D. Baker, Model-Based Tracking, in *Proc. British Machine Vision Conference*, Glasgow/UK, Sept. 24-26, 1991, pp. 310-318.
- [Worrall et al. 93] A.D. Worrall, G.D. Sullivan, K.D. Baker, Advances in Model-Based Traffic Vision, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 559-568.
- [Worrall et al. 94] A.D. Worrall, G.D. Sullivan, K.D. Baker, Pose Refinement of Active Models Using Forces in 3D, in J.-O. Eklundh (Ed.), *Proc. Third European Conference on Computer Vision (ECCV '94)*, Vol. I, Stockholm, Sweden, May

- 2-6, 1994, Lecture Notes in Computer Science **800**, Springer-Verlag, Berlin, Heidelberg, New York/NY and others, 1994, pp. 341-350.
- [Zhang *et al.* 93] S. Zhang, G.D. Sullivan, K.D. Baker, The Automatic Construction of a View-Independent Relational Model for 3-D Object Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-15:6** (1993) 531-544.
- [Zhang 93a] X. Zhang, Computation of Vehicle Trajectories Using a Neural Network, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 489-498.
- [Zhang 93b] Z. Zhang, Strategies for Tracking Tokens in a Cluttered Scene, in *Proc. British Machine Vision Conference*, Guildford/UK, Sept. 21-23, 1993, pp. 207-216.
- [Zhang 94] J. Zhang, *Bereichsbasierte Verfahren zur Straßenerkennung für die autonome Führung von Fahrzeugen*, Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Februar 1994. Erschienen in VDI-Forschungsberichte, Reihe **10**: Informatik/Kommunikationstechnik Nr. **298**, VDI-Verlag, Düsseldorf, 1994.
- [Zhang & Faugeras 92a] Z. Zhang, O.D. Faugeras, *3D Dynamic Scene Analysis*, Springer-Verlag, Berlin, Heidelberg, New York/NY, London, Paris, Tokyo, 1992.
- [Zhang & Faugeras 92b] Z. Zhang, O.D. Faugeras, Estimation of Displacements from Two 3-D Frames Obtained from Stereo, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-14** (1992) 1141-1156.
- [Zielke *et al.* 93] T. Zielke, M. Brauckmann, W. von Seelen, Intensity and Edge-Based Symmetry Detection with an Application to Car-Following, *CVGIP: Image Understanding* **58** (1993) 177-190.

Stichwortverzeichnis

- Ablauf 1, 8, 103–104
- Ablehnungsbereich 82–83, 139–142, 148–154
- Ableitungsoperatoren 25–27, 90
- Abstand 61–63, 80–84, 139–154, 156–159
- Abstimmungsparameter 31, 89–90
- Abstraktionsebenen 1, 6, 25, 103
- Abtastfrequenz 12, 20–21, 25
- Agens 103–105
- Agensgeschehen 104
- Aktualisierung 2–4
 - Fahrbahn 57, 60–72, 101–103, 157–159
 - Fahrzeug 23, 88–102, 159–161
- Akzeptionsbereich
 - Ablehnungsbereich
- Änderungsdetektion 11, 19
- Anpassung →Aktualisierung
- Ausprägung 3, 23–24, 55–57, 85, 101, 111
- Azimutwinkel
 - der Einfallrichtung des Lichtes 38, 46–47, 51–54
 - einer 3D-Richtung im KKS 47–48, 122–123, 129
- Ballungsanalyse 12, 25, 28–29, 71–73, 115
- Beschreibung
 - begriffliche 1, 4, 6, 22, 71, 103–107, 111–112
 - geometrische 1–7
- Bewegungs
 - detektion 9–14, 19–20, 25–29
 - feld 13
 - modell 4–5, 17, 23, 68, 89, 159
 - parameter 9–13, 16–17, 39
 - verben 56 103–107
- Bewertungsfunktion 89
- Bezugsobjekt 104
- Bild
 - aufnahme 2–3, 18–19, 43, 59
 - bereich 3, 10–14, 19, 35, 66
 - bereichshinweise 3, 5, 25
 - daten →Durlacher-Tor-Szene,
→Ettlinger-Tor-Szene,
→Frankfurter Kreuzung,
→Tankstellenszene
 - distanzvektor 20
 - ebene 35–37, 40, 49, 117–119, 125
 - folge →Durlacher-Tor-Szene,
→Ettlinger-Tor-Szene,
→Frankfurter Kreuzung,
→Tankstellenszene
 - folgenauswertung 1–7,
→Interpretationszyklus
 - gerade 118, 120–123, 135
 - gradient →Ableitungsoperatoren
 - kantensegment →Kantensegment
 - kontrast →Kontrast
 - koordinaten
 - von Geraden 31, 118–119, 135–136
 - von Punkten 117–118, →Fluchtpunkte
 - system 36
 - mitte →Hauptpunkt
 - punktverhältnis 48, 129
 - sphäre 38, 119, 126, 137
 - strom 17
 - struktur 30
 - zentrum 47
 - ursprungsgerade 45–54, 122–123, 128–134
- Brennweite 37–50, 117–119, 124–128, 134

- Bus
 -modell 73-76
 -modellanpassung 31-33, 91-92
 -trajektorie 103-104
- Datenassoziation 57, 63, 80-84, 135-154, 156, 158
- datengetrieben 3-5, 12, 25, 58
- Datensegment → Kantensegment
- Detektion 1, 4-5, 7, 9-15, 23-30, 58, 71-73, 91, 94, 109, 115
- Differenz
 -bilder 9-11, 14, 19
 -winkel 83, 146
- Diskurs-Repräsentationstheorie 113
- Distanz → Abstand
- Drehung 87, 140, 145
- Durlacher-Tor-Szene 14, 31, 96-100
- Ebenennormale 125, 133-134
- Eigenvektor 122, 135
- Eigenwert 31, 122, 135-139
- Einheits
 -eigenvektor 122
 -kugel 38, 40-42
- Episode 103
- Ettlinger-Tor-Szene 2, 27-33, 40, 48-49, 70, 72, 80-84, 90-95, 115, 148-154
- Fahrbahn
 -ebene 5-6, 16, 32, 35-39, 45-54, 60-63, 125, 133, 134
 -geschehen 104, 106-107
 -markierung 5, 50, 60-61, 65-66, 68, 137
 -modell
 geometrisches 22, 60-61, 65-66, 70-71
 begriffliches 71, 104, 106-107
 -modellanpassung
 Initialisierung 35-37, 45-51, 60, 62-63
 an Kantensegmente 58-72, 156-159
 an Kantenelemente 101-102
- Fahrspuren, 18, 27, 56, 58, 63, 71-73
- Fahrzeug
 -frontleuchten 75-77
 -zeugmodell
 2D 14
 3D 15-16, 56, 73-77, 112-113
 -pulks 112
 -typen 56
- Faltung 20, 86-91
- Fehler
 -fortpflanzung 138
 -rechnung 80-84, 135-154, 156-159
 -untersuchung 38-45, 120-123
- Ferngerade 118
- Fernpunkt 118-119
- Flächen
 -anpassung 85-102
 -modell → 3D-Fahrzeugmodell
- Fluchtlinie 42, 125
- Fluchtpunkt 5, 38-54, 119-134
 -definition 119-120
 -detektion 38-40
 -schätzung 120-123
 -übersichtstabelle 134
- Fluß → optischer Fluß
- Frankfurter Kreuzung 58-61, 70, 73, 96-98
- Fußgänger 11, 15, 17-18, 113
- Gauß-Newton-Verfahren 17, 89
- Gaußfunktion 85-88, 91, 94-95, 160, auch → Faltung
- Geometrie 36, 38, 117, 119
 Kamera- 34-37
 projektive 38, 117-119
- Geraden 118-119, 120-134
 -orientierung 133
 -schar 42, 48, 120
 -segment → Kantensegment
- Gesamtsystem 4-7, 111
- Geschehen 103-107, 112-113
- Geschichte 103
- Geschwindigkeit
 2D → optischer Fluß
 3D 68, 88, 144, 159

- Gradient →Abteilungsoperatoren,
→Faltung
- Gradienten
-betragsschwelle 30
-norm 85-88
-richtung 30
- Grauwert →Bild
- Grauwertvariation 27, 58
- Großkreise 38-40
- Güteschwelle 27
- Halbildauswertung 29-24, 28, 92, 100
- Haltestreifen 50
- Handlungsmöglichkeit 105
- Hauptfluchtpunkt 39, 41-48, 62, 124
- Hauptpunkt 97, 41-43, 122, 129
- Heuristiken 4-6, 14-15, 25, 27, 79
- Hindernisdetektion 14-15
- Hintergrund 4-6, 10, 12-14, 22, 30
- Hintergrundbild 10-11
- Horizont 39, 42, 125, 134
- Hough-Transformation 15, 38-39
- Hyperellipsoid 82, 139-140
- Initialisierung 2-5
Fahrzeug 25-30, 58-60, 71-73, 94-95
Fahrbahn 35-50, 60-63, 101
- Interpretationsbaum 15, 66, 79
- Interpretationszyklus 2-4,
nach [Koller 92] 24
Fahrzeug
 Initialisierung 26
 Kantensegmente 56
 Flächenanpassung 85
- Fahrbahn
 Initialisierung 57
 Kantensegmente 57
 Flächenanpassung 101
- Gesamtsystem 111
- Jakobimatrix 89, 144, 155-160
- Kalibrierkörper 43
- Kalibrierplatte 37, 39, 41-43
- Kalibrierung 35-50
- Kalibrierungsparameter 46-47
- Kalmanfilter 10, 13-14, 16-17, 21, 23,
31, 68, 88-102, 157-161
- Kamera
-aufnahme →Bild
-koordinatensystem 35-37
-parameter 35-37, 45-47
 auch →Kalibrierung
- Kanten
-element 30-31, 135-139
-segment 15-16, 23, 25-26, 31, 33,
60-66, 79-84, 135-154, 156-159
-segmentextraktion 25-26, 90, 85
-segmentzuordnung →Abstand
- Katasterplan 32, 35, 55
- Kippwinkel 39
- KKS →Kamerakoordinatensystem
- Klassifikation 9-10, 14, 21, 112
- Kontrast 29-30, 75-76, 90-95
- Koordinatensystem 35-37
- Kovarianzmatrix
 Fahrbahn 68, 156-157
 Fahrzeug 88-89, 159-161
 Kantenelement 137, 144
 Kantensegment 80-84, 135-139, 144
- Kreuzung →Fahrbahn
- Lageschätzung →Initialisierung,
→Aktualisierung,
→Prädiktion
- Lampen 75-77
- Lichteinfallrichtung 5, 10, 21, 35, 38,
45-46, 51-54
- Lichtstrahlen 51-52
- Mahalanobisdistanz →Abstand
- MAP-Schätzung 23, 88-89, 157-158
- Markierungsplan →Katasterplan
- Mehrfachzuordnung 66, 156-159
- Meßfunktion 65, 88-89, 144, 157-161
- Mittelpunkt-Repräsentation 31,
80-88, 143-144, 156-161
- Mittelstreifen 63-65
- Modell →Fahrbahnmodell,
→Fahrzeugmodell
- modellbasiert 2-5, 15, 20, 23, 31, 55-58
- Modellgradientennorm 87-88

- Modellkantensegment
 Fahrbahn 63-73, 156-159
 Fahrzeug 15-16, 79-89, 23, 135-154,
 159-161
 Modellverfeinerung 55-56, 60-66, 70-
 71, 73-77, 109, 112
- Neigwinkel 39
 Nibelungenplatz →Frankfurter Kreuzung
- Normalenvektor der Fahrbahnebene
 37, 42, 50, 52-53, 119, 125,
 127, 130-131, 133-134
- Objektbildkandidat 9-14, 19-20, 23-
 31, 58, 71-73, 91, 94-95
 Objektgeschehen 104
 Objektkandidat 23-24, 55-56, 71-73
 optischer Fluß 11-14, 19-20, 25-30,
 58, 60, 71-73, 112
- parallele Geraden →Geradenschar
 Parametereinstellung 25, 58, 79, 89-
 90, 110, 115
- Patiens 104
 Pixelkoordinatensystem 36
 Pixelseitenverhältnis 37, 41-43
 Polarwinkel →Azimutwinkel
 Polyedermodell →Fahrzeugmodell
 Prädiktion 2-4
 Fahrbahn 68, 159
 Fahrzeug 23, 88-90, 159
 begrifflich 105-106
 Projektion 2-4, 23-24, 32, 35-38
 Projektionsebene 38, 40, 118-119
 projektiv 38, 49, 117-119
 Prozeßrauschen 90
- Rotations
 -freiheitsgrad 37-39
 -matrix 37, 87, 134
 -parameter 37-39, 41-49, 127, 134
 -unabhängigkeit 145
 Rückprojektion 2-4, 23, 37, 71-73, 133
- Schar →Geradenschar
 Schattenwurf 10, 21, 51-54
- Schätzverfahren →Aktualisierung,
 →Prädiktion, →Kalibrierung,
 →Lichteinfallrichtung
- Schnittwinkel
 von Straßenzügen 37, 127-128, 133-
 134
 von Kantensegmenten 80-84, 143-154
 Schwellwertentscheidung 9-11, 30-31,
 58, 79, 115, 154
 Schwenkwinkel 39
 Segmentierungsverfahren
 →Ballungsanalyse
- Sehstrahl 118-119
 Situationsschema 105, 113
 Sphäre →Bildsphäre
 Startkovarianzmatrix 90
 Startwert →Initialisierung
 Straße →Fahrbahn
 Straßenbahn 22, 112-113
 Straßenkreuzung
 →Durlacher-Tor-Szene,
 →Ettlinger-Tor-Szene,
 →Frankfurter Kreuzung
- Straßenschild 31
 Szenen
 -bereich 3-4, 23-24
 -bereichshinweise 3
 -skizze 4, 16, 88
- Tank
 -platz 107
 -spur 107
 -stellengeschehen 103, 107, 113
 -stellenszene 103, 107
- Trajektorien
 Durlacher-Tor-Szene 100
 Ettlinger-Tor-Szene 32, 105
 Frankfurter Kreuzung 98
 Tankstellenszene 107
 Translationsparameter 37-38, 41-47, 50
- Varianz →Kovarianz
 Vektorfeld 11-12, 25-27
 Verbalphrasen 103
 Verben 103
 Verdeckungen 10-11, 31, 79, 136

- Verfolgung → Aktualisierung,
→ Prädiktion
- Verkehrsbetriebe 73-74
- Verkehrsgeschehen 1-2, 103-107
- Verkehrsschilder 35, 63, 113
- Videoaufnahme → Bildaufnahme
- Vollbild → Halbbild
- Weltkoordinatensystem 35-37
- Winkelbereich 46
- Winkeldifferenz → Schnittwinkel
- Winkelgeschwindigkeit 88, 144, 159
- WKS → Weltkoordinatensystem
- Zählverfahren 13, 16, 38-39
- Zeilenverschränkung 23-24, 90-95, 100
- Zentralprojektion 36, 118
- Zuordnung → Abstand
- Zusicherungsgrad 104
- Zustandsaktualisierung
→ Aktualisierung
- Zustandsbeschreibung 104
- Zustandsübergangsfunktion
→ Prädiktion
- Zustandsvektor
- Fahrbahn 60-61
- Fahrzeug 88

DISKI – Dissertationen zur Künstlichen Intelligenz

Band 1: Norbert Reithinger (Universität Saarbrücken): Eine parallele Architektur zur inkrementellen Generierung multimodaler Dialogbeiträge
ISBN 3-929037-01-7

Band 2: Claus Müller (Universität Karlsruhe): Verwendung von Bildauswertungsmethoden zur Erkennung und Lagebestimmung von generischen polyedrischen Objekten im Raum
ISBN 3-929037-02-5

Band 3: Andreas Günter (Universität Hamburg): Flexible Kontrolle in Expertensystemen zur Planung und Konfigurierung in technischen Domänen
ISBN 3-929037-03-3

Band 4: Ingo Syska (Universität Hamburg): Modulare Problemlösungsarchitekturen für Konstruktionssysteme
ISBN 3-929037-04-1

Band 5: Wolfgang Tank (TU Berlin): Modellierung von Expertise über Konfigurierungsaufgaben
ISBN 3-929037-05-X

Band 6: Günter Wöhlke (Universität Karlsruhe): Wissensbasierte Greifplanung für Mehrfinger-Roboterhände
ISBN 3-929037-06-8

Band 7: Knut Möller (Universität Bonn): Adaptive Roboterkontrolle mit konnektionistischen Systemen
ISBN 3-929037-07-6

Band 8: Andreas Strasser (TU München): Generierung domänenspezifischer Wissensrepräsentationssysteme und Transformation von Wissensbasen mit einer Anwendung in der Rechtsinformatik
ISBN 3-929037-08-4

Band 9: Longin Latecki (Universität Hamburg): Digitale und Allgemeine Topologie in der bildhaften Wissensrepräsentation
ISBN 3-929037-09-2

Band 10: Jutta Eusterbrock (Universität Stuttgart): Wissensbasierte Verfahren zur Synthese mathematischer Beweise: Eine kombinatorische Anwendung
ISBN 3-929037-10-6

Band 11: Andreas Hörmann (Universität Karlsruhe): Begleitende Montageablaufplanung für ein sensorgestütztes Zweiarm-Manipulatorsystem
ISBN 3-929037-11-4

Band 12: Franz Kummert (Universität Erlangen-Nürnberg): Flexible Steuerung eines sprachverstehenden Systems mit homogener Wissensbasis
ISBN 3-929037-12-2

Band 13: Dieter Koller (Universität Karlsruhe): Detektion, Verfolgung und Klassifikation bewegter Objekte in monokularen Bildfolgen am Beispiel von Straßenverkehrsszenen
ISBN 3-929037-13-0

Band 14: Thomas Martinetz (TU München): Selbstorganisierende neuronale Netzwerkmodelle zur Bewegungssteuerung
ISBN 3-929037-14-9

Band 15: Roman Cunis (Universität Hamburg): Das 3-stufige Frame-Repräsentationsschema – eine mehrdimensional modulare Basis für die Entwicklung von Expertensystemkernen
ISBN 3-929037-15-7

Band 16: Byoung-Tak Zhang (Universität Bonn): Lernen durch Genetisch-Neuronale Evolution: Aktive Anpassung an unbekannte Umgebungen mit selbstentwickelten parallelen Netzwerken
ISBN 3-929037-16-5

Band 17: Kai Zercher (TU München): Wissensintensives Lernen für zeitkritische technische Diagnoseaufgaben
ISBN 3-929037-17-3

Band 18: Gerd Neugebauer (TH Darmstadt): Pragmatische Programmsynthese
ISBN 3-929037-18-1

Band 19: Manfred Jeusfeld (Universität Passau): Änderungskontrolle in deduktiven Datenbanken
ISBN 3-929037-19-X

Band 20: Adelinde Uhrmacher (Universität Koblenz-Landau): EMSY – Ein Modellierungskonzept für ökologische und biologische Systeme unter besonderer Berücksichtigung ihrer dynamischen Veränderung
ISBN 3-929037-20-3

Band 21: Ulrich Aßmann (Universität Bonn): Parallele Modelle für Deduktionssysteme
ISBN 3-929037-21-1

Band 22: Tilo Messer (TU München): Wissensbasierte Synthese von Bildanalyseprogrammen
ISBN 3-929037-22-X

Band 23: Klaus-Dieter Althoff (Universität Kaiserslautern): Eine fallbasierte Lernkomponente als integraler Bestandteil der MOLTKE-Werkbank zur Diagnose technischer Systeme
ISBN 3-929037-23-8

Band 24: Klaus Obermayer (TU München): Adaptive Neuronale Netze und ihre Anwendung als Modelle der Entwicklung kortikaler Karten
ISBN 3-929037-24-6

Band 25: Wolfgang Ertel (TU München): Parallele Suche mit randomisiertem Wettbewerb in Inferenzsystemen
ISBN 3-929037-25-4

Band 26: Klaus Fischer (Universität Saarbrücken): Verteiltes und kooperatives Planen in einer flexiblen Fertigungsumgebung
ISBN 3-929037-26-2

Band 27: Josef Pauli (TU München): Erklärungs-basiertes Computer-Sehen von Bildfolgen
ISBN 3-929037-27-0

Band 28: Reinhard Prechtel (Universität Erlangen-Nürnberg): Erklärungen für komplexe Wissensbasen. ISBN 3-929037-28-9

Band 29: Christoph Lehner (Universität München): Grammatikentwicklung mit Constraint Logik Programmierung. Implementierung einer Grammatik für das Deutsche mit PROLOG III
ISBN 3-929037-29-7

Band 30: Petra Ludewig (Universität Osnabrück): Inkrementelle wörterbuchbasierte Wortschatzerweiterungen in sprachverarbeitenden Systemen. Entwurf einer konstruktiven Lexikonkonzeption
ISBN 3-929037-30-0

Band 31: Wolfgang Eckstein (TU München): Die Bildanalyseprache TRIAS
ISBN 3-929037-31-9

Band 32: Claudia Sommer (Universität Erlangen-Nürnberg): MoKon – Ein Ansatz zur Wissensbasierten Konfiguration von Variantenerzeugnissen
ISBN 3-929037-32-7

Band 33: Yong Cao (Universität Hamburg): Zur Darstellung und Verarbeitung von Wissen über Himmelsrichtungen
ISBN 3-929037-33-5

Band 34: Michael Spreng (Universität Karlsruhe): Situationsanalyse bei Kontakten während der Ausführung von Roboterbewegungen in unsicheren Umgebungen
ISBN 3-929037-34-3

Band 35: Hans-Christian Brüning (TU München): Erweiterung der Wissensbasierten CAD-Konstruktion um Restriktionsnetztechniken
ISBN 3-929037-35-1

Band 36: Haibin Liu (Universität Erlangen-Nürnberg): Oberflächenbasierte Segmentierung von Tiefenbildern
ISBN 3-929037-36-X

Band 37: Jürgen Sauer (Universität Oldenburg): Wissensbasiertes Lösen von Ablaufplanungsproblemen durch explizite Heuristiken
ISBN 3-929037-37-8

Band 38: Axel Köhne (Universität Karlsruhe): Integration von Aktionsplanung und Konfigurierung
ISBN 3-929037-38-6

Band 39: Karlhorst Klotz (TU München): Eine mehrschichtige Architektur zur Fehlerdiagnose und Fehlerbehebung bei der Entwicklung logischer Programme
ISBN 3-929037-39-4

Band 40: Barbara Hemforth (Universität Bochum): Kognitives Parsing: Repräsentation und Verarbeitung sprachlichen Wissens
ISBN 3-929037-40-8

Band 41: Michael Ley (Universität Trier): Ein Datenbankkern zur Speicherung variabel strukturierter Feature-Terme. Implementierungstechniken
ISBN 3-929037-41-6

Band 42: Günther Specht (TU München): Source-to-Source Transformationen zur Erklärung des Programmverhaltens bei deduktiven Datenbanken
ISBN 3-929037-42-4

Band 43: Josef Ingenerf (RWTH Aachen): Benutzeranpaßbare semantische Sprachanalyse und Begriffsrepräsentation für die medizinische Dokumentation
ISBN 3-929037-43-2

Band 44: Thomas Bayer (Universität Erlangen-Nürnberg): Ein modellgestütztes Analysesystem zum Bildverstehen strukturierter Dokumente
ISBN 3-929037-44-0

Band 45: Volker Steinhage (Universität Bonn): Verdeckungen und spezielle Sichten bei der Polyederrekonstruktion
ISBN 3-929037-45-9

Band 46: Gerhard Werling (Universität Karlsruhe): Produktorientierte automatische Planung von Prüfoperationen bei der robotergestützten Montage
ISBN 3-929037-46-7

Band 47: Albert Maier (Universität Trier): Einbettung von Konzepthierarchien in ein Deduktives Datenbanksystem
ISBN 3-929037-47-5

Band 48: Frank Maurer (Universität Kaiserslautern): Hypermediabasiertes Knowledge Engineering für verteilte wissensbasierte Systeme
ISBN 3-929037-48-3

Band 49: Lothar Simon (Universität Erlangen-Nürnberg): Dynamische, situationsbezogene Hypertext-Handbücher für komplexe Tätigkeiten
ISBN 3-929037-49-1

Band 50: Marion Mast (Universität Erlangen-Nürnberg): Ein Dialogmodul für ein Spracherkennungs- und Dialogsystem
ISBN 3-929037-50-5

Band 51: Joachim Posegga (Universität Karlsruhe): Deduktion mit Shannongraphen für Prädikatenlogik erster Stufe
ISBN 3-929037-51-3

Band 52: Stephan Mehl (Universität Koblenz-Landau): Dynamische semantische Netze – Zur Kontextabhängigkeit von Wortbedeutungen
ISBN 3-929037-52-1

Band 53: Jürgen Angele (Universität Karlsruhe): Operationalisierung des Modells der Expertise mit KARL
ISBN 3-929037-53-X

Band 54: Martin Schröder (Universität Hamburg): Erwartungsgestützte Analyse medizinischer Befundungstexte. Ein wissensbasiertes Modell zur Sprachverarbeitung
ISBN 3-929037-54-8

Band 55: Jochen Heinsohn (Universität Saarbrücken): ALCP – Ein hybrider Ansatz zur Modellierung von Unsicherheit in terminologischen Logiken
ISBN 3-929037-55-6

Band 56: Andrea Neufeld (Universität Karlsruhe): Validierung konzeptueller Schemata
ISBN 3-929037-56-4

Band 57: Werner Karbach (Universität Bielefeld): MODEL-K: Modellierung und Operationalisierung von Selbsteinschätzung und -steuerung durch Reflexion und Metawissen
ISBN 3-929037-57-2

Band 58: Andreas Winklhofer (TU München): Zeitrepräsentation und merkmalsgesteuerte Suche zur Terminplanung
ISBN 3-929037-58-0

Band 59: Jun Zhao (Universität Bremen): Qualitative Analyse im Rahmen qualitativen und modellbasierten Schließens
ISBN 3-929037-59-9

- Band 60:** Susanne Neubert (Universität Karlsruhe): Modellkonstruktion in MIKE. Methoden und Werkzeuge
ISBN 3-929037-60-2
- Band 61:** Karsten Berns (Universität Karlsruhe): Steuerungsansätze auf der Basis Neuronaler Netze für sechsbeinige Laufmaschinen
ISBN 3-929037-61-0
- Band 62:** Thorsten von Stein (Universität Hamburg): Wissensbasierte Analyse medizinischer Bilder – das Biotop-Verfahren
ISBN 3-929037-62-9
- Band 63:** Michael Otte (Universität Karlsruhe): Extraktion von linienförmigen Merkmalen und Ermittlung des optischen Flusses mit seinen Ableitungen aus Voll- und Halbbitsfolgen. ISBN 3-929037-63-7
- Band 64:** Klaus Grebner (Universität Bielefeld): Wissensbasierte Entwicklungsumgebung für Bildanalyse-systeme aus dem industriellen Bereich
ISBN 3-929037-64-5
- Band 65:** Jana Köhler (Universität Saarbrücken): Wiederverwendung von Plänen in deduktiven Planungssystemen. ISBN 3-929037-65-3
- Band 66:** Christoph Klauck (Universität Bremen): Eine Graphgrammatik zur Repräsentation und Erkennung von Features in CAD/CAM
ISBN 3-929037-66-1
- Band 67:** Markus A. Thies (Universität Saarbrücken): Planbasierte Hilfeverfahren für direkt-manipulative Systeme. Erkennung, Vervollständigung und Visualisierung von Interaktionsplänen
ISBN 3-929037-67-X
- Band 68:** Jürgen Kreuziger (Universität Karlsruhe): Eine Architektur zur Anwendung symbolischer Lernverfahren in der Robotik
ISBN 3-929037-68-8
- Band 69:** Bernd Müller (Universität Oldenburg): PPO – Eine objektorientierte Prolog-Erweiterung zur Entwicklung wissensbasierter Anwendungssysteme
ISBN 3-929037-69-6
- Band 70:** Ute Schmid (Technische Universität Berlin): Erwerb rekursiver Programmier-techniken als Induktion von Konzepten und Regeln
ISBN 3-929037-70-X
- Band 71:** Jörg R. J. Schirra (Universität Saarbrücken): Bildbeschreibung als Verbindung von visuellem und sprachlichem Raum. Eine interdisziplinäre Untersuchung von Bildvorstellungen in einem Höremodell
ISBN 3-929037-71-8
- Band 72:** Volker Gengenbach (Universität Karlsruhe): Einsatz von Rückkopplungen in der Bildauswertung bei einem Hand-Auge-System zur automatischen Demontage
ISBN 3-929037-72-6
- Band 73:** Jan Schloen (Universität Karlsruhe): Wissensbasierte Bewegungsausführung für die Montageautomatisierung mit Industrierobotern
ISBN 3-929037-73-4
- Band 74:** Friedrich Mädlar (Universität Osnabrück): Problemzerlegung durch Nadelöhrmengen. Ein modellbasierter Ansatz zur Akquisition von Kontrollwissen für Planungssysteme
ISBN 3-929037-74-2

Band 75: Gerhard Weiß (TU München): Distributed Machine Learning
ISBN 3-929037-75-0

Band 76: Michael Thielscher (TH Darmstadt): Automatisiertes Schließen über
Kausalbeziehungen mit SLDENF-Resolution
ISBN 3-929037-76-9

Band 77: Jianwei Zhang (Universität Karlsruhe): Ein integriertes Verfahren zur effizienten
Planung und Ausführung von Roboterbewegungen in unscharfen Umgebungen
ISBN 3-929037-77-7

Band 78: Enno Littmann (Universität Bielefeld): Strukturierung Neuronaler Netzwerke zwischen
Biologie und Anwendung
ISBN 3-929037-78-5

Band 79: Manfred Meyer (Universität Kaiserslautern): Finite Domain Constraints: Declarativity
meets Efficiency – Theory meets Application
ISBN 3-929037-79-3

Band 80: Thomas Kuhn (Universität Erlangen-Nürnberg): Die Erkennungsphase in einem
Dialogsystem
ISBN 3-929037-80-7

Band 81: Knut Hinkelmann: Transformation von Hornklausel-Wissensbasen: Verarbeitung
gleichen Wissens durch verschiedene Inferenzen
ISBN 3-929037-81-5

Band 82: Nils Lenke: Paraphrasen. Lösungen für antizipierte Leserprobleme bei der
automatischen Textgenerierung
ISBN 3-929037-82-3

Band 83: Dominik Henrich: Lastverteilung für feinkörnig parallelisiertes Branch-and-Bound
ISBN 3-929037-83-1

Band 84: Dieter Landes: Die Entwurfsphase in MIKE. Methode und Beschreibungssprache
ISBN 3-929037-84-X

Band 85: Thomas Pirlein: Wiederverwendung von Commonsense-Ontologien im Knowledge-
Engineering. Methoden und Werkzeuge
ISBN 3-929037-85-8

Band 86: Karsten Poeck: Konfigurierbare Problemlösungsmethoden am Beispiel der
Problemklassen Zuordnung und Diagnostik
ISBN 3-929037-86-6

Band 87: Wolfgang Melchert: Ein Ansatz zur Formalisierung der Echtzeit-Bildauswertung und
seine Nutzung zur automatischen Erzeugung lauffähiger Maschinenprogramme
ISBN 3-929037-87-4

Band 88: Henner Kollnig: Ermittlung von Verkehrsgeschehen durch Bildfolgenauswertung
ISBN 3-929037-88-2